Resource

Spatially resolved single-cell atlas unveils a distinct cellular signature of fatal lung COVID-19 in a Malawian population

Check for updates

A list of authors and their affiliations appears at the end of the paper

Postmortem single-cell studies have transformed understanding of lower respiratory tract diseases (LRTDs), including coronavirus disease 2019 (COVID-19), but there are minimal data from African settings where HIV, malaria and other environmental exposures may affect disease pathobiology and treatment targets. In this study, we used histology and high-dimensional imaging to characterize fatal lung disease in Malawian adults with (n = 9) and without (n = 7) COVID-19, and we generated single-cell transcriptomics data from lung, blood and nasal cells. Data integration with other cohorts showed a conserved COVID-19 histopathological signature, driven by contrasting immune and inflammatory mechanisms: in US, European and Asian cohorts, by type I/III interferon (IFN) responses, particularly in blood-derived monocytes, and in the Malawian cohort, by response to IFN-γ in lung-resident macrophages. HIV status had minimal impact on histology or immunopathology. Our study provides a data resource and highlights the importance of studying the cellular mechanisms of disease in underrepresented populations, indicating shared and distinct targets for treatment.

Progress toward a human cell atlas (HCA) using single-cell RNA-sequencing (scRNA-seq) and high-dimensional imaging is transforming understanding of cells and their states in health and disease and is rapidly becoming a major resource for the development of novel treatments and vaccines[1–3]. However, data within this atlas are heavily biased toward populations in the Northern Hemisphere. Populations in sub-Saharan Africa (SSA) are particularly underrepresented[4]. Genetic and environmental factors may lead to important differences in cell development and cell compositions in different organs, thus affecting cellular responses to diseases, vaccines and therapies[5,6]. Capturing data from SSA populations is critical to assure more equitable benefit from the treatment advances derived from the HCA.

Immunomodulation plays a critical role in coronavirus disease 2019 (COVID-19) outcomes. Single-cell data from lung tissue facilitated identification of specific immunomodulatory targets[6–12]. Apart from our high-dimensional imaging data from a Brazilian cohort[13], single-cell data are restricted to populations in the Northern Hemisphere, such as clinical trial data validating their efficacy[14–16]. For future outbreaks of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or related viruses, this knowledge gap needs to be addressed. Indeed, given fewer intensive care facilities, the benefit of immunomodulation for severe disease is even more important in SSA. Although immunomodulatory therapies can be lifesaving, they can also be harmful[16]. Immunomodulation has broadly focused on two opposing strategies: augmenting inflammatory responses to aid viral clearance or attenuating inflammatory response to reduce hyperinflammation. Extensive studies in Northern Hemisphere cohorts have established that, by the time patients present with life-threatening illness, viral loads are declining, hyperinflammation generally predominates and, thus, anti-inflammatory interventions are more effective[15,16]. Given evidence that repeated exposure to malaria and other parasitic infections can induce immune tolerance[17,18], and because parasitic infections occur at higher levels in SSA populations[19–22], we hypothesized that the immune balance may be different in patients in SSA. Although sometimes this clinical context may be protective, in those who progress to severe disease a tolerance-skewed response might blunt immune-mediated viral clearance, leading to a more viral-driven pathology. However, the reverse

✉e-mail: matthias.marti@uzh.ch; thomasdan.otto@glasgow.ac.uk; christopher.moxon@glasgow.ac.uk

is also possible. High pathogen exposure can induce an accelerated inflammatory response on re-exposure to pathogens[5]. Either scenario might impact cellular responses in the lung and have important implications to inform treatment choices in SSA populations. To address some of these knowledge gaps, we conducted an autopsy study in well-characterized patients at a large public hospital in Malawi, a low-income country in SSA with high rates of malaria, tuberculosis (TB) and HIV.

