

Persistent alveolar type 2 dysfunction and lung structural derangement in post-acute COVID-19

André F. Rendeiro^{1,2,6,*}, Hiranmayi Ravichandran^{2,5*}, Junbum Kim¹,
Alain C. Borczuk^{3,7,Ω}, Olivier Elemento^{1,2,Ω}, Robert E. Schwartz^{4,Ω}

¹ Institute for Computational Biomedicine, Department of Physiology and Biophysics,
Weill Cornell Medicine, New York, NY, USA

² Caryl and Israel Englander Institute for Precision Medicine,
Weill Cornell Medicine, New York, NY, USA

³ Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA

⁴ Division of Gastroenterology and Hepatology, Department of Medicine, Weill Cornell Medicine, New York, NY, USA

⁵ Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

⁶ Current address: CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Lazarettgasse 14 AKH BT 25.3, 1090, Vienna, Austria

⁷ Current address: Northwell Health, Department of Pathology and Laboratory Medicine, Greenvale, NY

* Co-first authors

Ω Co-senior authors

Correspondence: res2005@med.cornell.edu

Keywords: COVID-19, PASC, infection biology, single-cell analysis, spatial-omics

Abstract

SARS-CoV-2 infection can manifest as a wide range of respiratory and systemic symptoms well after the acute phase of infection in over 50% of patients. Key questions remain on the long-term effects of infection on tissue pathology in recovered COVID-19 patients. To address these questions we performed multiplexed imaging of post-mortem lung tissue from 12 individuals who died post-acute COVID-19 (PC) and compare them to lung tissue from patients who died during the acute phase of COVID-19, or patients who died with idiopathic pulmonary fibrosis (IPF), and otherwise healthy lung tissue. We find evidence of viral presence in the lung up to 359 days after the acute phase of disease, including in patients with negative nasopharyngeal swab tests. The lung of PC patients are characterized by the accumulation of senescent alveolar type 2 cells, fibrosis with hypervascularization of peribronchial areas and alveolar septa, as the most pronounced pathophysiological features. At the cellular level, lung disease of PC patients, while distinct, shares pathological features with the chronic pulmonary disease of IPF. which may help rationalize interventions for PC patients. Altogether, this study provides an important foundation for the understanding of the long-term effects of SARS-CoV-2 pulmonary infection at the microanatomical, cellular, and molecular level.

Main text

Coronavirus disease 2019 (COVID-19) is now recognized as a systemic disease affecting the lung and several extrapulmonary organs^{1,2}. Its effects can be prolonged after acute infection developing into a syndrome of post-acute sequelae of COVID-19 (PASC)³⁻⁶. The cumulative frequency of PASC symptoms such as fatigue, cough, joint and chest pain, and dyspnea from patients that survive acute COVID-19⁷⁻¹⁰ is a considerable burden in loss of health and quality of life and impacts upwards of 50% of patients following infection^{11,12}. Recent studies of health records have tallied the cumulative burden of PASC symptoms, revealing increased risk for developing comorbidities systemically, in a manner dependent on the severity of the acute phase of disease¹¹. Blood profiling of PASC patients revealed lasting immune, serological and hormonal differences in comparison with convalescent COVID-19 patients without PASC¹³. However, the pathophysiological roots of the pervasiveness of dyspnea and other symptoms that manifest at the respiratory level in individuals suffering from post-acute COVID-19 is not well understood. We had previously observed increased fibrosis¹⁴ as well as widespread infection and injury of alveolar type 2 (AT-2) cells in the lungs of severe COVID-19 patients¹⁵. We now hypothesize that long-term sequelae of SARS-CoV-2 infection may share similarities with chronic lung disease such as usual interstitial pneumonia associated with idiopathic pulmonary fibrosis (UIP/IPF). Indeed, accumulating evidence suggests a role of AT-2 dysfunction driving fibrosis in UIP/IPF^{16,17}. To test this hypothesis, we used multiplexed imaging to conduct a multiscale comparative analysis of post-mortem lung samples of patients of acute COVID-19, patients that died after the acute phase of COVID-19, as well as post-mortem samples of IPF patients, thus seeking to uncover potentially shared pathophysiological features from the microanatomical to the single-cell level.

Persistent SARS-CoV-2 presence associated with AT-2 cell senescence

To shed light into the process of lung recovery after COVID-19, we investigated a cohort of 12 patients with respiratory symptoms who died incidentally after an acute period of COVID-19 (post-COVID, PC) (**Figure 1a, Supplementary Table 1**). We divided the PC group based on the last nasopharyngeal (NP) test before the time of death as positive (PC-pos) or negative (PC-neg). The PC-neg group of patients had a clinical course with a mean of 77 days, and a mean of 4 negative COVID-19 NP tests before death (range 2-8) (**Figure 1b**). The PC-pos group had a clinical course with a mean of 314 days, with an acute infection period in early 2020 (**Figure 1b**). As reference groups, we included samples from individuals who died without lung disease (n = 2), patients with UIP/IPF (n = 2), and patients who died during the acute phase of COVID-19 (n = 4) (**Figure 1a**). From this latter group, we included patients who died after 1 and 5 days (early COVID-19), and 35 and 36 days (late COVID-19) after symptom onset (**Figure 1a**).

To investigate the cellular and microanatomical pathophysiology of lung tissue from the patients, we designed a panel of 39 antibodies for imaging mass cytometry¹⁸ (**Supplementary Table 2**). We then generated 175 multiplexed images with 1 μm resolution, comprising 221 mm^2 of tissue, from which 396,343 single cells could be identified with high-confidence (**Figure 1a, Supplementary Figure 1**). To augment the cells in our images with structural context of the tissue, we delineated boundaries for microanatomical domains such as airways, vessels, muscle, and alveolar region in the multiplexed images (**Figure 1c, Supplementary Figure 2**). At the microanatomical level, we noticed significantly less space dedicated to gas exchange and blood circulation for all pathological sample groups including PC patients when compared with healthy lung (**Figure 1c-d, Supplementary Table 3**). We also noticed increased fibrosis (with the exception of early acute COVID-19), which we had previously reported¹⁴. Surprisingly, the level of microanatomical pathology in PC patients was comparable to both early and late groups of acute COVID-19 and UIP/IPF (**Figure 1c-d**).

To better understand the cellular changes in PC patients, we leveraged the similarity of marker

expression between single-cells to establish 43 cell clusters which we further categorized into 22 cell types depending on marker intensity and microanatomical context (**Figure 1e, Supplementary Figure 3**). The inclusion of markers specific to various types of epithelial cells in the lung such as SFTPA, SFTPC, and SCGB1A1, together with microanatomical annotation allowed us to resolve cellular identity to a higher degree than previously possible¹⁴, particularly for epithelial and endothelial cell types. We observed both COVID-induced changes for several cell types, as well as commonalities between acute COVID-19, post-COVID-19 or UIP/IPF (**Supplementary Figure 4**). For instance, neutrophils - which are a hallmark in lungs of early acute COVID-19 patients^{14,19,20} - are significantly higher in PC patients in comparison with late acute COVID-19 patients. Conversely, predominantly peribronchial, interstitial macrophages - which are the hallmark of late acute COVID-19^{14,15} - are significantly increased in all disease groups in comparison with healthy lung (**Figure 1f, Supplementary Figure 4**). Similar to the UIP/IPF group, we also observed increased density of fibroblasts in PC in comparison with healthy lung, as well as higher density of vascular endothelial cells predominantly in the airway walls and CD4+ T-cells. The latter two cell types were increased in PC not only in comparison to healthy lung, but also to acute COVID-19 and regardless of the status of last NP test of the patient (**Figure 1f, Supplementary Figure 4**). These observations suggest persistent changes in both the immune and structural compartments of the lung long after acute infection and associated with long-term post-acute COVID-19 sequelae.

In some samples we also observed a higher density of alveolar type 2 (AT-2) cells in PC, particularly in the PC-neg group (**Figure 1f, Supplementary Figure 4, Supplementary Table 3**). Since AT-2 cells are the cell type most infected by SARS-CoV-2 in the lung^{14,15}, and their injury leads to accumulation of dysfunctional cells¹⁵, we checked whether AT-2 cells of PC samples had detectable virus. Immunoreactivity for SARS-CoV-2 Spike protein was found most predominantly in surfactant C⁺ (SFTPC) AT-2 cells of samples from the PC-neg group (**Figure 1g-h, Supplementary Figure 5, Supplementary Figure 6a-b**), to a degree comparable to late COVID-19. Furthermore, in AT-2 cells predominantly of late COVID-19 and PC-neg, we also detected the highest levels of the inflammatory marker IL-6 (**Figure 1g**), which is upregulated in SARS-CoV-2 infected cells¹⁴.

SARS-CoV-2 infection is responsible for the injury and accumulation of dysfunctional AT-2 cells¹⁵, which are responsible to maintain and regenerate the alveolar niche of the lung²¹⁻²⁴. Accumulating evidence points to the depletion or senescence of AT-2 cells in IPF as responsible for the fibrotic phenotype of the pathology^{16,17}. We investigated whether AT-2 cells were bearing markers of cellular senescence. Indeed, the senescence-associated p16^{INK4A 25-27} and urokinase-type plasminogen activator receptor (uPAR)^{28,29} were significantly increased in AT-2 cells from late COVID-19, and PC-neg patients in a pattern matching the rate of SARS-CoV-2 infection in each group (**Figure 1i, Supplementary Table 3**). Furthermore, we also observed significant increases of the senescence markers for these patient groups in fibroblasts, mesenchymal and vascular endothelial cells (**Supplementary Figure 5b**). Finally, we also observed ectopic AT-2 cells in the lumen of airways and alveoli (**Figure 1e**). Such anomalous structural derangement has been observed in chronic lung disease as well^{30,31}.

