

The spatial landscape of lung pathology during COVID-19 progression

<https://doi.org/10.1038/s41586-021-03475-6>

Received: 25 October 2020

Accepted: 19 March 2021

Published online: 29 March 2021

 Check for updates

André F. Rendeiro^{1,2,10}, Hiranmayi Ravichandran^{2,3,4,10}, Yaron Bram⁵, Vasuretha Chandar⁵, Junbum Kim¹, Cem Meydan^{1,4}, Jiwoon Park⁴, Jonathan Foox^{1,4}, Tyler Hether⁶, Sarah Warren⁶, Youngmi Kim⁶, Jason Reeves⁶, Steven Salvatore⁷, Christopher E. Mason^{1,3,4,8}, Eric C. Swanson⁹, Alain C. Borczuk^{7,11}, Olivier Elemento^{1,2,3,4,11} & Robert E. Schwartz^{4,5,11}

Recent studies have provided insights into the pathology of and immune response to COVID-19^{1–8}. However, a thorough investigation of the interplay between infected cells and the immune system at sites of infection has been lacking. Here we use high-parameter imaging mass cytometry⁹ that targets the expression of 36 proteins to investigate the cellular composition and spatial architecture of acute lung injury in humans (including injuries derived from SARS-CoV-2 infection) at single-cell resolution. These spatially resolved single-cell data unravel the disordered structure of the infected and injured lung, alongside the distribution of extensive immune infiltration. Neutrophil and macrophage infiltration are hallmarks of bacterial pneumonia and COVID-19, respectively. We provide evidence that SARS-CoV-2 infects predominantly alveolar epithelial cells and induces a localized hyperinflammatory cell state that is associated with lung damage. We leverage the temporal range of fatal outcomes of COVID-19 in relation to the onset of symptoms, which reveals increased macrophage extravasation and increased numbers of mesenchymal cells and fibroblasts concomitant with increased proximity between these cell types as the disease progresses—possibly as a result of attempts to repair the damaged lung tissue. Our data enable us to develop a biologically interpretable landscape of lung pathology from a structural, immunological and clinical standpoint. We use this landscape to characterize the pathophysiology of the human lung from its macroscopic presentation to the single-cell level, which provides an important basis for understanding COVID-19 and lung pathology in general.

SARS-CoV-2 is the coronavirus that causes COVID-19, which has become a global pandemic: as of February 2021, over 100 million people have been infected and there have been more than 2 million fatalities^{10,11}. A growing body of evidence indicates that the severity of COVID-19 is driven by an inflammatory syndrome caused by hyperactivation of the immune system^{8,12} in an attempt to clear the virus. Persistent inflammation can result in damage to lung tissue¹³, the exudation of pulmonary-oedema fluid that leads to dyspnoea, and acute respiratory distress syndrome (ARDS)^{14,15}. Immune profiling in peripheral blood^{1–3,5,7,16} or bronchoalveolar lavage fluid¹⁷ have revealed major changes in the immune system; excessive neutrophil activation¹⁸, lymphopenia³ and aberrant responses of the adaptive immune system² are among the most prominent changes. However, thorough analysis of infected tissue and the immune system in a spatial context has only recently been started^{4,19–21} and is currently lacking for most infected organs, including the lung. Although most patients with severe COVID-19 develop ARDS,

administering routine clinical supportive care as for other ARDS does not entirely aid in patient recovery. The degree to which SARS-CoV-2 infection and the immune response to COVID-19 resemble or differ from other insults in the lung is therefore unclear. To elucidate the cellular composition, spatial context and interplay between immune and structural cell types during SARS-CoV-2 infection in the lung, we performed imaging mass cytometry (IMC) in post-mortem lung tissue from patients with COVID-19 or with other lung infections that cause ARDS, and in otherwise-healthy individuals.

Pathophysiology of lungs in patients with COVID-19

We investigated a cohort of 23 patients that included individuals who died with ARDS after influenza ($n=2$) or bacterial infection ($n=4$), with acute bacterial pneumonia ($n=3$) or with COVID-19 respiratory distress syndrome ($n=10$), as well as individuals who died without lung disease

¹Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA. ²Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA.

³WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA. ⁴Department of Physiology, Biophysics and Systems Biology, Weill Cornell Medicine, New York, NY, USA. ⁵Division of Gastroenterology and Hepatology, Department of Medicine, Weill Cornell Medicine, New York, NY, USA. ⁶NanoString Technologies, Inc, Seattle, WA, USA. ⁷Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA. ⁸New York Genome Center, New York, NY, USA. ⁹Fluidigm Inc., Markham, Ontario, Canada. ¹⁰These authors contributed equally: André F. Rendeiro, Hiranmayi Ravichandran. ¹¹These authors jointly supervised this work: Alain C. Borczuk, Olivier Elemento, Robert E. Schwartz. ^{e-mail:} alb9003@med.cornell.edu; ole2001@med.cornell.edu; res2025@med.cornell.edu

Article

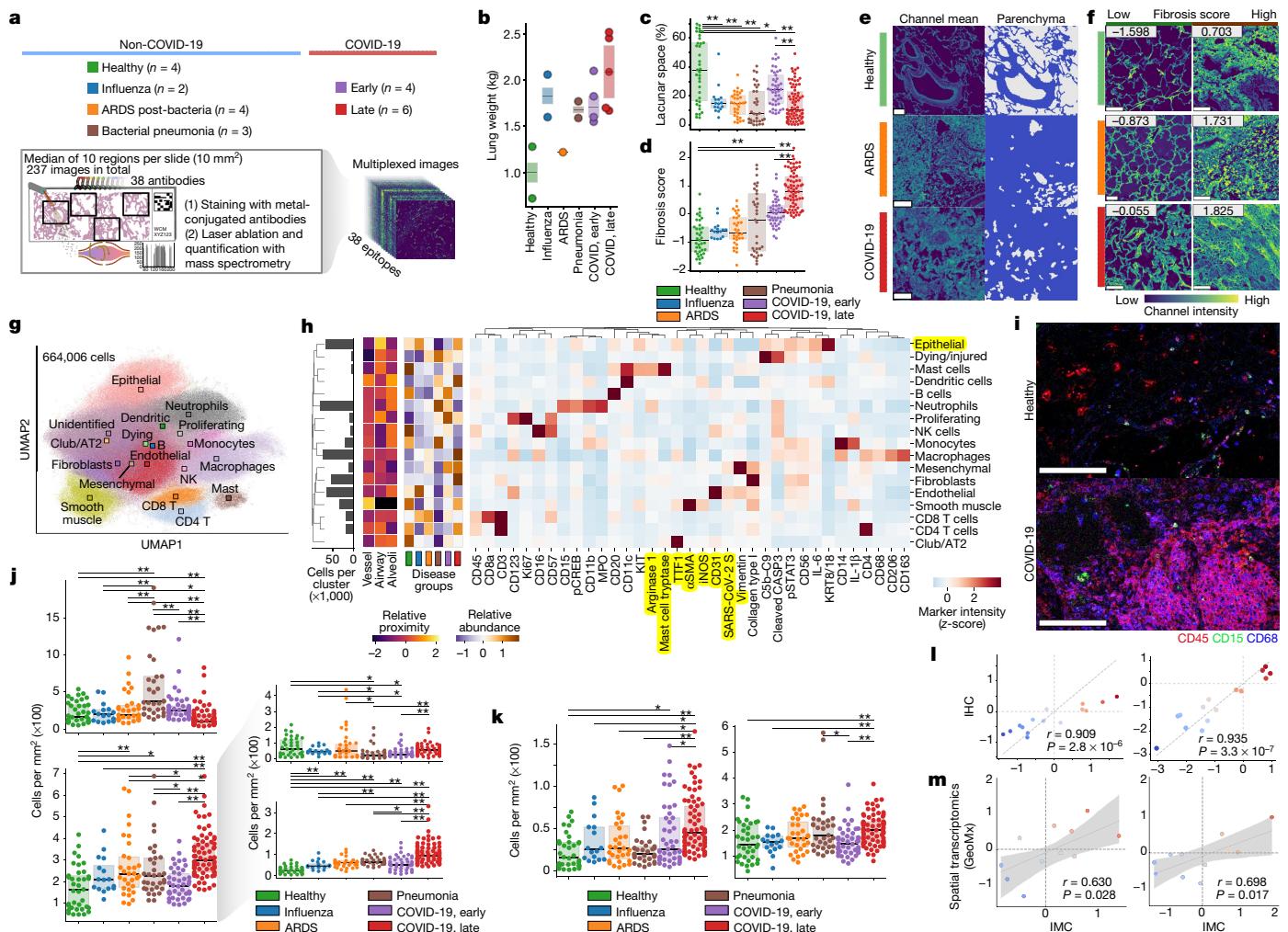


Fig. 1 | Structural and immunological disorder of lung infection.
a, Composition of lung-infection cohort, and schematic procedure to acquire highly multiplexed spatially resolved data with IMC from post-mortem lung samples. **b**, Total lung weight per disease group measured at autopsy. $n=16$ biologically independent samples. **c**, Lacunar space for each acquired IMC image as a percentage of image area. **d**, Fibrosis score for each acquired IMC image. **e**, Representative images illustrating the lacunar and parenchymal structure of healthy lungs, and lungs from patients with ARDS or COVID-19. **f**, Collagen in images from healthy lungs and lungs from patients with ARDS or COVID-19 and the associated fibrosis score. Images with lowest and highest fibrosis scores are depicted. **g**, Uniform manifold approximation and projection (UMAP) of all cells, and the metacluster of each cell. Centroids are shown as squares. **h**, Mean intensity of each marker in each metacluster. Histogram indicates metacluster abundance. Heat maps on left indicate relative proximity to lung structures or abundance per disease group. AT2, alveolar type 2 cells; KRT8/8, KRT8 and KRT18; NK, natural killer. **i**, Spatial distribution of immune cells in healthy lungs and lungs from a patient with COVID-19. **j**, Left, abundance of neutrophils (top) and macrophages (bottom) in each disease group. Right, macrophages divided into alveolar (top) and interstitial (bottom) subsets. **k**, Abundance of mesenchymal cells (left) and fibroblasts (right) in each disease group. **l**, Amount of change (effect size) pairwise between all disease groups ($n=15$) in MPO (left) and CD163 (right) markers between IMC (x axis) and immunohistochemistry (IHC) (y axis). **m**, Amount of change between late and early COVID-19 groups, pairwise for each cell type ($n=24$), as estimated by IMC (x axis) and targeted spatial transcriptomics (y axis) for the same (left) and independent (right) cohorts. For **c**, **d**, **j**, **l**, **m**, $n=237$ images from 27 biologically independent samples. $**P<0.01$; $*P<0.05$, two-sided Mann–Whitney *U* test, pairwise between groups, Benjamini–Hochberg false-discovery rate (FDR) adjustment. In **o**, **p**, **r**, Pearson correlation coefficient; P , two-sided P value; shade indicates 95% confidence interval. Scale bars, 100 μm (**e**, **f**, **k**). Box plots show interquartile range (25th to 75th percentiles) with centre line as the median (50th percentile).

distribution of immune cells in healthy lungs and lungs from a patient with COVID-19. **j**, Left, abundance of neutrophils (top) and macrophages (bottom) in each disease group. Right, macrophages divided into alveolar (top) and interstitial (bottom) subsets. **k**, Abundance of mesenchymal cells (left) and fibroblasts (right) in each disease group. **l**, Amount of change (effect size) pairwise between all disease groups ($n=15$) in MPO (left) and CD163 (right) markers between IMC (x axis) and immunohistochemistry (IHC) (y axis). **m**, Amount of change between late and early COVID-19 groups, pairwise for each cell type ($n=24$), as estimated by IMC (x axis) and targeted spatial transcriptomics (y axis) for the same (left) and independent (right) cohorts. For **c**, **d**, **j**, **l**, **m**, $n=237$ images from 27 biologically independent samples. $**P<0.01$; $*P<0.05$, two-sided Mann–Whitney *U* test, pairwise between groups, Benjamini–Hochberg false-discovery rate (FDR) adjustment. In **o**, **p**, **r**, Pearson correlation coefficient; P , two-sided P value; shade indicates 95% confidence interval. Scale bars, 100 μm (**e**, **f**, **k**). Box plots show interquartile range (25th to 75th percentiles) with centre line as the median (50th percentile).

($n=4$) and from whom post-mortem lung tissue was available (Fig. 1a, Extended Data Fig. 1a, b, Supplementary Table 1). To better understand and capture anatomical manifestations of the progression of lung disease, we divided patients with COVID-19 into those with ‘early’ and ‘late’ disease, depending on whether death occurred before or after 30 days from the start of respiratory symptoms, respectively (Supplementary Table 1). To comprehensively investigate the cellular environment and spatial organization of the lung, we designed a metal-labelled antibody panel for IMC that was composed of 36 biomarkers, and used it to generate 237 highly multiplexed images at 1- μm resolution: in total, we profiled 332 mm^2 of tissue and identified 664,006 single cells across all specimens. IMC leverages laser ablation based on inductively coupled

plasma mass spectrometry of lanthanide-metal-tagged antibodies from tissues for the quantitative detection of epitope abundance in a spatially resolved manner (Fig. 1a, Supplementary Table 2). Our panel included phenotypic markers of endothelial, epithelial, mesenchymal and immune cells, functional markers (activation, inflammation and cell death), and an antibody specific to the spike (S) protein of SARS-CoV-2. We used the IMC data to quantify the histopathology of the lung under infection (Methods), as we observed that the post-mortem lungs showed a considerable increase in weight across all pathologies (Fig. 1b). Consistent with the gain of weight during infection, the lacunar space of infected lungs was significantly reduced from 41.1% in the healthy lung to median ranges of 28.72% and 15.3% in the lungs