Results

A conserved histological signature of COVID-19 in Malawian patients

We recruited patients with fatal illness aged 45–75 years who were admitted to Queen Elizabeth Central Hospital (QECH), Blantyre, from October 2020 to July 2021 and stratified them into three groups based on clinical criteria: (1) COVID-19 acute respiratory distress syndrome (ARDS) ($n = 9$); (2) lower respiratory tract disease (LRTD) ($n = 5$) with ARDS of diverse non-COVID-19 etiology; and (3) non-LRTD ($n = 2$) (Fig. 1a,b, Extended Data Table 1, Supplementary Table 1 and Methods). Most patients with COVID-19 were overweight or obese (78%), and four had type 2 diabetes (44%). Patients with LRTD and non-LRTD patients were generally underweight. HIV infection was common across groups: five patients with COVID-19 (56%), three patients with LRTD (60%) and two non-LRTD patients (67%) had been living with HIV. In three patients, this diagnosis was not known during life; the other six had been on antiretroviral treatment, although drug availability was limited during the pandemic. All patients had low CD4 counts (median, 134 cells per mm³) (Extended Data Table 1 and Supplementary Tables 1 and 2).

Using minimally invasive autopsy[23–25], we obtained lung, liver and brain samples in 16 patients, bone marrow in 15 patients and spleen in eight patients. A pathologist read hematoxylin and eosin (H&E)-stained tissue slides alongside patients' history and antemortem laboratory results. In the lung, the pathologist identified classical features of COVID-19 (refs. 26–32), which were less frequent in patients with LRTD (Supplementary Fig. 1 and Supplementary Table 3). COVID-19-specific changes were absent in other organs, focusing our further investigations on the lung. Then, two additional pathologists, blinded to diagnosis, scored the lung pathology in all 16 patients using more detailed semi-quantitative criteria[13]. In our patients with COVID-19, type II pneumocyte hyperplasia, vascular congestion, syncytia, granulation of tissue and lymphocyte infiltration were more common and severe than in the non-COVID-19 LRTD group (Extended Data Fig. 1a). No significant histopathology differences were observed due to HIV status (Extended Data Fig. 1b).

Lack of international consensus in COVID-19 lung pathology criteria, and of studies with systematic scoring, prevented quantitative comparison with other cohorts to assess similarities and differences. Therefore, we compared proportions of pulmonary lesion types with a study that combined cohorts from Europe and the United States (US)[27] and with our published Brazilian cohort[13] (Fig. 1c and Extended Data Fig. 1c). Acute alveolar changes, defined by neutrophil infiltration and fibrin deposition, were more frequent in the Malawian and Brazilian cohorts than in the US cohort. 'Chronic' alveolar changes with monocytes, macrophages or fibrosis were detected more frequently in the US and Malawian cohorts. In the Malawian cohort, 'chronic' disease was predominantly characterized by macrophage and monocytes with less fibrosis than in the US and Brazilian cohorts. Thus, despite a short duration from illness to death and demographic differences, patients in our Malawian cohort exhibited classical COVID-19 lung pathology but with a macrophage predominance in alveolar lesions.

Resident macrophages predominate in COVID-19 and neutrophils in LRTD

To assess pathology at the cellular level, tissue microarrays (TMAs) from 130 representative regions of interest (ROIs) from nine patients with COVID-19, three patients with LRTD and two non-LRTD patients, containing specific pathological lesions or normal lung areas, were analyzed by

imaging mass cytometry (IMC). We used a 39-antibody panel optimized for staining in lung tissue[13]. After cell segmentation and quality control, 76,369 cells were annotated from 118 ROIs and classified into subtypes (Fig. 2a, Extended Data Fig. 2a–d and Supplementary Table 4).

In our Malawian cohort, neutrophils (CD66b^pos CD11b^pos CD14^neg/low) were significantly more numerous in the patients with LRTD (49.6%) than in non-LRTD patients (21.1%) or in patients with COVID-19 (16.1%, $P \leq 0.001$) (Fig. 2b,d, Extended Data Fig. 2e and Supplementary Table 4). Reciprocally, macrophages were increased in patients with COVID-19 (44.1%) compared to patients with LRTD (30.4%) and non-LRTD patients (23.6%; $P \leq 0.0001$, Fig. 2b,c, Extended Data Fig. 2e and Supplementary Table 4). In contrast to data in prior published US and European cohorts[32,33], these were predominantly tissue-resident alveolar macrophages (CD206^high CD163^high Iba1^low MHCII^low CD14^neg) with a lower proportion of monocyte-derived CD14^high/int macrophages.