To orthogonally validate the presence of SARS-CoV-2 Spike protein, we performed *in situ* hybridization, immunohistochemistry, and PCR of formalin-fixed paraffin embedded lung tissue for SARS-CoV-2. We observed that cell types with the highest signal in the imaging

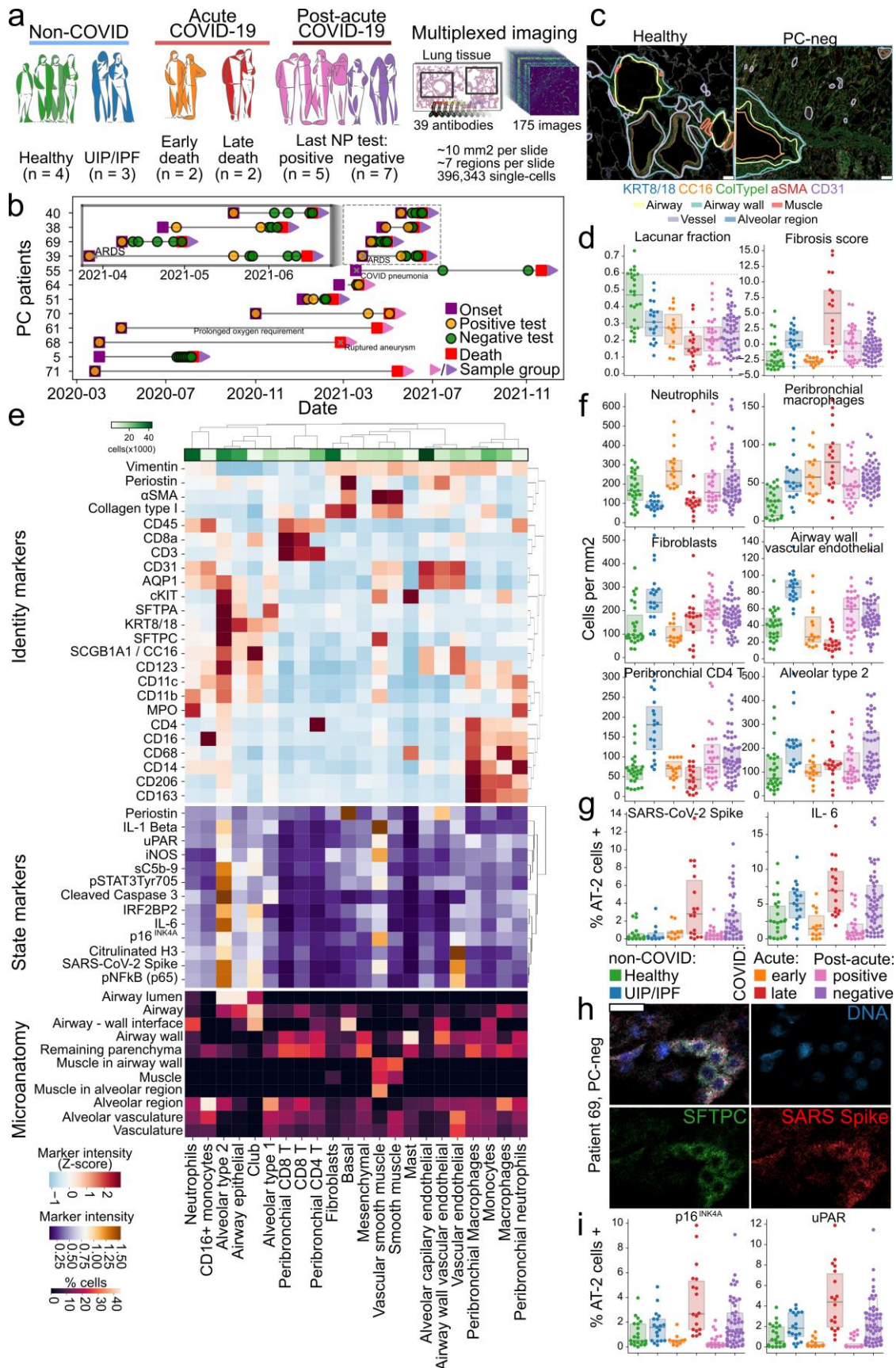


Figure 1: Long-term presence of SARS-CoV-2 in the lungs of post-acute COVID-19 patients.

a) Conceptual description of the cohort of patients in the study and process of acquisition of imaging mass cytometry data. **b)** Timelines of PC patients including major events during the courses of disease. **c)** Representative images of lung tissue imaged with imaging mass cytometry for a healthy lung sample (left), and PC-neg with outline of microanatomical domains overlaid. Scale bar represents 100 μ m. **d)** Quantification of the lacunar fraction (left), and fibrosis score (right) for each image grouped by sample

group. **e)** Clustered heatmap of cell types identified across all images in the IMC dataset. The upper part contains markers associated with cell type identity, the middle with cellular state, and the lower part of the heatmap represents the relative enrichment of each cell type in the delineated microanatomical domains of the lung. **f)** Absolute abundance (cell density per square millimeter) of selected cell types across sample groups. **g)** Percentage of AT-2 cells from different sample groups which are positive for SARS-CoV-2 Spike protein (left), or IL-6 (right). **h)** Illustration of infected AT-2 cells in the lung of a patient from the PC-neg group. The upper left image represents the merge of all remaining three. Scale bar represents 25 μm . **i)** Percentage of AT-2 cells from different sample groups which are positive for p16 (left), or uPAR (right).

data also had the largest difference between samples labeled as positive by the orthogonal methods, showing positive, significant agreement between multiplexed imaging and standard methods of tissue analysis (**Supplementary Figure 6c**). The agreement between orthogonal assays reinforces the validity of our observations of specific and persistent changes in microanatomy, cellular composition, and expression in the lungs of individuals who died post-acute COVID-19.

Derangement of vascular and epithelial lung structure in PC patients

We observed similarity in specific aspects of lung organization and composition between acute COVID-19 and patients who died post-acute COVID-19 (increased neutrophils and macrophages), as well between UIP/IPF and PC (increased fibroblasts and AT-2 injury) (**Figure 1d-g,i**). We sought to more systematically compare the disease entities by observing the global similarity between samples (**Figure 2a-b**, **Supplementary Figure 7**), and disease groups (**Figure 2c**). UIP/IPF and late COVID-19 were outstanding groups, whereas healthy lung, acute COVID-19 and PC were more similar to each other, with more heterogeneous groups (**Figure 2a-b**). We nonetheless observed that PC individuals whose last NP test was positive (PC-pos) were most globally similar to individuals with UIP/IPF, whereas the PC-neg group was more globally similar to the early COVID-19 group. This could be due to persistent active viral-driven disease or differences in time since index infection between the groups, which could have different levels of fibrosis.

To shed light into the features underlying the differences and similarities between disease groups, we first contrasted the cellular composition of the samples between them (**Figure 2d**). As previously described, we found increased density of peribronchial macrophages and monocytes in acute COVID-19 in comparison with healthy lung¹⁴, while vascular remodeling³², CD4+ T-cells³³ and mast cells³⁴ are hallmarks of IPF (**Figure 2d**, left). We found overall agreement between changes in cell composition of post-acute and acute COVID-19 with CD163⁺/CD206⁺ myeloid cells dominating changes for both disease groups in comparison to healthy lung (**Figure 2d**, center, **Figure 2e**). However, we found that vascular endothelial cells predominantly in airway walls were increased in density only in PC (**Figure 2d**, center), suggesting that as in IPF, vascular remodeling may be of pathophysiological relevance. Indeed, we observed not only widespread ectopic appearance of microvasculature but also significantly more smooth muscle cells in peribronchial regions and alveolar septa (**Figure 2f-g**). The appearance of both ectopic AT2-cells (**Figure 1d**) and microvasculature highlights the increased derangement of lung structure in PC. Furthermore, since neutrophil deployment of extracellular traps (NET) is known to drive vascular inflammation^{35,36} and neutrophils in PC remain increased in comparison with healthy lungs, we sought to quantify the amount of NETs in the lung. We observed that the total amount of extracellular citrullinated H3, a marker of neutrophil NETs³⁷ was significantly higher in PC when compared with healthy lung (**Figure 2h-i**). These results highlight the degree of long-term cellular and structural derangements in post-acute COVID-19 patients.

Lung healing through regeneration or repair is a complex process unfolding over the timescale of months³⁸. It is therefore plausible that patients who died during the acute phase of COVID-19 compared with post-acute patients have distinct features due to the continuous healing process, in addition to potential specific processes related with either acute or post-acute COVID-19. We sought to deconvolute time- and pathology-specific processes in our