of individuals with influenza and late COVID-19, respectively (Fig. 1c, e); the most pronounced change was seen in the alveolar epithelium (Extended Data Fig. 1c, d). As collagen deposition is a known mediator of both normal and dysregulated tissue repair during recovery from infection²², we quantified the extent and intensity of collagen deposition into a fibrosis score that was inspired by the Ashcroft score²³ (Methods). The fibrosis score was significantly higher for lung pathologies than for the healthy lung, especially for the two COVID-19 groups (Fig. 1d, f, Extended Data Fig. 1e–g). We constructed a spatially resolved single-cell atlas to understand the cellular composition of the lung during various insults (Methods). We projected the 664,006 single cells from all disease groups into a two-dimensional space (Fig. 1g, Extended Data Fig. 1h) and clustered them on the basis of their phenotype (Fig. 1h, Extended Data Fig. 2a), which resulted in a single-cell phenotypic atlas for the human lung. We identified 36 clusters, which we organized into 17 metaclusters on the basis of predominant markers, overall phenotypes and proximity to lung structures (Fig. 1h, Extended Data Fig. 2a). This atlas was dominated by abundant structural cell types, including KRT8⁺KRT18⁺ alveolar epithelial cells, α -smooth muscle actin⁺ (α SMA) cells that line the vasculature and immune cells such as CD15⁺CD11b⁺ polymorphonuclear neutrophils and CD68⁺ macrophages (Extended Data Fig. 2c). Although the broad compartments of lung structural cells and immune cells did not show large changes in absolute numbers between the patient groups, the specific internal composition of the structural cells of the lung and the immune system differed extensively (Extended Data Fig. 2d). We observed increased immune infiltration in the lungs of patients with COVID-19 as compared to the healthy lung, but to a degree that was comparable with other lung infections (Extended Data Fig. 2d). Within the specific immune components, we observed a prominent increase in infiltration of myeloid cells in the lungs of patients with COVID-19 (as compared with the healthy lung), but to a lesser extent than was seen in the lungs of patients with bacterial pneumonia (Fig. 1i, Extended Data Fig. 3a, b). We performed a more detailed examination of the phenotypic diversity of myeloid cells in respect to their location in the lung (Extended Data Fig. 3c–e), which revealed that CD14⁺CD16⁺CD206⁺CD163⁺CD123⁺ interstitial macrophages—which were probably recruited from peripheral blood—displayed the greatest increase in the lungs of patients with COVID-19 (particularly in the late COVID-19 group) as compared with the healthy lung (Fig. 1j), and the highest expression of IL-1 β in monocytes in the lungs of patients in the early COVID-19 group (Extended Data Fig. 3d). Although neutrophil levels are similar between the lungs of patients in the early COVID-19 group and healthy lungs, they are present in significantly lower absolute numbers in late COVID-19 (Fig. 1j, Extended Data Fig. 3a, b); this is in stark contrast to the lungs of individuals with bacterial pneumonia, which contain the highest numbers of neutrophils across all disease groups. Populations of macrophages were particularly increased in the lungs of patients with COVID-19, as compared to all other disease states (Fig. 1j, Extended Data Fig. 3a, b). We also observed that CD8⁺ T cells were significantly increased in lungs of individuals with ARDS not associated with COVID-19, but depleted in bacterial pneumonia, in comparison with the healthy lung (Extended Data Fig. 3a, b). To further functionally characterize the immune system in healthy lungs and the lungs of individuals with COVID-19, we performed IMC with an immune panel of 39 markers on a subset of samples (2 healthy lungs and 4 lungs from patients with COVID-19) (Extended Data Fig. 4). In comparison with healthy lungs, we observed increased levels of the alarmin calprotectin (S100A9) across several cell types in the lungs of individuals with COVID-19—most prominently, in macrophages and neutrophils, but also in alveolar epithelial cells (Extended Data Fig. 4d–f). Alveolar epithelial cells also expressed increased levels of HLA-DR in COVID-19. Beyond the immune compartment, we observed a shift in the stromal compartment of the lung in COVID-19, with a significant reduction in absolute numbers of endothelial cells and an increase in mesenchymal cells and fibroblasts in lungs of patients in the late COVID-19 group

(Fig. 1k, Extended Data Fig. 3a, b). The increase in fibroblast abundance with COVID-19 is consistent with the increased fibrosis score that we observed in the lungs of patients with COVID-19 (Fig. 1d, Extended Data Fig. 5a). To orthogonally validate our findings, we performed immunohistochemical staining of lung tissue from the same donors for two markers and found excellent agreement between the relative frequency of cells that were positive for the markers in IMC and immunohistochemistry (Fig. 1l, Extended Data Fig. 5b–h), as well in the ability to estimate the size of the lacunar space of lungs in the various disease states (Extended Data Fig. 5i–k). We also observed good agreement between the changes in cell-type composition between disease groups using targeted spatial transcriptomics²⁴, both for matched samples in the same cohort (Fig. 1m, Extended Data Fig. 6a–d) and samples in an independent study²⁵ (Fig. 1m, Extended Data Fig. 6e–h).

Tissue damage by widespread inflammation

Our phenotypic single-cell atlas of the lung contains clusters that are defined by cell-identity markers and markers of cell state across all of the diseases that we evaluated. We developed an unsupervised classification of cells according to the abundance of each marker as a complementary approach (Extended Data Fig. 7a), and observed a high concordance with the phenotypic clusters (Extended Data Fig. 7b). Using this approach, we observed a high specificity of our SARS-CoV-2 S antibody to tissue samples from patients with COVID-19 (Fig. 2a, Extended Data Fig. 7c). Among all cell types, alveolar epithelial cells displayed the highest rate of SARS-CoV-2 S positivity (Fig. 2b). These alveolar epithelial cells were also highly positive for the phosphorylated signal transducer and activator of transcription 3 (pSTAT3), the receptor tyrosine kinase and proto-oncogene KIT and contained increased levels of interleukin 6 (IL-6), arginase 1, the apoptosis marker cleaved caspase 3 (CASP3) and the assembled complement membrane attack complex C5b–C9 (Fig. 2c, d). Given that S⁺ alveolar epithelial cells in the same regions showed no increase in the levels of functional markers, this probably indicates viral-specific alterations of the cellular state. Increased IL-6 and pSTAT3 levels were also seen in the lungs of individuals with influenza and pneumonia as compared with healthy lung, but were not seen in ARDS that was not associated with COVID-19 (Extended Data Fig. 7d). However, high levels of cleaved CASP3 and C5b–C9 were exclusive to S⁺ alveolar epithelial cells from individuals with COVID-19, and indicate the initiation of apoptosis and complement-mediated host immune defence, respectively, which led to increased damage to the alveolar lining. Although the alveolar epithelium was the predominant cell type that was positive for S, we also found that a mean of 7.8% and 2.7% of macrophages and neutrophils, respectively, were positive for S across all images of lungs of patients with COVID-19; some regions were up to 38.6% and 43.6% S⁺, respectively (Extended Data Fig. 7e, f). Consistent with our observations in S⁺ epithelial cells, both macrophages and neutrophils exhibited higher levels of cleaved CASP3, pSTAT3 and IL-6 but—unlike the alveolar epithelial cells—these cells showed no positive staining for C5b–C9 (Extended Data Fig. 7g–j). However, KIT was specifically upregulated in macrophages and not in neutrophils (Extended Data Fig. 7g, h). This non-epithelial cell marker profile phenotype seen in S⁺ cells was also seen in other cell types, albeit at a much lower frequency (Extended Data Fig. 7k). We also observed high heterogeneity in the localization of S⁺ cells, often within the same 1-mm² tissue region (Fig. 2e). Although we observed interactions between S⁺ epithelial cells and immune and nonimmune cells (Fig. 2e top right), other S⁺ cells did not interact with these cells at all or seemed to be encapsulated in structures that precluded interactions with other cell types (Fig. 2e bottom right). To generate a quantitative map of cellular interactions, we quantified proximal interactions between and within cell types for each image and generated disease-specific interaction maps (Extended Data Fig. 8a–c, Methods). Comparing interactions between the healthy lung and lungs from patients with

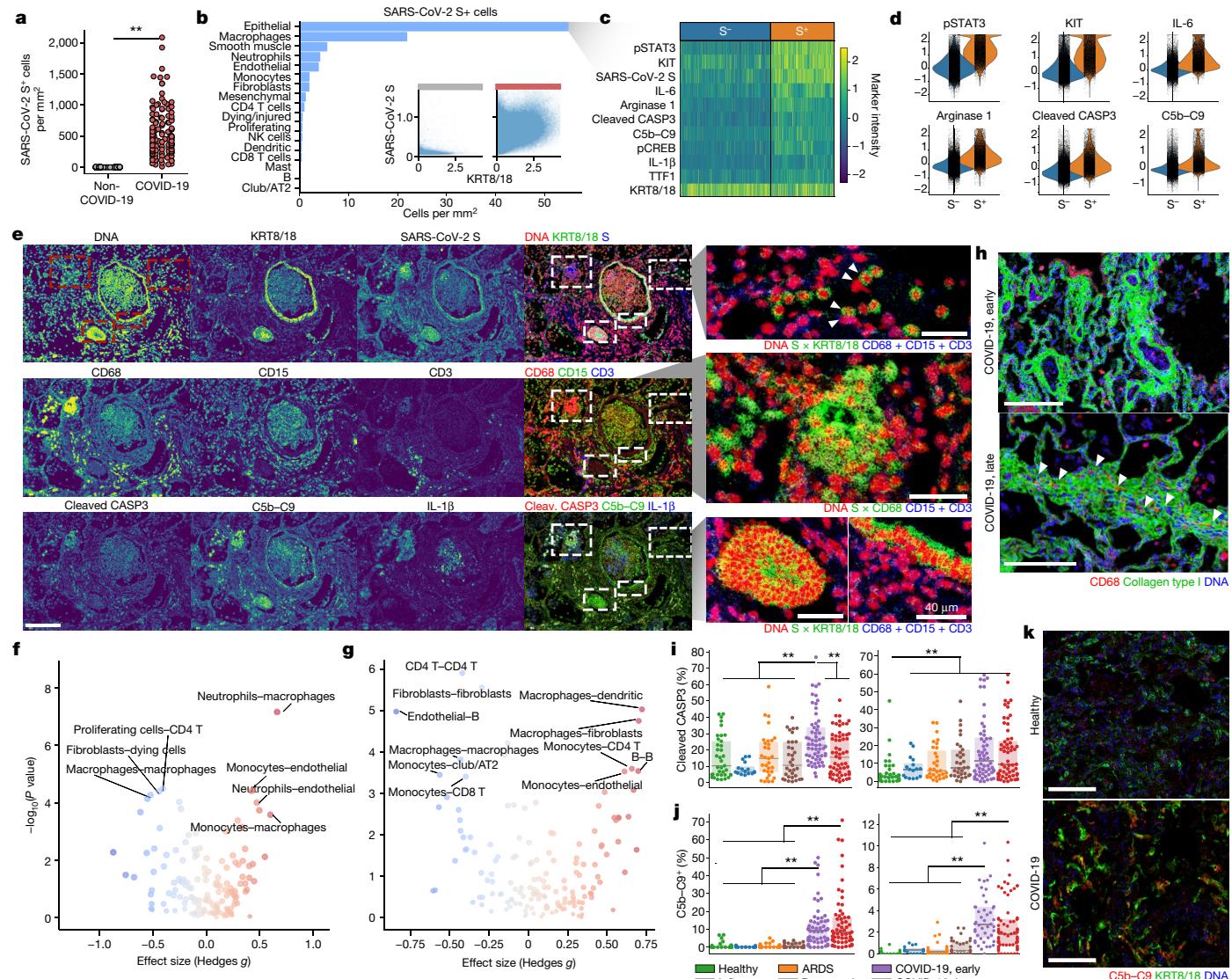


Fig. 2 | Cellular tropism of SARS-CoV-2 infection. **a**, Absolute abundance of S⁺ cells for lungs of patients without COVID-19 (grey) or with COVID-19 (red). **b**, Distribution of S⁺ cells across metaclusters in COVID-19. Inset displays intensity of KRT8/18 and S⁺ for single cells from non-COVID-19 (left) and COVID-19 (right) groups. **c**, Phenotype of alveolar epithelial cells in COVID-19, depending on levels of S⁻ and S⁺. **d**, Intensity of differential markers between cells dependent on S⁻ and S⁺ levels. **e**, Distribution of S signal in a spatial context. Structural, cell-type-specific and functional markers are displayed alone or in combination. For the green channel in the images in the rightmost column, the S channel was multiplied with KRT8/18 or CD68 to highlight T cells that are positive for both markers. Scale bar, 200 µm (main panels), 50 µm (magnified images on right (unless otherwise indicated)). **f**, **g**, Differential interactions in

COVID-19, we observed increased interactions between neutrophils and macrophages and decreased interactions within macrophages; in late COVID-19, the intra-cell-type interactions in macrophages, fibroblasts and CD4⁺ T cells decreased further, and were accompanied by an increase in interactions between macrophages and fibroblasts or dendritic cells (Fig. 2f, g). When spatially contextualizing these interactions, we observed that macrophages preferentially interacted with fibroblasts in the alveolar walls, which suggests a contribution to fibrosis and the thickening of the alveolar wall in late COVID-19 (Fig. 2h). As epithelial cells as a whole did not show any particular change in interactions in the lungs of patients with COVID-19 as compared with healthy lung, we investigated whether S⁺ epithelial cells differed in cellular

healthy lung and lungs of patients with COVID-19 (**f**) or between early and late COVID-19 (**g**). **h**, Fibroblasts and macrophages from early and late COVID-19. Scale bars, 200 µm. **i**, **j**, Proportion of cleaved CASP3⁺ macrophages (left) or neutrophils (right) (**i**), and C5b-C9⁺ epithelial (left) or endothelial (right) cells (**j**), for each disease group. **k**, Deposition of C5b-C9 in epithelial cells in healthy lung and lungs from patients with COVID-19. Scale bars, 100 µm. In **a**, **i**, **j**, $n = 237$ images from 27 biologically independent samples; $**P < 0.01$; $*P < 0.05$, two-sided Mann–Whitney *U* test, pairwise between groups, Benjamini–Hochberg FDR adjustment. In **f**, **g**, *P* values are from two-sided Mann–Whitney *U* test with Benjamini–Hochberg FDR adjustment. Box plots show interquartile range (25th–75th percentiles); centre line is median (50th percentile).

interactions to their S⁻ counterparts in the lungs of patients with COVID-19 (Extended Data Fig. 8d–i). Across all cell types, there was a trend for S⁺ cells to have reduced cellular interactions (Extended Data Fig. 8d–f). S⁺ alveolar epithelial cells in particular lacked interactions with other cell types, as compared with S⁻ cells (Extended Data Fig. 8g–i). We observed progressively more cells with markers of cell death (particularly macrophages and neutrophils with cleaved CASP3) (Fig. 2i), whereas epithelial and endothelial cells preferentially had C5b–C9 (Fig. 2j, k)—which probably indicates alveolar damage. Using spatial transcriptomics data, we observed that pathways of inflammatory response (such as interferon and interleukin signalling) were increased in the lungs of patients with COVID-19 as compared with healthy lung—particularly

interactions to their S⁻ counterparts in the lungs of patients with COVID-19 (Extended Data Fig. 8d–i). Across all cell types, there was a trend for S⁺ cells to have reduced cellular interactions (Extended Data Fig. 8d–f). S⁺ alveolar epithelial cells in particular lacked interactions with other cell types, as compared with S⁻ cells (Extended Data Fig. 8g–i). We observed progressively more cells with markers of cell death (particularly macrophages and neutrophils with cleaved CASP3) (Fig. 2i), whereas epithelial and endothelial cells preferentially had C5b–C9 (Fig. 2j, k)—which probably indicates alveolar damage. Using spatial transcriptomics data, we observed that pathways of inflammatory response (such as interferon and interleukin signalling) were increased in the lungs of patients with COVID-19 as compared with healthy lung—particularly