No consistent differences were observed in T cell numbers among the COVID-19, LRTD and non-LRTD disease groups, but, among patients with COVID-19, there was an expansion in regulatory T cells and proliferating T cells and a decrease in the ratio of effector memory (CD45RO^high) to naive (CD45RO^low) CD8 T cells (Fig. 2b and Supplementary Table 4). B cell numbers were not markedly different in COVID-19, although our panel had few markers to characterize B cells (Fig. 2b). Consistent with vascular pathology visualized by histology (fibrin deposition and thrombosis), there was increased endothelial cell activation in patients with COVID-19 compared to patients with LRTD and non-LRTD patients (Fig. 2b). Alveolar macrophages were the most common SARS-CoV-2^+ immune cell, followed by Arg^high neutrophils and interstitial macrophages (Fig. 2b and Supplementary Table 4). In the stromal compartment, type II pneumocytes (AT2) and epithelial cells were the most frequent SARS-CoV-2^+ cells. We found no SARS-CoV-2^+ endothelial cells or fibroblasts. Surprisingly, total numbers of SARS-CoV-2^+ cells were lower in HIV^+ patients (Extended Data Fig. 2d and Supplementary Table 4).

Exploiting the spatial and cellular resolution of IMC, we characterized cellular compositions of lesion types (Extended Data Fig. 2f) and then quantified lesion type levels by group (Fig. 2e). Type II pneumocyte hyperplasia was specific to the COVID-19 group. Diffuse alveolar damage occurred in both LRTD and COVID-19 but had different compositions, indicating different pathological processes: in LRTD, with neutrophil-driven fibrinopurulent inflammation; in COVID-19, a more heterogeneous immune cell composition, dominated by the presence of macrophages, except fibrin-containing lesions, which were neutrophilic. Together, these data implicate macrophages in alveolar damage and neutrophils in vascular damage and coagulopathic processes.

Common and unique myeloid compositions across cohorts

To systematically compare data across cohorts, we integrated IMC data from the Malawian patients with COVID-19 ($n = 9$) with our Brazilian cohort ($n = 11$) that employed the same antibody panel[13] and a US cohort[34] ($n = 10$) that used several of the same markers (Fig. 2f). Many similarities in cell proportions were observed among the three cohorts but also important differences (Fig. 2g,h and Supplementary Table 4). In the myeloid compartment, the Malawian and Brazilian cohorts were dominated by high levels of alveolar and interstitial macrophages compared to the US patients (Malawi 44.1%; Brazil 51.9%; US 26.3%). Reciprocally, the US cohort had the highest proportion of neutrophils (34.1%), and the Malawian cohort had the fewest (16.7%; Brazil 21.4%; $P < 4.25 \times 10^{-9}$). The proportion of B cells was also significantly higher in the Malawian cohort (6.89%; Brazilian cohort <0.1%; US cohort 0.7%; $P = 1.85 \times 10^{-16}$). In the stromal compartments, there was a lower proportion of fibroblasts in the Malawian cohort, in keeping with lower levels of fibrosis on histology.

SARS-CoV-2 antigen gives an indication of the quantity of viral material, although it does not distinguish replicating virus. The US cohort had the highest number of SARS-CoV-2^+ immune cells (23.7%; Malawi 7.6%; Brazil 8.3%). These were principally monocytes and neutrophils in the US cohort versus CD206^high tissue-resident alveolar macrophages

a

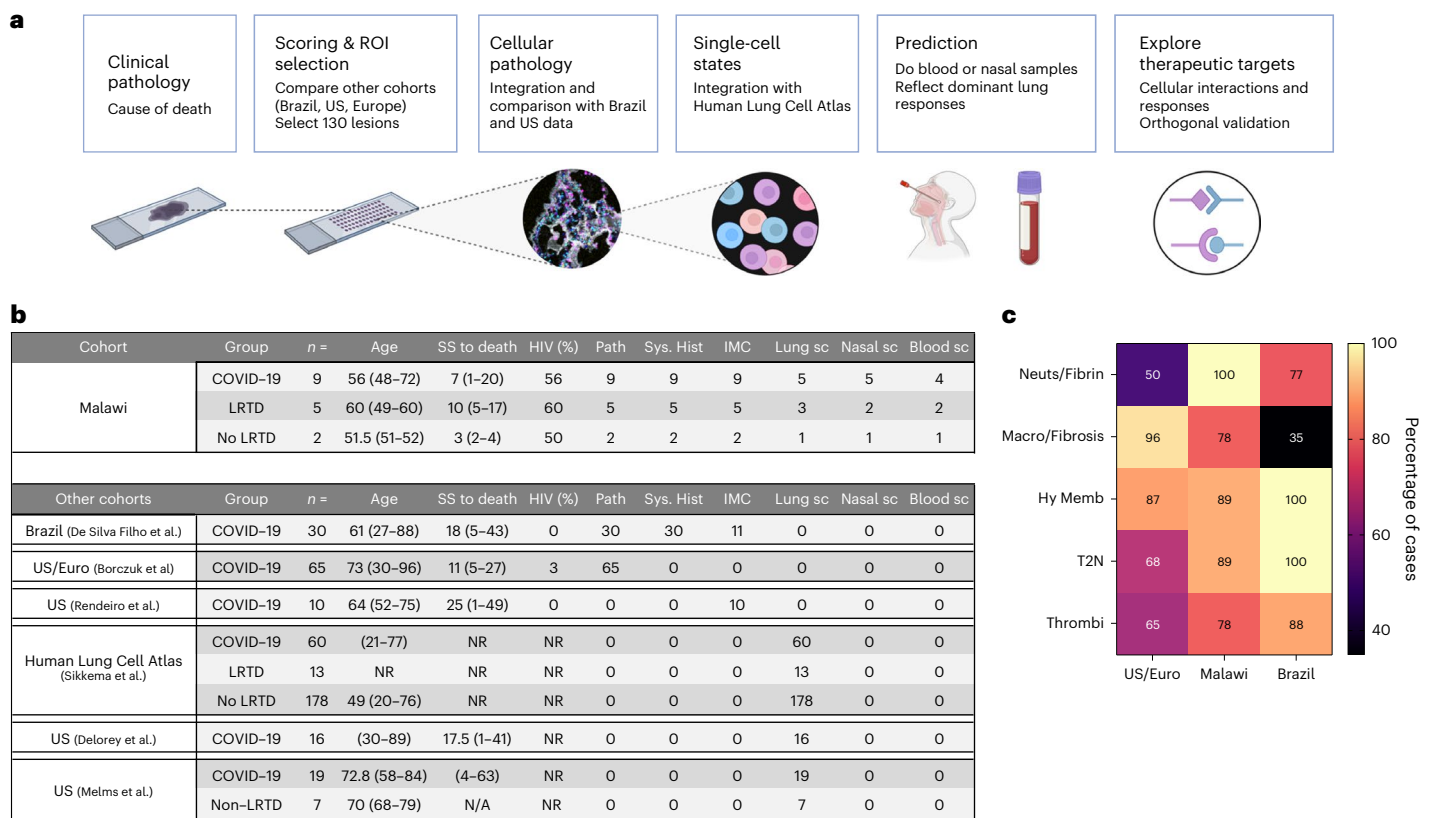| Clinical pathology	Scoring & ROI selection	Cellular pathology	Single-cell states	Prediction	Explore therapeutic targets
Cause of death	Compare other cohorts (Brazil, US, Europe) Select 130 lesions	Integration and comparison with Brazil and US data	Integration with Human Lung Cell Atlas	Do blood or nasal samples Reflect dominant lung responses	Cellular interactions and responses Orthogonal validation

b

Cohort	Group	n =	Age	SS to death	HIV (%)	Path	Sys. Hist	IMC	Lung sc	Nasal sc	Blood sc
Malawi	COVID–19	9	56 (48–72)	7 (1–20)	56	9	9	9	5	5	4
	LRTD	5	60 (49–60)	10 (5–17)	60	5	5	5	3	2	2
	No LRTD	2	51.5 (51–52)	3 (2–4)	50	2	2	2	1	1	1

Other cohorts	Group	n =	Age	SS to death	HIV (%)	Path	Sys. Hist	IMC	Lung sc	Nasal sc	Blood sc
Brazil (De Silva Filho et al.)	COVID–19	30	61 (27–88)	18 (5–43)	0	30	30	11	0	0	0
US/Euro (Borczuk et al)	COVID–19	65	73 (30–96)	11 (5–27)	3	65	0	0	0	0	0
US (Rendeiro et al.)	COVID–19	10	64 (52–75)	25 (1–49)	0	0	0	10	0	0	0
Human Lung Cell Atlas (Sikkema et al.)	COVID–19	60	(21–77)	NR	NR	0	0	0	60	0	0
	LRTD	13	NR	NR	NR	0	0	0	13	0	0
	No LRTD	178	49 (20–76)	NR	NR	0	0	0	178	0	0
US (Delorey et al.)	COVID–19	16	(30–89)	17.5 (1–41)	NR	0	0	0	16	0	0
US (Melms et al.)	COVID–19	19	72.8 (58–84)	(4–63)	NR	0	0	0	19	0	0
	Non–LRTD	7	70 (68–79)	N/A	NR	0	0	0	7	0	0

c

Fig. 1 | Study overview, overview of our cohort and comparator cohorts and histological lesion comparison with other cohorts. a, Overview of study approach, created with BioRender.com. **b**, Summary of the characteristics of our Malawian cohort versus published cohorts that we have used for different comparisons. **c**, Heatmap shows the proportion of patients in the three cohorts (US/European, Malawian and Brazilian) who have each given lesion type. SS to death, symptom start to death in days; Path, pathology, denotes the number of patients included in each cohort in which postmortem pathological features are described; Sys. Hist., systematic histopathology, denotes the number of patients included in each cohort with scoring of the frequency and severity of different lesions scored based on pre-defined criteria; IMC, imaging mass cytometry, denotes the number of patients with data for this; Lung sc, lung cell single-cell RNA-seq, denotes the number of patients with scRNA-seq data from lung tissue; Nasal sc, nasal cell single-cell RNA-seq, denotes the number of patients with scRNA-seq data from nasal tissue; Blood sc, blood cell single-cell RNA-seq, denotes the number of patients with these data. Hy Memb, hyline membranes; Macro, macrophages; N/A, not applicable; NR, not recorded; Neuts, neutrophilis; T2N, type II pneumocyte hyperplasia.

and neutrophils in the Malawian cohort and interstitial macrophages in the Brazilian cohort (Fig. 2g and Supplementary Table 4). In the stromal compartment, SARS-CoV-2 was detected in epithelial cells in all three cohorts but was significantly lower in the Malawian cohort (Fig. 2h; 5.8%, $P = 1.18 \times 10^{-10}$) compared to the Brazilian (13.1%) and US (12.1%) cohorts.

A possible explanation for different myeloid compositions in Malawian versus Brazilian and US patients is illness duration. Previous COVID-19 studies demonstrated that patients dying within 2 weeks of illness onset (early death) have different immune responses from those dying after 2 weeks (late death)[13,33]. Only one Malawian case was late death, and the median illness duration before death was shorter than in US and Brazilian cohorts (Fig. 1b). Therefore, instead, we compared early versus late death US patients. If illness duration was a major driver of cell compositions, myeloid cell proportions in early death US patients should be more like the Malawian patients (lower neutrophils and monocytes, higher macrophages). Instead, early death US patients had an even higher proportion of neutrophils (40.4% US early, 13.7% US late, 13.4% Malawi) and monocytes (11.1% US early, 7.0% US late, 3.8% Malawi) and a lower proportion of lung-resident macrophages (15.0% early versus 40.1% late, versus 44.1% Malawi early) (Extended Data Fig. 3a and Supplementary Table 4). Furthermore, on a dimension reduction plot, samples clustered by population (Malawian, US and Brazilian) rather than by illness duration (early versus late) (Extended Data Fig. 3c). SARS-CoV-2 variant may also be an important driver of variance, as all US and Brazilian patients were of the ancestral variant,

whereas Malawian patients were a mixture of Beta and Delta variant. However, cell proportions between patients with Beta or Delta variants in the Malawian patients were similar (Extended Data Fig. 3b). Furthermore, on a dimension reduction plot, we observed grouping by population, not viral variant (Extended Data Fig. 3d), suggesting that population is the main driver of lung immune composition.