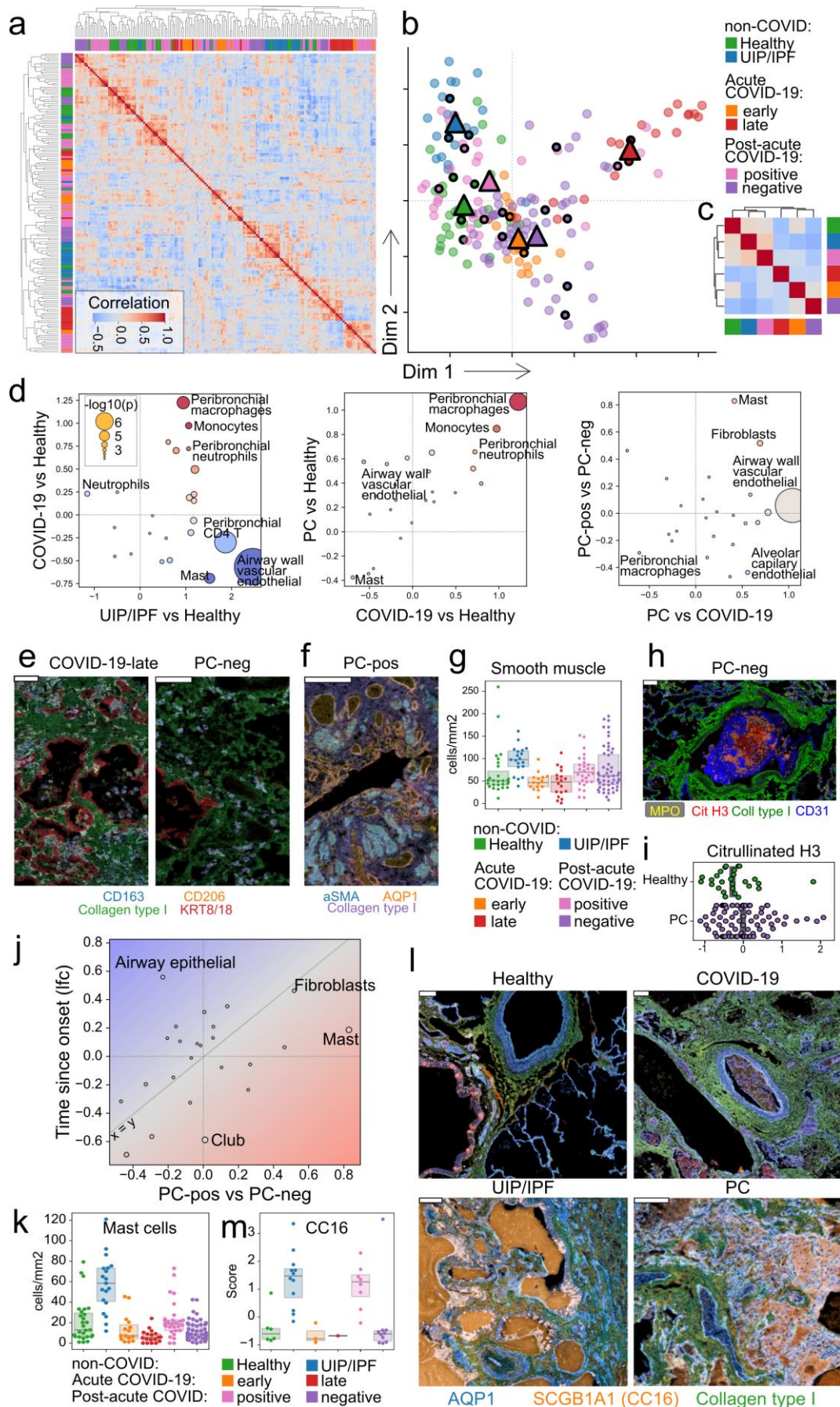


Figure 2: Comparative microanatomical analysis of lung identifies lasting structural derangement.
 a) Clustering heatmap of pairwise correlation between images based on the abundance of cell types. Image group identity is

marked by colors in rows and columns with legend in b). **b)** Two-dimensional projection of the same data as in a) using Spectral Embedding. **c)** Same as a) but grouped by disease group. **d)** Scatter plot of effect sizes estimated from comparing various disease groups. Point size is related to the significance of change, while color displays agreement between comparisons (agreement in red, disagreement in blue tones). **e)** Representative images of macrophages in acute COVID and PASC. **f)** Image highlighting vascular network and muscularization of lung tissue in PASC sample. **g)** Absolute abundance of smooth muscle cells in lung tissue from various disease groups. **h)** Exemplary image of blood clot where MPO⁺ neutrophils can be seen in the lumen with Citrullinated H3 from NETs in proximity. **i)** Quantification of extracellular citrullinated H3 signal in healthy lung and PASC. **j)** Scatter plot of estimated effect sizes of cell type abundance changes compared between patients in the PC-pos and -neg groups (x-axis), and within PC patients along the time since disease onset (y-axis). **k)** Absolute abundance of mast cells in the lungs of patients from the various disease groups. **l)** Representative images of secreted CC16 for various disease groups. AQP1 and Collagen shown as reference. **m)** Amount of extracellular CC16 (SCGB1A1) across disease groups for images containing airways.

multiplexed imaging data by focusing only on the group of post-acute COVID-19 patients and contrasting changes associated with the time since the acute phase of disease across all patients, and differences between the two PC subgroups (**Figure 2j**). This analysis revealed an overall agreement between the cellular changes associated with the two PC groups and the time since disease onset, with increases in fibroblasts over time consistent with the fibrotic-mediated scarring during lung repair. However, we also detected changes in cell abundance between PC groups which were independent of the time since symptom onset. Mast cells were more abundant in the PC-pos group than PC-neg independently of the time, in some cases to a similar degree as in UIP/IPF (**Figure 2k**). We also detected changes in the airway epithelium which were mostly driven by time rather than differences between PC groups. However, a common feature of IPF is the ectopic appearance of airway epithelial cells (bronchiolization), and their secretion of CC16/SCGB1A1 particularly in Club cells³⁹⁻⁴². We had observed that similarly to IPF, some images from PC patients showed extensive secreted CC16 in peribronchial regions (**Figure 2l**). We therefore quantified the amount of secreted SCGB1A1 and found levels in the PC-pos group comparable to IPF and unlike that seen in healthy lung or acute COVID-19 (**Figure 2m**). These observations highlight a globally aberrant microanatomical and cellular state of lungs of post-acute COVID-19 patients and identifies similarities with pathological features of UIP/IPF particularly for patients with prolonged active viral disease.

Discussion

Our analysis of post-mortem lung tissue from the microanatomical to single-cell level with multiplexed imaging provides a thorough examination of acute infection and host response, as well as the chronic pathological processes that manifest at long-term after acute infection.

In post-acute COVID-19 (PC) patients followed for a period of up to 399 days, we identified surprisingly persistent loss of lung lacunarity (as a proxy for pulmonary capacity), fibrosis, and perturbed immune and structural cellular components of the lung. Of particular note, we identified persistent presence of SARS-CoV-2 epitopes in the lung of PC patients, which did not necessarily agree with the status of the last nasopharyngeal test of the patients. It however agrees with recent reports showing viral factors present in circulation⁴³ and in the gut⁴⁴ of PASC patients. We also observed that SARS-CoV-2 presence is linked with the activation of senescence in AT-2, mesenchymal, endothelial, and fibroblast cells as detected by the upregulation of p16^{INK4A}, uPAR, and IL-6 which is part of the senescence associated secretory profile (SASP). SARS-CoV-2 induction of senescence in AT-2 cells had been suggested in acute disease⁴⁵⁻⁴⁷, but its long-term persistence after the acute disease raises the question of whether AT-2 senescence provides a reservoir for SARS-CoV-2 - something that would require future validation using viral replication assays. Furthermore, the fact that mesenchymal cells in addition to AT-2s display increased senescence suggests a “double-hit” for the maintenance of epithelial cells in the lung since they can promote AT-2 self-renewal^{48,49} and relieve AT-2 senescence⁵⁰.

Overall, we observed that PC cases showed distinct features in comparison to healthy lung or

acute COVID-19, with extensive derangement of the vascular network characterized by increased microvascularization and NET-induced vascular damage. Thrombosis as well as intussusceptive and sprouting angiogenesis have been found in acute COVID-19⁵¹ at higher levels than in Influenza patients. The fact that in our data we find the level of vascular derangement higher in post-acute than in acute COVID-19 suggests that vascular remodeling is a continuously acting process during post-acute COVID disease.

While UIP/IPF and sequelae of SARS-CoV-2 infection develop at largely different time scales (years and months respectively), we identified several common aspects between the diseases such as increased fibrosis and fibroblast abundance, vascular remodeling, increase of Mast cells, and upregulation of CC16/SCGB1A1 secretion. Mast cells play a pivotal role in IPF, with particular importance to the establishment of the fibrotic phenotype^{52,53}. Mast cell involvement in PASC has been suggested based on comparative analysis of symptomatic questionnaires⁵⁴ and molecular analysis of circulatory factors⁵⁵. The secreted Club cell marker CC16/SCGB1A1 is overexpressed in IPF⁴² but it was recently linked to negative regulation of the microbial response of alveolar macrophages. This feedback may however not function similarly in the host response to COVID-19 which is characterized by high influx of interstitial macrophages to the lung and may not be responsive to CC16 inhibition.

Overall, our comparative analysis suggests that while pathologically distinct entities, prolonged lung disease after acute COVID-19 may depend on the use of common features as IPF. Our study's use of tissue from fatal cases of post-acute COVID-19 patients may not allow complete extrapolation to the full diversity of the PASC syndrome. It does however represent an extremely timely and valuable resource to study lung pathology of acute and post-acute COVID-19 from the microanatomical to single cell level and highlights persistent pathological derangement of lung structure in patients with prolonged lung disease post-acute COVID-19.

Methods

Data reporting

No statistical methods were used to predetermine sample size due to the limited availability of samples from post-acute COVID-19 patients.

Tissue acquisition and preparation

Tissue samples were provided by the Weill Cornell Medicine Department of Pathology using protocols approved by the Tissue Procurement Facility of Weill Cornell Medicine. Experiments using samples from human subjects were conducted in accordance with local regulations and with the approval of the IRB at the Weill Cornell Medicine. The autopsy samples were collected under protocol 20-04021814.

Lung tissues were fixed via 10% neutral buffered formalin inflation, sectioned and fixed for 24 hours prior to processing and embedding into paraffin blocks. Freshly cut five micron sections were mounted onto charged slides.

Antibody panel design

We designed an antibody panel capturing different immune and structural compartments of the lung. Antibody clones were validated through immunofluorescence and chromogenic staining and verified by a pathologist. 100 µg of purified antibody in BSA and Azide free format was conjugated using the MaxPar X8 multimetal labeling kit (Fluidigm) as per the manufacturer's protocol. Sections were ablated on an Hyperion Imaging System (Fluidigm). To confirm the antibody binding specificity after conjugation and to identify the optimal dilution for each custom conjugated antibody, sections from the different patient tissues cohorts were stained. These sections were then ablated on Fluidigm Hyperion Imaging System and visualized using MCD Viewer for an expected staining pattern and optimal dilution that presented with good signal-to-noise ratio for each channel. For channels with visible spillover into the neighboring channels, a higher dilution factor was adopted when staining the cohort tissues.

Imaging mass cytometry

Freshly cut 5-micron thick FFPE sections were stored at 4°C for a day before staining. Slides were first incubated for 1 hour at 60°C on a slide warmer followed by dewaxing in fresh CitriSolv (Decon Labs) twice for 10 minutes, rehydrated in descending series of 100%, 95%, 80%, and 75% ethanol for 5 minutes each. After 5 minutes of MilliQ water wash, the slides were treated with antigen retrieval solution (Tris-EDTA pH 9.2) for 30 minutes at 96°C. Slides were cooled to room temperature (RT), washed twice in TBS and blocked for 1.5 hours in SuperBlock Solution (ThermoFischer), followed by overnight incubation at 4°C with the prepared antibody cocktail containing all metal-labeled antibodies in a humid chamber (**Supplementary Table 2**). Next day, slides were washed twice in 0.2% Triton X-100 in PBS, and twice in TBS. DNA staining was performed using Intercalator-Iridium in PBS solution for 30 minutes in a humid chamber at room temperature. Slides were washed with MilliQ water and air dried prior to ablation. The instrument was calibrated using a tuning slide to optimize the sensitivity of detection range. Hematoxylin and Eosin (H&E) stained slides were used to guide the selection of regions representative of the whole range of lung pathology. All ablations were performed with a laser frequency of 200 Hz. Tuning was performed intermittently to ensure the signal detection integrity was within the detectable range.

Immunohistochemistry

Automated immunohistochemistry on a Leica Bond III instrument was performed on five micron tissue sections using antibodies for myeloperoxidase (Clone 59A5, Leica ready to use antibody, without antigen retrieval) and CD163 (Clone MRQ-26, Leica ready to use antibody, antigen retrieval 20 minutes, high pH) using 3,3'-Diaminobenzidine chromogen. For each slide, a grid (5 x 5 grid, 0.4 cm x 0.4 cm boxes) was placed on the section and 5 alveolar, 2 vascular and 2 airway regions were randomly selected using random number generated X,Y coordinates and that region of interest (using a 20X objective) photographed.

Preprocessing of imaging mass cytometry data

Imaging mass cytometry data was preprocessed as previously described¹⁴ with some modifications. Briefly, MCD files exported from the Hyperion instrument were converted to OME-TIFF format. To perform cell segmentation, we trained a pixel classifier using ilastik⁵⁶ (version 1.3.3). We manually labeled pixels in random crops of images as belonging to one of three classes: nuclei – the area marked by signal in the DNA and Histone H3 channels; cytoplasm – the area immediately surrounding the nuclei and overlapping with signal in cytoplasmic channels; and background – pixels with low signal across all channels. The trained model was applied to the original images and output class probabilities for each pixel were then used to segment the images using DeepCell⁵⁷ (version 0.12.1).

Macroscopic assessment of imaging mass cytometry data

Calculation of the lacunar space in the images was done as previously described¹⁴. Briefly, we used the mean of all channels in each image stack after performing histogram equalization per channel and binarized the mean image with Otsu's method. We then performed morphological dilation and closing operations with a disk of 5 μm diameter, filled holes in the image, and removed objects smaller than 625 μm^2 (25×25). The lacunar fraction of the image is the fraction of the image with positive signal.

To quantify fibrosis, we used the same definition as previously described¹⁴, except that we used both the Collagen Type I and Periostin channels. The final score is the average of the scores for both channels weighted by the average intensity of the channels across images.

Annotation of microanatomical domains

To provide microanatomical context to the single-cells, we exported an JPEG image with the Aquaporin 1, Collagen type I, and alpha-smooth muscle actin channels, and used labelme⁵⁸ to draw polygons representing specific areas of lung microanatomy as done before^{59,60}. Polygons were converted into binary masks and intersected with the position of the cell segmentations for annotation of microanatomical domains per single-cell.

Cell type identification

To identify cell types in an unsupervised fashion, we first quantified the intensity of each channel in each segmented cell not overlapping image borders. In addition, for each cell we computed various morphological features. Values were log-transformed, scaled per image with a cap at -3 and 3, and globally scaled. We used Combat⁶¹ to remove technical signal associated with specific samples, performed Principal Component Analysis (PCA), built a k-nearest neighbor graph using batch-balanced k-nearest neighbors (bbknn)⁶², performed Uniform Manifold Approximation and Projection (UMAP)⁶³ (version 0.5.3), and finally clustered cells with the Leiden algorithm⁶⁴ (*leidenalg* package, version 0.8.10) with resolution 3.5. All steps were performed with Scanpy⁶⁵ (version 1.9.1). Each cluster was labeled with a cell type label based on the abundance of signal from the channels and the enrichment in each microanatomical domain. Only cell types with total abundance of at least 1% of total are displayed.

Dimensionality reduction

We used various dimensionality reduction and manifold learning methods such as Multidimensional Scaling, Principal Component Analysis, Spectral Embedding, and Isomap from the scikit-learn package (version 1.0.2). The inputs were cell type abundance measurements for each region of interest where the abundance was normalized in percentage of total cells found per observation, and by area acquired. The two matrices were then concatenated, centered and scaled. Furthermore, in order to compare images with distinct microanatomical context, we regressed out the composition of each image based on the microanatomical annotations using the “regress_out” function in Scanpy. The same resulting matrix was also used for computation of pairwise correlations between images.

Statistical testing

Unless otherwise mentioned, we used pairwise Mann-Whitney U-tests from the Pingouin package⁶⁶ (version 0.5.2) to compare the abundance of cell types, positivity of markers, and abundance of interactions between groups of images. Reported effect size values are the Hedges g value. P-values were then adjusted for multiple testing with the Benjamini-Hochberg false discovery rate (FDR) method⁶⁷. The results of testing are available in **Supplementary Table 3**.

Data availability

Multiplexed imaging data are available at the following publicly available repository: <https://doi.org/10.5281/zenodo.7060271>

Code availability

Source code used in the analysis is available at the following publicly available repository: <https://github.com/ElementoLab/post-covid-imc>

Acknowledgements

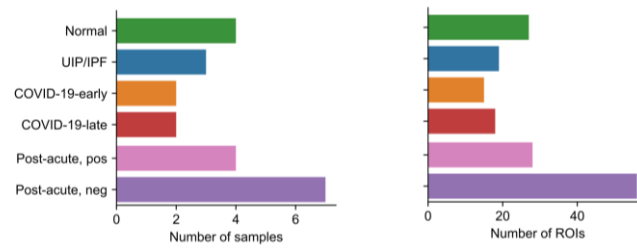
We thank the patients and their families for their contributions to our study. This work was partially supported by the WorldQuant Initiative for Quantitative Prediction. A.F.R. was supported by a NCI T32CA203702 grant. O.E. is supported by Volastra, Janssen and Eli Lilly research grants, NIH grants UL1TR002384 and R01CA194547, and Leukemia and Lymphoma Society SCOR 7012-16, SCOR 7021-20 and SCOR 180078-02 grants. R.E.S. is supported by NIH grants NIAID 2R01AI107301 and NIDDK R01DK121072 and the American Heart Association. R.E.S. is supported as Irma Hirschl Trust Research Award Scholar. We thank the Englander Institute of Precision Medicine Mass Cytometry Core facility for processing, labeling and acquiring samples for IMC imaging, and the translational research laboratory (Department of Pathology, Weill Cornell Medicine) for histology and IHC support.

Competing Interests

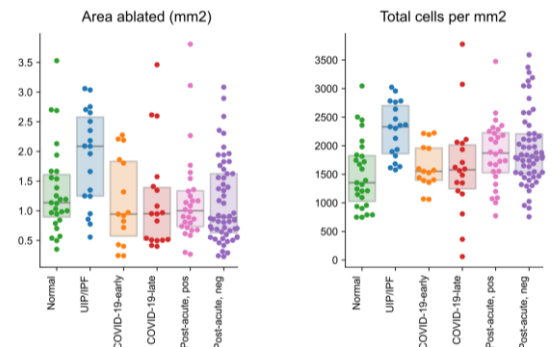
R.E.S. is on the scientific advisory board of Miromatrix Inc and Lime Therapeutics and is a paid consultant and speaker for Alnylam Inc. O.E. is scientific advisor and equity holder in Freenome, Owkin, Volastra Therapeutics and OneThree Biotech. The remaining authors declare no competing financial interests.

Supplementary Figures

a

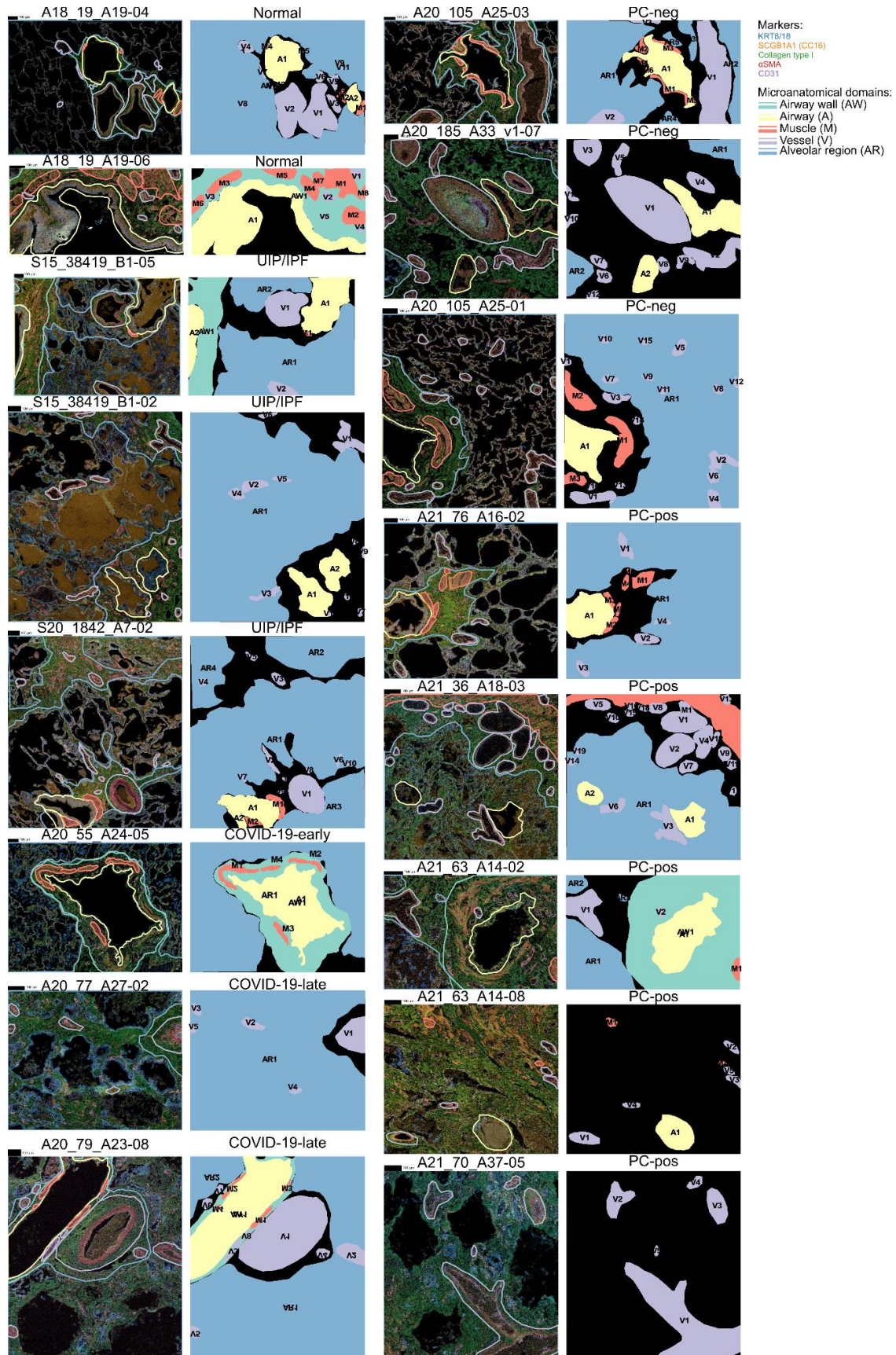


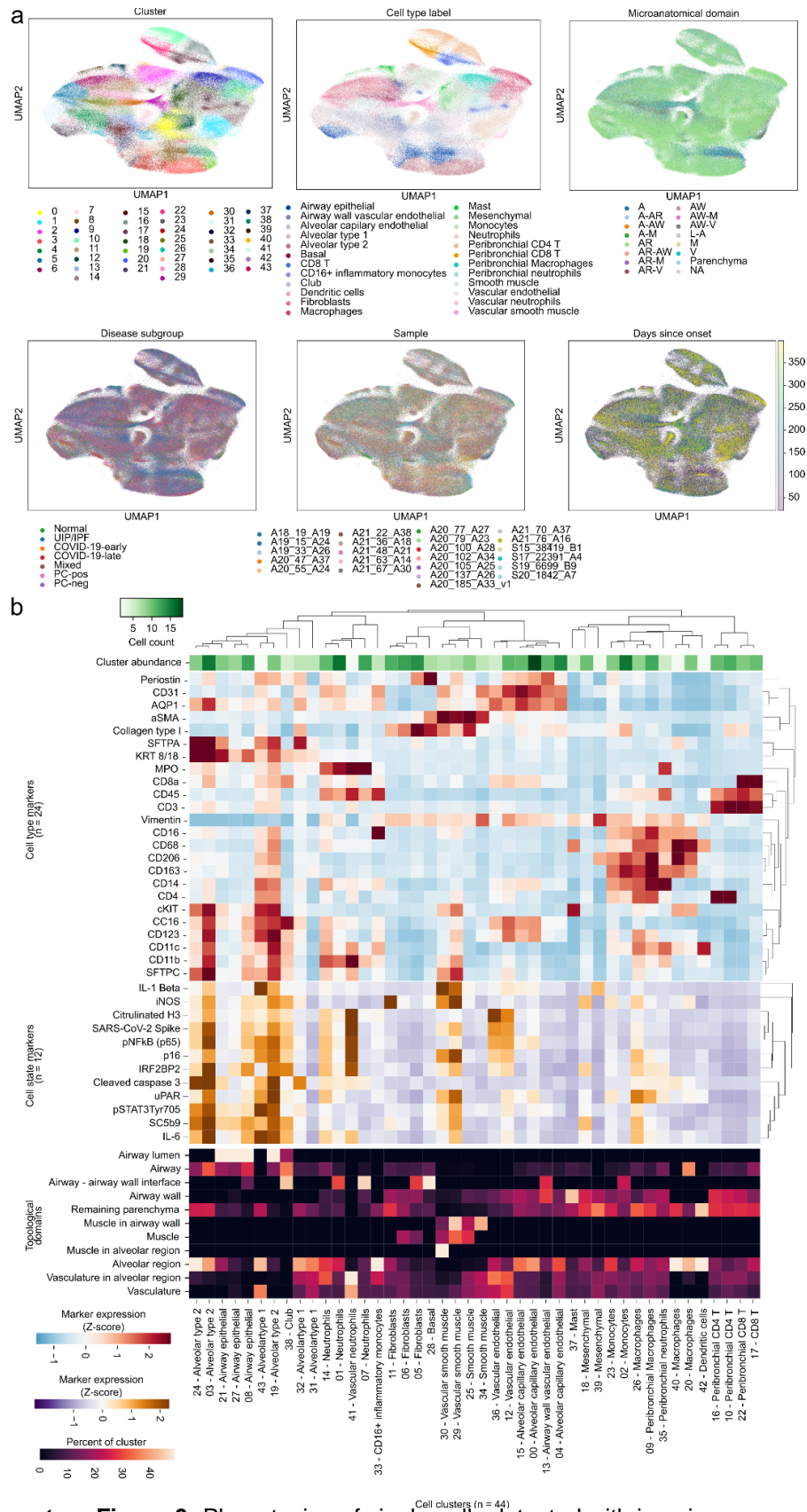
b



Supplementary Figure 1: Technical properties of acquired multiplexed image data.

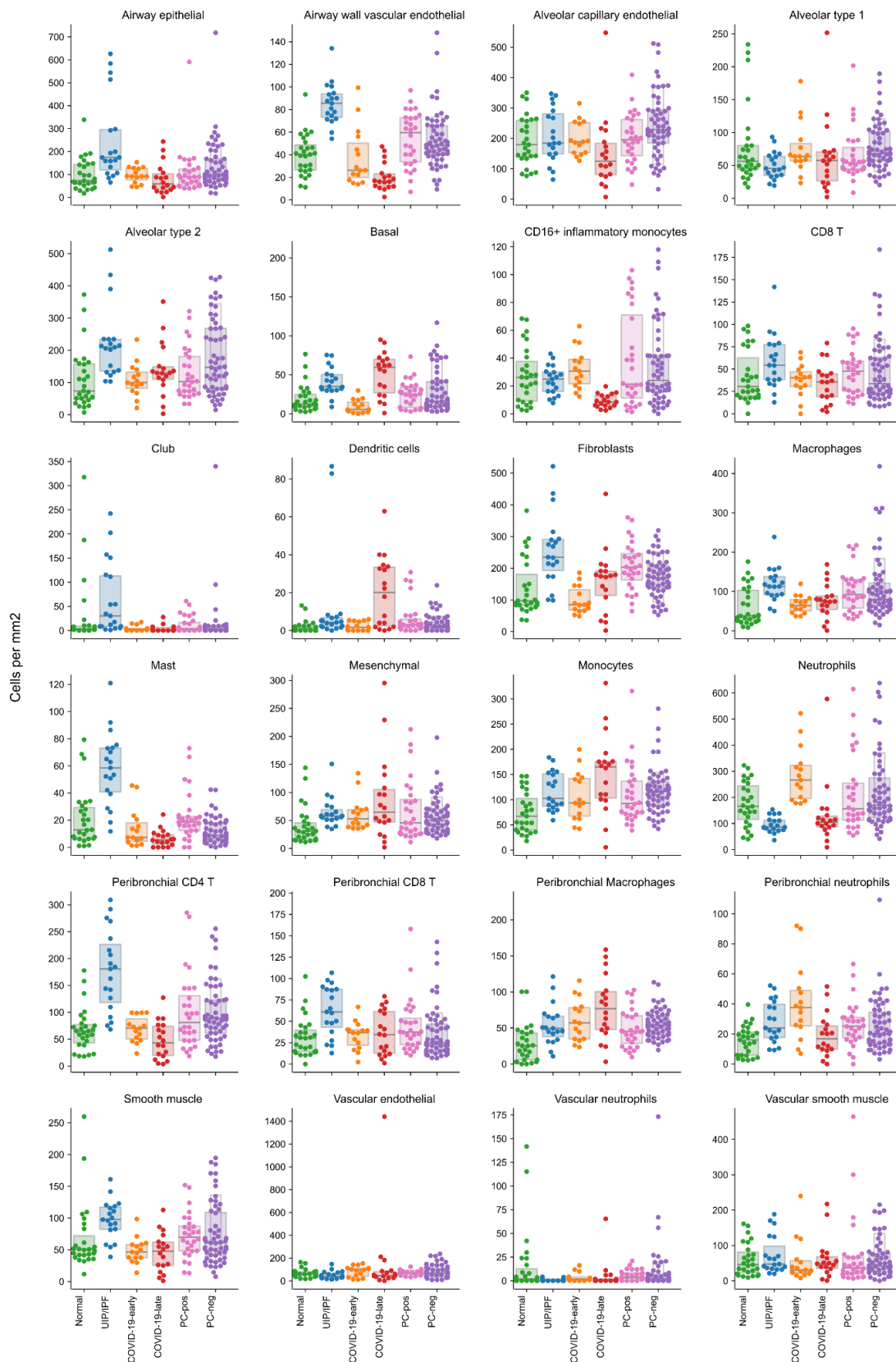
a) Distribution of number of samples or regions of interest (ROI) per sample group. b) Image size (area ablated with imaging mass cytometry) (left), and cellular density (right) of images.



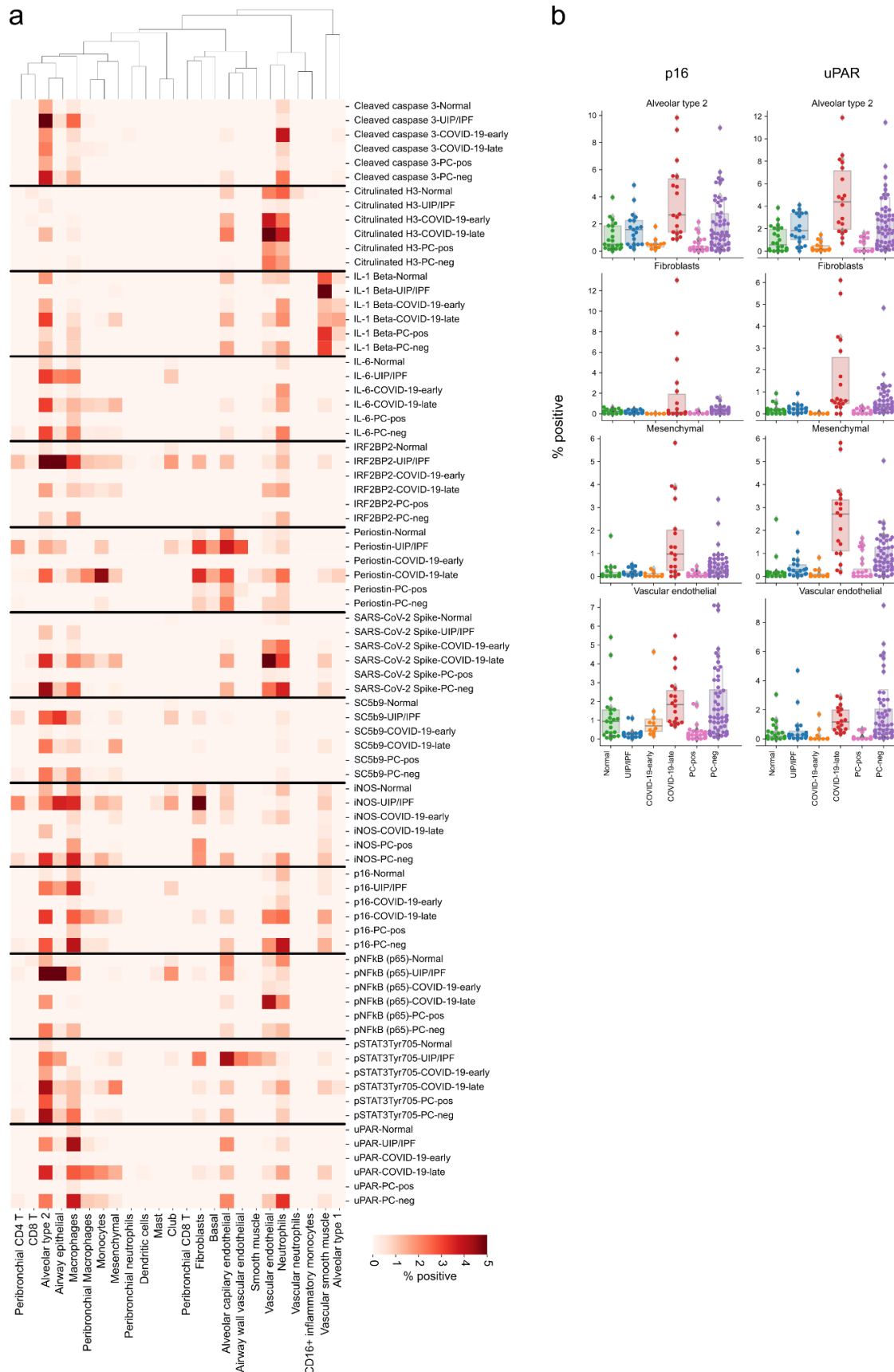


Supplementary Figure 3: Phenotyping of single-cells detected with imaging mass cytometry.

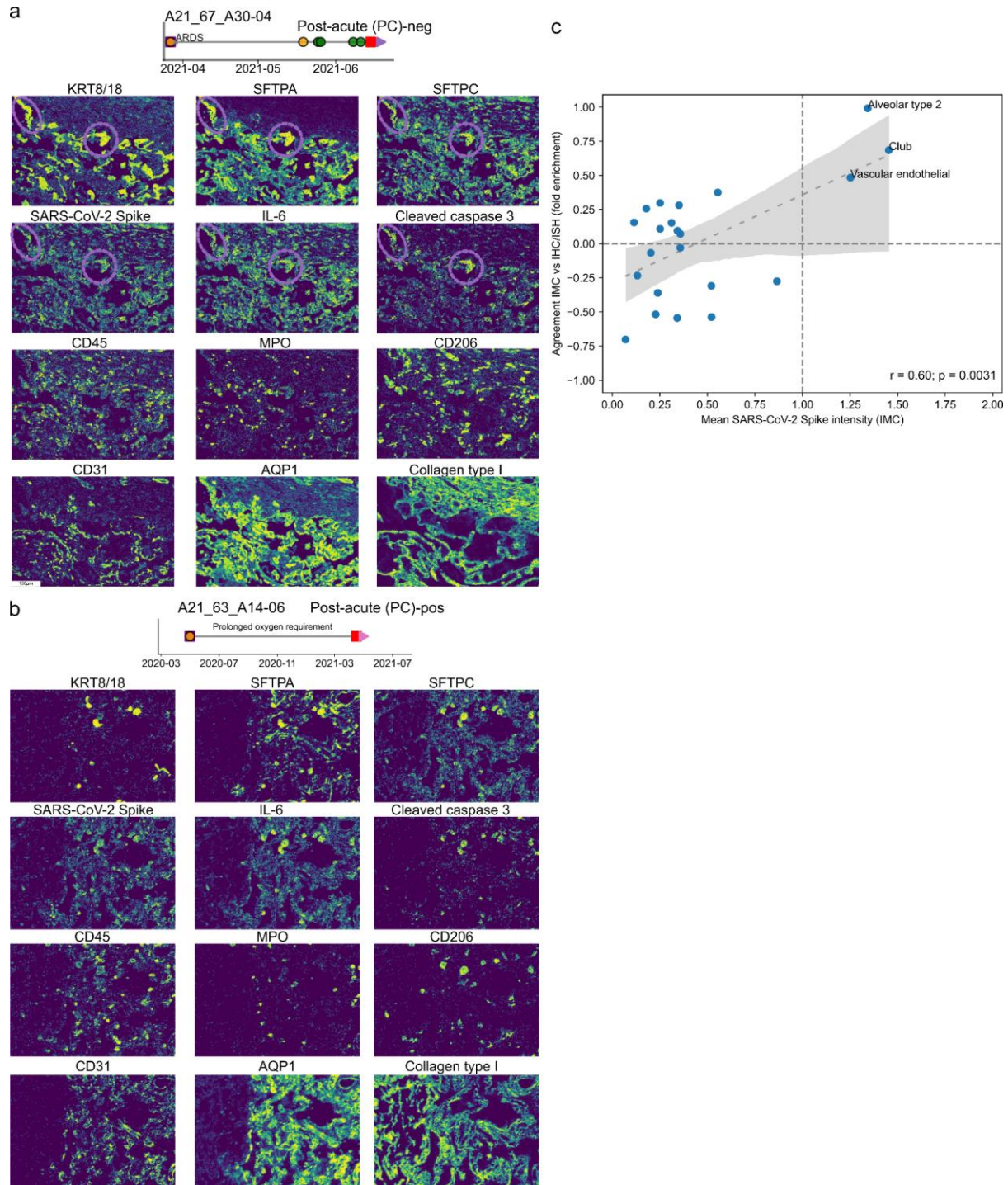
a) UMAP plot of single-cells colored by the clinical attributes of its sample, the microanatomical context of the cell, and by the clusters. b) Clustered heatmap of the discovered 44 clusters with the top part representing the average intensity of cell state markers, and the bottom the relative intensity of enrichment in microanatomical structures for each cell cluster.



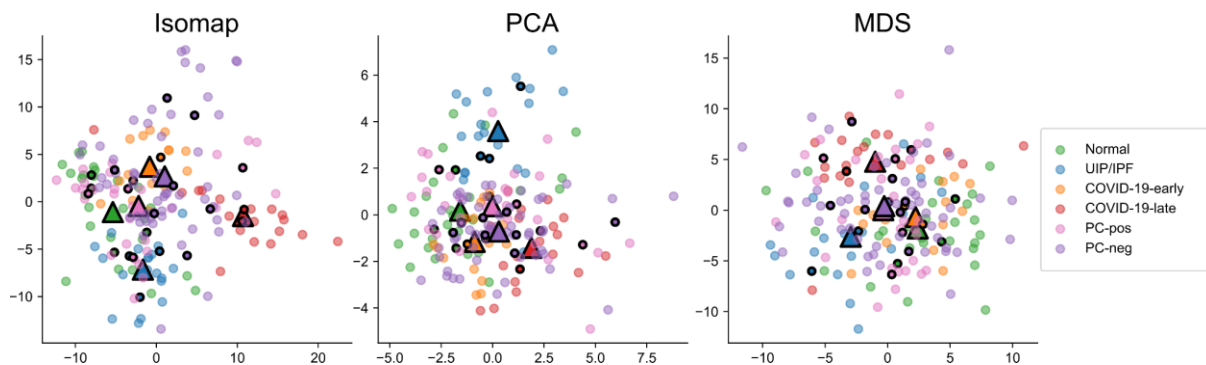
Supplementary Figure 4: Changes in cell type abundance between sample groups.
Cell density (cells per square millimeter) of each cell type for each sample group.



Supplementary Figure 5: Changes in expression of cell state markers between sample groups. Percentage of cells positive for a marker for each cell type in each sample group.



Supplementary Figure 6: Detection and validation of SARS-CoV-2 proteins in tissue.
a-b) Selected images illustrating SARS-CoV-2 signal in lung tissue from a PC-neg (a) and PC-pos (b) patient. **c)** Agreement between orthogonal methods and multiplexed imaging in quantification of SARS-CoV-2 detection.



Supplementary Figure 7: Global similarity between sample groups given by various dimensionality reduction methods.

For Principal Component Analysis only the first two dimensions are shown.

Supplementary Tables

Supplementary Table 1. Demographic and clinical details of samples included in analysis.

ID	Group	Fever	Cough	Shortness of breath	Days intubated	Days of disease	Days in hospital	Days since onset to death	PLT/ mL	D-dimer (mg/L)	WBC	LY %	PMN %	CRP mg/L	ESR mm/h
3	COVID-19-early	Y	Y	Y	0	5	0		NA	NA	NA	NA	NA	NA	NA
10	COVID-19-early	Y	Y	Y	0	11	5		400	33122	9.5	23	68	24.6	NA
28	COVID-19-late	Y	N	Y	32	36	36		348	4848	15.8	0.9	97.1		
30	COVID-19-late	39	Y	Y	29	35	35		126	5128	11.5	0.5	91.7		
64	PC				30	30	30	64	213	270	5.66	16	78.5	16.07	71
38	PC-neg	38.2	N	Y		44		44	212	754	8.3	17	74.3	1.3	NA
39	PC-neg		Y	Y	post covid	post covid	post covid	81	271	3996	9.3	3.9	86.5	4.1	
40	PC-neg	N	N	N	post covid	post covid	post covid	29							
5	PC-neg	36	N	N	28	post covid	28	120	304	1860	35.7	43	42	NA	NA
55	PC-neg				NA	NA	NA	200	NA	NA	NA	NA	NA	NA	NA
51	PC-neg	Y	Y	Y	30	50	43	42	300	3816	16.3	5.5	90.1	15.6	na
69	PC-neg	Y	N	N	0	5	22	22	269	2307	9.8	6.7	83		66
68	PC-pos	Y	Y	N	0	NA	0	350	361	1067	3.8	12	74.6	15	121
61	PC-pos	Y	Y	Y	NA	NA	NA	350							
70	PC-pos	N	Y	Y	22	40	30	155	150	18538	6.4	12	82	5.27	72
71	PC-pos	N	N	Y	0	18	18	399	165	5460	8.28	28	58	7	100

Supplementary Table 2. Markers used in imaging mass cytometry analysis of lung tissue.

Epitope	Metal tag	Antibody vendor	Antibody catalog	Antibody clone	Antibody lot	Antibody dilution
Histone H3	In113	Cell Signaling Technology	4499BF	D1H2	17	1:400
aSMA	Pr141	Fluidigm	3141017D	1A4	651904	1:400
iNOS	Nd142	Abcam	ab239990	SP126	GR3288157-2	1:50
cKIT	Nd143	Novus Biologicals	NBP2-71076	OT12B12	F001	1:50
CD206	Nd144	Cell Signaling Technology	91992BF	E2L9N	2	1:100
uPAR	Nd145	Abcam	ab251557	EPR20191	GR3377419-3	1:50
CD16	Nd146	Abcam	ab215977	EPR16784	GR3377429-4	1:50
CD163	Sm147	Invitrogen	MA1-82342	EDHu-1	UL2904541	1:100
CD14	Nd148	Cell Signaling Technology	56082BF	D7A2T	2	1:50
CD11b	Sm149	Abcam	ab238794	SP331	GR3245427-1	1:50

Epitope	Metal tag	Antibody vendor	Antibody catalog	Antibody clone	Antibody lot	Antibody dilution
p16	Nd150	Abcam	ab282285	EPR24167-43	GR3395434-2	1:50
CD31	Eu151	Abcam	ab207090	EPR3094	GR3229164-12	1:50
CD45	Sm152	Cell Signaling Technology	13917BF	D9M8I	8	1:100
SARS-CoV-2 Spike	Eu153	GeneTex	GTX135356	Rabbit Polyclonal	43957	1:50
Citrullinated H3	Sm154	Abcam	ab5103	Rabbit polyclonal	GR3430937-1	1:50
SC5b9	Gd155	Quidel Corporation	A239	Murine monoclonal	153280	1:25
CD4	Gd156	Abcam	ab181724	EPR6855	GR3285644-12	1:50
pSTAT3Tyr705	Gd158	Cell Signaling Technology	9145BF	D3A7	39	1:50
IRF2BP2	Tb159	Novus Biologicals	NBP2-81995	Polyclonal	5601-1001	1:50
IL-6	Gd160	Novus Biologicals	NBP2-71027	OTI3G9	F004	1:25
Periostin	Dy161	Millipore Sigma	MABS1183	5B2.3	NA	1:50
CD8a	Dy162	eBioscience	14-0085-82	C8/144B	2132595	1:100
SCGB1A1(CC16)	Dy163	BioVendor	RD181022220-01	Mouse Polyclonal	NA	1:200
AQP1	Dy164	Abcam	ab178352	EPR11588(B)	GR3330184-1	1:100
pNFkB (p65)	Ho165	BD Pharmingen	pS529	558393	9052785	1:25
IL-1 Beta	Er166	Cell Signaling Technology	12242BF	3A6	2	1:50
CD123	Er167	Novus Biologicals	NBP2-79851	IL3RA/2947R	3563-4PABX210615	1:50
Ki67	Er168	Fluidigm	3168022D	B56	2201558-10	1:50
Collagen type I	Tm169	Fluidigm	3169023D	Polyclonal	2501807	1:400
CD3	Er170	Fluidigm	3170019D	Polyclonal. C-Terminal	2202221-16	1:100
SFTPA	Yb171	Abcam	ab51891	6F10	GR216448-14	1:150
Cleaved caspase 3	Yb172	Fluidigm	3172027D	5A1E	412006	1:50
MPO	Yb173	Cell Signaling Technology	14569BF	E1E7I	2	1:100
KRT8/18	Yb174	Fluidigm	3174022D	C51	0372007	1:100
SFTPC	Lu175	Abcam	ab90716	Rabbit polyclonal	GR3392370-1	1:200
CD11c	Yb176	Abcam	ab216655	EP1347Y	GR3340299-2	1:50
DNA1	Ir191	Fluidigm	201192B	DNA intercalator		1:300
DNA2	Ir193	Fluidigm	201192B	DNA intercalator		1:300
Vimentin	Pt195	Cell Signaling Technology	5741BF	D21H3	9	1:100
CD68	Pt196	Invitrogen	14-0688-82	KP1	2342153	1:100

Supplementary Table 3. Statistical testing for differences between sample groups. Testing was performed for differences in lacunarity, fibrosis, cellular abundance and marker positivity. For cell type abundance comparisons values tested are absolute frequencies (cells per square millimeter). The table is available as an external file to the manuscript.

References

1. Mehta, O. P., Bhandari, P., Raut, A., Kacimi, S. E. O. & Huy, N. T. Coronavirus Disease (COVID-19): Comprehensive Review of Clinical Presentation. *Front Public Health* **8**, 582932 (2020).
2. Gupta, A. *et al.* Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).
3. Choutka, J., Jansari, V., Hornig, M. & Iwasaki, A. Unexplained post-acute infection syndromes. *Nat. Med.* **28**, 911–923 (2022).
4. Dryden, M. *et al.* Post-COVID-19 condition 3 months after hospitalisation with SARS-CoV-2 in South Africa: a prospective cohort study. *Lancet Glob. Health* **10**, e1247–e1256 (2022).
5. Bahmer, T. *et al.* Severity, predictors and clinical correlates of post-COVID syndrome (PCS) in Germany: A prospective, multi-centre, population-based cohort study. *EClinicalMedicine* **51**, 101549 (2022).
6. Nalbandian, A. *et al.* Post-acute COVID-19 syndrome. *Nat. Med.* **27**, 601–615 (2021).
7. Carfi, A., Bernabei, R., Landi, F. & Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* **324**, 603–605 (2020).

8. Chopra, V., Flanders, S. A., O'Malley, M., Malani, A. N. & Prescott, H. C. Sixty-Day Outcomes Among Patients Hospitalized With COVID-19. *Ann. Intern. Med.* **174**, 576–578 (2021).
9. Halpin, S. J. *et al.* Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: A cross-sectional evaluation. *J. Med. Virol.* **93**, 1013–1022 (2021).
10. Garrigues, E. *et al.* Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. *J. Infect.* **81**, e4–e6 (2020).
11. Al-Aly, Z., Xie, Y. & Bowe, B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* **594**, 259–264 (2021).
12. Groff, D. *et al.* Short-term and Long-term Rates of Postacute Sequelae of SARS-CoV-2 Infection: A Systematic Review. *JAMA Netw Open* **4**, e2128568–e2128568 (2021).
13. Klein, J. *et al.* Distinguishing features of Long COVID identified through immune profiling. *medRxiv* 2022.08.09.22278592 (2022) doi:10.1101/2022.08.09.22278592.
14. Rendeiro, A. F. *et al.* The spatial landscape of lung pathology during COVID-19 progression. *Nature* **593**, 564–569 (2021).
15. Melms, J. C. *et al.* A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).
16. Parimon, T., Yao, C., Stripp, B. R., Noble, P. W. & Chen, P. Alveolar Epithelial Type II Cells as Drivers of Lung Fibrosis in Idiopathic Pulmonary Fibrosis. *Int. J. Mol. Sci.* **21**, (2020).
17. Yao, C. *et al.* Senescence of Alveolar Type 2 Cells Drives Progressive Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* **203**, 707–717 (2021).
18. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
19. Zuo, Y. *et al.* Prothrombotic autoantibodies in serum from patients hospitalized with COVID-19. *Sci. Transl. Med.* **12**, (2020).
20. Veras, F. P. *et al.* SARS-CoV-2-triggered neutrophil extracellular traps mediate COVID-19 pathology. *J. Exp. Med.* **217**, (2020).
21. Barkauskas, C. E. *et al.* Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
22. Hogan, B. L. M. *et al.* Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. *Cell Stem Cell* **15**, 123–138 (2014).
23. Zacharias, W. J. *et al.* Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* **555**, 251–255 (2018).
24. Nabhan, A. N., Brownfield, D. G., Harbury, P. B., Krasnow, M. A. & Desai, T. J. Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* **359**, 1118–1123 (2018).
25. Tumor suppressors and oncogenes in cellular senescence☆. *Exp. Gerontol.* **35**, 317–329 (2000).
26. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras Provokes Premature Cell Senescence Associated with Accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).
27. Stein, G. H., Drullinger, L. F., Soulard, A. & Dulić, V. Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts. *Mol. Cell. Biol.* **19**, 2109–2117 (1999).
28. Smith, H. W. & Marshall, C. J. Regulation of cell signalling by uPAR. *Nat. Rev. Mol. Cell Biol.* **11**, 23–36 (2010).
29. Amor, C. *et al.* Senolytic CAR T cells reverse senescence-associated pathologies. *Nature* **583**, 127–132 (2020).
30. Watanabe, N. *et al.* Anomalous Epithelial Variations and Ectopic Inflammatory Response in Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Cell Mol. Biol.* (2022) doi:10.1165/rcmb.2021-0555OC.
31. Adams, T. S. *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* **6**, eaba1983 (2020).
32. Gaikwad, A. V. *et al.* Vascular remodelling in idiopathic pulmonary fibrosis patients and its detrimental effect on lung physiology: potential role of endothelial-to-mesenchymal transition. *ERJ Open Research* **8**, (2022).
33. Adegunsoye, A. *et al.* Skewed Lung CCR4 to CCR6 CD4 T Cell Ratio in Idiopathic Pulmonary Fibrosis Is Associated with Pulmonary Function. *Front. Immunol.* **7**, 516 (2016).
34. Kawanami, O., Ferrans, V. J., Fulmer, J. D. & Crystal, R. G. Ultrastructure of pulmonary mast cells in patients with fibrotic lung disorders. *Lab. Invest.* **40**, 717–734 (1979).
35. Gómez-Moreno, D., Adrover, J. M. & Hidalgo, A. Neutrophils as effectors of vascular inflammation. *Eur. J. Clin. Invest.* **48 Suppl 2**, e12940 (2018).
36. Pieterse, E. *et al.* Neutrophil Extracellular Traps Drive Endothelial-to-Mesenchymal Transition.

- Arterioscler. Thromb. Vasc. Biol.* (2017) doi:10.1161/ATVBAHA.117.309002.
37. Leshner, M. *et al.* PAD4 mediated histone hypercitrullination induces heterochromatin decondensation and chromatin unfolding to form neutrophil extracellular trap-like structures. *Front. Immunol.* **3**, 307 (2012).
 38. Lucas, A., Yasa, J. & Lucas, M. Regeneration and repair in the healing lung. *Clin Transl Immunology* **9**, e1152 (2020).
 39. Lederer, D. J. & Martinez, F. J. Idiopathic Pulmonary Fibrosis. *N. Engl. J. Med.* **378**, 1811–1823 (2018).
 40. Reynaud, P. *et al.* Club cell loss as a feature of bronchiolization in ILD. *Front. Immunol.* **12**, 630096 (2021).
 41. Stancil, I. T. *et al.* Pulmonary fibrosis distal airway epithelia are dynamically and structurally dysfunctional. *Nat. Commun.* **12**, 1–11 (2021).
 42. Zuo, W.-L. *et al.* Dysregulation of club cell biology in idiopathic pulmonary fibrosis. *PLoS One* **15**, e0237529 (2020).
 43. Craddock, V. *et al.* Persistent Circulation of Soluble/EV-Linked Spike Protein and Viral RNA in Individuals with Post-Acute Sequelae of COVID-19. (2022) doi:10.2139/ssrn.4186787.
 44. Zollner, A. *et al.* Postacute COVID-19 is Characterized by Gut Viral Antigen Persistence in Inflammatory Bowel Diseases. *Gastroenterology* **163**, 495–506.e8 (2022).
 45. Tripathi, U. *et al.* SARS-CoV-2 causes senescence in human cells and exacerbates the senescence-associated secretory phenotype through TLR-3. *Aging* **13**, 21838–21854 (2021).
 46. Evangelou, K. *et al.* Pulmonary infection by SARS-CoV-2 induces senescence accompanied by an inflammatory phenotype in severe COVID-19: possible implications for viral mutagenesis. *Eur. Respir. J.* **60**, 2102951 (2022).
 47. Tsuji, S. *et al.* SARS-CoV-2 infection triggers paracrine senescence and leads to a sustained senescence-associated inflammatory response. *Nat Aging* **2**, 115–124 (2022).
 48. Zepp, J. A. *et al.* Distinct Mesenchymal Lineages and Niches Promote Epithelial Self-Renewal and Myofibrogenesis in the Lung. *Cell* **170**, 1134–1148.e10 (2017).
 49. Taghizadeh, S. *et al.* Characterization in mice of the resident mesenchymal niche maintaining AT2 stem cell proliferation in homeostasis and disease. *Stem Cells* **39**, 1382–1394 (2021).
 50. Lai, X. *et al.* Mesenchymal stromal cells attenuate alveolar type 2 cells senescence through regulating NAMPT-mediated NAD metabolism. *Stem Cell Res. Ther.* **13**, 12 (2022).
 51. Ackermann, M. *et al.* Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N. Engl. J. Med.* NEJMoa2015432 (2020) doi:10.1056/NEJMoa2015432.
 52. Overed-Sayer, C. *et al.* Inhibition of mast cells: a novel mechanism by which nintedanib may elicit anti-fibrotic effects. *Thorax* **75**, 754–763 (2020).
 53. Veerappan, A. *et al.* Mast cells: a pivotal role in pulmonary fibrosis. *DNA Cell Biol.* **32**, 206–218 (2013).
 54. Weinstock, L. B. *et al.* Mast cell activation symptoms are prevalent in Long-COVID. *Int. J. Infect. Dis.* **112**, 217–226 (2021).
 55. Wechsler, J. B., Butuci, M., Wong, A., Kamboj, A. P. & Youngblood, B. A. Mast cell activation is associated with post-acute COVID-19 syndrome. *Allergy* **77**, 1288–1291 (2022).
 56. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).
 57. Bannon, D. *et al.* DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes. *Nat. Methods* (2021) doi:10.1038/s41592-020-01023-0.
 58. Wada, K. *labelme: Image Polygonal Annotation with Python (polygon, rectangle, circle, line, point and image-level flag annotation)*. (Github).
 59. Rustam, S. *et al.* A Unique Cellular Organization of Human Distal Airways and Its Disarray in Chronic Obstructive Pulmonary Disease. *bioRxiv* 2022.03.16.484543 (2022) doi:10.1101/2022.03.16.484543.
 60. Kim, J. *et al.* Unsupervised discovery of tissue architecture in multiplexed imaging. *bioRxiv* 2022.03.15.484534 (2022) doi:10.1101/2022.03.15.484534.
 61. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
 62. Polański, K. *et al.* BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964–965 (2020).
 63. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* vol. 3 861 Preprint at <https://doi.org/10.21105/joss.00861> (2018).

64. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* vol. 9 Preprint at <https://doi.org/10.1038/s41598-019-41695-z> (2019).
65. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
66. Vallat, R. Pingouin: statistics in Python. *JOSS* **3**, 1026 (2018).
67. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).

Figure 0 **bioRxiv** preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission. [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

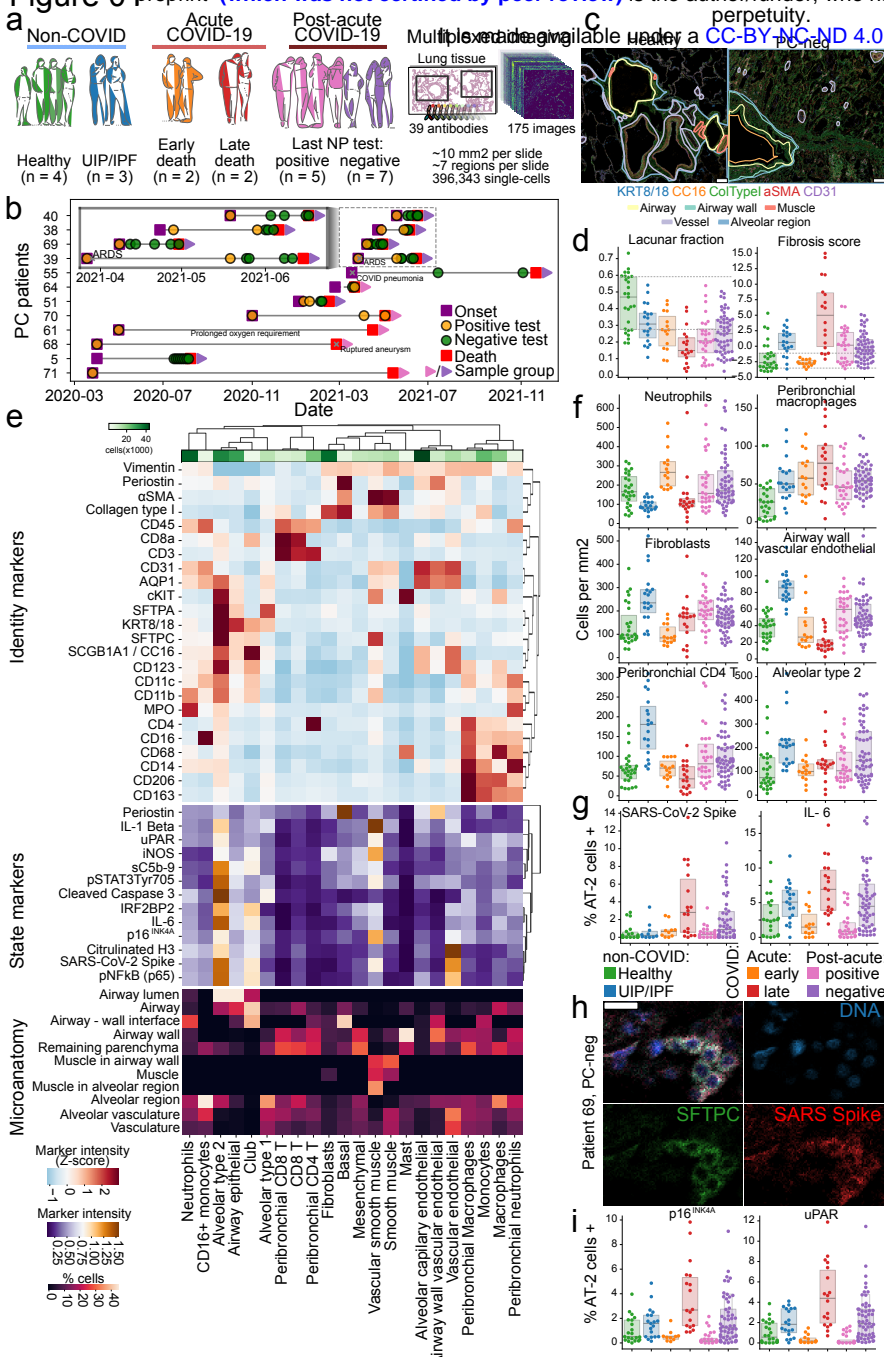


Figure 02