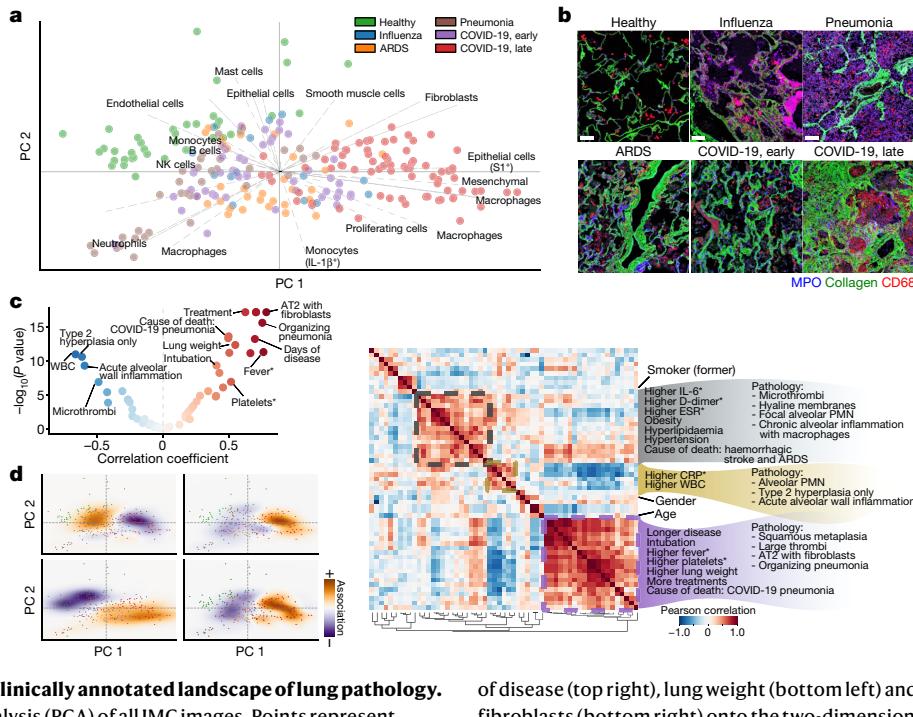


Fig. 3 | A data-driven and clinically annotated landscape of lung pathology.

a, Principal component analysis (PCA) of all IMC images. Points represent images, and are coloured by disease group. Arrows are vectors for each cell cluster, and indicate the area in which each cell type is most abundant.

b, Microanatomy and immune content of the disease groups. Scale bar, 100 μ m. **c**, Volcano plot showing strength of association between clinical parameters and principal component (PC)1, and significance. WBC, white blood cell.

d, Projections of white blood cell count (measured at admission) (top left), days

of disease (top right), lung weight (bottom left) and alveolar type-2 cells with fibroblasts (bottom right) onto the two-dimensional PCA space. **e**, Similarity of landscape of IMC data. Pairwise correlation of demographic, clinical and pathological variables in the association with the principal components. Matrix rows and columns are the same. Highlighted groups of variables reflect hierarchically clustered groups of variables explaining the IMC data. In **c–e**, an asterisk indicates that the clinical parameter was measured at admission.

in the alveolar and airway compartments in early COVID-19 (Extended Data Fig. 9a–e). However, pathways related to angiogenesis, myogenesis and the epithelial-to-mesenchymal transition were increased in the lungs of individuals with COVID-19 as compared with healthy lungs, which increased progressively in late COVID-19 (Extended Data Fig. 9b). In accordance with the IMC data, we also observed that coagulation, complement activation and apoptosis pathways were upregulated in alveolar areas and in blood vessels in late COVID-19 (Extended Data Fig. 9c). This suggests that, after an early period of disease that is dominated by inflammatory responses to SARS-CoV-2, late COVID-19 in the lung may be driven by pathogen-independent mechanisms that are a consequence of an immune response with an aberrant resolution.

An interpretable landscape of lung pathology

Building on the cell types we identified, their functional status and their interactions, we sought to define a landscape of lung pathology from the data to form an unbiased view of the multicellular architecture of lung tissue during infection (Fig. 3a, Extended Data Fig. 9f–j, Methods; available at <http://covid-imc.eipm-research.org/>). The major axes of this landscape, which is based on a principal component analysis, demonstrate the distribution of samples and the major drivers of the establishment of the landscape. This confers biological interpretability to the landscape, as the underlying cellular composition at each given point is readily identifiable. Although the landscape was determined in an unsupervised manner (without knowledge of sample groups), it largely recapitulates the disease ontogeny of the samples: it is dominated by the difference between samples of healthy lung and lungs of patients with COVID-19 who succumbed after prolonged disease (Fig. 3a). This is exemplified by the abundance of immune cell types such as macrophages or neutrophils (which are most abundant in the lungs of patients with COVID-19 or pneumonia, respectively),

but also by collagen deposition in lungs from patients with COVID-19 (Fig. 3b). We largely recapitulated the structure of this landscape when using complementary methods (Extended Data Fig. 9h–j). To add a layer of clinical interpretability to the landscape, we performed an association analysis between its axes and known demographic, clinical and pathological factors of patients with lung infections (Fig. 3c, Extended Data Fig. 10a–d, Methods). We observed strong associations primarily between clinical factors and the first principal component (Extended Data Fig. 10a): specifically, a significant positive association between the distribution of samples in the first principal component axis with the presence of alveolar type 2 cells with fibroblasts, organizing pneumonia, the number of days that elapsed since beginning of symptoms, hospitalization and intubation, and lung weight at death (Fig. 3c, Extended Data Fig. 10a). We also observed a significant association between a reduction in white blood cell counts and the major axis associated with disease progression (principal component 1) (Fig. 3c, Extended Data Fig. 10a). The values of clinical factors overlaid with their respective images in the landscape confer a convenient way of interpreting the associations, effectively rendering a landscape annotated with clinical information both biologically (Fig. 3a) and clinically interpretable (Fig. 3d, Extended Data Fig. 10e). To develop the clinical interpretation of the landscape, we further related the associated clinical, demographic and pathological variables by how similar they are in explaining the IMC data (Fig. 3e, Extended Data Fig. 10b). We found that variables were organized into three large blocks: (1) high C-reactive protein and white blood cell count at presentation, as well as pathology characterized by acute inflammation of the alveolar wall; (2) high values of IL-6, erythrocyte sedimentation rate and D-dimer at presentation, comorbidities such as obesity and hypertension, lung pathology characterized by microthrombi and chronic alveolar inflammation with macrophages, and haemorrhagic stroke as the cause of death; and (3) prolonged disease and associated interventions such as

Article

intubation and treatment, with pathology characterized by squamous hyperplasia, large thrombi, organizing pneumonia and alveolar type-2 cells associated with fibroblasts. These groups probably represent a progressive range of pathology that is associated with extremely acute disease that results in early death (1) to a chronic manifestation of prolonged disease (3). Beyond these dominant clusters, we found that demographic and behavioural variables (such as age, gender or smoking) did not strongly associate with the larger groups, which suggests they have little influence in the pathology of lethal disease associated with SARS-CoV-2 lung infection. The similarity between variables in the IMC data differs considerably from that obtained from simple co-occurrence of the variables (Extended Data Fig. 10c, d), which provides evidence for the added value of high-content multiplexed imaging of lung tissue in infectious disease.

Discussion

Our spatially resolved single-cell analyses of post-mortem lung tissue from patients with COVID-19 or other lung infections provides a comprehensive examination of the response of the human lung to infection, from the macroscopic to the single-cell level. Across all diseases, we observed a significant reduction in alveolar lacunar space, increased immune infiltration and cell death by apoptosis, as compared with healthy lungs. We also noted that neutrophil infiltration—although it is increased in the lungs of patients with ARDS or early COVID-19 compared with normal lung—is a hallmark of bacterial pneumonia, and that a high degree of inflammation, infiltration of interstitial macrophages, complement activation and fibrosis is particular to the lungs of individuals with late COVID-19. Our analysis agrees with recent reports that indicate that the type of pathophysiological response to SARS-CoV-2 infection may not be entirely different from ARDS that is unrelated to COVID-19¹⁴, but contradicts reports that suggest that the hyperinflammatory phenotype (as assessed by cytokine levels in peripheral blood) is not specific to COVID-19¹⁵. Our observation that S⁺ alveolar epithelial cells do not differentially interact with cells of the immune system (despite extensive immune infiltration in the lung) potentially highlights the lack of an ‘on-target’ immunological response, and the high amount of complement activation in lung tissue from patients with COVID-19 probably results in indiscriminate ‘off-target’ tissue damage—exacerbating COVID-19 and continuing the cycle of inflammation. The increased presence of IL-1 β ⁺ monocytes in the lungs of patients with early COVID-19 suggests a mechanism for neutrophil recruitment to the lung. Neutrophil recruitment was highest in the lungs of patients with bacterial pneumonia (the only group that we studied with an active disease of bacterial pathogen origin). The differential pathogen recognition between viral and bacterial infection in the lung could explain the differences in chemokine secretion in the ensuing immune response. However, despite sharing a viral pathogenic origin with influenza, COVID-19 (specifically the expansion of mesenchymal cells and fibroblasts, particularly in late COVID-19) probably reflects a response to the extensive tissue damage from complement activation. Despite this, the high mortality rate of COVID-19 is at odds with productive recovery from tissue damage and healing, which suggests the need for further investigation into complement-activation-induced damage to the lung, additional immunological factors (such as neutrophil extracellular traps) and microthrombi formation¹⁹. This raises the possibility that early immunological interventions that suppress excessive complement activation could have a therapeutic benefit. Our biologically interpretable and clinically annotated landscape of

lung pathology provides a framework for a data-driven, spatially aware understanding of lung pathology, and will be an important resource for the study of COVID-19 and other lung infections.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03475-6>.

1. Laing, A. G. et al. A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* **26**, 1623–1635 (2020).
2. Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
3. Mathew, D. et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
4. Nienhold, R. et al. Two distinct immunopathological profiles in autopsy lungs of COVID-19. *Nat. Commun.* **11**, 5086 (2020).
5. Wang, W. et al. High-dimensional immune profiling by mass cytometry revealed immunosuppression and dysfunction of immunity in COVID-19 patients. *Cell. Mol. Immunol.* **17**, 650–652 (2020).
6. Untermaier, A. et al. Single-cell omics reveals dyssynchrony of the innate and adaptive immune system in progressive COVID-19. Preprint at <https://doi.org/10.1101/2020.07.16.20153437> (2020).
7. Takahashi, T. et al. Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature* **588**, 315–320 (2020).
8. Hadjadj, J. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).
9. Giesen, C. et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
10. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
11. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
12. Manson, J. J. et al. COVID-19-associated hyperinflammation and escalation of patient care: a retrospective longitudinal cohort study. *Lancet Rheumatol.* **2**, e594–e602 (2020).
13. Tay, M. Z., Poh, C. M., Rénia, L., MacAry, P. A. & Ng, L. F. P. The trinity of COVID-19: immunity, inflammation and intervention. *Nat. Rev. Immunol.* **20**, 363–374 (2020).
14. Grasselli, G. et al. Pathophysiology of COVID-19-associated acute respiratory distress syndrome: a multicentre prospective observational study. *Lancet Respir. Med.* **8**, 1201–1208 (2020).
15. Sinha, P. et al. Prevalence of phenotypes of acute respiratory distress syndrome in critically ill patients with COVID-19: a prospective observational study. *Lancet Respir. Med.* **8**, 1209–1218 (2020).
16. Arunachalam, P. S. et al. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210–1220 (2020).
17. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
18. Reyes, L. et al. A type I IFN, prothrombotic hyperinflammatory neutrophil signature is distinct for COVID-19 ARDS. Preprint at <https://doi.org/10.1101/2020.09.15.20195305> (2020).
19. Bradley, B. T. et al. Histopathology and ultrastructural findings of fatal COVID-19 infections in Washington State: a case series. *Lancet* **396**, 320–332 (2020).
20. Bengsch, B. et al. Deep spatial profiling of COVID19 brains reveals neuroinflammation by compartmentalized local immune cell interactions and targets for intervention. Preprint at <https://doi.org/10.21203/rs.3.rs-63687/v1> (2020).
21. Tian, S. et al. Pathological study of the 2019 novel coronavirus disease (COVID-19) through postmortem core biopsies. *Mod. Pathol.* **33**, 1007–1014 (2020).
22. Duffield, J. S., Luperhe, M., Thannickal, V. J. & Wynn, T. A. Host responses in tissue repair and fibrosis. *Annu. Rev. Pathol.* **8**, 241–276 (2013).
23. Ashcroft, T., Simpson, J. M. & Timbrell, V. Simple method of estimating severity of pulmonary fibrosis on a numerical scale. *J. Clin. Pathol.* **41**, 467–470 (1988).
24. Merritt, C. R. et al. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat. Biotechnol.* **38**, 586–599 (2020).
25. Desai, N. et al. Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. *Nat. Commun.* **11**, 6319 (2020).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Data reporting

No statistical methods were used to predetermine sample size. Image acquisition, segmentation, quantification and clustering were blinded to patient identifiers and clinical metadata.

Human studies

Tissue samples were provided by the Weill Cornell Medicine Department of Pathology. The Tissue Procurement Facility operates under Institutional Review Board (IRB) approved protocol and follows guidelines set by Health Insurance Portability and Accountability Act. 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. Autopsy consent was obtained from the families of the patients.

Tissue section preparation

Lung tissues were fixed via 10% neutral buffered formalin inflation, sectioned and fixed for 24 h before processing and embedding into paraffin blocks. Freshly cut 5- μm sections were mounted onto charged slides.

Antibody panel design and validation

We designed an antibody panel to capture different immune and stromal compartments of the lung. Antibody clones were extensively validated through immunofluorescence and chromogenic staining and verified by a pathologist. Once the clone was approved, 100 μg of purified antibody in BSA and azide free format was procured and conjugated using the MaxPar X8 multimetal labelling kit (Fluidigm) as per the manufacturer's protocol. To confirm the antibody binding specificity after conjugation and to identify the optimal dilution for each custom conjugated antibody, sections from healthy lung, and bacterial pneumonia, non-COVID-19 ARDS and SARS-CoV-2-infected lung 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 neighbouring channels, a higher dilution factor was adopted when staining the cohort tissues.

IMC

On the basis of the clinical and pathological characteristics and quality of the preserved tissues, suitable representative fresh cut 4- μm -thick FFPE sections were acquired from the Department of Pathology of Weill Cornell Medicine for IMC staining. The tissues were stored at 4 °C for a day before staining. Slides were first incubated for 1 h at 60 °C on a slide warmer followed by dewaxing in fresh CitriSolv (Decon Labs) twice for 10 min, rehydrated in descending series of 100%, 95%, 80% and 75% ethanol for 5 min each. After 5 min of MilliQ water wash, the slides were treated with antigen retrieval solution (Tris-EDTA pH 9.2) for 30 min at 96 °C. Slides were cooled to room temperature, washed twice in TBS and blocked for 1.5 h in SuperBlock Solution (ThermoFischer), followed by overnight incubation at 4 °C with the prepared antibody cocktail containing all 36 metal-labelled antibodies (Supplementary Table 2). The 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 min in a humid chamber at room temperature. Slides were washed with MilliQ water and air-dried before ablation.

The instrument was calibrated using a tuning slide to optimize the sensitivity of the detection range. Haematoxylin and eosin-stained slides were used to guide the selection of regions of interest (ROIs) containing alveolar parenchyma, airways and thrombotic vessels to obtain regions that were 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. A total of 240 image stacks were ablated. The raw MCD files were exported for further downstream processing.

IHC

Automated IHC on a Leica Bond III instrument was performed on 5- μm 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 min, high pH) using 3,3' diaminobenzidine chromogen. For each slide, a grid (5 × 5 grid, 0.4-cm × 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 ROI (using a 20× objective) was photographed.

Targeted spatial transcriptomics using GeoMx

In brief, selected cases of lung injury associated with COVID-19 (4 patients with early COVID-19 and 4 patients with late COVID-19), bacterial pneumonia (2) and healthy lung (3) from the IMC cohort were evaluated using the GeoMx²⁴ COVID-19 Immune Response Atlas with approximately 1,850 RNA targets. Spatial transcriptomics analysis included up to 24 ROIs per tissue. Alveolar, airway and vascular compartments were analysed.

For GeoMx DSP slide preparation, we followed the GeoMx DSP slide preparation user manual (MAN-10087-04). In brief, tissue slides were baked in a drying oven at 60 °C for 1 h and then loaded to Leica Biosystems BOND RX FFPE for deparaffinization and rehydration. After the target retrieval step, tissues were treated with proteinase K solution to expose RNA targets followed by fixation with 10% NBF. After all tissue pretreatments were done, tissue slides were unloaded from the Leica Biosystems BOND RX and incubated overnight with RNA probe mix (COVID-19 Immune Response Atlas; a pool of *in situ* hybridization probes with UV photocleavable oligonucleotide barcodes). The next day, tissues were washed and stained with tissue visualization markers: CD68-647 at 1:400 (Novus Bio, NBP2-34736AF647), CD45-594 at 1:10 (NanoString Technologies), panCK-532 at 1:20 (NanoString Technologies) and/or SYTO 13 at 1:10 (Thermo Scientific S7575).

For GeoMx DSP sample collections, we followed the GeoMx DSP instrument user manual (MAN-10088-03). In brief, tissue slides were loaded on the GeoMx DSP instrument and then scanned to visualize whole-tissue images. For each tissue sample, a board-certified pathologist selected 24 total ROIs from 3 types of functional tissue: vascular zone, large airway and alveoli zone. Each ROI was subdivided into compartments on the basis of fluorescent cell-specific markers, and serial UV illumination of each compartment was used to sequentially collect the probe barcodes from the different cell types.

Each GeoMx DSP sample plus nontemplate controls (NTCs) was uniquely indexed using the i5 × i7 dual-indexing system of Illumina. Four μl of a GeoMx DSP sample was used in a PCR reaction with 1 μM of i5 primer, 1 μM i7 primer and 1× NSTG PCR Master Mix. Thermocycler conditions were 37 °C for 30 min, 50 °C for 10 min, 95 °C for 3 min, 18 cycles of 95 °C for 15 s, 65 °C for 60 s, 68 °C for 30 s, and final extension of 68 °C for 5 min. PCR reactions were purified with two rounds of AMPure XP beads (Beckman Coulter) at 1.2× bead-to-sample ratio. Libraries were paired-end sequenced (2 × 75) on a NextSeq550 generating up to 400 million aligned reads in total.

Processing and filtering of the raw next-generation sequencing data was performed on the DNA sample libraries that were sequenced, producing about 1.3 billion reads. NextSeq-derived FASTQ files for each sample were compiled for each compartment using the bcl2fastq program of Illumina, and then demultiplexed and converted to digital count conversion (DCC) files using the GeoMx DnD pipeline (v.1) of Nanostring. These DCC files were then converted to an expression count matrix using a custom Python script. A minimum of 10,000 reads were

Article

required for each non-NTC sample (2 compartments removed). Probes were checked for outlier status by implementing a global Grubbs's outlier test with alpha set to 0.01. The counts for all remaining probes for a given target were then collapsed into a single metric by taking the geometric mean of probe counts. A count of 1 was added to any probe that yielded 0 counts before the geometric mean was taken. For each sample, an RNA-probe-pool-specific negative probe normalization factor was generated on the basis of the geometric mean of negative probes in each pool. To ensure good data quality, we calculated the 75th percentile of the gene counts (that is, geometric mean across all non-outlier probes for a given gene) for each ROI, and normalized to the geometric mean of the 75th percentile across all ROIs to give the upper quartile (Q3) normalization factors for each ROI. The distribution of these Q3 normalization factors were then checked for outliers defined as any ROI greater than two s.d. from the mean log₂-transformed Q3 normalization factor. This criterion removed 15 ROIs that fell below the range and 1 ROI that fell above the range.

Preprocessing IMC data

IMC data were preprocessed as previously described²⁶ with some modifications. In brief, image data were extracted from MCD files acquired with the Fluidigm Hyperion instrument. Hot pixels were removed using a fixed threshold, the image was amplified two times, Gaussian smoothing was applied and, from each image, a square 500-pixel crop was saved as a HDF5 file for image segmentation. Segmentation of cells in the image was performed with ilastik²⁷ (version 1.3.3) by manually labelling pixels 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). Ilastik uses the labelled pixels to train a random forest classifier using features derived from the image and its derivatives. Features used were the Laplacian of the Gaussian, Gaussian gradient magnitude, difference of Gaussians, structure tensor eigenvalues and the Hessian of Gaussian eigenvalues, each of which had Gaussian kernels of widths from 0.3 to 10 (37 features in total). The outputs of prediction are class probabilities for each pixel, which were used to segment the image using CellProfiler²⁸ (version 3.1.8) with the IdentifyPrimaryObjects module. This was followed by the IdentifySecondaryObjects module, in which the identified nuclei are used to seed an expansion of the cell area to the area with the sum of the nuclear and cytoplasmatic probability map, and finally gaps in the identified cells are filled.

We assessed the quality of each acquired channel by computing a set of metrics for each channel across all images: the mean and squared coefficient of variation of each channel in the whole image and, in the area with cells, a difference between those values in the cells and the whole image (foreground versus background signal), an estimate of noise variance²⁹, a robust wavelet-based estimator of Gaussian noise s.d.^{29,30}, the fractal dimension (Minkowski–Bouligand approximation using the box counting method) and lacunarity of the image³¹. Across all 240 ablated images three were discarded based on these metrics and visual inspection.

Computing lacunarity and fibrosis score

To identify lacunae in the images, we used the mean of all channels in each image stack after performing histogram equalization per channel (`skimage.exposure.equalize_hist`). Images were thresholded with Otsu's method (`skimage.filters.threshold_otsu`), successively dilated and closed (`ski.morphology.binary_dilation/ski.morphology.closing`) with a disk of 5-μm diameter to remove objects without holes and—for the objects with holes—objects within the hole were removed on the negative image (`scipy.ndimage.binary_fill_holes`) and only objects with area larger than 625 pixels (25²) were kept (`skimage.morphology.remove_small_objects`). To provide biological context for the single-cell clusters we identified, we further classified each of the lacuna of healthy

lungs into one of three classes: blood vessels (arteries and veins), airways and alveoli. Vessels showed a very thin lining of endothelial cells, followed by a thick layer of smooth muscle cells that are α-SMA⁺; the airway epithelium is lined by KRT8⁺KRT18⁺ cells; and alveoli are covered in alveolar epithelial cells that have various degrees of CD31, vimentin, KRT8 and KRT18. On the basis of this, we developed a semisupervised strategy for lacuna classification that had two stages: first each of the lacuna objects was dilated by a 15-pixel disk and the mean intensity of the channels above was quantified only in the dilated area, and these values were Z-score-transformed per image. We used these values in three ways, in which each provided a vote towards a lacuna being one of the three classes: absolute intensity, Z-score-transformed intensity, ratio of α-SMA to KRT8 and KRT18. For each, a set of rules was enforced in which lacunae with higher values in α-SMA and low in KRT8 and KRT18 were labelled as vessels, and those with higher KRT8 and KRT18 were labelled as airways; the remaining lacunae were labelled as alveoli. Essentially, absolute and relative intensity of the markers determine the class, and the ratio of the two is the tiebreaker in case of disagreement. In a second phase, the suggested labels were reviewed by an expert and overruled if needed. In general, we found that the rules above were accurate with only a systematic bias to underclassify vessels (hence the need for supervision).

To develop a score for fibrosis, we were inspired by the Ashcroft score²³ in which the fraction of fibrotic tissue that occupies each image is translated into a score in a Likert scale. We quantified the fraction of the image occupied by collagen type I as thresholded by the Otsu method (`skimage.filters.threshold_otsu`), but in addition quantified the density of collagen per area unit by using the spectral counts given by the IMC data. The final score is the mean of a Z-score of the fraction covered by collagen and a Z-score of its intensity.

Cell-type identification

To identify cell types in an unsupervised fashion, we first quantified the intensity of each channel in each segmented cell that did not overlap image borders. In addition, for each cell we computed the morphological features 'area', 'perimeter', 'major_axis_length', 'eccentricity' and 'solidity' (`skimage.measure.regionprops_table`). Values were Z-scored per image and cells with area values above -1.5, solidity above -1 and eccentricity below 1 were kept. In addition, we calculated the sum of $\log(1+x)$ signals in the IMC channels and kept cells with values between 2 and 7. We used Scanpy³² (version 1.6.0) to perform a PCA, compute a neighbour graph on the PCA latent space, compute a UMAP¹⁰ embedding (`umap package, version 0.4.6`) and cluster the cells with the Leiden algorithm³³ with resolution 1.0 (`leidenalg package, version 0.8.1`). Each cluster was manually labelled with a broad ontogeny and the channels that were most abundant in each cluster. These broad labels formed the basis of the metaclusters used to aggregate clusters on the basis of cell type and regardless of cellular state. Clusters without enrichment for any particular marker were not aggregated.

To obtain an easy way to quantify the fraction of cells positive for a given marker, we used univariate Gaussian mixture models using scikit-learn³⁴ (version 0.23.0). For each channel, we performed model selection with models with two to six mixtures, selected the model on the basis of the Davies–Bouldin index³⁵ and labelled a cell as positive for a given channel if its value was in the top mixture (in cases in which the selected model had only two mixtures) or the top two mixtures (if the selected model had more).

The analysis of the immune-centric IMC panel was performed in the same manner as the larger dataset, with the exception that the thresholds for cell filtering based on area, DNA intercalator intensity and solidity were performed automatically with Gaussian mixture models as described in the previous paragraph. Differential marker abundance was tested with a two-sided Wald test between healthy lung samples and samples from patients with COVID-19, and adjusted for multiple comparisons with the Benjamini–Hochberg FDR.

Quantification of cellular interactions

To quantify the degree and importance of intra- and inter-cell-type interactions, we started by constructing a region adjacency graph representing the interactions between cells, in which the edges are weighted by the Euclidean distances between cells, using scikit-image³⁶ (version 0.17.2). A pairwise adjacency matrix between cell clusters was computed using networkx³⁷ (version 2.5). To get a degree of confidence on cellular interactions given the cell type abundance in each image, we permuted the cell cluster assignments 1,000 times and computed the difference between the log-normalized frequency of cell-type interactions in the real data versus the permuted (interaction scores). For visualization, we generated chord plots by aggregating the interaction scores of the images from each disease group or subgroups by the fraction of images with an interaction score was higher than 1. To discover differential interactions specific to a subgroup, we tested whether the distribution of interaction scores between disease groups or subgroups for each pairwise cell-type combination was different as described in ‘Analysis of IHC data’.

Analysis of IHC data

IHC images were segmented with Stardist³⁸ (version 0.6.1) with the 2D_versatile_he pretrained model and normalization to the unit space after capping the intensity to the 3rd and 98th percentiles. The image was decomposed into a haematoxylin and diaminobenzidine intensity channels using a preset colour space for the stains from the scikit-image package³⁶ (`skimage.colour.hdx_from_rgb`) and the intensity of each nuclei in both channels was calculated by a mean reduction. As the H-DAB and IHC signal in general does not linearly reflect the molecular stoichiometry of chemical reactions, we discretized the signal into positive and negative fractions per image by using Gaussian mixture models with model selection in the same way as for the IMC data. Cells were declared positive for a stain if they belonged to the group with highest signal.

For the quantification of image lacunarity in IHC data we first normalized illumination in YCrCb colour space by applying a Gaussian blur with 5 pixels of s.d. per 300×300 size image and subtracting this background from the brightness channel (Y or luma). Images were converted back to RGB colour space and the inverted mean of the channels was segmented using the Otsu method. The lacunarity of the image is the fraction of the image covered in background. This was performed only for MPO-stained images within the alveolar space and not with airways or vessels.

Analysis of targeted spatial transcriptomics data using GeoMx

To transform the gene set space into a cell-type space that would be compatible with the cell types identified in the IMC dataset, we performed single-sample gene set enrichment analysis (ssGSEA) using GSEAPY (version 0.10.2). For each ROI, gene set signatures used were from the molecular signatures database (MSigDB, version 7.2); in particular, the cell-type signature gene sets (group C8) were used. ssGSEA enrichment values were Z-scored, and signatures containing the keywords ‘Epithelial’, ‘Mesenchym’, ‘Fibroblast’, ‘Smooth_muscle’, ‘Club’, ‘CD4_T’, ‘CD8_T’, ‘NK_cell’, ‘Macrophage’, ‘Monocyte’, ‘Neutrophil’, ‘B_cell’, ‘Mast’ and ‘Dendritic’ were grouped and averaged, generating an enrichment score for each ROI in each cell type. If a group was composed of less than three signatures it was discarded. For the pathway-based functional analysis of lung tissue, ssGSEA was performed on the Hallmark group of gene set signatures from MSigDB (group H) without further aggregation.

PCA of lung pathology and association with clinical parameters

To create the unsupervised landscape of lung pathology, we used PCA on a matrix of cell counts per cell cluster (features) and per image (observations), after previous normalization by total, scaling and

centring, using the Scanpy implementation. The feature loadings were plotted on the same dimensions as the observations by scaling them by a constant factor of 20. The correlation coefficient between each principal component and the continuous clinical variables was used as a relative measure of direction and strength of association. However, the significance of association was assessed by permuting the clinical variables 10^6 times and using the mean and s.d. of the correlation coefficients from the permuted data as location and scale parameters, respectively, of a normal distribution, from which the $2 \times \text{CDF}(|x|)$ was calculated as a two-tailed empirical P value. All principal components and clinical factors were permuted and the empirical P values were adjusted with the Benjamini–Hochberg FDR method. We used the signed empirical, FDR-corrected P values as an effectively regularized measure of association between principal components and clinical factors. To project the clinical parameters into the PCA space, we fit Gaussian kernel density estimator functions to the distribution of the images in the first principal components, in one case without and another with the numeric values of the clinical variables as weights. The difference in predicted densities in the two-dimensional space between the weighted and unweighted kernels was used as a visual aid to identify regions in the latent space with relatively higher or lower fraction of samples with that clinical parameter.

Statistics and reproducibility

Unless otherwise stated in the figure legends, statistical testing was performed pairwise between groups with a two-sided Mann–Whitney U test, and adjusted for multiple comparisons with the Benjamini–Hochberg FDR method using pingouin³⁹ (version 0.3.7). Estimated values of central tendency, effect sizes and P values are provided in Supplementary Tables 3, 4. In the box plots in all figures, the box is the interquartile range (25th to 75th percentiles) and the centre line is the median (50th percentile). Experiments were not repeated independently, owing to the use of unique autopsy material.

The following software versions were used: Python, version 3.8.2; numpy⁴⁰, version 1.18.5; scipy⁴¹, version 1.4.1; scikit-image³⁶, version 0.17.2; networkx³⁷, version 2.5; Scanpy, version 1.6.0; pingouin³⁹, version 0.3.7; CellProfiler²⁸, version 3.1.8; and Stardist³⁸, version 0.6.1.

Reporting summary

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

Data availability

IMC data are available at <https://doi.org/10.5281/zenodo.4110560>, <https://doi.org/10.5281/zenodo.4139442> and <https://doi.org/10.5281/zenodo.4637034>. IHC data are available at <https://doi.org/10.5281/zenodo.4633905>. Targeted spatial transcriptomics data are available at <https://doi.org/10.5281/zenodo.4635285>. Any other relevant data are available from the corresponding author upon reasonable request.

Code availability

The source code for data preprocessing and analysis is available at <https://github.com/ElementoLab/covid-imc> and <https://github.com/ElementoLab/covid-imc-viz>.

26. Jackson, H. W. et al. The single-cell pathology landscape of breast cancer. *Nature* **578**, 615–620 (2020).
27. Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).
28. McQuin, C. et al. CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol.* **16**, e2005970 (2018).
29. Immerkær, J. Fast noise variance estimation. *Comput. Vis. Image Underst.* **64**, 300–302 (1996).
30. Donoho, D. L. & Johnstone, I. M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–455 (1994).

Article

31. Kit, O. & Lüdeke, M. Automated detection of slum area change in Hyderabad, India using multitemporal satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **83**, 130–137 (2013).
32. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
33. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
34. Pedregosa, F. & Varoquaux, G. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
35. Davies, D. L. & Bouldin, D. W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI* **1**, 224–227 (1979).
36. van der Walt, S. et al. scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).
37. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In Proc. 7th Python in Science Conference (SciPy2008) (eds. Varoquaux, G. et al.) 11–15 (2008).
38. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (eds Frangi, A. F. et al.) 265–273 (Springer, 2018).
39. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).
40. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
41. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Acknowledgements We thank the patients and their families for their contributions to our study. We apologize to colleagues whose work we could not cite owing to editorial constraints. This work was partially supported by the WorldQuant Initiative for Quantitative Prediction. A.F.R. is 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 NCI R01CA234614, NIAID 2R01AI107301, NIDDK R01DK121072 and 1R03DK117252. 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, labelling and acquiring samples for IMC imaging, and the translational research laboratory (Department of Pathology, Weill Cornell Medicine) for histology and IHC support. We thank Zenodo for storing the raw data associated with this Article with a higher usage quota.

Author contributions A.F.R., H.R., O.E., A.C.B. and R.E.S. planned the study; A.C.B., Y.B., V.C., R.E.S. and S.S. provided and evaluated samples and clinical data; H.R. and E.C.S. performed IMC; T.H., S.W., Y.K. and J.R. performed the digital spatial transcriptomics with contributions from C.M., J.P., J.F. and C.E.M. A.F.R. performed analysis of the data, with contributions from J.K.; O.E., and R.E.S. supervised the research. A.F.R., H.R., O.E., A.C.B. and R.E.S. wrote the manuscript with contributions from all authors.

Competing interests O.E. is scientific advisor and equity holder in Freenome, Owkin, Volastra Therapeutics and OneThree Biotech. R.E.S. is on the scientific advisory board of Miromatrix Inc and is a consultant and speaker for Alnylam Inc. C.E.M. is a cofounder of Biotia and Onegevity Health. E.C.S. is an employee of Fluidigm. T.H., S.W., Y.K. and J.R. are employees of Nanostring Inc. The remaining authors declare no competing financial interests.

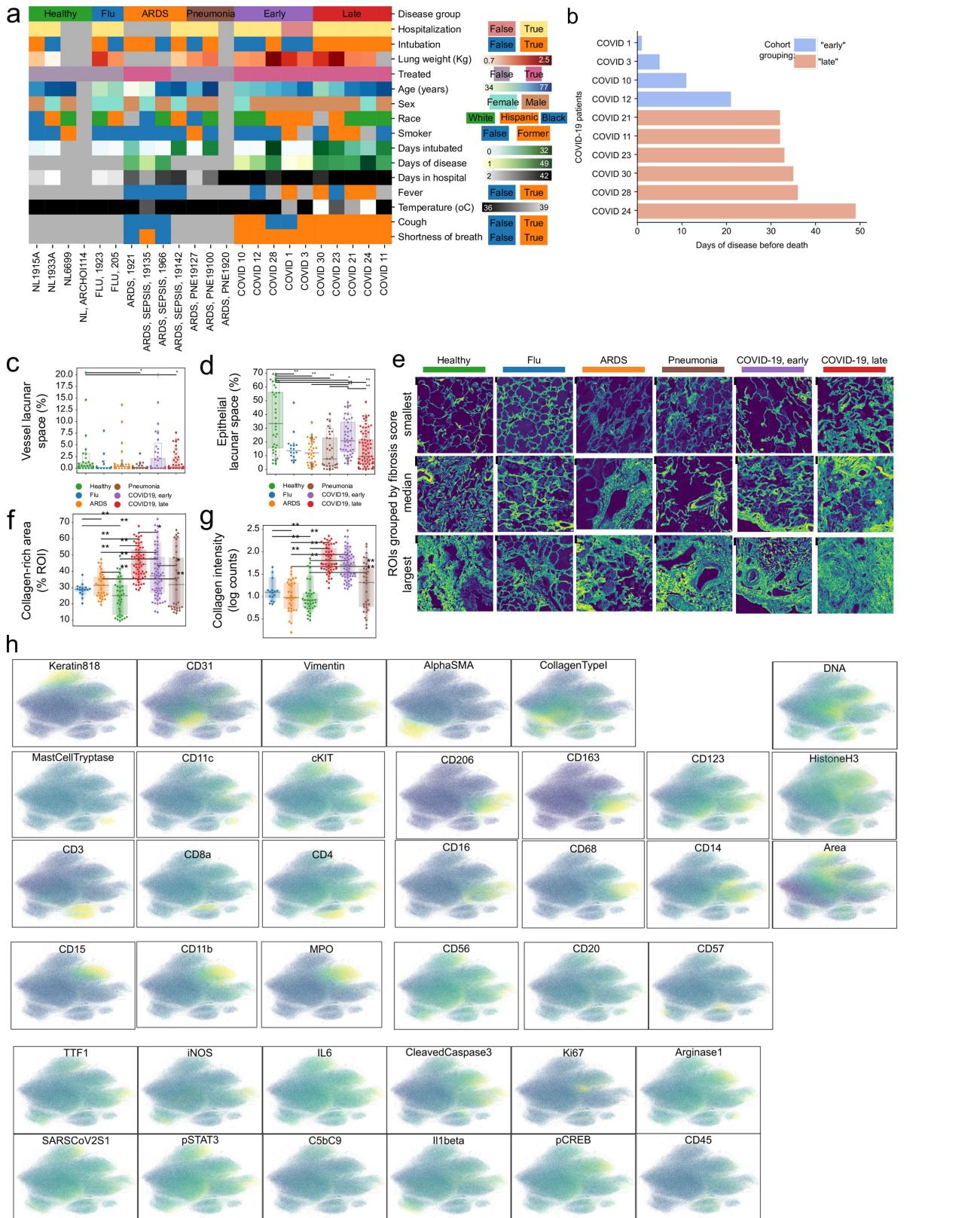
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-03475-6>.

Correspondence and requests for materials should be addressed to A.C.B., O.E. or R.E.S.

Peer review information *Nature* thanks John Alcorn, Ross Levine and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



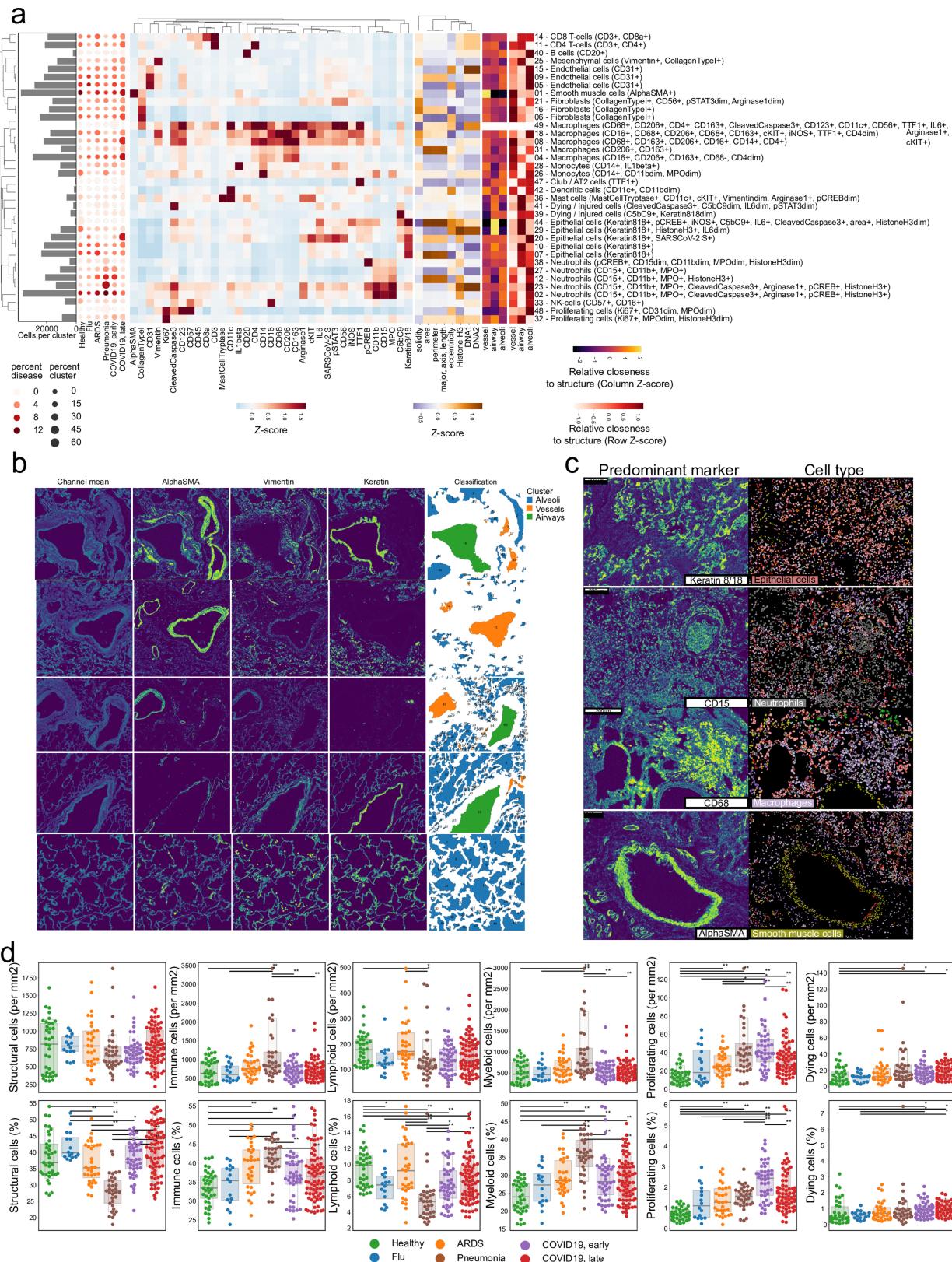
Extended Data Fig. 1 | Macroscopic pathology of the SARS-CoV-2-infected lung.

a, Heat map depicting the values of each individual for all clinical and demographic variables. Grey colour indicates missing or non-applicable values. **b**, Time of death relative to start of symptoms in patients with COVID-19. **c, d**, Percentage of lacunar space attributed to vessel (**c**) or epithelial (**d**) space per image, grouped by disease. **e**, Collagen type I in images of healthy lungs or lungs from patients with lung pathology, and the associated fibrosis score.

Images with the lowest, median and highest fibrosis scores are depicted.

f, Percentage of image covered in collagen type I for each image, grouped by disease group. **g**, Mean intensity of collagen type I in lung IMC images, grouped by disease group. **h**, UMAP projection of all single cells, in which cells are coloured by the intensity of each channel. In **c, d, f, g**, $^{**}P < 0.01$; $^{*}P < 0.05$, two-sided Mann–Whitney *U*-test, pairwise between groups, Benjamini–Hochberg FDR adjustment.

Article

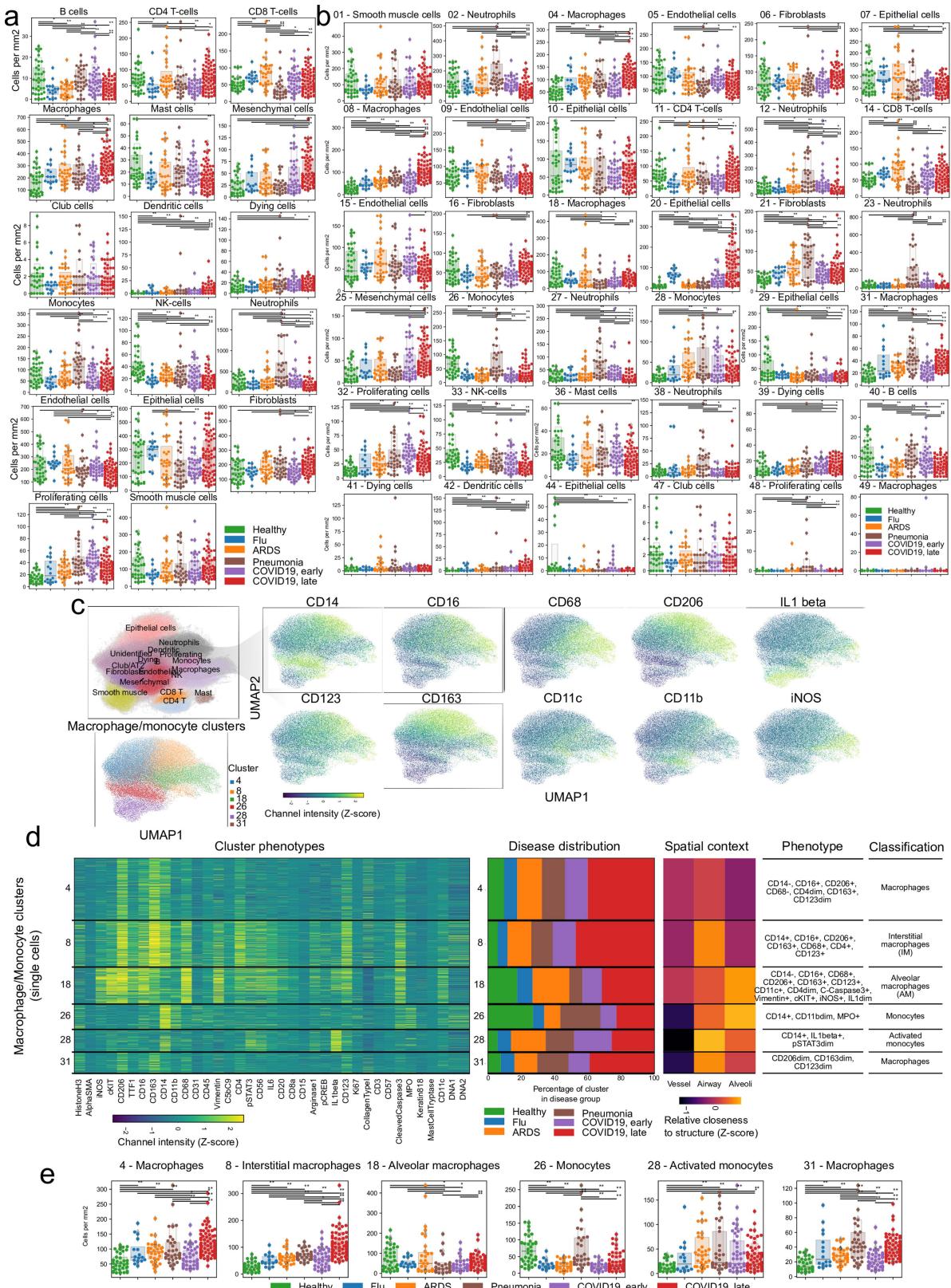


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Clustering analysis and quantification of structural cell types across disease types. **a**, Hierarchically clustered heat map of discovered clusters (rows) and the mean intensity of each channel (columns) for each. The histogram on the left represents the absolute abundance of each cluster across all images. The dot plot represents the relative abundance of each cluster in each disease group. **b**, Classification of lung lacunae. Representative images of healthy lung with the mean of all channels, and channels that are important for discerning between vessels, airways and alveoli. The rightmost column represents the final classification of lacunae

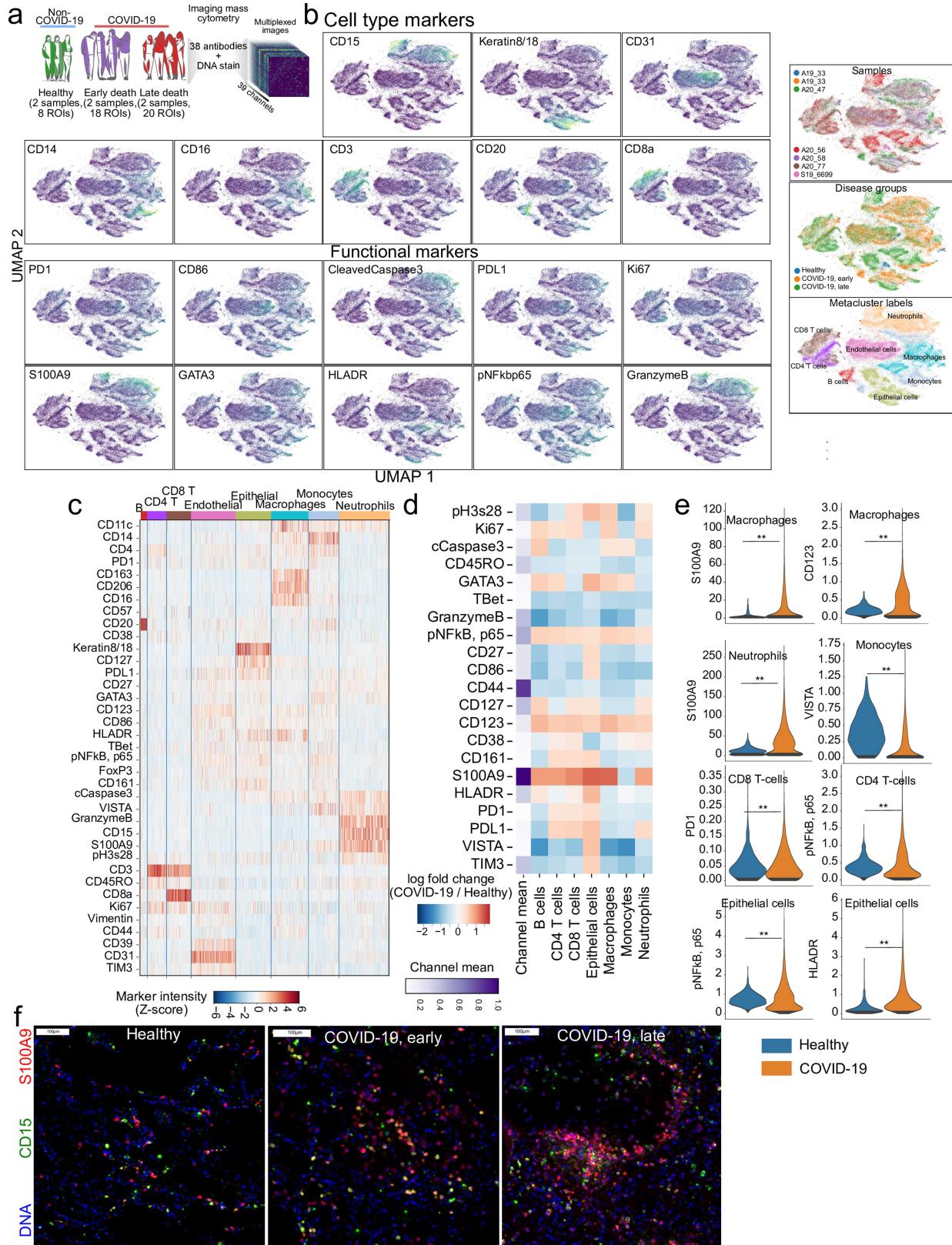
into each of the three classes of structures. **c**, Representative spatial context of three metaclusters (rows). Left column, spatial distribution of the predominant marker for each metacluster; right column, segmented cells coloured by the metacluster to which they were assigned. **d**, Global abundance of structural and immune cells. Absolute (top panels) and relative (bottom panels) abundance of groups of cells, depending on disease group. ** $P < 0.01$; * $P < 0.05$, two-sided Mann–Whitney U -test, pairwise between groups, Benjamini–Hochberg FDR adjustment.

Article



Extended Data Fig. 3 | Analysis of cellular phenotypes reveals alterations in the myeloid and epithelial compartments. **a, b**, Absolute abundance of metaclusters (**a**) or clusters (**b**) per image, grouped by disease group. **c–e**, Diversity of myeloid cells in the lung. **c**, UMAP representation of myeloid cells and the prominent markers associated with them. **d**, Phenotypic markers,

spatial context and abundance in disease groups for each of the six myeloid clusters. **e**, Abundance of each myeloid cluster in the disease groups. Each point represents the abundance of that cluster in a given ROI. In **a, b, e**, ** $P < 0.01$; * $P < 0.05$, two-sided Mann–Whitney *U*-test, pairwise between groups, Benjamini–Hochberg FDR adjustment.



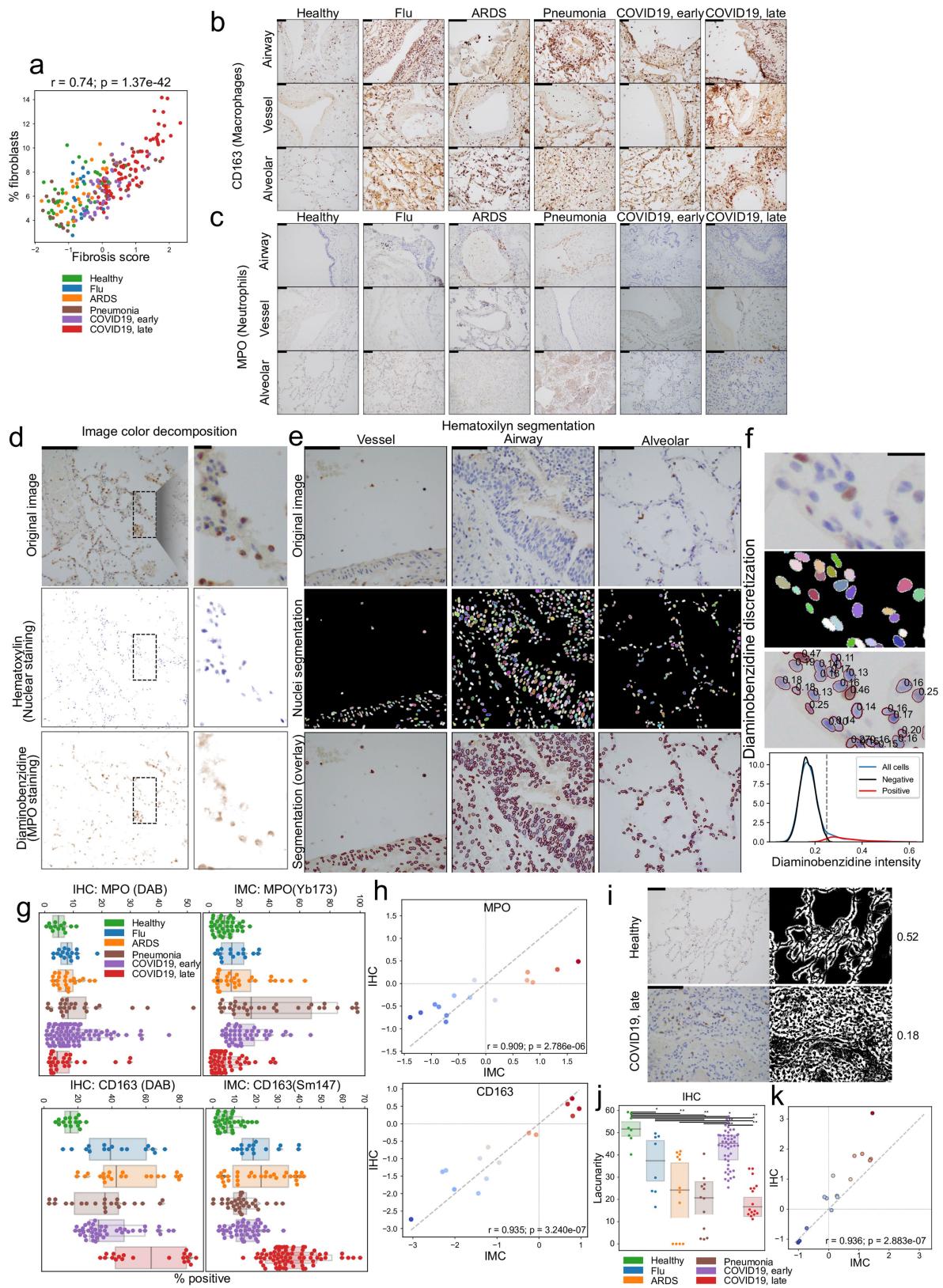
Extended Data Fig. 4 | Immune cell analysis using an orthogonal probe set.

a, Schematic of IMC panel, and its application to healthy lung tissue samples and lung tissue samples from patients with COVID-19. **b**, UMAP representation of single cells, coloured by the intensity of the marker for cell-type-defining markers and functional markers. **c**, Expression of markers for each single cell grouped by the cell type. **d**, Differential expression of functional markers

between lungs of patients with COVID-19 and healthy lung for each cell type.

e, Expression range represented as violin plots for the selected cell types from **b**, between lungs of patients with COVID-19 and healthy lung. ** $P < 0.01$ two-sided Wald test, Benjamini–Hochberg FDR adjustment. **f**, Representative images of S100A9 and CD15 marker expression in healthy lung and lungs of patients with COVID-19. Scale bars, 100 μ m.

Article

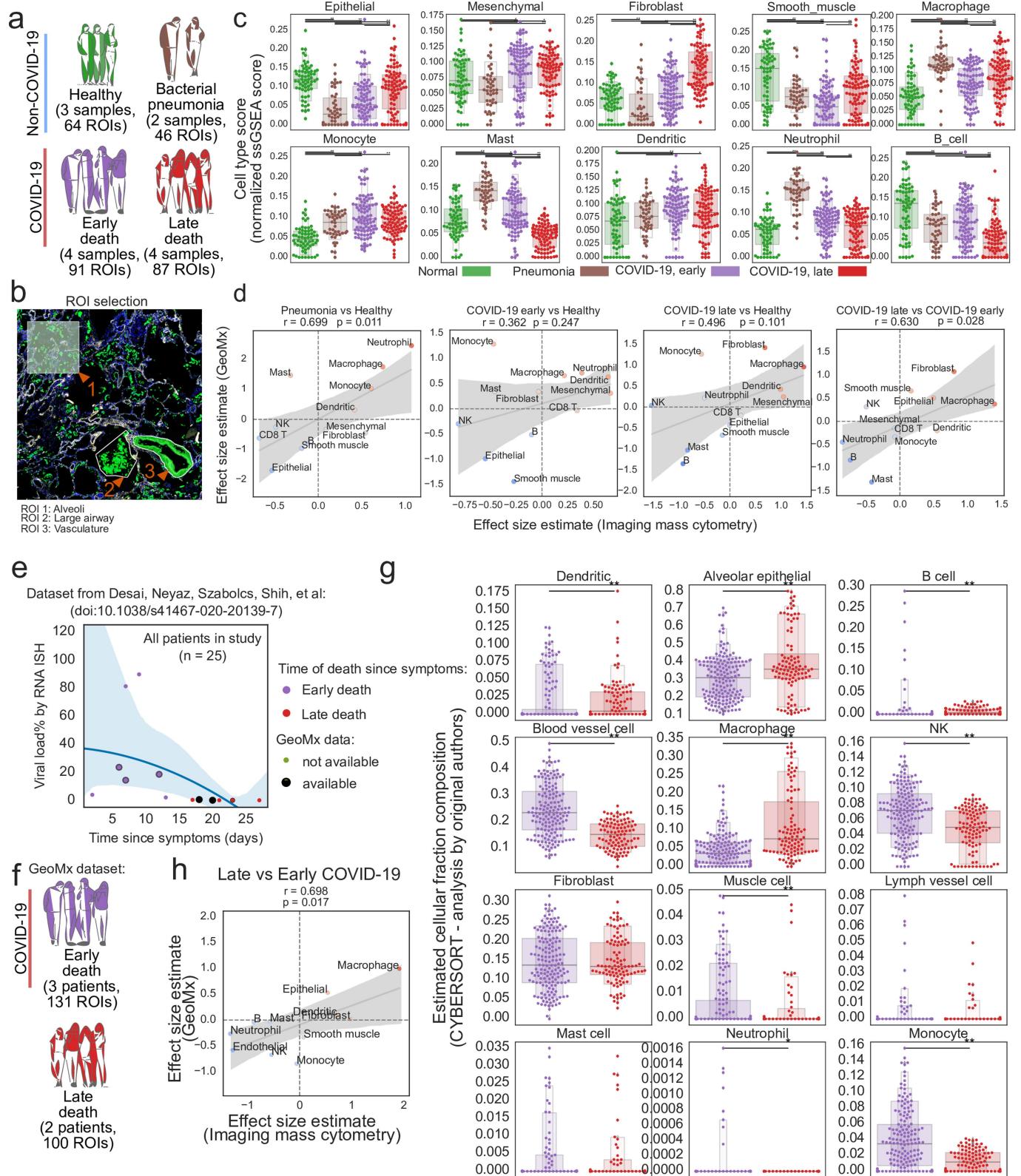


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Validation of IMC findings with IHC. **a**, Relationship between fibrosis score and fibroblast metacluster abundance, visualized as a scatter plot. **b, c**, IHC for two markers across all disease groups. Haematoxylin-diaminobenzidine staining of CD163 (**b**) or MPO (**c**) in tissue from healthy and diseased lung, matching the patients in the IMC cohort. Scale bars, 200 μm . **d**, Analysis of IHC data. Example images demonstrating the process of colour decomposition that underlies the separation of the haematoxylin (nuclei) and diaminobenzidine (CD163 or MPO) (H-DAB) in lung tissues. Scale bars, 400 μm (long bars), 50 μm (short bars). **e**, Example images demonstrating the process of nuclei segmentation. Left column, original images in RGB space; middle column, resulting segmentation in which each nucleus has a random colour and the background is black; and right column, borders of segmented nuclei in red overlain on the original image. Scale bars, 200 μm . **f**, Example image section, demonstrating the process of quantification of the diaminobenzidine stain. Top, original image in RGB space; second panel, nuclei segmentation; third panel, numeric value of the DAB stain for each nucleus; and bottom, histogram of nuclei intensity in DAB stain, modelled as a Gaussian mixture with two components used to discretize nuclei into negative or positive for DAB on

the basis of a threshold that best separates the two mixtures. Scale bars, 50 μm . **g**, Percentage of cells within an image that are positive for the respective DAB stain in IHC (left column) or positive for the respective marker in IMC data (right column). **h**, Comparison of the estimated effect sizes of change between disease groups estimated from IMC (xaxis) or IHC (yaxis) data for the two stains. The Pearson correlation coefficient and its significance are indicated. **i**, Analysis of lacunae with IHC data. Representative images of the amount of lacuna space in IHC data for healthy lungs and lungs of patients with late COVID-19. For each image, the original image and the segmented background space is shown along with a value denoting the amount of lacuna space for the image (fraction of the image without cells that represents the alveolar or capillary space). Scale bars, 200 μm . **j**, Quantification of lacuna space across MPO images in IHC. ** $P < 0.01$; * $P < 0.05$, two-sided Mann–Whitney U -test, pairwise between groups, Benjamini–Hochberg FDR adjustment. **k**, Comparison of the estimated effect sizes of change in the amount of lacuna space between disease groups, estimated from IMC (xaxis) or IHC (y-axis) data. For **a, h, k, r**, Pearson coefficient; P , its two-tailed P value.

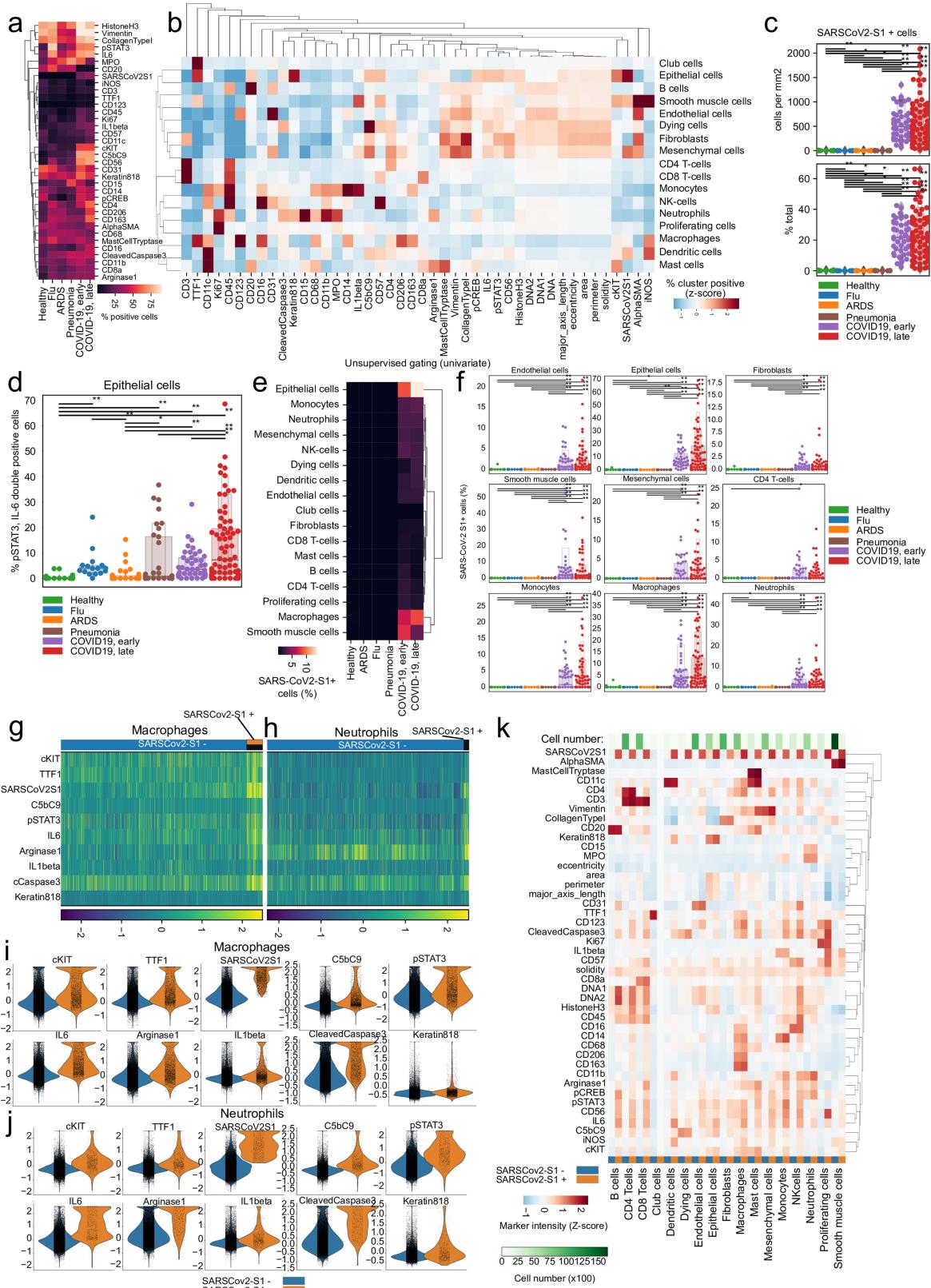
Article



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Profiling of lung tissue with targeted spatial transcriptomics using GeoMx. **a**, Experimental design of GeoMx dataset. **b**, Representation of the procedure of choosing a ROI within the lung tissue to capture with GeoMx. **c**, Enrichment of cell-type-specific gene set signatures for various cell types that match the IMC data, across disease groups. **d**, Comparison of the estimated changes in cell-type abundance with IMC (xaxis) and gene set signatures in GeoMx (yaxis). **e**, Viral load, dependent on the time of death relative to the onset of COVID-19 symptoms in an independent cohort. COVID-19 samples were categorized into ‘early’ or ‘late’ death depending on whether death occurred before or after 15 days after the

onset of symptoms, respectively. **f**, Schematic of the cohort of patients for whom GeoMx data are available (in total, 5 patients and 231 ROIs). **g**, Estimated fractions of cell-type composition, using the CYBERSORT program, between early and late COVID-19 from the original publication²⁵. **h**, Comparison of the estimated changes in cell-type abundance with IMC (xaxis) and GeoMx (yaxis) between late and early COVID-19. In **d**, **h**, *r*, Pearson coefficient, *P*, its two tailed *P*value; shaded area indicates 95th confidence interval. In **c**, **g**, ***P*<0.01; **P*<0.05, two-sided Mann–Whitney *U*-test, pairwise between groups, Benjamini–Hochberg FDR adjustment.

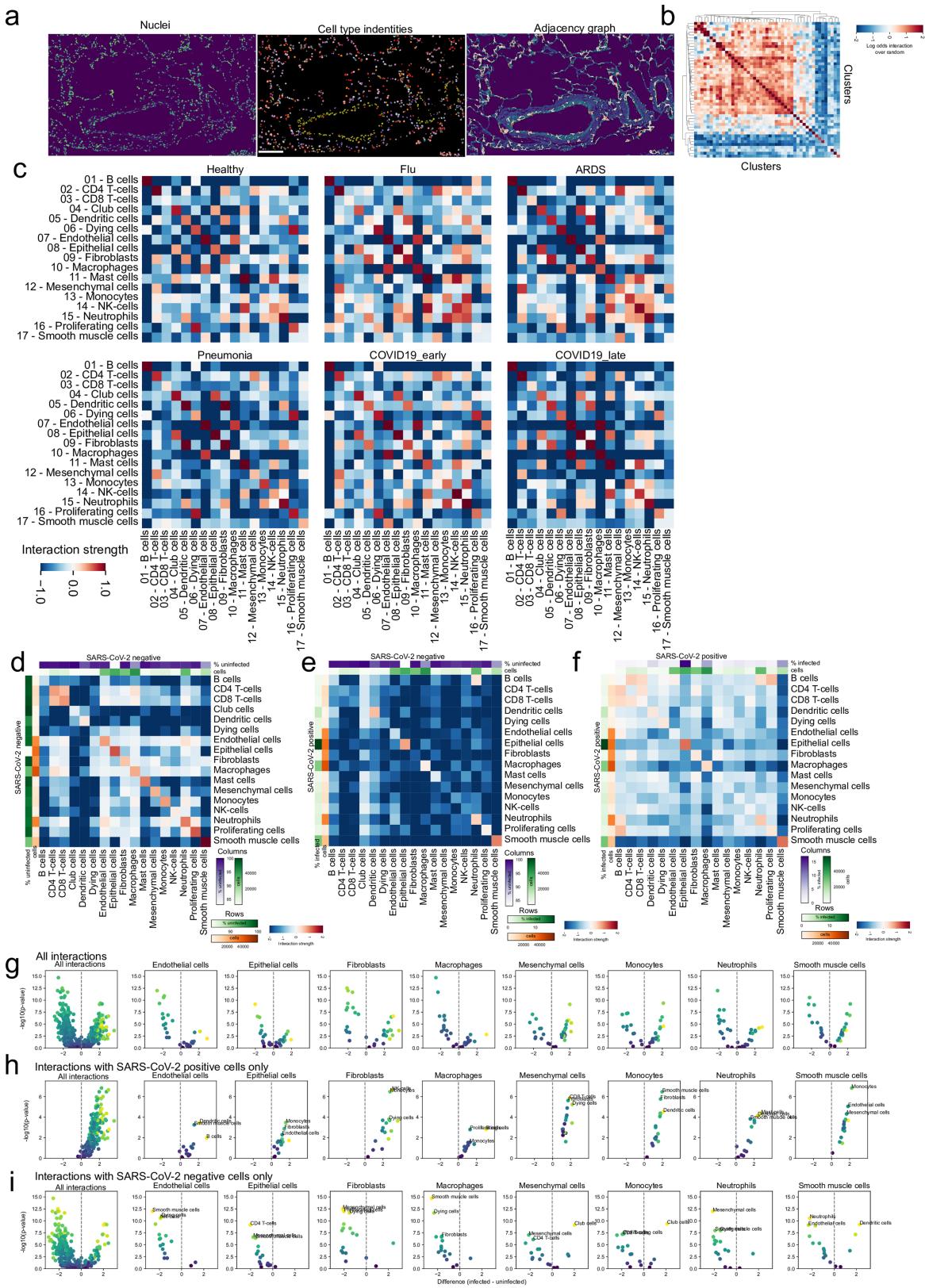


Extended Data Fig. 7 | Functional state of myeloid cells across disease cohorts. **a**, Percentage of cells positive for each IMC channel, as classified by univariate Gaussian mixture models per disease group. **b**, Percentage of channel-positive cells for each metacluster. Values represent a column-wise Z-score. **c**, Absolute (top) and relative (bottom) frequency of SARS-CoV-2S⁺ cells per disease group. **d**, Proportional abundance of SARS-CoV-2S⁺, IL-6⁺ and pSTAT3⁺ cells across disease groups. **e**, Proportional amount of SARS-CoV-2S⁺ cells grouped by metacluster and disease group. **f**, Proportional frequencies of

cells positive for SARS-CoV-2S⁺ cells by metacluster and disease group.

g, h, Heat map of single macrophages or neutrophils (**g**) (columns) and functional markers (**h**) (rows) with cells grouped by SARS-CoV-2S positivity.

i, j, Intensity of IMC channels per single cell, depending on SARS-CoV-2S positivity for macrophages (**i**) and neutrophils (**j**). **k**, Mean channel intensity for all metaclusters dependent on SARS-CoV-2S positivity. In **c, d, f**, **P<0.01; *P<0.05, two-sided Mann-Whitney U-test, pairwise between groups, Benjamini-Hochberg FDR adjustment.

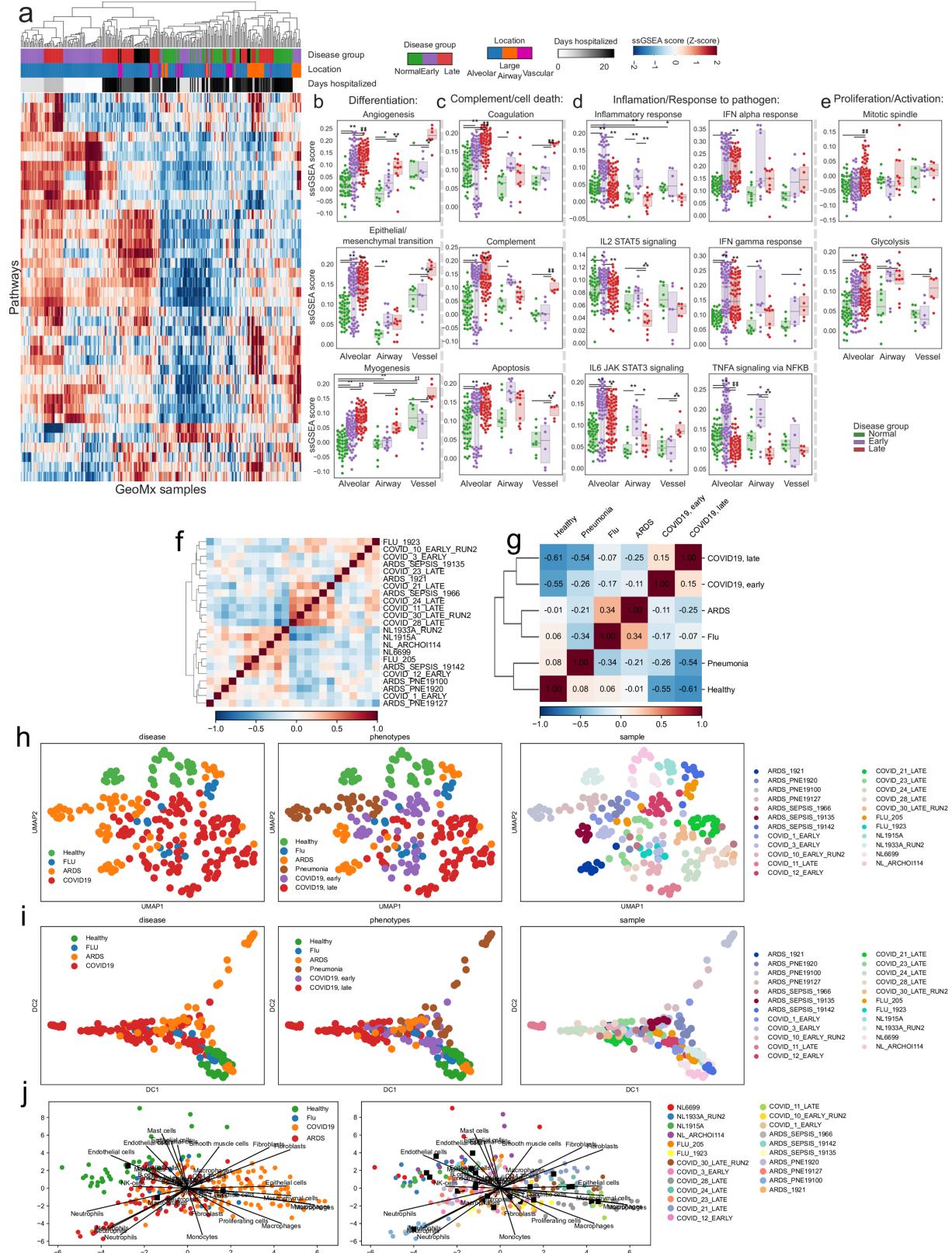


Extended Data Fig. 8 | See next page for caption.

Article

Extended Data Fig. 8 | Mapping cellular interactions between SARS-CoV2-infected and non-infected compartments. **a**, Example description of the derivation of a region adjacency graph for a given lung IMC image. Left, DNA channel that marks nuclei; middle, identified metaclusters; and right, region adjacency graph represented as edges between adjacent cells. Scale bar, 100 μm . **b**, Observed values of pairwise cluster interactions over the expected values for the same cellular interactions for the image in **a**. **c**, Pairwise interactions between metaclusters, aggregated by the mean value across images depending on the disease group. **d–f**, Pairwise cellular interactions

between metaclusters, depending on SARS-CoV-2S positivity in uninfected cells (**d**), between SARS-CoV-2S⁺ and S⁻ cells (**e**) and between infected cells (**f**). **g–i**, Statistical testing of differential interactions of infected cells and other cell types, and uninfected cells and other cell types, depending on SARS-CoV-2S positivity of the second cell type. **g**, Both SARS-CoV-2S⁻ and S⁺ cells. **h**, Only SARS-CoV-2S⁺ cells. **i**, Only SARS-CoV-2S⁻ cells. The rows display a volcano plot in which the x axis displays the difference in interaction between SARS-CoV-2S⁺ and S⁻ cells and the y axis displays the $-\log_{10}$ -transformed Mann–Whitney *U*-test, FDR-adjusted *P* value.



Extended Data Fig. 9 | Pathway enrichment analysis in spatial transcriptomics and IMC datasets. **a**, Enrichment scores for hallmark pathways in MsigDB across all ROIs in the GeoMx dataset. **b–e**, Enrichment score of selected pathways from **a**, across disease groups but dependent on the location within the lung from which they were obtained. ** $P < 0.01$; * $P < 0.05$,

two-sided Mann–Whitney U -test, pairwise between groups, Benjamini–Hochberg FDR adjustment. **f, g**, Pairwise Pearson correlation of cell-type abundances between IMC samples (**f**) or disease groups (**g**). **h–j**, UMAP (**h**), diffusion map (**i**) or PCA (**j**) projection of IMC images, coloured by disease group, subgroup or sample identifier.

Article



Extended Data Fig. 10 | Association analysis between known demographic, clinical and pathological factors of patients with lung infections.

a, Correlation coefficients (left) or FDR-adjusted P values (middle) and signed p -values (right), demonstrating the association between demographic, pathological and clinical factors, and principal components. **b**, Pairwise correlation of demographic, pathological and clinical factors across all principal components. Matrix was clustered using average linking and the Pearson correlation as the distance metric. Values used were signed

FDR-adjusted P values. **c**, Pairwise correlation of demographic, pathological and clinical factors across samples (co-occurrence). Matrix was clustered using average linking and the Pearson correlation as the distance metric. **d**, As in **b**, **c**; the top triangular matrix is from **b** and the bottom is from **c**. The order of the rows and columns is the same as in **b**. **e**, Projection of clinical factors onto the pathology landscape. Kernel density estimation for various clinical and demographic factors weighted by the factor values, unweighted or their difference.

Reporting Summary

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

Statistics

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

n/a Confirmed

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

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

Software and code

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

Data collection We used the Hyperion CyTOF Software (Fluidigm), version 7.0.8493 and MCDViewer (Fluidigm), version 1.0.560.2.

Data analysis We used Python version 3.8.2, Numpy version 1.18.5, Scipy version 1.4.1, scikit-image 0.17.2, networkx version 2.5, Scanpy version 1.6.0, pingouin version 0.3.7, CellProfiler version 0.3.1, Stardist version 0.6.1.
Source code of the analysis scripts is available as a Github repository: <https://github.com/ElementoLab/covid-imc>

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

Data

Policy information about [availability of data](#)

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

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

Raw imaging mass cytometry data is available in the following Zenodo repository: <https://doi.org/10.5281/zenodo.4110560>

We used the molecular signatures database (MSigDB) (<http://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) version 7.2.

We used a dataset of targeted spatial transcriptomics from Desai et al (doi:10.1038/s41467-020-20139-7), available at <https://doi.org/10.1038/s41467-020-20139-7>

Field-specific reporting

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

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

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

Life sciences study design

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

Sample size

No statistical methods were used to predetermine sample size. This was due to the limited availability of tissue from autopsies. We profiled available samples which tissue quality was appropriate for imaging mass cytometry, and we imaged a large portion of tissue (332mm² in total) in order to increase the robustness of the observations. To account for the temporal axis of COVID-19 disease we profiled 10 patients, whereas for the other groups which is 2-3 times more than the other groups.

Data exclusions

3 images out of 240 were excluded due to failing quality control. Images excluded were ROI 1, 9, and 10 from sample "20200701_COVID_11_LATE".

Replication

Due to the use of autopsy tissue, each patient was only profiled once. No replication was performed.

Randomization

Sample allocation to groups was given by the underlying disease for each patient. In the case of COVID-19, two groups were formed dependent on whether death occurred 14 days before or after 30 days from the start of respiratory symptoms.

Blinding

For all experiments, investigators were blinded to patient identifiers during data acquisition and analysis.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Imaging mass cytometry, lung COVID-19 panel:
anti-Alpha-Smooth Muscle Actin, clone 1A4, Fluidigm, cat. num. 3141017D, lot id: 651904; anti-Arginase-1, clone D4E3M, Cell Signaling Technology, cat. num. 93668BF, lot id: 5; anti-CD117/C-KIT, clone OTI2B12, Novus Biologicals, cat. num. NBP2-71076, lot id: F001; anti-CD11b, clone SP331, Abcam, cat. num. ab238794, lot id: GR3245427-1; anti-CD11c, clone EP1347Y, Abcam, cat. num. ab216655, lot id: GR3210349-14; anti-CD123, clone IL3RA/2947R, Novus Biologicals, cat. num. NBP2-79851, lot id: 3563-4PABX200206; anti-CD14, clone D7A2T, Cell Signaling Technology, cat. num. 56082BF, lot id: 2; anti-CD15, clone W6D3, BioLegend, cat. num. 323035, lot id: B267076; anti-CD16, clone EPR16784, Fluidigm, cat. num. 3146020D, lot id: 1011904; anti-CD163, clone EDHu-1, Fluidigm, cat. num. 3147021D, lot id: 1631904; anti-CD20, clone L26, Novus Biologicals, cat. num. NBP2-80486, lot id: 931-2PABX200423; anti-CD206, clone E2L9N, Cell Signaling Technology, cat. num. 91992BF, lot id: 2; anti-CD3, clone Polyclonal, C-Terminal, Fluidigm, cat. num. 3170019D, lot id: 241809; anti-CD31, clone EP3095, Abcam, cat. num. ab216459, lot id: GR3296596-1; anti-CD4, clone OTI5D9, Novus Biologicals, cat. num. NBP2-70357, lot id: W001; anti-CD45, clone D9M8I, Fluidigm, cat. num. 3152018D, lot id: 2891711; anti-CD56, clone MEM-188, BioLegend, cat. num. 304602, lot id: B245303; anti-CD57, clone HNK-1, Fluidigm, cat. num. 359602, lot id: B287684; anti-CD68, clone KP1, Abcam, cat. num. ab233172, lot id: GR3290526-3; anti-CD8a, clone C8/144B, eBioscience, cat. num. 14-0085-82, lot id: 2132595; anti-Cleaved Caspase-3 (Asp175), clone 5A1E, Fluidigm, cat. num. 3172027D, lot id: 412006; anti-Collagen type I, clone Polyclonal, Fluidigm, cat. num. 3169023D, lot id: 2501807; anti-Histone H3, clone D1H2, Cell Signaling Technology, cat. num. 44998BF, lot id: 17; anti-IL-1 β , clone 3A6, Cell Signaling Technology, cat. num. 12242BF, lot id: 2; anti-IL-6, clone OTI3G9, Novus Biologicals, cat. num. NBP2-71027, lot id: F001; anti-iNOS, clone SP126, Abcam, cat. num. ab239990, lot id: GR3245763-1; anti-Keratin 8/18, clone C51, Fluidigm, cat. num. 3174022D, lot id: 181806; anti-Ki-67, clone B56, Fluidigm, cat. num. 3168022D, lot id: 2411806; anti-Mast cell Tryptase, clone AA1, BioLegend, cat. num. 369402, lot id: B287682; anti-Myeloperoxidase (MPO), clone E1E7I, Cell Signaling Technology, cat. num. 14569BF, lot id: 2; anti-Phospho-CREB

(Ser133), clone 87G3, Cell Signaling Technology, cat. num. 9198BF, lot id: 17; anti-Phospho-Stat3 (Tyr705), clone D3A7, Cell Signaling Technology, cat. num. 9145BF, lot id: 39; anti-SARS-CoV 2 Spike S1 protein, clone Rabbit Polyclonal, GeneTex, cat. num. GTX135356, lot id: 43932; anti-SC5b-9, clone Murine monoclonal, Quidel Corporation, cat. num. A239, lot id: 153280; anti-TTF1, clone 8G7G3/1 + NX2.1/690, Novus Biologicals, cat. num. NBP2-47773, lot id: 7080-3PABX200409; anti-Vimentin, clone D21H3, Fluidigm, cat. num. 3154014A, lot id: 591704.

Imaging mass cytometry, activation panel:

anti-pHH3s28 (BioLegend, clone HTA28, cat. num. 641002); anti-CD38 (Abcam, clone EPR4106, cat. num. ab226034); anti-CD45RO (BioLegend, clone UCHL1, cat. num. 304239); anti-TIM3 (Cell Signaling Technology, clone D5D5R, cat. num. 45208BF); anti-T-bet (Cell Signaling Technology, clone D6N8B, cat. num. 13232BF); anti-PD-1 (Cell Signaling Technology, clone D4W2J, cat. num. 86163BF); anti-CD86 (Cell Signaling Technology, clone E2G8P, cat. num. 91882BF); anti-CD44 (BioLegend, clone IM7, cat. num. 103001); anti-FoxP3 (Invitrogen, clone 236A/E7, cat. num. 1-4777-82); anti-CD161/CCR6 (Abcam, clone OTI1D8, cat. num. ab273666); anti-CD39 (Abcam, clone EPR20627, cat. num. ab236038); anti-VISTA (Fluidigm, clone D1L2G, cat. num. 3160025D); anti-GATA3 (Cell Signaling Technology, clone D13C9, cat. num. 5852BF); anti-HLA-DR (Abcam, clone EPR3692, cat. num. ab215985); anti-p-Nfkbp65 (BD Pharmingen, clone pS529, cat. num. 558393); anti-GranzymeB (Fluidigm, clone EPR20129-217, cat. num. 3167021D); anti-CD127 (Fluidigm, clone EPR2955(2), cat. num. 3168026D); anti-CD27 (Abcam, clone EPR8569, cat. num. ab256583); anti-S100A9 (Abcam, clone EPR3555, cat. num. ab227570); anti-PD-L1 (Abcam, clone SP142, cat. num. ab236238).

Immunohistochemistry:

Myeloperoxidase (Leica, clone 59A5); CD163 (Leica, clone MRQ-26).

Validation

Antibody clones compatible for human FFPE tissue processing were chosen based on extensive validation data-sheet from the manufacturer. In addition, these antibodies were validated post metal-conjugation for expected staining patterns and optimized for dilutions through a pilot run on IMC on normal lung tissue and COVID lung tissue.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Patient characteristics reflect the age, gender, and race representative of patients admitted with the same conditions.

Recruitment

Patients were cared for at New York Presbyterian/Weill Cornell Medicine Hospital. COVID-19 cases were collected between the months of April and June 2020. Inclusion criteria were dependent on consent to perform an autopsy and good preservation of tissue as formalin-fixed paraffin embedded blocks.

Ethics oversight

Consent for autopsy was obtained from the patients' families. Research was approved by the Institutional Review Board of Weill Cornell Medical College.

Note that full information on the approval of the study protocol must also be provided in the manuscript.