Contrary to our initial hypothesis of a tolerized immune response in SSA populations, we found a highly inflammatory response and low levels of viral antigen in the Malawian patients versus other cohorts. The prominence of alveolar macrophages in lung lesions and enrichment of CD206^{high} tissue-resident macrophages in Malawian patients prompted further investigation of the inflammatory response in these cells.

scRNA-seq reveals an interferon-gamma-dominated lung macrophage response

To explore cellular responses in the lung at greater depth in Malawian patients, we used single-nuclei and single-cell RNA-seq in four patients with COVID-19, three patients with LRTD and one non-LRTD patient. Integrating 66,882 cells resulted in 16 cell clusters composed of a mixture of immune and stromal cells (Fig. 3a, Supplementary Fig. 2 and Supplementary Tables 5–7).

We detected few SARS-CoV-2 reads, suggesting that, at time of death, there was minimal replicating virus (Supplementary Fig. 4), consistent with our IMC data supporting inflammatory rather than direct viral-driven pathogenetic mechanisms.

We then undertook finer annotation of immune cell (Fig. 3b) and stromal cell (Fig. 3c) pools. Consistent with IMC data, we identified alveolar, interstitial and monocyte-derived macrophages and mature and immature neutrophils. Stromal cells included adventitial and alveolar fibroblasts, type I and type II pneumocytes (AT1 and AT2) and basal, secretory and ciliated epithelial cells. Cell proportions should be interpreted with caution given the few patients per group, but they showed cell diversity expansion in the COVID-19 and LRTD groups not observed or absent in the LRTD group (Extended Data Fig. 4a,b).

Principal differences in COVID-19 compared to LRTD were in myeloid cells, particularly alveolar macrophages (Fig. 3d–f), with few differences in lymphocytes, dendritic cells or stromal cells (Supplementary Tables 6 and 7). In alveolar macrophages, top differentially regulated genes included markers of tissue residency (*C1QC* and *C1QB*)[35] and factors shown to mediate lung fibrosis (*CCL18*)[36] and apoptosis (*S100 6*)[37] and myeloid activation and recruitment (*SPP1*)[38]. Interferon-gamma (IFN-γ) response protein (*IFI30*) and major histocompatibility complex (MHC) proteins (*HLA-DRA* and *HLA-DRB1*) were all upregulated, indicating response to IFN-γ.

This IFN-γ-dominant response contrasts with type I and type III dominant IFN responses shown to be critical in pathogenesis in Northern Hemisphere COVID-19 cohorts[16,39]. Given the prominence of alveolar macrophages in the immune response and in alveolar damage identified by IMC, we analyzed alveolar macrophage IFN response pathways. IFN-γ modules were expressed in a high proportion of cells (Fig. 3e) and strongly upregulated in COVID-19 compared to LRTD (log fold change, 0.1136; Fig. 3f). IL6/JAK/STAT pathway was also expressed, but to a lower extent, and the difference from LRTD was less clear (log fold change, 0.0418). In contrast, IFN-α, IFN-β and IFN-λ were minimally expressed, without clear differences from LRTD. This increased IFN-γ response could be due either to increased IFN-γ production or to increased responsiveness in macrophages. Using a pseudobulk approach, *IFNG* (IFN-γ gene) in T cells was not different between patients with COVID-19 and patients with LRTD (Extended Data Fig. 4c). In contrast, IFN-γ response genes were consistently upregulated in alveolar macrophages (Extended Data Fig. 4d), together implying that the increased IFN-γ response in patients with COVID-19 is due to a heightened response propensity of lung-resident macrophages rather than simply a heightened inflammatory response. In support of this, across other myeloid cells, IFN responses were heterogeneous, and TNF response was upregulated in the LRTD group in several cell types (Extended Data Fig. 4e).

Contrasting IFN responses between Malawian and other cohorts

To compare IFN responses with Northern Hemisphere cohorts, we integrated our Malawian single-cell data with multi-cohort COVID-19 (five cohorts, 60 patients), LRTD (one cohort, 13 patients) and non-LRTD (23 cohorts, 178 patients) data from the Human Lung Cell Atlas[10] (HLCA) (Fig. 4a; cohorts summarized in Fig. 1b).

Pathways indicative of IFN-γ response were increased across all cell types in the Malawian cohort (Fig. 4b, orange arrow) and were particularly upregulated in alveolar macrophages (Fig. 4c). Furthermore, *IFNG* (IFN-γ gene) was specifically increased in the Malawian cohort in CD4 and CD8 T cells versus HLCA COVID-19 and non-LRTD groups (Extended Data Fig. 5a). Other inflammatory pathways showed a mixture of upregulation and downregulation in the Malawian cohort compared to HLCA cohorts, including IL6/JAK/STAT (Fig. 4b, green arrow) and TNF-NFKB (Fig. 4b, blue arrow)−key targets for therapies being used in COVID-19. Many of the other IFN response genes were more upregulated in the HLCA cohorts or had a heterogenous distribution across cells, although, notably, monocyte-derived macrophages generally had a higher IFN response in HLCA COVID-19 cohorts (Extended Data Fig. 5a).

As with IMC data, we explored whether bias in illness duration explains differences in IFN responses among cohorts. Specifically, we compared IFN gene module scores in early versus late death patients in a US study with these metadata available. If the population-specific profiles were a function of illness stage, then, in the US cohort, we would expect higher IFN-γ levels in early death and higher IFN-α,β,λ levels in late death. Instead, IFN-α,β,λ responses were significantly stronger in early death, whereas the IFN-γ response was not different (Extended Data Fig. 5b).

Although many inflammatory pathways were shared between Malawian and Northern Hemisphere cohorts, the Malawian cohort exhibited amplified IFN-γ responses in lung-resident macrophages.

Nasal cell responses may be a useful proxy for lung cell responses

Although the lung is the principal organ involved in severe COVID-19 disease, it would be useful to know if we can predict lung responses using nasal or blood samples that can readily be obtained during life.

We performed scRNA-seq on nasal cells in eight patients (five COVID-19, two LRTD and one non-LRTD) and peripheral blood mononuclear cells (PBMCs) in seven patients (four COVID-19, two LRTD and one non-LRTD). We recovered 8,098 nasal cells that mapped to 10 clusters composing immune and stromal cells and 13,350 blood cells (Fig. 5a,b and Supplementary Fig. 3). Nasal macrophages had several differentially expressed (DE) genes in patients with COVID-19 versus patients with LRTD that mirrored alveolar macrophage responses (*SPP1*, *LGALS1* and *TMSB10*), including IFN-γ response genes (*HLA-DPB1*, *HLA-DQA1* and *C1QB*) (Fig. 5c). There was also *IFNG* (IFN-γ gene) upregulation in T cells in patients with COVID-19 (Fig. 5d). Pathway analysis also showed higher levels of IFN-γ response in macrophages and T cells (Fig. 5e). In blood, there was upregulation of inflammatory (*AREG*) and vascular damage (*NDRG1*) genes in COVID-19 in monocytes but no upregulation of *IFNG* or IFN-γ response genes (Fig. 5f). Hence, in our small cohort,

Fig. 2 | IMC reveals an immunopathological landscape of COVID-19 in Malawian patients driven by alveolar macrophages. a, UMAP embedding of the cell types identified in the lung samples by IMC, after supervised assignment to major cell types. Each major cell type was clustered, and resulting clusters were annotated and merged to extract the final set of cell types. Color key for cell types is on the right-hand side of **b**. Frequency of the immune cell types was identified in the postmortem lung samples by IMC according to clinical groups. The stacked bar plot shows the averaged frequency of the cell types by grouping the values from ROIs according to the clinical groups. Dashed lines highlight principal differences in major cell populations between COVID-19 and other respiratory disease groups. **c**, Representative denoised IMC images from one of 84 ROIs for patients with COVID-19 show abundant CD206^high macrophages (yellow) and few neutrophils (CD66b, red) and monocytes (CD14, purple). Scale bar, 140 μm. **d**, Representative denoised IMC images from one of 19 ROIs for a non-COVID-19 LRTD case show abundant neutrophils (CD66b, red) and few CD206⁺ macrophages (yellow). Scale bar, 140 μm. **e**, Frequency of histopathological

lesions based on matched H&E and IMC analysis of postmortem lung samples from the different clinical groups. The cellular composition and frequency of different cell types are indicated in Extended Data Fig. 3f. