nature portfolio

https://doi.org/10.1038/s41586-021-03570-8

Supplementary information

COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets

In the format provided by the authors and unedited

Supplementary Information Guide

COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets

Toni M. Delorey[1*], Carly G. K. Ziegler[2,3,4,5,6,7*], Graham Heimberg[1*], Rachelly Normand[2,8,9,10,11*], Yiming Yang[1,8*], Åsa Segerstolpe[1*], Domenic Abbondanza[1,2*], Stephen J. Fleming[12,13*], Ayshwarya Subramanian[1*], Daniel T. Montoro[2*], Karthik A. Jagadeesh[1*], Kushal K. Dey[14*], Pritha Sen[2,8,15,16*], Michal Slyper[1*], Yered H. Pita-Juárez[2,10,17,18,19*], Devan Phillips[1*], Jana Biermann[20,21*], Zohar Bloom-Ackermann[22], Nick Barkas[12], Andrea Ganna[23,24], James Gomez[22], Johannes C. Melms[20,21], Igor Katsyv[25], Erica Normandin[2,10], Pourya Naderi[10,17,18], Yury V. Popov[10,26,27], Siddharth S. Raju[2,28,29], Sebastian Niezen[10,26,27], Linus T.-Y. Tsai[2,10,26,30,31], Katherine J. Siddle[2,32], Malika Sud[1], Victoria M. Tran[22], Shamsudheen K. Vellarikkal[2,33], Yiping Wang[20,21], Liat Amir-Zilberstein[1], Deepak S. Atri[2,33], Joseph Beechem[34], Olga R. Brook[35], Jonathan Chen[2,36], Prajan Divakar[34], Phylicia Dorceus[1], Jesse M. Engreitz[2,37], Adam Essene[26,30,31], Donna M. Fitzgerald[38], Robin Fropf[34], Steven Gazal[39], Joshua Gould[12], John Grzyb[40], Tyler Harvey[1], Jonathan Hecht[10,17], Tyler Hether[34], Judit Jané-Valbuena[1], Michael Leney-Greene[2], Hui Ma[1,8], Cristin McCabe[1], Daniel E. McLoughlin[38], Eric M. Miller[34], Christoph Muus[2,41], Mari Niemi[23], Robert Padera[40,42,43], Liuliu Pan[34], Deepti Pant[26,30,31], Carmel Pe'er[1], Jenna Pfiffner-Borges[2], Christopher J. Pinto[16,38], Jacob Plaisted[40], Jason Reeves[34], Marty Ross[34], Melissa Rudy[2], Erroll H. Rueckert[34], Michelle Siciliano[40], Alexander Sturm[22], Ellen Todres[1], Avinash Waghray[44,45], Sarah Warren[34], Shuting Zhang[22], Daniel R. Zollinger[34], Lisa Cosimi[46], Rajat M. Gupta[2,33], Nir Hacohen[2,9,47], Hanina Hibshoosh[25], Winston Hide[10,17,18,19], Alkes L. Price[14], Jayaraj Rajagopal[38], Purushothama Rao Tata[48], Stefan Riedel[10,17], Gyongyi Szabo[2,10,26], Timothy L. Tickle[1,12], Patrick T. Ellinor[49†], Deborah Hung[22,50,51†], Pardis C. Sabeti[2,32,52,53,54†], Richard Novak[55†], Robert Rogers[26,56†], Donald E. Ingber[41,55,57†], Z. Gordon Jiang[10,26,27†], Dejan Juric[16,38†], Mehrtash Babadi[12,13†], Samouil L. Farhi[1,2†], Benjamin Izar[20,21,58,59†], James R. Stone[36†], Ioannis S. Vlachos[2,10,17,18,19†], Isaac H. Solomon[40†], Orr Ashenberg[1†], Caroline B.M. Porter[1†], Bo Li[1,8,16†], Alex K. Shalek[2,3,4,5,6,7,10,44,60,61,62†], Alexandra-Chloé Villani[2,8,9,16†], Orit Rozenblatt-Rosen[1,63†], Aviv Regev[1,5,53,63†]

1 Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA, USA

2 Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

3 Program in Health Sciences & Technology, Harvard Medical School & Massachusetts Institute of Technology, Boston, MA 02115, USA

4 Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

5 Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

6 Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA

7 Harvard Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA

8 Center for Immunology and Inflammatory Diseases, Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

9 Center for Cancer Research, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

10 Harvard Medical School, Boston, MA 02115, USA

11 Massachusetts Institute of Technology, Cambridge, MA 02139, USA

12 Data Sciences Platform, Broad Institute of MIT and Harvard, Cambridge, MA 02142

13 Precision Cardiology Laboratory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

14 Department of Epidemiology, Harvard School of Public Health

15 Division of Infectious Diseases, Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

16 Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

17 Department of Pathology, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA

18 Harvard Medical School Initiative for RNA Medicine, Boston, MA 02115, USA

19 Cancer Research Institute, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA

20 Department of Medicine, Division of Hematology/Oncology, Columbia University Irving Medical Center, New York, NY

21 Columbia Center for Translational Immunology, New York, NY

22 Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

23 Institute for Molecular Medicine Finland, Helsinki, Finland

24 Analytical & Translational Genetics Unit, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA

25 Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY

26 Department of Medicine, Beth Israel Deaconess Medical Center, MA 02115, USA

27 Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

28 Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

29 FAS Center for Systems Biology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

30 Division of Endocrinology, Diabetes, and Metabolism, Beth Israel Deaconess Medical Center, Boston, MA 02115

31 Boston Nutrition and Obesity Research Center Functional Genomics and Bioinformatics Core Boston, MA 02115, USA

32 Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

33 Divisions of Cardiovascular Medicine and Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

34 NanoString Technologies Inc., Seattle, WA 98109, USA

35 Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

36 Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA

37 Department of Genetics and BASE Initiative, Stanford University School of Medicine

38 Massachusetts General Hospital Cancer Center, Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

39 Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

40 Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115

41 John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138

42 Harvard-MIT Division of Health Sciences and Technology, Cambridge MA

43 Department of Pathology, Harvard Medical School, Boston, MA 02115, USA

44 Harvard Stem Cell Institute, Cambridge, MA, USA

45 Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

46 Infectious Diseases Division, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

47 Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

48 Department of Cell Biology, Duke University School of Medicine, Durham, NC, USA

49 Cardiovascular Disease Initiative, The Broad Institute of MIT and Harvard, Cambridge, MA

50 Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

51 Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, MA 02114, USA

52 Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

53 Howard Hughes Medical Institute, Chevy Chase, MD, USA

54 Massachusetts Consortium on Pathogen Readiness, Boston, MA, USA

55 Wyss Institute for Biologically Inspired Engineering, Harvard University

56 Massachusetts General Hospital, MA 02114, USA

57 Vascular Biology Program and Department of Surgery, Boston Children's Hospital, Harvard Medical School, Boston, MA USA

58 Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY

59 Program for Mathematical Genomics, Columbia University Irving Medical Center, New York, NY

60 Program in Computational & Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

61 Program in Immunology, Harvard Medical School, Boston, MA 02115, USA

62 Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

63 Current address: Genentech, 1 DNA Way, South San Francisco, CA, USA

[*] these authors contributed equally

[†]these authors jointly supervised this work, correspondence to:
shalek@mit.edu, avillani@mgh.harvard.edu, ivlachos@bidmc.harvard.edu, orit.r.rosen@gmail.com, covid19-autopsy-northeast@broadinstitute.org, aviv.regev.sc@gmail.com

Table of Contents

METHODS 5

REFERENCES (METHODS) 47

SUPPLEMENTARY INFORMATION TABLES 52

Methods

Human donor samples

Samples in this study underwent IRB review and approval at the institutions where they were originally collected. Specifically, Dana-Farber Cancer Institute approved protocol 13-416, Partners/Massachusetts General Hospital (MGH) and Brigham and Women's Hospital (BWH) approved protocols: 2020P000804, 2020P000849, and 2015P002215; Beth Israel Deaconess Medical Center (BIDMC) approved protocols 2020P000406 and 2020P000418; New York Presbyterian Hospital/Columbia University Medical Center (NYP) approved protocols IRB-AAAT0785, IRB-AAAB2667, and IRB-AAAS7370. Secondary analysis of samples at the Broad Institute was covered under Massachusetts Institute of Technology (MIT) IRB protocols 1603505962 and 1612793224, or the NHSR (not-involving-human-subjects research) protocol ORSP-3635. No subject recruitment or ascertainment was performed as part of the Broad protocol. Donor identities were encoded at the hospitals prior to shipping to or sharing with the Broad Institute for sample processing or data analysis, respectively.

Tissue collection

We assembled a cross-body COVID-19 autopsy biobank of 420 autopsy specimens (Hospital sites A-C: MGH, BWH, and BIDMC), spanning 11 organs and 17 donors and used it to generate a single cell atlas of COVID-19 lung, kidney, liver and heart, and a lung spatial atlas from a subset of donors, including additional samples from hospital site D: NYP). Exclusion criteria included a post mortem interval > 24 hours and HIV infection. Donor ages ranged from 30-35 to > 89 years of age (**Extended Data Fig. 1a**, **Supplementary Information Table 1**). In addition, we included unpublished snRNA-seq data (M.S., A.R. and O.R.R., unpublished data) from three 50-59 year old healthy donor autopsies, two male and one female, collected from archived frozen samples obtained years prior to the COVID-

19 pandemic. We also included unpublished scRNA-seq data (M.S., A.R. and O.R.R., unpublished data) from nine healthy donors, ages 57 (F), 59 (F), 20 (M), 42 (F), 0.3 (M), 66 (F), 57 (F), 18 (F), 66 (F), and 46 (M)[25].

All patients at the Massachusetts General Hospital (MGH) who succumbed to SARS-CoV-2 infection, as confirmed by the qRT-PCR assays performed on nasopharyngeal swab specimens, were eligible for clinical autopsy upon consent by their healthcare proxy or next of kin. A subset of these patients were also enrolled in the MGH Rapid Autopsy Protocol if they had a history of known or suspected malignancy. Their clinical data and research specimens were collected in accordance with Dana Farber/Harvard Cancer Center Institutional Review Board-approved protocol 13-416. Autopsies at MGH were performed in a negative pressure isolation room by personnel equipped with powered air-purifying or N95 respirators. Organs were removed from the body *en bloc*, and subsequently dissected for individual organ examination, including weighing and photographing. Representative samples of each lung lobe, left ventricle of the heart, liver, kidney, and brain (for donor D6) were placed in collection tubes, which were then placed inside a cooler for immediate transport to the Broad Institute where all tissues were processed fresh.

All patients at Brigham and Women's Hospital (BWH) who succumbed to SARS-CoV-2 infection, confirmed by pre- or peri-mortem nasopharyngeal swab qRT-PCR testing, were eligible for clinical autopsy upon consent by their healthcare proxy or next of kin. Autopsies at Brigham and Women's Hospital were performed in a negative pressure isolation room by personnel equipped with powered air-purifying or N95 respirators. Clinical data and research specimens were collected in accordance with the Mass General Brigham Human Research Committee-approved protocol 2015P002215. Organs were removed from the body *en bloc*, and subsequently dissected for individual organ examination, including

weighing and photographing. Representative samples of each lung lobe, trachea, left ventricle of the heart, aorta, coronary arteries, liver, kidney, lymph nodes, and spleen, as well as nasal and oral swab/scrapings were placed in 25 mL of RPMI-1640 media with 25 mM HEPES and L-glutamine (ThermoFisher Scientific) + 10% heat inactivated FBS (ThermoFisher Scientific) in 50 mL falcon tubes (VWR International Ltd). Collection tubes containing tissue were then placed inside a cooler for immediate transport to the Broad Institute. Additional tissue specimens from lung and heart were fixed in 10% formalin, processed and paraffin embedded using standard protocols. Five μm-thick slides were prepared from the FFPE tissue blocks and transferred to the Broad Institute.

Autopsies at Beth Israel Deaconess Medical Center were performed on patients who expired in the intensive care unit with SARS-CoV-2 infection by following a minimally invasive approach. Consent for autopsy and research was obtained from the healthcare proxy or the next of kin. Ultrasound-guided biopsies were performed within three hours of death on a gurney in the hospital morgue. The body was not removed from the bag until staff was wearing standard precautions and an N95 mask. Biopsies were obtained through 13G coaxial guide with 14G core biopsy and 20 mm sample length. Multiple core biopsies were acquired from each organ by tilting the coaxial needle a few degrees in different directions[51]. Cores for snRNA-Seq and spatial investigations were flash-frozen and stored at -80˚C until assayed or embedded in paraffin, respectively.

Inclusion criteria for autopsies from COVID-19 donors cared for at New York Presbyterian Hospital/Columbia University Medical Center included real-time reverse transcription polymerase chain reaction (RT-PCR) confirmed infection, consent to perform rapid autopsy and post mortem intervals <10 hours. Appropriate consent was obtained from donors or the donors' next of kin. All procedures performed on donor samples were in accordance with the ethical standards of the IRB and

the Helsinki Declaration and its later amendments. Frozen control tissues were assessed by a pulmonary pathologist and represent "uninvolved" regions of biobanked tumor resections. Donor characteristics reflect the age, gender, and race representation of patients admitted to New York Presbyterian Hospital/Columbia University Medical Center with COVID-19. Control samples were selected to reflect median age distribution of COVID-19 cases included in the study and match the gender distribution.

Tissue processing

Tissues from deceased COVID-19 donors were processed immediately upon arrival in a Biosafety Level (BSL) 3 lab space. Lung was collected from all donors, and when possible, we also collected one or more specimens of kidney, spleen, trachea, peribronchial/subcarinal lymph node, muscle, and oral mucosa. Brain tissue was collected for the single donor who presented with neurological symptoms. Upon receiving specimens at the Broad, we dissected tissue from each organ and collected multiple biopsies from each in a BSL3 facility using a specimen processing pipeline (**Fig. 1a**) that included: (1) dissociating biopsies into single-cell suspensions followed by immediate scRNA-Seq profiling; (2) viably freezing biopsies for future scRNA-Seq; and (3) flash freezing biopsies for future snRNA-Seq, bulk RNA-Seq, and viral qPCR. While tissue from 17 COVID-19 donors was preserved in our biobank, we did not profile samples from D9 by sc/snRNA-seq. On the day that issue from D9 was allocated, specimens from two other donors were received simultaneously. To efficiently preserve tissue across all three allocations, samples from D9 were banked in larger sections but not dissected into biopsy sized specimens. Thus, D9 was analyzed only by spatial RNA profiling.

Tissues were washed prior to dissection in cold PBS and dissected in sterile Petri dishes. Grossly necrotic tissue areas were excluded. As phenotypes varied across organs, representative biopsies were

collected from each tissue sample. Biopsies (each approximately 0.5 cm^3) were cut and either directly dissociated for scRNA-Seq or frozen for future processing (each described below).

For processing details on tissues collected at Hospital site D (**Extended Data Fig.1a**), see Melms, *et al.* (companion manuscript).

Biopsy collection for snRNA-Seq and scRNA-Seq

For snRNA-Seq, three dry biopsies per tissue were placed in one cryovial and flash frozen on dry ice before being placed at -80°C for long term storage. Two to four flash frozen tubes were collected per tissue.

For scRNA-Seq of viably frozen tissue, tissue was viably frozen by placing 4 biopsies in 500 mL of CryoStor CS10 freezing media (STEMCELL Technologies). Tubes were inverted and then placed on ice for up to 30 minutes. Samples were then stored at -80°C for future processing.

For scRNA-Seq of fresh tissue (*i.e.*, samples processed on the day of collection), single biopsies were placed in a dissociation buffer and processed as described below.

Fresh lung tissue processing and cryopreserved tissue processing for scRNA-Seq

For each sample, 1 mL of lung dissociation medium (100 µg/mL of DNAse I, Roche; 1.25 mg/mL of Pronase, Roche; 9.2 µg/mL of Elastase, Worthington; 100 µg/mL of Dispase II, Roche; 1.5 mg/mL of Collagenase A, Roche; 100 µg/mL of Collagenase IV, Worthington) was aliquoted in a 15 mL Falcon tube (VWR International Ltd) and kept on ice until use. For each tissue, 2 wells of a 24 well plate were

prepared and contained 2 mL RPMI-1640 media with 25 mM HEPES and L-glutamine (ThermoFisher Scientific) + 10% heat inactivated FBS (ThermoFisher Scientific).

For viably frozen tissue, before sample thawing, tubes with lung dissociation media and the 24 well plate containing media were placed at 37ºC to warm. One biopsy piece was transferred to one well of a 24 well tissue culture plate (Corning) and moved back and forth in the liquid to wash. The biopsy was then placed in a second well containing media for 1 minute before moving to a 1.5 mL Eppendorf tube containing warmed lung dissociation media. While the tube was in a rack, sterile scissors were used to mince the tissue until most pieces were < 1 mm. The Eppendorf tubes containing minced tissues were then placed in a thermomixer (Eppendorf) and incubated for 25 minutes at 300 RPM, 37ºC. After 10-15 minutes, the sample was briefly triturated with a P1000 pipette, and returned to the thermomixer. After 25 minutes, 600 µL cold RPMI-1640 + 10% FBS + 1 mM EDTA (Invitrogen) was added to quench the digestion. Using a wide bore P1000 pipette tip (Rainin), the tissue was triturated~ 20 times. The dissociated tissue was then placed on ice. One 50 mL conical per tissue preparation was placed in a rack and on ice, and a 100 µm filter was prepared by wetting with 1 mL cold RPMI-1640 + 10% FBS + 1 mM EDTA. Using a wide-bore P1000 pipette tip, digested tissue was passed through the filter. The filter was washed with 4 mL cold RPMI-1640 + 10% FBS. Tubes were spun at 300g x 8 minutes at 4ºC. Supernatant was removed and pellet was resuspended in 2 mL ACK lysis buffer, placed on ice for 2-3 minutes. The lysis was quenched in 2 mL RPMI-1640 + 10% FBS, and the sample was spun at 300g x 8 minutes at 4ºC. After centrifugation, most of the supernatant was removed and the sample pellet was resuspended in ~1 mL of RPMI-1640 + 10% FBS. Cells were counted and immediately loaded on the 10x Chromium controller (10x Genomics) for partitioning single cells into droplets.

Flash-frozen tissue processing for snRNA-Seq

All sample handling steps were performed on ice. TST and ST buffers were prepared fresh as previously described[13,52]. A 2x stock of salt-Tris solution (ST buffer) containing 292 mM NaCl (ThermoFisher Scientific), 20 mM Tris-HCl pH 7.5 (ThermoFisher Scientific), 2 mM $CaCl_2$ (VWR International Ltd) and 42 mM $MgCl_2$ (Sigma Aldrich) in ultrapure water was made and used to prepare 1xST and TST. TST was then prepared with 1 mL of 2x ST buffer, 60 µL of 1% Tween-20 (Sigma Aldrich), 10 µL of 2% BSA (New England Biolabs), 930 µL of nuclease-free water and supplemented with 0.2 U/uL RNaseIN Plus (Promega) and 0.1U/uL Superasin (ThermoFisher Scientific). 1xST buffer was prepared by diluting 2x ST with ultrapure water (ThermoFisher Scientific) in a ratio of 1:1, and was supplemented with 0.1% RNase inh (Enzymatics). 1 mL of 1xST without RNase was also prepared for sample dilution prior to 10x chip loading. Single frozen biopsy pieces were kept on dry ice until immediately prior to dissociation. With clean forceps, a single frozen biopsy was placed into a gentleMACS C tube on ice (Miltenyi Biotec) containing 2 mL of TST buffer. gentleMACS C tubes were then placed on the gentleMACS Dissociator (Miltenyi Biotec) and tissue was homogenized by running the program "m_spleen_01" one to three times, until tissue was fully dissociated. A 40 µm filter (VWR International Ltd) was placed on a 50 mL falcon tube (VWR International Ltd) and pre-wet with 1 mL of 1xST. Homogenized tissue was then transferred to the 40 µm filter and washed with 2 mL of 1xST buffer containing RNase inhibitors. Samples were then centrifuged at 500g for 10 minutes at 4ºC. Sample supernatant was removed and the pellet was resuspended in 100 µL - 500 µL 1xST buffer (without RNase inhibitor). Nuclei were counted and immediately loaded on the 10x Chromium controller (10x Genomics) for single nucleus partitioning into droplets.

Single-cell and single-nucleus RNA-Seq

For each sample, 8,000-15,000 cells or nuclei were loaded in one channel of a Chromium Chip (10x genomics). 3' v3 chemistry was used to profile lung nuclei. 5'v1.1 chemistry was used to profile freshly dissociated lung cells, as well as cells obtained from cryopreserved lung tissue. 3' v3.1 chemistry was used to process all other tissues. cDNA and gene expression libraries were generated according to the manufacturer's instructions (10x Genomics). cDNA and gene expression library fragment sizes were assessed with a DNA High Sensitivity Bioanalyzer Chip (Agilent). cDNA and gene expression libraries were quantified using the Qubit dsDNA High Sensitivity assay kit (ThermoFisher Scientific). Gene expression libraries were multiplexed and sequenced on the Nextseq 500 (Illumina) with a 150 cycle kit and the following read structure: Read 1: 28 cycles, Read 2: 96 cycles, Index Read 1: 8 cycles, or on the Novaseq (S2 flow cell; Illumina) with a 100 cycle kit and the following read structure: Read 1: 28 cycles, Read 2: 91 cycles, Index Read 1: 8 cycles.

Bulk RNA-Seq of tissue for high *vs*. low/no viral load

With clean forceps, a single frozen biopsy was placed into a gentleMACS M tube (Miltenyi Biotec) containing 600 µL of 2x DNA/RNA shield (Zymo Research). gentleMACS M tubes were then placed on the gentleMACS Dissociator (Miltenyi Biotec) and tissue was homogenized by running the program "RNA_02". Samples were centrifuged at 2,000xg for 1 minute and moved to new tubes. As DNA/RNA shield inactivates SARS-CoV-2, the samples could then be moved out of the BSL3 lab space. RNA was extracted using the quick-RNA MagBead kit (Zymo Research). RNA was proteinase K treated according to manufacturer's instructions. For some donors (D13-D17), we had access to a single frozen lung biopsy. Bulk nuclei were isolated (as described above) and flash frozen in RLT + 1 % 1 BME. To isolate total RNA, we added ~ 600 mL of 2x DNA/RNA shield. After proteinase K treatment, one

volume of water was added to dilute the DNA/RNA shield. Total RNA was quantified via the Nanodrop (ThermoFisher Scientific).

cDNA and gene expression libraries were constructed with Smart-Seq2 chemistry and the Nextera XT DNA library prep kit (Illumina), as previously described[53,54]. DNA and gene expression library fragment sizes were assessed with a DNA High Sensitivity Bioanalyzer Chip (Agilent). cDNA and gene expression libraries were quantified using the Qubit dsDNA High Sensitivity assay kit (ThermoFisher Scientific). Gene expression libraries were multiplexed and sequenced on the Nextseq 500 (Illumina) with a 75 cycle kit and the following read structure: Read 1: 38 cycles, Read 2: 38cycles, Index Read 1: 8 cycles, Index Read 2: 8 cycles.

Differential gene expression analysis was performed on lung samples with high viral load (D7,D10,D11) *vs.* low/no viral load (D4, D6, D8, D15, D17) **(Supplementary Information Table 5).**

Bulk RNA-Seq of tissue for viral genome assembly

Bulk RNA sequencing libraries were created for lung samples across donors, one heart sample (D3) and one brain sample (D6; a deceased donor who had neurological symptoms) for viral genomic analysis. Briefly, 5 µL of each RNA sample underwent RNAse H-based ribosomal RNA depletion[55], then libraries were constructed using an Illumina TruSeq Stranded Total RNA Library Construction kit (20020596). This ligation-based library construction method utilized adapters (1.5 uM each) that contained a unique molecular identifier on the upstream of the i7[56]. Some libraries with lower viral loads were selected for targeted enrichment of SARS-CoV-2 sequences using a probe set designed to capture sequences from several respiratory pathogens, including SARS-CoV-2; this design was adapted from a published method[57]. Targeted enrichment was performed using a Twist Bioscience Target

Enrichment workflow (https://www.twistbioscience.com/sites/default/files/resources/2019-11/Protocol_NGS_HybridizationTE_31OCT19_Rev1.pdf) (**Supplementary Information Table 5**). All libraries were pooled and sequenced to a read depth of at least 2 million paired end reads on an Illumina MiSeq V3 and a NovaSeq SP.

To maximize the number of subjects that we obtained viral genomes from, multiplexed amplicon sequencing was also performed on select samples from which we had not assembled a complete viral genome using bulk RNA sequencing with targeted capture. The experimental pipeline[58] was based on the ARTIC consortium protocol, utilizing the V3 primer scheme[59].

Very few SARS-CoV-2 reads were detected in bulk RNA-seq for the heart (166 reads) and no reads were recovered from the brain sample (0 reads), even with targeted enrichment for viral reads. This result is consistent with the low viral loads by RT-qPCR in each case: 65 copies/μL in the heart sample, and 4 copies/μL in the brain sample, which is below the limit of detection. Because we were not able to assemble a genome, we compared the SARS-CoV-2 reads from the D3 heart sample to the consensus genome assembled from the lung of the same subject, but still could not reliably identify intrahost variants given such low coverage. As we performed RNA-seq from biopsy-sized pieces of tissue, it is possible that we performed bulk or snRNA-Seq on regions of heart, liver and kidney tissue which did not contain virus.

Viral quantification by RT-qPCR

For each bulk RNA sample, SARS-CoV-2 RNA was quantified by an RT-qPCR assay targeting the nucleocapsid protein (modified from the CDC N1 assay, as described previously[60]). The assay was performed on samples (diluted 1:3) in triplicate, along with negative controls and a synthetic DNA

standard curve. Reported quantifications, calculated from the standard curve, are mean values across triplicates, and are corrected by the dilution factor (**Supplementary Information Table 5**).

Sample processing for RNAscope® and DSP analysis

FFPE blocks were cut onto Superfrost Plus microscope slides at the BWH Histopathology Core and shipped to the Broad Institute to be stored at -80°C.

For the RNAscope® assay, staining was performed according to the manufacturer protocol using the RNAscope® Multiplex Fluorescent Reagent Kit v2 kit (ACD Bio, 323100) with the V-nCoV2019-S-C3 probe (ACD Bio, 848561-C3) and the TSA Plus Cyanine 5 fluorophore (Akoya Biosciences, NEL745001KT). For samples D13-17, serial sections were fluorescently stained with RNAscope Probes - V-nCoV2019-S –C3 d, *ACE2* and *TMPRSS2* (ACDBio) and GeoMx Nuclear Stain (NanoString Technologies, GMX-MORPH-NUC-12) in blue. Additionally, for the same samples, serial sections were DAB stained with 2 µg/mL of anti-SARS-CoV-2 spike glycoprotein antibody (Abcam, ab272504). Slides were scanned at 20x directly on the NanoString GeoMx instrument to facilitate alignment with subsequent WTA/CTA/Protein scans. For the subjects D13-17, a semi-quantitative (5-tier) viral abundance score (Viral Score) was used based on RNAScope SARS-CoV-2 staining to identify ROIs exhibiting diverse viral loads. The score assignment per AOI was performed by board-certified pathologists.

For the NanoString GeoMx DSP RNA assays, slides were prepared following the Manual RNA Slide Preparation Protocol in the GeoMx DSP Slide Preparation User Manual (NanoString, MAN-10087-03 for software v1.4). Briefly, slides were baked at 60°C for 45 minutes. Deparaffinization and rehydration were performed in xylene for 3x5 minutes, 100% ethanol for 2x5 minutes, 95% ethanol for 1x5 minutes

and once in 1x PBS (Sigma Aldrich, P5368-10PAK) for 1 minute. Antigen retrieval was performed in 1x Tris EDTA pH 9.0 (Sigma Aldrich, SRE0063) in a Tinto Retriever Pressure Cooker (BioSB, BSB 7008) for 20 minutes at 100°C. Thereafter, the slides were washed 1x PBS for 5 minutes. Tissue RNA targets were exposed by incubating the slides with 1mg/mL proteinase K (Ambion, 2546) in 1x PBS at 37°C for 15 minutes. Slides were washed in 1x PBS for 5 minutes, then immediately placed in 10% Neutral Buffered Formalin (EMS Diasum, 15740-04) for 5 minutes, followed by incubation in NBF Stop Buffer (1.48M Tris Base, 563mM Glycine) for 2x5 minutes, and once in 1x PBS for 5 minutes. *In situ* hybridizations with the GeoMx® Cancer Transcriptome Atlas Panel (CTA, 1,811 targets) or an early-access GeoMx® Whole Transcriptome Atlas Panel (WTA, 18,335 targets), each supplemented with 26 human genes associated with SARS-CoV-2 biology (**Supplementary Information Table 6**) were performed at 4 nM final probe concentration in Buffer R (NanoString). Both probe sets were spiked with a panel of COVID-19 RNA plus probes (NanoString) at a final concentration of 4 nM. One at a time, slides were dried of excess 1x PBS, set in a hybridization chamber lined with Kimwipes wetted with 2x SSC, and covered with 200 µL prepared Probe Hybridization solution. HybriSlips (Grace Biolabs, 714022) were gently applied to each slide and they were left to incubate at 37°C overnight. The HybriSlips were removed by dipping the slides in 2x SSC (Sigma Aldrich, S6639)/0.1% Tween20. To remove unbound probes the slides were washed twice in 50% Formamide (ThermoFisher, AM9342)/2x SSC at 37°C for 25 minutes, followed by two washes in 2x SSC for 2 minutes. Slides were blocked in 200 µL Buffer W (NanoString), placed in a humidity chamber and incubated at room temperature for 30 minutes. Morphology marker solution was prepared for 4 slides at a time: 22 µL SYTO13 (NanoString), 5.5 µL Pan-cytokeratin morphology marker (NanoString), 7.5 µL CD45 morphology marker (NanoString), 2.2 µl Alexa Fluor 647 alpha-Smooth Muscle Actin (Novus Bio, IC1420R, clone 1A4) and 187 µL Buffer W for a total volume of 220 µL/slide. For the slides comprising CD68 staining (D13-D17), morphology marker solution was prepared for 4 slides at a time using 22 µL

of SYTO83 (ThermoFisher, S11364), 11µL of CD68-594 (Novus Bio, NBP2-34736AF647), 11µL of CD45-647 (Novus Bio, NBP2-34527AF647), and 5.5µL of PanCK-488 (eBioscience, 53-9003-82), and 187 µL Buffer W for a total volume of 220 µL/slide. One at a time, slides were dried of excess Buffer W, set in a humidity chamber, covered with 200 µL morphology marker solution, and left to incubate at room temperature for 2 hours. Slides were washed twice with 2x SSC for 5 minutes and were immediately loaded onto the NanoString DSP instrument.

For the NanoString GeoMx DSP Protein Assay, slides were prepared following the Manual FFPE Protein Slide Preparation Protocol in the GeoMx DSP Slide Preparation User Manual (NanoString, MAN-10087-03 for software v1.4). Briefly, slides were baked at 60°C for 45 minutes. Slides were deparaffinized in xylene for 3x5 minutes and rehydrated in 100% ethanol for 2x5 minutes, 95% ethanol for 2x5 minutes and twice in DEPC water for 5 minutes. Antigen retrieval was performed with 1x Citrate Buffer pH 6.0 (Vector Laboratories, H-3300-250) in a Tinto Retriever Pressure Cooker (BioSB, BSB 7008) for 15 minutes at high pressure and high temperature. Pressure was released and the slides were removed to room temperature for 25 minutes. Slides were then washed twice in 1x TBS-T (Cell Signaling Technologies, 9997S). Hydrophobic barriers were drawn around each tissue section and the tissue blocked with 200 µL Buffer W (NanoString) for 1 hour at room temperature in a dark humidity chamber. Antibody solution consisting of eight protein panels (77-plex total) was prepared by mixing 8 µL of each panel per slide of: Immune Cell Profiling Panel Core (NanoString), IO Drug Target Panel (NanoString), Immune Activation Status Panel (Nanostring), Immune Cell Typing Panel (NanoString), Cell Death Panel (NanoString), MAPK Signaling Panel (NanoString), PI3K/AKT Signaling Panel (NanoString) and COVID-19 Custom Panel (Abcam), 5 µL Pancytokeratin morphology marker (NanoString), 7 µL CD45 morphology marker (NanoString), and 123.5 µL Buffer W for a total volume of 200 µL/slide. Slides were dried of excess Buffer W and incubated with 200 µL antibody solution in

a dark humidity chamber at 4°C overnight. Slides were washed three times in 1x TBS/0.1% Tween-20 for 10 minutes before being incubated with 200 µL 4% paraformaldehyde (Electron Microscopy Sciences, 50-980-487) in a dark humidity chamber at room temperature for 30 minutes. Nuclear stain solution was prepared for 4 slides at a time: 20 µLSYTO13 (Nanostring) and 180 µL 1x TBS (Cell Signaling Technologies, 12498S) for a total volume of 200 µL/slide. Slides were washed twice in 1x TBS-T for 5 minutes before incubation with nuclear stain solution in a dark humidity chamber at room temperature for 15 minutes. Slides were washed two times in 1x TBS-T and were loaded onto the DSP. Slides that were not immediately loaded onto the DSP were mounted with FluoromountG mounting media (SouthernBiotech, 0100-01), coverslipped, and stored in a dark humidity chamber at 4°C until they were ready to be processed. When ready, slides were soaked in 1x TBS-T for 30 minutes or until the coverslip fell off, and washed once in 1X TBS-T for 5 minutes before being loaded onto the DSP.

GeoMx DSP instrument and ROI selection

For most spatial samples, slides were loaded into the slide holder of the GeoMx DSP machine and secured by closing the slide tray clamp. Each slide was covered with 2 mL of buffer S and the underside of the slides cleaned with 70% ethanol before loading the slide tray into the DSP. The scanning areas were set to only scan the tissues and the tissue boundaries were determined by adjusting the x and y scanning areas and the sensitivity. Each slide was scanned with a 20x objective and scan parameters: 100 ms FITC/525 nm, 300 ms Cy3/568 nm, 300 ms Texas Red/615 nm. 12-25 rectangular 700 µm x 800 µm region of interests (ROIs) were placed based on selection and assessment by a pathologist per tissue. When applicable, a segmentation mask was applied to ROIs selecting $PanCK^+$ regions and $SYTO13^+$ areas of interest (AOI). Each channel threshold for segments was manually adjusted to maximize inclusion of the AOI signal areas and minimize non-signal areas. In addition, a general setting of erode 1-2 µm, N-dilate 2 µm, hole size 160 μm^2 and particle size 50 was used. After approval of the

ROIs, the GeoMx DSP photocleaves the UV cleavable barcoded linker of the bound RNA probes and collects the individual segmented areas into separate wells in the DSP collection plate. For donors D13-D17, we performed a similar protocol, selecting WTA ROIs by tissue type (bronchial, epithelium, artery, alveoli) and DAB staining serial sections with an anti-SARS-CoV-2 spike glycoprotein antibody and RNAscope probes for SAR-CoV-2.

GeoMx protein nCounter pooling and preparation

DSP collection sample plates were dried down by sealing the plates with aeroseal and placed at room temperature overnight. Wells were resuspended with 7 µL DEPC-treated water, incubated for 10 minutes at room temperature and briefly spun down. Probe working pools were made corresponding to the number of rows of DSP aspirate in each collection. COVID-19 Custom Probe R was diluted 1:5 in 10 µL DEPC-treated water to achieve a working concentration. For full DSP collection plates, the Probe R working pool was prepared by mixing 8 µL of the following Probe Rs corresponding to the antibody panels used above: Immune Cell Profiling Panel Core, IO Drug Target Panel, Immune Activation Status Panel, Immune Cell Typing Panel, Cell Death Panel, MAPK Signaling Panel, PI3K/AKT Signaling Panel, COVID-19 Panel (1:5 working concentration) and 2 µL of DEPC-treated water for a total volume of 66 µL. The Probe U working pool was prepared by mixing 8 µL Probe U master stock and 58 µL DEPC-treated water for a total volume of 66 µL. For collections filling 2-3 rows of the 96-well DSP collection plate and requiring 2-3 Hyb Codes, the Probe R working pool was prepared by mixing 4 µL of the following Probe Rs corresponding to the antibody panels used above: Immune Cell Profiling Panel Core , IO Drug Target Panel, Immune Activation Status Panel, Immune Cell Typing Panel, Cell Death Panel, MAPK Signaling Panel, PI3K/AKT Signaling Panel, COVID-19 Panel (1:5 working concentration) and 1 µL of DEPC-treated water for a total volume of 33 µL. The Probe U working pool was prepared by mixing 4 µL Probe U master stock and 29 µL DEPC-treated water for a total volume

of 33 µL. The final probe/buffer pool was prepared for the number of rows of DSP aspirate in each collection and a corresponding number of Hyb Codes required. For collections filling 8 rows of the 96-well DSP collection plate and requiring 8 Hyb Codes, the probe/buffer pool was prepared by mixing 64 µL Probe R working pool, 64 µL Probe U working pool, and 640 µLHybridization Buffer for a final volume of 768 µl. For collections filling 2 rows of the 96-well DSP collection plate and requiring 2 Hyb Codes, the probe/buffer pool was prepared by mixing 16 µL Probe R working pool, 16 µL Probe U working pool, and 160 µl Hybridization Buffer for a final volume of 192 µL. From these pools, 84 µL were aliquoted to each of the required Hyb Code tubes. 8 µL of each Hyb Code was distributed across a new hybridization 96-well plate row wise and 7 µL DSP aspirates from the collection plate added to the corresponding wells in the hybridization plate. Hybridization plates were heat sealed, briefly spun down and loaded onto a thermal cycler at 67°C with the lid at 75°C for 16-24 hrs. Hybridization plates were cooled on ice and briefly spun down before pooling.

The amount of hybridized product from each plate to pool together was determined based on the total segment area (μm^2) per column from the Lab Worksheet file produced by the DSP instrument following the volumes listed in Table 6 in the nCounter Readout User Manual (NanoString, MAN-10089-04 for software v2.0 JULY 2020). Hybridized products from each well were pooled column-wise into a 12-tube strip, briefly spun down, and loaded onto the nCounter MAX/FLEX Prep Station, run on the "High Sensitivity" setting. The prepared sample cartridge was then loaded onto the nCounter Digital Analyzer (NanoString) along with its respective Cartridge Definition File (CDF) from the DSP instrument, which contains vital indexing information, and was run with the setting provided by the CDF file. Once finished, the sample cartridge was wrapped in aluminum foil and stored at 4°C. The RCC file output by the Digital Analyzer was loaded onto the DSP instrument for analysis.

GeoMx RNA Illumina library preparation

DSP collection sample plates were dried down by sealing the plates with aeroseal and placed at room temperature overnight. Following resuspension of the wells with 10 µL of DEPC-treated water the plates were incubated for 10 minutes at room temperature before briefly spun down. 96-well PCR plates were prepared by mixing 2 µL PCR mix (NanoString), 4 µL of index primer mix (NanoString) and 4 µL of DSP sample. The following PCR program was used to amplify the Illumina sequencing compatible libraries; 37°C for 30 min, 50°C for 10 min, 95°C for 3 min, followed by 18 cycles of (95°C for 15 sec, 65°C for 1 min, 68°C for 30 sec), 68°C for 5 minutes and a final hold at 4°C. The indexed libraries were pooled with a 8-channel pipette by combining 4 µL per well from the 12 columns into 8-well strip tubes. The combined 48 µl pools in the 8-well strip were incubated with 58 µL AMPure XP beads (Beckman Coulter) (1.2x bead to sample ratio) for 5 min. The 8-well strip was placed on a magnetic stand for 5 minutes before removal of the supernatant. The beads were then washed twice with 200 µl of 80% ethanol and dried for 1 minute before being eluted with 10 µL elution buffer (10mM Tris-HCl pH 8, 0.05% Tween-20). The 8 samples were combined into two pools of 40 µL each and incubated with 48 µL AMPure buffer (2.5M NaCl, 20% PEG 8000) for 5 min. Following magnetic incubation for 5 minutes and removal of supernatants the beads were washed twice with 200 µL of 80% ethanol and dried for 1 minute before eluted with 17 µL elution buffer. The two pools were combined to a final elute of 34 µL. Sequencing library size was assessed with a DNA high sensitivity bioanalyzer assay (Agilent Technologies) and the expected library size of 161 bp was observed. Library concentration was assessed with Qubit dsDNA high sensitivity assay (Thermo Fisher Scientific).

Total target counts per DSP collection plate for sequencing were calculated from the total sampled areas (μm^2) reported in the DSP generated Lab Worksheet. For CTA and WTA libraries, the target sequencing depths were 30 and 150 counts/μm^2, respectively. Each library was diluted to 4-10 nM and combined

to reach the estimated counts/μm^2 per library in the final pool. All sequencing libraries were generated with unique indexes allowing pooled sequencing. We combined two of the CTA and five of the WTA libraries for sequencing with an Illumina NovaSeq S4 platform at a loading concentration of 365 pM with 5% PhiX. Two of the CTA collection plate sequencing libraries were sequenced with Illumina NextSeq500 HO 75 cycle kits at a loading concentration of 1.6 pM with 5% PhiX. The sequencing parameters used were: read 1, 27 cycles; read 2, 27 cycles; index 1, 8 cycles; index 2, 8 cycles.

Computational methods

Sc/snRNA-Seq expression matrices

Cell Ranger mkfastq (10x Genomics) was used to demultiplex the raw sequencing reads, and Cell Ranger count on Terra using the cellranger_workflow in Cumulus[15] was used to align sequencing reads and generate a counts matrix. Reads were aligned to a custom-built Human GRCh38 and SARS-CoV-2 RNA reference. ScRNA-Seq reads were aligned to "GRCh38_and_SARSCoV2", and snRNA-Seq reads were aligned to "GRCh38_premrna_and_SARSCoV2". The GRCh38 premrna reference captures reads mapping to both exons or introns, and was built as previously described[13]. The SARS-CoV-2 viral sequence (FASTA file) and accompanying gene annotation and structure (GTF file) are as previously described[61]. The GTF file was edited to include only CDS regions, with added regions for the 5' UTR ("SARSCoV2_5prime"), 3' UTR ("SARSCoV2_3prime"), and anywhere within the Negative Strand ("SARSCoV2_NegStrand") of SARS-CoV-2. Trailing A's at the 3' end of the virus were excluded from the SARSCoV2 fasta file, as we found these to drive spurious viral alignment in pre-COVID-19 samples.

Ambient RNA removal from expression matrices

CellBender remove-background[14] was run (on Terra) to remove ambient RNA and other technical artifacts from the count matrices. The workflow is available publicly as cellbender/remove-background

(snapshot 11) and documented on the CellBender Github repository as v0.2.0: https://github.com/broadinstitute/CellBender.

This latest version of CellBender remove-background cleans up count matrices using a principled model of noise generation in scRNA-Seq. The parameters "expected-cells" and "total-droplets-included" were chosen for each dataset based on the total UMI per cell *vs*. cell barcode curve in accordance with CellBender documentation. Other inputs were left at their default values. The false positive rate parameter "fpr" was set to 0.01, 0.05, and 0.1. For downstream analyses we used the 'FPR_0.01_filtered.h5' file.

As a comparison, we ran the method DecontX[62] on a single sample, using only the droplets which CellBender remove-background identified as non-empty. Removal of the surfactant genes from ambient RNA was qualitatively similar in DecontX, though CellBender remove-background removed more counts **(Extended Data Fig.1h)**.

Quality control

Following CellBender, individual samples were processed using Cumulus, including filtering out cells/nuclei with fewer than 400 UMIs, 200 genes, or greater than 20% of UMIs mapped to mitochondrial genes. Note that we report cells/nuclei with high mitochondrial reads, because while these may reflect technical artifacts they may also reflect pathological events, such metabolically active or necrotic cells. We retained all cells that pass these quality metrics, and without a restriction on number of cells per sample. We have indicated samples with less than 150 cells in **Supplementary Information Table 1.**

We examined the correlation between the post mortem interval (PMI), intermittent mandatory ventilation (IMV) duration, and time from symptom onset to death, with three QC metrics for snRNA-Seq: the number of UMIs per nucleus, the number of genes per nucleus, and the fraction of UMIs mapped to mitochondrial genes. Among these, there was a weak association that was nominally significant, but not by FDR, between PMI and number of genes (Spearman rank correlation = 0.53, p-value=0.04, FDR = 0.10).

Sc/snRNA-Seq sample integration

Sc/sn RNA-Seq data from individual samples were combined into a single expression matrix and analyzed using Scanpy[63] with identical pre-processing steps — aggregation, PCA, and batch correction using Harmony-Pytorch[64]. First, transcript counts were clipped and then normalized. UMI counts were clipped at 13, the value of the 99th percentile of non-zero counts. Clipped count data were then normalized so that UMI counts per cell/nucleus summed to 10,000 and then logged, resulting in log(1+10,000*UMIs / total UMIs) for each cell/nuclei, "logtp10k".

Next, highly variable genes were identified by Scanpy's highly_variable_genes function applied separately to data from each profiling modality: cryopreserved (cells), fresh tissue (cells), and fresh-frozen (nuclei), and retained genes that were highly variable across two or more methods.

Next, data dimensionality was reduced to 75 top principal components by PCA, followed by Harmony[64,65] for batch correction, where each sample was considered a separate batch. Neighbors were computed using the Harmony corrected PCA coefficients and UMAP and Leiden clusters (below) were computed using the resulting nearest neighbors' graph.

Clustering

Leiden clustering was performed on a *k*-nearest neighbor graph (*k*=100) with two different resolutions: 1.3 and 2. Resolution 1.3 was sufficient for coarse annotations of clusters, but both resolutions are included in the metadata on the Single Cell Portal (see **Data Availability**). Clustering was performed using Pegasus[15] with default parameters (except the resolution as mentioned above). Following batch correction[65], the 106,792 lung cells/nuclei formed distinct clusters that were generally well-mixed across donors and protocols (**Extended Data Fig. 2e-g**, 3.8% adjusted rand index between sample and cluster at 3.8%).

Automated cell annotation

Individual cell profiles were annotated by classification using a logistic regression model trained on six publicly available lung or bronchial alveolar lavage scRNA-Seq datasets with published manual labels, all profiled using the 10x Genomics platform. These datasets included healthy lung[19] (and *unpublished*), pulmonary fibrosis[66,67], and COVID-19[68] (and *unpublished*) samples. Two additional datasets were withheld from training to evaluate the predictive model[69,70] and yielded ~80% accuracy by cross validation on two held out datasets (**Extended Data Fig. 2h** and *data not shown*). Since annotations between published datasets differed in resolution and nomenclature, we manually determined a single reference label set. We selected the most granular cell type labels that appeared in at least two datasets in the training set. Published cell type labels were mapped to that reference list or those cells were removed from training if they did not match a term. Because the model is trained on full gene expression profiles, it does not require cell clustering, prior knowledge of marker genes, or highly variable gene selection.

To ensure consistent gene expression quantification across studies, training datasets were re-indexed to a single reference gene list and re-normalized from counts. Before training, data from all datasets were scaled together with the Scanpy scale function with zero_center=False.

An L_2 penalized logistic regression model was trained using SciKit-Learn's LogisticRegression function with C=0.1, max_iter=30, and multiclass=ovr. Performance on train and test sets was quantified by determining the frequency that predictions matched provided labels. If the predicted label was a cell type not included in the provided labels for that dataset, that cell was excluded from the accuracy calculation.

The learned model coefficients (**Supplementary Information Table 2**) represent the contribution of each of the 18,000 included genes to the prediction of all 28 cell types. These coefficients identify many known cell type markers. For example, the top 18 most important markers for Regulatory T cells include *CD3G*, *CD3D*, *CTLA4*, *ICOS*, *CD28*, *IL2RA* and *FOXP3*, the top 12 most predictive genes for AT2 cells are surfactant genes *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPC*, and *SFTPD*, and the top five smooth muscle cell genes are *ACTG2*, *ACTA2*, *CNN1*, *MYH11*, *DES*. While these genes are learned from training on other datasets, they are co-expressed within cell type specific modules in our lung samples.

Kidney, heart and liver were classified using a similar approach after ambient removal and basic quality control as described above. For classifying kidney nuclei, a published meta-atlas[71] was leveraged as a training dataset. For liver and heart, we used publicly available 10x datasets for training (**Supplementary Information Table 2**).

Annotation by manual curation of lung

Manual annotation of clusters was conducted in two main steps: (1) identifying main lineages in the global clustering, (2) identifying subpopulations of cells by sub-clustering cells from each lineage separately.

Identifying main lineages

Lung cell clusters were manually annotated by identifying elevated expression of pre-known markers in UMAP plots as well as visualizing mean Z-score expression of gene sets characterizing known populations. Immune cell types were identified using canonical markers (**Extended Data Fig. 4a**), whereas in order to identify the rest of the lineages we used both published[13,72,73] and *unpublished* gene sets (**Supplementary Information Table 12**).

Immune and endothelial cell subsets: sub-clustering

After identifying the main cell lineages, we studied each lineage by sub-clustering it separately from the rest of the cells, to enable identification of subtle cell states.

For each of the three immune cell types and the endothelial cells we re-computed highly variable genes, performed PCA, batch corrected with Harmony, generated a *k*-nearest neighbor graph (k=100), and performed Leiden clustering on the graph, all with Pegasus[15].

We tested different options for the number of highly variable genes and PCs, followed by manual inspection of the resulting UMAP, and choosing the parameters where cell clusters were evenly distributed and devoid of finger-like structure. We assessed the stability of clusters with rand-index[74,75], and chose the highest resolution that showed a rand_index value > 0.85.

The final parameters of number of highly-variable genes, number of PCs and resolution were:

T and NK cells - (1500, 5, 0.7)

B and Plasma cells - (150, 4, 0.3)

Endothelial cells - (500, 15, 0.5)

Myeloid cells - (1500, 5, 0.9)

We tested each cluster for differentially expressed genes, and merged clusters that showed no difference. One cluster in the original T and NK sub-clustering was identified as a B cell subset and was reported in that lineage.

Harmonization at the lineage level merged different samples evenly across clusters, but cells and nuclei data were separated in all the sub-clusterings, making the cell data merged in one cluster, with or without additional nuclei in the same cluster (**Extended Data Fig. 4c**).

Immune and endothelial cell subsets: identifying unique differentially expressed genes per sub-cluster

Differentially expressed (DE) genes were determined for each final sub-cluster by two methods: (1) single cell differential analysis implemented in Pegasus (Mann-Whitney U test, two sided) and (2) pseudo-bulk analysis, summing UMIs from all cells in each cluster per donor, and then performing DE analysis between clusters by fitting a linear regression using the functions "lmFit" and "eBayes" from the limma R package[76] (Results from method 1 are available in **Supplementary Information Table 13**).

To summarize the top DE genes in a heatmap (**Extended Data Fig. 4b**) we set different cutoffs to the statistics measured in each lineage to get the following number of DE genes per lineage: (a) the top 25 DEGs from method 1, (b) all DE genes above a threshold value of AUC from method 1, (c) all DE genes from with adjusted p-value <0.05 and fold-change above a threshold. The threshold used for each lineage and the total number of DEGs found are as follows:

Endothelial cells: 128 DE genes, minimum AUC=0.8, minimal log fold-change=4.2.

T+NK cells: 91 DE genes, minimum AUC=0.8, minimal log fold-change=3.

Myeloid: 111 DE genes, minimum AUC=0.8, minimal log fold-change=4.

B+Plasma cells: 78 DE genes, minimum AUC=0.75, minimal log fold-change=2.

Immune and endothelial cell subsets: sub-cluster annotations

Annotations of the unique sub-populations that were identified in the sub-clustering process were done by using legacy knowledge of these sub-populations through looking at unique gene markers associated with AUC statistics log fold change (as described above) per lineage. Myeloid cells and B and plasma cells were split into states based on highly expressed genes in each sub-cluster.

The endothelial cell cluster featured a grouping of predicted doublets in the general (main lineage) clustering. These cells do not cluster together in the sub-clustering, and therefore were separately rmanually annotated by the cluster they belong to. In the T+NK sub-clustering, there were three different clusters that were manually annotated as doublets, as they showed expression of genes from at least two different separated lineages.

To annotate the endothelial cells we overlaid the 30 top DE genes of identified endothelial cells from a healthy lung cell atlas[19] onto the UMAP of our endothelial data.

In the remaining four subpopulations, we did not detect known gene signatures that correspond to healthy human lung ECs, so we imputed the endothelial subpopulations using Adaptively-thresholded Low Rank Approximation (ALRA; implemented with Seurat v3.2.1) to compensate for EC signature dropout. After imputation, we were able to annotate the remaining EC subpopulations: cluster 1-3. The remaining EC subpopulation (cluster 4) has a gene expression signature that overlaps with multiple EC subpopulations in the healthy lung reference data.

Identifying and annotating epithelial cell and fibroblast subsets

For the lung epithelial and fibroblast lineages, we harmonized the cells using LIGER[77], a non-negative matrix factorization approach that allows cells to be described as occupying multiple different cell states. For the epithelial and fibroblast lineages, we used 8 and 7 factors, respectively, in the factorization. Each factor describes a different cell state, and to describe these states, we examined the top 100 genes loading each factor (**Supplementary Information Table 14**). We identified the PATS cell state using the gene signature derived from alveolar organoids in Supplementary Table 1 of Kobayashi *et al*[21].

Within the epithelial lineage, we further sub-clustered the KRT8/PATS/ADI/DATPs (cluster 7 in **Fig. 2b**) and airway epithelial cells (cluster 3 in **Fig. 2b**), and for each we applied LIGER with 4 factors. We identified intrapulmonary basal-like progenitor cells (IPBLPs) and airway basal cells by expression of the marker genes TP63, KRT5, and KRT8, and by testing for differentially expressed genes using the Mann-Whitney U test (**Supplementary Information Table 3**). To identify expression differences between IPBLPs and airway basal cells, we compared these two cell subsets using the same test (**Supplementary Information Table 3**).

Annotation by manual curation of heart, kidney and liver

For heart, liver and kidney, we used standard approaches of graph-based clustering on the integrated datasets (Seurat v3.2.1 for kidney and liver; Scanpy v1.6.1 for heart) to first derive high-level groupings followed by differential gene expression to determine cluster enriched genes.

For kidney and liver, we defined the endothelial and immune compartments as predominantly *PECAM1+/CD31+* and *PTPRC+/CD45+* clusters respectively from the high-level assignment, and

subjected them to iterative clustering, differential expression and post-hoc annotation to determine granular subsets. Similar iterative clustering was performed for the high-level mesenchymal groups in both tissues, proximal tubular cells in the kidney and hepatocytes in the liver.

For heart, Leiden clustering at a resolution of 1.5 yielded clusters whose marker genes were compared to the heart atlas from Tucker *et al.*[78] Cell classes were generally well-mixed among donors (**Extended data Fig. 10b,c,e,f,h,i)**. Where applicable, technical or putatively low-quality clusters are indicated as such to allow for permissive downstream analysis by the interested reader. For plotting, R packages ggplot2 v3.3.2, dplyr 0.8.0.1, reshape2 v1.4.3 and cowplot v1.1.0 were used.

Doublet detection of snRNA-seq data for lung, heart, kidney and liver

We identified doublets using a two-step procedure for each tissue type based on ambient RNA removed gene count matrices. First, we independently identify doublets for each sample by calculating Scrublet-like[79] doublet scores and then automatically determining a doublet score cutoff. Second, we integrated and clustered all samples from the sample tissue type and further marked any cluster that has more than 70% of cells identified as doublets as a doublet cluster. All cells within doublet clusters were marked as doublets.

We used a recent re-implementation of the Scrublet algorithm in Pegasus to calculate Scrublet-like doublet scores per sample. The original Scrublet algorithm consists of three major steps: preprocessing, doublet simulation and doublet score calculation using a k-nearest-neighbor (kNN) classifier. In the preprocessing step, we filtered low quality cells using the same criteria discussed in the **quality control** section, normalized each cell to TP100K, log-transformed expression values to log(TP100K+1) and identified highly variable genes using Pegasus[15]. We then standardized each highly variable gene based

on the normalized count matrix (not log-transformed) and computed the first 30 principal components. In the doublet score calculation step, we built the kNN graph using Pegasus' kNN building function.

We automatically determined a doublet score cutoff for each sample by log-transforming the Scrublet-like doublet scores for simulated doublets, estimating a smoothed density curve based on log-transformed doublet scores using kernel density estimation, identifying an appropriate cutoff by inspecting the local maxima and signed curvature values of the density curve, and finally transforming the cutoff from the log space by taking the exponential. Any cells with a doublet score larger than the cutoff were identified as doublets. For more details, please refer to documentation on Github[80].

The previous sample-level doublet detection procedure might miss some doublets. Inspired by previous work[81,82], we integrated quality-controlled samples for each tissue type, normalized each cell to TP100K and log transformed the expression value (log(TP100K+1), selected highly variable genes, computed the first 50 principal components (PC), corrected batch effects based on the PCs using Harmony-PyTorch, and clustered cells using the Louvain algorithm with a resolution of 1.3. We then tested if each cluster was statistically significantly enriched for doublets using Fisher's exact test and controlling False Discovery Rate at 5%. Finally, among clusters that are significantly enriched for doublets, we picked those with more than 70% of cells identified as doublets and marked all cells in these clusters as doublets.

Note that we report possible "doublets" because while these may reflect not only technical artifacts but also pathological events, such as dying or virally infected cells being phagocytosed (for doublets; *e.g.*, T+endothelial/macrophage cells).

Determination of significant changes of cell type proportions

To identify significant changes in cell proportions between COVID-19 and healthy lung, we used a Dirichlet-multinomial regression approach, as previously described[83], to find differences in cell type proportion that addresses the limitation that proportions must sum to 1 and are not independent of each other. As such, when looking for differences between healthy *vs*. COVID-19, the tests account for all cell types simultaneously, and a significant p-value tells us that even if the proportion of a given cell type was expected to, for example, decrease given the increase in other cell types, it has decreased significantly *more* than expected.

Cell proportion differences between COVID-19 and healthy heart tissue were assessed using scCODA[84] v0.1.1.post1) using the model '~ disease', with the reference cell type defined as fibroblasts, as determined using the method proposed by Brill *et al*.[85]

Deconvolution of bulk RNA-Seq lung samples

We performed composition estimation for the bulk lung samples using the MuSiC R package[86]. Ten bootstrap estimates, each using an independent sampling of 10,000 single cells, stratified by annotated cell type, to use as reference, were used to perform estimation. We confirmed that the subsampling of the reference to 10,000 cells was appropriate and did not introduce a specific sample composition bias by comparing the predicted sample composition with successively lower number of reference cells in the range of 10,000 to 5,000 cells (**Extended Data Fig. 3b**).

Healthy reference comparisons and differential expression

Due to the lack of matched healthy tissue samples, we supplemented our data with publicly available healthy and COVID-19 positive lung tissue data (**Supplementary Information Table 2)**. These

datasets were all scRNA-Seq analyzed with 10x Genomics to minimize technical variance between comparisons. Using this aggregated resource, we performed a meta-differential expression analysis to enable healthy vs. COVID-19 comparisons. To identify cell intrinsic changes in expression associated with severe COVID-19, we created a lung meta-atlas of ~1,280,000 cells and nuclei from 10 healthy lung sc/snRNA-Seq studies with sc/snRNA-seq data from healthy lungs and four studies from COVID-19 patients (this study, Chua *et al*.[68], Liao *et al*[70], Xu *et al.*[87]) (**Supplementary Information Tables 2 & 3**), assigned each cell to one of the 28 classes in our severe COVID-19 lung atlas, and used two different models to identify differentially expressed genes between COVID-19 and healthy tissues for each (**Supplementary Information Table 4**).

We tested for differential expression between COVID-19 and other lung samples using two linear regression models applied to each individual gene with the automated cell annotations.

In one model, a single cell model for each cell type was built using the "statsmodels" module in Python and included Boolean covariates for snRNA-Seq *vs*. scRNA-Seq and 10x V3 kit *vs.* other kits as well as a continuous covariate for UMI counts for each cell and a random intercept (single cell gene expression ~ c_1 + c_2*(nuclei indicator) + c_3*(10X v3 indicator) + c_3*(number of UMIs) + c_4*(COVID-19 indicator)). Since each dataset had different numbers of cells, we sampled 2,000 cells of a given cell type with replacement from the 10 healthy datasets and four COVID-19 datasets weighted specifically to balance the disease states and studies.

Separately, we also applied another linear model to pseudobulk expression profiles from the same studies to gain confidence in our differential expression test. This model was designed similarly to identify which genes in which cell types change expression levels in a COVID-19 infection

(pseudobulked gene expression ~ c1 + c2*(COVID-19 indicator) + c3*Study). We applied Hold-Sidak multiple test correction (in Python "statsmodels") with an error rate of alpha=0.05 to test which COVID-19 coefficients were predictive of gene expression levels in each model.

Differential expression analysis for heart data included 22 left ventricular samples from COVID-19 (this study), 11 left ventricular samples from 7 transplant donors from Tucker *et al.*[78], and 18 left ventricular samples from 14 individuals from Litviňuková *et al.* [88], totaling 40,880 COVID-19 and 169,352 healthy nuclei. Samples were meta-analyzed by automatic cell-type classification (above) for differences between COVID-19 cases and healthy controls. Differential expression testing was carried out using limma and voom, after summing cells per cluster per individual donor (as in Lun and Marioni[89]), and a linear model "~ disease + study + version10x + sex" was fit to obtain the desired COVID-19 versus healthy contrast, separately for each cell type. Results were compared with the above statsmodels linear regression, and log2-fold-changes and p-values for both methods are reported in **Supplementary Information Table 10**. Gene set enrichment was performed using fGSEA (R package)[90] ranking by limma t-statistic, and run using MSigDB gene sets C5: GO Biological Process and C5: GO Molecular Function. Full results are reported in **Supplementary Information Table 10**.

Correction for ambient SARS-CoV-2 UMI

To address viral ambient RNA, we adapted several methods (CellBender[14] and those described in Kotliar, Lin *et al.*[91] and Cao *et al.*[92]) to create a conservative criterion for assigning a single cell or single nucleus as SARS-CoV-2 RNA+ or RNA-. We used a binomial test to determine the probability that the cell contained more SARS-CoV-2 UMIs than expected by ambient contamination, given the fractional abundance of SARS-CoV-2 aligning UMIs per cell, the abundance of SARS-CoV-2 aligning UMIs in the ambient pool, and the estimated percent ambient contamination of the single cell/nucleus. The

fractional abundance of SARS-CoV-2 aligning UMIs per cell was defined as the number of UMIs to all viral genomic features (including the negative strand) divided by the total number of UMIs aligning to either SARS-CoV-2 or GRCh38 per cell. The abundance of SARS-CoV-2 aligning UMIs in the ambient pool was defined as the sum of all UMIs in the pre-CellBender output (from CellRanger counts) within discarded cells determined as "empty" or "low quality", and was therefore determined on a per-sample basis. The estimated percent ambient contamination was defined as the ratio of the total UMIs for each single cell before *vs.* after CellBender. An exact binomial test was applied to each single cell using R given these parameters. P-values were adjusted using a Bonferroni correction and $p < 0.01$ was considered significant. Cells were assigned as "SARS-CoV-2 RNA+", "SARS-CoV-2 Ambient" (if they had any UMIs to SARS-CoV-2 but were not significantly higher than the ambient pool) and "SARS-CoV-2 RNA-" (no UMIs to SARS-CoV-2).

Distinction between positive and negative strand viral RNA

To putatively distinguish between possibly non-productive virions from productive viral infections we examined viral splice variants (sub-genomic species) or alignments to the negative strand, observing rare alignments to the negative strand, which may indicate host cells containing actively-replicating viruses. However, we caution from making definitive inferences from these observations, because the "absence" of viral replication intermediates among samples from later time points may also simply reflect differential capture of viral RNA in these samples and we could not disentangle the effects of a low abundance of viral reads relative to human fragments (due to an overall diminished abundance of viral reads) with low frequency of negative strand or spliced intermediates.

Differential expression between SARS-CoV-2 RNA+ *vs.* RNA- cells

To test for host genes associated with the presence of SARS-CoV-2 RNA, we used the following approach to mitigate potential biases or confounding factors due to low cell numbers, differences in cell quality, or donor-to-donor variability: (**1**) we removed any cells classified as "SARS-CoV-2 ambient" from our analysis; (**2**) we analyzed DE genes only within cell types that contained at least 10 SARS-CoV-2 RNA+ cells; (**3**) within a given cell type, we only considered donors where at least 2 cells were SARS-CoV-2 RNA+, and selected SARS-CoV-2 RNA- cells only from these donors; (**4**) we subsampled the SARS-CoV-2 RNA- cells to match complexity distributions to the SARS-CoV-2 RNA+ as previously described[25](briefly, cells were partitioned into 10 bins based on complexity (defined by the $\log_{10}$(# genes/cell)), and the SARS-CoV-2 RNA- cells were randomly subsampled to match the distribution in the SARS-CoV-2 RNA+ cells).

We tested differential expression using DESeq2[93], which fits a generalized linear model (GLM) to each gene's expression, and included % mitochondrial genes per cell as a continuous covariate and the donor as a discrete covariate. Genes were considered DE if they had an FDR q-value < 0.05. Gene set enrichment testing with GSEA[39,40] was completed by defining a pre-ranked list of genes based on the $\log_2$(fold change) calculated from DESeq2 between SARS-CoV-2 RNA- cells and SARS-CoV-2 RNA+ cells of a given cell type. Hallmark, C2, and C5 gene sets were tested for enrichment, and an FDR q-value < 0.05 was considered significant.

Differential expression analysis of bulk RNA-Seq profiles from adjacent lung biopsies comparing samples with high *vs.* low/no viral load

Differential expression (DE) analysis was performed after FASTQ files were aligned against Hg19 using STAR (v2.6.0c)[94]. RSEM[95] (v1.2.8) was used to generate a count matrix and DE genes were

identified in R[96] using DESeq2[93]. A log2 FC cutoff of 1.4 and an FDR adjusted p-value < 0.05 was used as a threshold for significance. GSEA[39,40] was used to detect gene ontology terms associated with genes significantly upregulated or downregulated in samples with high viral load (number of SARS-CoV-2 copies/ng RNA by RT-qPCR).

Viral Enrichment score

Inspired by Viral Track[97], a program that identifies viral UMIs in human single cells and enumerates them per meta-cell, we computed a viral enrichment score per cluster. We could not employ the enrichment score of ViralTrack given the low number of SARS-CoV-2 RNA+ cells in our data, such that entire clusters, rather than higher resolution meta cells, were required for sufficient power.

The enrichment score for cluster C in a clustering of cells is computed as follows:

Enrichment(C)=log((Observed(Vcells in C)+epsilon)/(Expected(Vcells in C)+epsilon)) =log((Vcells in C)+epsilon)/((Vcells in total * X_c)+epsilon)

where:

Vcells are SARS-CoV-2 RNA+ cells

X_c is the proportions of the total number of cells in cluster C out of the total number of cells in the given clustering (either all cells or within lineage).

$epsilon$=0.0001

We assessed the significance of each enrichment score empirically by permuting the data 100 times by randomly assigning the same number of SARS-CoV-2 RNA+ annotations to cells, such that the overall proportion of SARS-CoV-2 RNA+ cells was the same, computing the enrichment score per cluster, and calculating empirical p-value as the fraction of the permutations that showed a similar or higher enrichment score compared to the real result. We calculated an FDR as implemented in the function 'fdrcorrection' in the Python package statmodels[98].

Generation of expression matrices from NanoString GeoMx CTA, WTA and protein data

For CTA and WTA (RNA) data, sequencing reads from NovaSeq or NextSeq was compiled into FASTQ files corresponding to each AOI using bcl2fastq. FASTQ files were demultiplexed and converted to Digital Count Conversion files using NanoString's GeoMx NGS DnD Pipeline. The resulting DCC files were converted to an expression count matrix. AOIs (2 for CTA, 2 for WTA) were removed from the analysis if sequencing reads totaled less than 10,000. Target counts were generated by taking the geometric mean of all probe counts against the same gene target in each well. Individual CTA probes were dropped if they had counts less than 10% of the target counts (2 CTA) or if they were global Grubb's outliers at alpha = 0.01 (6 for CTA). These probe screens were not performed on the WTA assay as each gene is sampled by a single probe. As a quality control step, sequencing saturation was calculated by dividing the number of non-unique sequencing reads by the total number of reads for each AOI. AOIs with sequence saturation below 50% were dropped from analysis (2 CTA, 0 WTA). The 75th percentile (Q3) of the target counts of each probe pool in each AOI were calculated and normalized to the geometric mean of the 75th percentiles across all AOIs to give normalization factors. Target counts were divided by these normalization factors to give normalized expression counts.

WTA and COVID-19 Spike-in GeoMx DSP samples from D13-17 donors were processed similarly as above with a few modifications. Both WTA and COVID-19 Spike-in panels were sequenced on an Illumina NovaSeq platform, aligned to probe references and demultiplexed using Nanostring's DnD pipeline (as above). The COVID-19 Spike-in panel contains multiple probes per target gene, and counts that were not excluded by outlier detection were collapsed into a single value by calculating the geometric mean. In any cases where the probe counts were zero, the probe count was set to one prior to taking the geometric mean. Target counts for both the WTA and COVID-19 Spike-in genes were then

normalized by multiplying each count by its pool-specific negative probe normalization factor. Genes were further removed if their mean expression was greater than 100. Following these AOI-based and feature-based filtering steps, the data were normalized to the geometric mean of the 75th percentile as above.

To measure DSP protein expression, indexing oligonucleotides were UV photocleaved from the profiled antibodies and deposited into a 96-well plate. Oligonucleotides were then hybridized to fluorescent barcodes and run through the nCounter MAX analysis system to generate raw digital counts. The raw counts were inputted into the DSP for calibration (adjusting for oligo-barcode binding efficiency) and for generation of a protein expression matrix; this matrix was subsequently normalized to internal spike-in positive controls (ERCC) to account for system variation. The calibrated and ERRC-normalized expression matrix was then normalized to the geometric mean of housekeeping genes[99] as follows: first, the geometric mean of three housekeeping genes, histone *H6*, *GAPDH*, and *S6*, was calculated for each AOI; next, normalization factors were calculated as the geometric mean in each AOI dividing by the geometric mean of geometric means across all AOIs; finally, raw expression values were divided by the calculated normalization factors to give normalized protein expression levels.

Log-transformed matrices were obtained from normalized CTA, WTA and protein expression matrices by applying the function $f(x) = log\ (x + 1)$ to each element of the matrices and were used for downstream analysis. We note that protein DSP data showed heterogeneity in immune cell marker levels across ROIs and donors, but in many alveolar ROIs, PanCK and CD45 immunofluorescence stains showed significant mucus staining, which could affect antibody-based DSP data. Such off-target mucus binding was not observed in RNAscope images, and is thus less likely to be a concern in oligonucleotide probe-based WTA and CTA data.

Gene signature score calculation

A signature score summarizes the expression levels of a set of functionally-related genes for each AOI while controlling for gene abundance related technical noise[72]. Signature scores were calculated on the log-transformed expression matrices using Pegasus'[15] new signature score calculation algorithm[100]. A gene set consisting of COVID-19 *S* and *ORF1ab* genes was used to calculate the SARS-CoV-2 signature score.

To score cell signatures we used a 'reconciled' annotation between the automated ('predictions') and manual annotations ('Cluster'), as determined by the following rules. If a cell is assigned more than one label according to the rules, the first rule that applies to that cell determines its label. AT1: 'leiden_res_2' clusters number 10 and 26 (also consistent with 'predictions'). AT2: 'leiden_res_2' clusters number 2 and 3 (consistent with 'predictions'). B cell, plasma cell: Cells assigned as these types in 'predictions'. CD4+ T cell: Cells predicted as either 'CD4+ T cell' or 'Treg' in 'predictions'. CD8+ T cell: Cells predicted as 'CD8+ T cell' in 'predictions' and annotated as 'T+NK' in 'Cluster'. NK cell: Cells predicted as 'NK cell' in 'predictions'. Macrophage: Cells annotated as 'Myeloid' in 'Cluster'. Mast cell: Cells annotated as 'MAST' in 'manual coarse annotation'. Ciliated cell: Cells annotated as 'Ciliated' in 'Cluster'. Secretory cell: Cells annotated as 'Secretory' in 'Cluster'. Vascular endothelial cell: Cells predicted as 'vascular endothelial' in 'predictions'. Lymphatic endothelial cell: Cells predicted as 'lymphatic endothelial' in 'predictions'. Pericyte: Cells predicted as 'pericyte' in 'predictions'. Smooth muscle cell: Cells predicted as 'smc' in 'predictions'. Fibroblast: Cells predicted as 'fibroblast' in 'predictions'. Myofibroblast: Cells predicted as 'myofibroblast' in 'predictions'. Mesothelial cell: Cells predicted as 'mesothelial' in 'predictions'.

Cell type specific markers for WTA and CTA deconvolution

To define cell markers for deconvolution, Pegasus was used to conduct differential expression (DE) analysis on the snRNA-Seq data using the 'reconciled' annotation (above). For each cell type, Mann-Whitney U tests were conducted between cells annotated as this cell type and all other cells and with an FDR control at 0.05. DE genes were ranked by their Area Under the Receiver Operating Characteristics (AUROC) scores (**Supplementary Information Table 7**). Next, 10 marker genes were selected for each cell type by the combination of the DE results and published lung cell type markers[13,25,69]. We additionally picked 6 and 7 published marker genes for basal cells and neutrophils, respectively. These cell type markers were used for **Fig. 4b** and are available in **Supplementary Information Table 8**. Many of these markers were not available for the CTA data, which comprises only ~1,800 genes. Thus, for cell types with less than 3 markers in the CTA data, we added genes based on the DE results to ensure at least 3 markers. These markers were used for **Extended Data Fig. 8h** and are available in **Supplementary Information Table 8.**

Differential expression analysis between subjects and controls for WTA and CTA data

All AOIs in WTA data annotated as alveoli PanCK$^+$ and alveoli PanCK$^-$ from donors D18, D19, D20, D21, D8, D9, D10, D11 and D12 were grouped together as the donor group; all AOIs in WTA data annotated as alveoli PanCK$^+$ and alveoli PanCK$^-$ from controls D22, D23 and D24 were grouped as the control group. Then the limma package[76] was used to perform differential expression analysis between the two groups (**Fig. 4c, Extended Data Fig. 8i, Supplementary Information Table 6**). Specifically, Q3-normalized count data, including two negative control probes, was transformed to TMM-normalized[101] $\log_2$-counts per million. The transformed data was fit with a linear model, considering the indicator functions of donor and control AOI, and empirical Bayes moderation was used to estimate

the contrast S - C, where S stands for donor data and C for control. A similar approach is taken for alveoli CTA (**Extended Data Fig. 8j,k, Supplementary Information Table 6**).

Gene set enrichment analysis (GSEA) between subjects and controls for WTA and CTA data

GSEA analysis was conducted using fGSEA (R package)[90] with the MSigDB Hallmark gene set (https://www.gsea-msigdb.org/gsea/msigdb/download_file.jsp?filePath=/msigdb/release/7.1/h.all.v7.1.symbols.gmt), with gene ranks preset by their $\log_2$(fold-change) in a DE analysis, and the size of a gene set to test ranging between 15 and 500. Terms that are upregulated (corresponds to COVID-19 samples) and downregulated (corresponds to healthy samples) were plotted separately, with an FDR q-value threshold of 0.05 **(Fig. 4c, Extended Data Fig. 8i-k**).

CTA and WTA correlation analysis

Correlation analysis was performed across 1,784 genes that were shared between the CTA and WTA data and 306 AOIs where both CTA and WTA data were collected. For each AOI, the Pearson's correlation coefficient between its corresponding CTA and WTA count vectors of the common genes. We excluded 10 AOIs where the WTA counts for the common genes were a constant vector of small numbers (resulting in NaN correlation), resulting in 296 AOI pairs used in the analysis. Lower correlations between CTA and WTA AOIs were partly due to physical distance between serial sections (**Extended Data Fig. 8g**).

Differential expression between SARS-CoV-2 high and low AOIs

WTA alveolar ROIs were first categorized by their SARS-CoV-2 virus signature scores (**Fig. 4d, Extended Data Fig. 8l**) as SARS-CoV-2-high (score>1.30), SARS-CoV-2-medium

(0.75<=score<=1.30), and SARS-CoV-2-low (score<0.75). The limma package was used to perform differential expression analysis comparing SARS-CoV-2-high *vs.* SARS-CoV-2-low AOIs (separately for $PanCK^+$ and $PanCK^-$ AOIs; **Fig. 4e, Extended Data Fig. 8m, Supplementary Information Table 9**), where Q3-normalized count data was transformed to TMM-normalized log_2-counts per million. The transformed data was fit into a linear model with indicator functions of SARS-CoV-2-high, SARS-CoV-2-medium, and SARS-CoV-2-low, respectively:

$log_2Y = b1 * I(X=\text{SARS-CoV-2-high}) + b2 * I(X=\text{SARS-CoV-2-low}) + b3 * I(X=\text{SARS-CoV-2-medium})$,

where X is the AOI's SARS-CoV-2 virus level, log_2Y is the transformed count data on each gene, and $I(X=\cdot)$ is the binary indicator function on the corresponding SARS-CoV-2 virus level.

Empirical Bayes moderation was then used to estimate the contrast SARS-CoV-2-high – SARS-CoV-2-low from this model.

Differential expression between inflamed and normal appearing alveoli

A total of 50 AOIs from lung biopsies (subjects D13-D17) were annotated as inflamed alveoli and 17 AOIs as normal-appearing alveoli within the same slides following review by board certified pathologists at the time of ROI selection (**Extended Data Fig. 9a**). Inflamed AOIs were selected based on histomorphology. Specifically, inflamed alveoli were defined as areas of alveolar tissue, excluding medium-sized bronchioles and large vessels, with intra-alveolar inflammatory cells ($CD45^+$ or $CD68^+$ cells) with or without fibroblastic foci. These areas generally also showed enlarged and detached type II pneumocytes and hyaline membrane formation. Normal-appearing alveoli were selected as areas of alveolar tissue from the same section. For reference, AOIs of bronchial epithelium (n=8) and of arterial vessels (n=5) were also selected. Q3-normalized count data were transformed to log_2-counts per million and fit into a generalized linear model (Expression ~ Donor + ROI Group). The donor-normalized

residuals were utilized to generate plots following dimensionality reduction (UMAP, PCA). Differential expression analysis between inflamed (n=50) and normal appearing alveoli (n=17) was performed by incorporating the $\log_2$-counts per million into a generalized linear model and using contrasts to identify significant comparisons. All models were fitted using the quasi-likelihood negative binomial generalized log-linear model implemented in edgeR[102] and TMM-normalized expression levels across ROIs were extracted for visualizations

Viral phylogenetic analysis

SARS-CoV-2 genomes were assembled using the open-source software viral-ngs version 2.0.21, implemented on the Terra platform (app.terra.bio), using workflows that are publicly available via the Dockstore Tool Registry Service (dockstore.org/organizations/BroadInstitute/collections/pgs). Briefly, reference-based assembly (assemble_refbased) was performed using the SARS-CoV-2 reference NC_045512.2. For multiplexed amplicon sequencing libraries, primer sites were trimmed prior to assembly. We constructed a phylogenetic maximum likelihood (ML) tree using the Augur pipeline (augur_with_assemblies) including 9 unique genomes assembled with greater than 90% genome coverage and 772 previously reported genomes from unique individuals from Massachusetts[103]. We used Kraken2[104] to search for the presence of 20 common respiratory viruses of interest (adenovirus, HCoV-229E, HCoV-HKU1, HCoV-NL63, betacoronavirus 1, parainfluenza 1, parainfluenza 2, parainfluenza 3, Parainfluenza 4, enterovirus A, enterovirus B, enterovirus C, enterovirus D, influenza A, influenza B, human metapneumovirus, respiratory syncytial virus, SARS-CoV, MERS-CoV, human rhinovirus). We used the classify_single and merge_metagenomics workflow with reference files as previously described[103].

Integrated COVID-19 scRNA-Seq and GWAS analysis

The COVID-19 GWAS data were obtained from COVID-19 Host Genetics Initiative's meta-analysis[45] (round 4 freeze, Oct 20, 2020) for two phenotypes: severe COVID-19 (B2: hospitalized COVID-19 versus the population) and COVID-19 (C2: COVID-19 versus population) (https://www.covid19hg.org/results/). Both these phenotypes had limited sample size, in particular the number of cases, but the severe COVID-19 phenotype showed significant non-zero heritability (B2: N.cases=8,638, N.controls= 1,736,547, h^2 (liability scale): 0.06, Z-score (h^2) = 4.8, and C2: N.cases= 30,937, N.controls= 1,471,815, h^2 (liability scale): 0.01, Z-score (h^2) = 2.7).

We identified 26 genes in a 5kb region to the 6 regions with genome-wide significant association signals (chromosomes 3p, 3q, 9, 12, 19 and 21 loci (https://www.covid19hg.org/results/), **Supplementary Information Table 11**), but excluded *ABO* from the reported results, because its expression cannot be reliably reported in sc/snRNA-Seq, given red blood cell contamination. Second, we identified 65 genes using SNP-set aggregation of GWAS signals using gene-level MAGMA with two different strategies to annotate SNPs to a gene - SNPs within the gene body (intron and exon) and SNPs in enhancers linked to genes using Roadmap[105] and ABC S2G strategies[106,107] (RoadmapUABC). To assess the significance for each gene, we performed a permutation test of MAGMA z-score for each gene with respect to MAGMA z-scores for all genes with expression in at least 1 cell in the lung tissue and then performed multiple testing correction (FDR < 0.05). We performed a supplementary analysis where the significance level is determined by a null model determined by a permutation test of each gene on all other genes expressed in at least 25 cells in the lung (84% of all genes). We observed 53 of the 65 genes to remain significant even after this permutation test.

Cell type gene programs were constructed from the scRNA-Seq data for each annotated cell type by performing a differential expression between the cell type on interest and the rest of the cells using a Wilcoxon rank sum test. Disease progression programs were constructed by a differential expression of disease and healthy cells of the same cell type, either at single-cell or pseudobulk level, as described in the "Healthy reference comparisons and differential expression" section earlier in the Methods. We also investigated gene programs based on genes that are differentially expressed in infected cells compared to uninfected, or genes that covary with the *ACE2* and *TMPRSS2* genes, but these programs showed no disease informative signal. We combined these cell type programs with the (RoadmapUABC). S2G strategy to generate SNP annotations. For each gene score X and S2G strategy Y, we define a combined annotation X × Y by assigning to each SNP the maximum gene score among genes linked to that SNP (or 0 for SNPs with no linked genes); this generalizes the standard approach of constructing annotations from gene scores using window-based strategies[108,109].

References (Methods)

51. Brook, O. R. *et al.* Feasibility and safety of ultrasound-guided minimally invasive autopsy in COVID-19 patients. *Abdom Radiol (NY)* (2020) doi:10.1007/s00261-020-02753-7.
52. Drokhlyansky, E. *et al.* The Human and Mouse Enteric Nervous System at Single-Cell Resolution. *Cell* vol. 182 1606–1622.e23 (2020).
53. Muñoz, J. F. *et al.* Coordinated host-pathogen transcriptional dynamics revealed using sorted subpopulations and single macrophages infected with Candida albicans. *Nat. Commun.* **10**, 1607 (2019).
54. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
55. Matranga, C. B. *et al.* Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. *Genome Biol.* **15**, 519 (2014).
56. MacConaill, L. E. *et al.* Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index

cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics* **19**, 30 (2018).

57. Metsky, H. C. *et al.* Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. *Nat. Biotechnol.* **37**, 160–168 (2019).
58. Lagerborg, K. A. *et al.* DNA spike-ins enable confident interpretation of SARS-CoV-2 genomic data from amplicon-based sequencing. *bioRxiv* (2021) doi:10.1101/2021.03.16.435654.
59. Tyson, J. R. *et al.* Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *Cold Spring Harbor Laboratory* 2020.09.04.283077 (2020) doi:10.1101/2020.09.04.283077.
60. Beigel, J. H. *et al.* Remdesivir for the Treatment of Covid-19 - Final Report. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2007764.
61. Kim, D. *et al.* The Architecture of SARS-CoV-2 Transcriptome. *Cell* **181**, 914–921.e10 (2020).
62. Yang, S. *et al.* Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol.* **21**, 57 (2020).
63. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
64. *harmony-pytorch*. (Github).
65. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
66. Adams, T. S. *et al.* Single Cell RNA-seq reveals ectopic and aberrant lung resident cell populations in Idiopathic Pulmonary Fibrosis. *Cold Spring Harbor Laboratory* 759902 (2019) doi:10.1101/759902.
67. Habermann, A. C. *et al.* Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science Advances* **6**, eaba1972 (2020).
68. Chua, R. L. *et al.* COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* **38**, 970–979 (2020).
69. 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).

70. Liao, M. *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
71. Subramanian, A. *et al.* RAAS blockade, kidney disease, and expression of ACE2, the entry receptor for SARS-CoV-2, in kidney epithelial and endothelial cells. *Cold Spring Harbor Laboratory* 2020.06.23.167098 (2020) doi:10.1101/2020.06.23.167098.
72. Jerby-Arnon, L. *et al.* A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* **175**, 984–997.e24 (2018).
73. Schiller, H. B. *et al.* The Human Lung Cell Atlas: a high-resolution reference map of the human lung in health and disease. *Am. J. Respir. Cell Mol. Biol.* **61**, 31–41 (2019).
74. Yeung, K. Y. & Ruzzo, W. L. Details of the adjusted rand index and clustering algorithms, supplement to the paper an empirical study on principal component analysis for clustering gene expression data. *Bioinformatics* **17**, 763–774 (2001).
75. Reyes, M. *et al.* An immune-cell signature of bacterial sepsis. *Nat. Med.* **26**, 333–340 (2020).
76. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
77. Welch, J. D. *et al.* Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* **177**, 1873–1887.e17 (2019).
78. Tucker, N. R. *et al.* Transcriptional and Cellular Diversity of the Human Heart. *Circulation* (2020) doi:10.1161/CIRCULATIONAHA.119.045401.
79. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281–291.e9 (2019).
80. Li, B. *Doublet detection in Pegasus*. https://github.com/klarman-cell-observatory/pegasus/blob/6c89b4d3e5116ecdc011e2e5010919ee9f2ee6c1/doublet_detection.pdf (2020).
81. Popescu, D.-M. *et al.* Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).
82. Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*

566, 490–495 (2019).

83. Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714–730.e22 (2019).
84. Büttner, M., Ostner, J., Müller, C. L., Theis, F. J. & Schubert, B. scCODA: A Bayesian model for compositional single-cell data analysis. *Cold Spring Harbor Laboratory* 2020.12.14.422688 (2020) doi:10.1101/2020.12.14.422688.
85. Brill, B., Amir, A. & Heller, R. Testing for differential abundance in compositional counts data, with application to microbiome studies. *arXiv [q-bio.GN]* (2019).
86. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat. Commun.* **10**, 380 (2019).
87. Xu, G. *et al.* Persistent viral activity, cytokine storm, and lung fibrosis in a case of severe COVID-19. *Clin. Transl. Med.* **10**, e224 (2020).
88. Litviňuková, M. *et al.* Cells of the adult human heart. *Nature* **588**, 466–472 (2020).
89. Lun, A. T. L. & Marioni, J. C. Overcoming confounding plate effects in differential expression analyses of single-cell RNA-seq data. *Biostatistics* **18**, 451–464 (2017).
90. Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. *Cold Spring Harbor Laboratory* 060012 (2019) doi:10.1101/060012.
91. Kotliar, D. *et al.* Single-Cell Profiling of Ebola Virus Disease In Vivo Reveals Viral and Host Dynamics. *Cell* (2020) doi:10.1016/j.cell.2020.10.002.
92. Cao, Y. *et al.* Single-cell analysis of upper airway cells reveals host-viral dynamics in influenza infected adults. *Cold Spring Harbor Laboratory* 2020.04.15.042978 (2020) doi:10.1101/2020.04.15.042978.
93. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
94. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
95. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

96. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).

97. Bost, P. *et al.* Host-Viral Infection Maps Reveal Signatures of Severe COVID-19 Patients. *Cell* 181, 1475–1488.e12 (2020).

98. Seabold, S. & Perktold, J. Statsmodels: Econometric and statistical modeling with python. in *Proceedings of the 9th Python in Science Conference* vol. 57 61 (Austin, TX, 2010).

99. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

100. Li, B. *A streamlined method for signature score calculation*. https://github.com/klarman-cell-observatory/pegasus/blob/6c89b4d3e5116ecdc011e2e5010919ee9f2ee6c1/signature_score.pdf (2020).

101. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

102. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

103. Lemieux, J. *et al.* Phylogenetic analysis of SARS-CoV-2 in the Boston area highlights the role of recurrent importation and superspreading events. *medRxiv* (2020) doi:10.1101/2020.08.23.20178236.

104. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).

105. Liu, Y., Sarkar, A., Kheradpour, P., Ernst, J. & Kellis, M. Evidence of reduced recombination rate in human regulatory domains. *Genome Biol.* **18**, 193 (2017).

106. Fulco, C. P. *et al.* Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).

107. Nasser, J. *et al.* Genome-wide maps of enhancer regulation connect risk variants to disease genes. *Cold Spring Harbor Laboratory* 2020.09.01.278093 (2020) doi:10.1101/2020.09.01.278093.

108. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

109. Zhu, X. & Stephens, M. Large-scale genome-wide enrichment analyses identify new trait-associated genes

and pathways across 31 human phenotypes. *Nat. Commun.* **9**, 4361 (2018).

110. Ravindra, N. G. *et al.* Single-cell longitudinal analysis of SARS-CoV-2 infection in human bronchial epithelial cells. *bioRxiv* (2020) doi:10.1101/2020.05.06.081695.

Supplementary Information Tables

Supplementary Information Table 1. Clinical meta-data for all donors and extended sample information.

(a) **Extended metadata** for donors included in the study, tab *donor_metadata*.

(b) A legend specifying which **donor is represented in each figure**, tab: *donors_in_figures*.

(c) **A list of samples with low cell/nucleus counts**, tab: *samples_low_cell_numbers*. Cells from these donor samples are still included in the atlases.

Supplementary Information Table 2. Studies included for building tissue meta-atlases and lung classifier coefficients.

(a) **Meta-atlas studies.** A list of studies used to build meta-atlases for cell type classification and lung differential expression analyses, their origin, and how they were used in this study. Tab: *datasets*.

(b) **Lung classifier coefficients.** Lung cell type classifier coefficients from one-vs-rest Logistic Regression model scoring each gene's contribution to classification decision. Tab: *coefficients*.

Supplementary Information Table 3. Differentially expressed genes within KRT8+/PATS cell subsets and between the IPBLP and basal cell subsets.

(a) **Differentially expressed genes within KRT8+/PATS cells.** We applied LIGER with 4 factors to the KRT8/PATS/ADI/DATPs (cluster 7 (light green) in **Fig. 2b**), assigned each cell to the factor with the greatest loading, and tested for differentially expressed genes (two-sided Mann–Whitney U test) across the 4 resulting clusters. Cluster 4 represents IPBLP cells (**Fig. 2f**, **Extended Data Fig. 5d**). The top 100 genes are ranked by their p-values ("p_val" column) within each cluster ("cluster" column); "p_val_adj" column includes Bonferroni adjusted p-values. Tab: *epithelial_cluster7_topDE_genes.*

(b) **Differentially expressed genes of the KRT8+/PATS subcluster of IPBLP cells vs. the basal cell subcluster of the airway epithelial cells.** We compare gene expression in the IPBLP cells to the basal cells in the airway epithelial cells (**Extended Data Fig. 3d**). Differentially expressed genes (two-sided Mann–Whitney U test) are ranked by their p-values with an FWER control of 0.05 ("p_val_adj" column). Tab: *DE_genes_IPLBP_vs_basal.*

Supplementary Information Table 4. Differentially expressed genes between healthy vs. COVID-19 lung (sc/snRNA-Seq) and between lung biopsies high vs. low/no viral load (bulkRNA-Seq).

(a) **Healthy *vs.* COVID-19 differentially expressed genes.** For differential expression, we used two linear regression models predicting gene expression levels as a function ofBoolean covariates for disease state (single cell and pseudobulk model), snRNA-Seq *vs*. scRNA-Seq (single cell and pseudobulk model) and 10x V3 kit *vs.* other kits (single cell and pseudobulk model) as well as a continuous covariate for UMI counts (single cell model) for each cell and a random intercept. We then applied multiple hypothesis correction with an error rate of alpha=0.05 to test which COVID-19 coefficients were significantly predictive of gene

expression levels in each model. Results for both the single cell and pseudobulk models are in tab *sc_sn_pseudobulk_DE*.

(b) **Genes upregulated or downregulated in bulk RNA-seq of lung biopsies with high *vs.* low/no viral load**. Statistical testing was completed with a negative binomial generalized linear model in DESeq2. FDR-corrected p-value ('padj'), Wald statistic ('stat'), and the log2 fold change are reported. Upregulated and downregulated genes are in tabs: *BULK_upreg_high_viral* and *BULK_downreg_high_viral*.

Supplementary Information Table 5. Analyses of SARS-CoV-2 RNA+ cells by cell type and viral sequencing

(a) **Cell-Type specific differentially expressed genes between SARS-CoV-2 RNA+ *vs.* SARS-CoV-2 RNA- cells**. Differential gene expression output statistics following comparison of SARS-CoV-2 RNA+ cells *vs*. SARS-CoV-2 RNA- cells. Only cell types with significant DE genes are listed (tabs *DE_genes_myeloid, DE_genes_epithelial, DE_genes_inflamm._monocytes, DE_genes_mac_PPARGhiCD5Lhi*). Statistical testing via the DESeq2 R package which implements a negative binomial generalized linear model. FDR, Wald statistic, and the log2fold change are reported for each comparison and each gene.

(b) **Viral enrichment score p-value, FDR and viral+ cell counts by cell type.** We report the p-value, FDR and the number of SARS-CoV2 RNA+ cells for main lineage cell types and for myeloid and endothelial sub-clusters, (for the seven donors for which we detected SARS-CoV-2+ cells). We report results for donors analyzed jointly (tab: *viral_enrichment_7donors*) and for each donor separately (tab: *viral_enrichment_by_donor*). P-values were computed by a permutation test and were adjusted by FDR.

(c) Viral qPCR and viral sequencing. Summary statistics of SARS-CoV-2 RT-qPCR and sequencing. Bulk RNA sequencing was performed on each sample, often in conjunction with targeted capture to enrich for SARS-CoV-2 sequences. The length of the unambiguous genome assembled and mean depth of coverage across it are reported, and highlighted in green if this genome constituted >90% unambiguous bps assembled. Multiplexed amplicon sequencing was also performed on select samples; unambiguous assembly lengths are reported where appropriate, and highlighted in green if the genome constituted >90% unambiguous bps assembled. Where multiple samples from the same donor assembled complete viral genomes, we address how they compared to each other and specify which genome was used in further analyses. Tab: *viral_qpcr_bulkRNAseq*.

(d) Metatranscriptomic count matrix. Matrix of counts for all taxonomic classifications across samples (25 samples, 2 negative controls) using Kraken2. A subset of viral species from this table - targeted by the hybrid capture panel - were used to enrich viral material in several samples (**Extended Data Fig. 6n**). Tab: *metatranscriptomics_counts*.

Supplementary Information Table 6. Spatial assay probe list and DE genes between COVID-19 and healthy donors

(a) **Genes and protein markers available in CTA, WTA and protein assays.** Tabs: *CTA_targets, WTA_targets, protein_targets*.

(b) **Differential expression genes between COVID-19 and healthy autopsy of both WTA and CTA data in alveolar AOIs.** Tabs with names ending in "*.up*" give significant up-regulated genes on subject data and tabs with names ending in "*.down*" list significant genes compared to healthy lung controls. Two-sided tests (via limma) were applied for differential expression analysis and false discovery rate was controlled at $\alpha = 0.05$. Only genes with adjusted p-value

less than 0.05 are listed; "*PanCK*" in tab names refers to PanCK$^+$ AOIs, while "*Syto13*" refers to PanCK$^-$ AOIs.

Supplementary Information Table 7. Differential expression results of snRNA-Seq based on the reconciled annotation. Each sheet lists up-regulated or down-regulated genes by comparing nuclei inside the corresponding cluster with those outside, under false discovery rate control at 0.05 using two-sided Mann-Whitney U test. Genes are sorted by AUROC values in descending order. These differentially expressed genes were used to select cell type markers for WTA and CTA deconvolution (listed in **Supplementary Information Table 8)**.

Supplementary Information Table 8. Cell type markers used for cell type deconvolution of WTA and CTA data in lung spatial data. Tabs: *cell_type_markers WTA* and *cell_type_markers_CTA*.

Supplementary Information Table 9. Spatial differentially expressed genes in SARS-CoV-2 high *vs.* low and inflamed *vs.* non-inflamed

(a) **Differential expression results and pathway enrichment analysis for COVID-19 lung.** We compared: normal-appearing alveoli *vs.* inflamed alveoli and normal-appearing alveoli *vs.* bronchial epithelium, for spatial transcriptomic samples of donors D13-17. Differential expression analysis between inflamed (n=50) and normal appearing alveoli (n=17) was performed by incorporating $\log_2$-counts per million into a quasi-likelihood negative binomial generalized log-linear model implemented in edgeR102 and using contrasts to identify significant comparisons. TMM-normalized expression levels across ROIs were extracted for visualizations. Tabs: *InflamedVsNormalAlveoliDE*, *InflamedVsNormalAlveoliGSEA*, *BronchialVsNormalAlveoliDE*, *BronchialVsNormalAlveoliGSEA*.

(b) **SARS-CoV-2 high *vs.* SARS-CoV-2 low.** Differential expression genes on COVID-19 donor tissues of WTA data in alveolar AOIs**.** Tabs with names ending in "*.up*" give significant up-regulated genes on SARS-CoV-2 high data and tabs with names ending in ".down" give significant genes on SARS-CoV-2 low data. Two-sided tests (via limma) were applied for differential expression analysis and false discovery rate was controlled at $\alpha = 0.05$. Only genes with adjusted p-value less than 0.05 are listed; "*PanCK*" in tab names refers to $PanCK^+$ AOIs, while "*Syto13*" refers to $PanCK^-$ AOIs.

Supplementary Information Table 10. COVID-19 heart differential expression analysis and GSEA results

(a) **Heart differential expression analysis results.** Differential expression analysis between COVID-19 and healthy left ventricle samples was computed using the limma + voom approach. "Cluster" column indicates the cell type tested. Columns "test.group" and "comparison" denote disease status. "*logFC*" denotes log2 fold-change, positive being up-regulated in COVID-19, and "*adj.P.Val*" is the Benjamini-Hochberg adjusted p-value. Column "*test.group.cell.mean.counts*" is the average count per cell (CellBender count data) of the given gene in the test group and cluster denoted. Column "*test.group.frac.expr>0*" gives the fraction of cells in the given cluster and test group with nonzero expression of the given gene. Column "*bkg.prob*" is a heuristic for the probability that a given result might be influenced by ambient RNA contamination. Columns that start with "*conservative_pseudobulk_*" provide an alternative analysis. Tab: *heart_DE_results*.

(b) **Heart gene set enrichment analysis results (GSEA).** GSEA was performed using the differential expression results (see **(a)**), testing MSigDB gene sets C5: GO Biological Process and C5: GO Molecular Function. Column "*cluster*" denotes the cell type, while "*test.group*" and

"*comparison*" denote the conditions being tested. Column "*pathway*" gives the full GO pathway name. Other columns are produced by fGSEA. Tab: *heart_GSEA_results*

Supplementary Information Table 11. COVID-19 GWAS gene list, differential expression analysis and Sc-linker heritability results

(a) **Gene descriptions for genes relevant to COVID-19 GWAS.** We list 26 genes (with gene descriptions) identified by the COVID-19 Human Genetics Initiative based on their proximity to GWAS peaks, and 65 genes identified by MAGMA analyses using 0kb and enhancer map S2G strategies. Tabs: *GWAS_curated_genes* and *MAGMA_genes*.

(b) **DE analysis of COVID-19 relevant genes for different program classes.** The Z-scores from differential expression analysis of cell type program (cells in a cell type *vs.* other cells) and DE analysis of disease progression program (cells in disease cell type *vs.* cells in healthy cell type) for the 26 curated GWAS genes and the 65 MAGMA implicated genes from the genes described in **(a)**. Tabs: *GWAS_genes_CellTypeProgram*, *GWAS_genes_DiseaseVsHealthy*, *MAGMA_genes_CellTypeProgram*, *MAGMA_genes_DiseaseVsHealthy*.

(c) **Sc-linker heritability results for gene programs.** Enrichment and p-value of enrichment for annotations derived from cell type and disease progression programs in four different tissues (lung, liver, kidney and heart) for COVID-19 and severe COVID-19 phenotypes. The annotations are generated by combining the gene programs with the Roadmap∪ABC enhancer S2G strategy. Results only reported for programs that attain nominal significance in enrichment (p value < 0.05). All analyses are conditional on the 86 baseline-LD (v2.1) model annotations. The enrichment p-value is computed using block jackknife. Tab: *sclinker_heritability_programs*.

(d) **Nominate functionally important genes in each cell type.** For each cell type, we report genes with high grade (> 0.8) in COVID-19/severe COVID-19 disease enriched gene program (in sc-linker analysis) that are additionally implicated by GWAS curated genes or MAGMA. Tab: *functionally_important_genes*.

Supplementary Information Table 12. Gene signatures used for high-level manual annotation of COVID-19 lung cells.

Supplementary Information Table 13. List of differential gene expression between sub-clusters of myeloid cells, T and NK cells, B and plasma cells, and endothelial cells in COVID-19 lung, as computed by Pegasus. Each tab includes the name of the sub-cluster tested, "*up*"in the tab name indicates upregulated genes and "down" in the tab name indicates downregulated genes. Differentially expressed genes were calculated using a two-sided Mann-Whitney U test.

Supplementary Information Table 14. List of top genes for LIGER factors of epithelial and fibroblast cells in COVID-19 lung. LIGER uses a non-negative matrix factorization approach that allows cells to be described as occupying multiple different cell states; we report the top 100 genes from shared LIGER factors each in epithelial cells and fibroblasts. Tabs: *epithelial*, *fibroblast*.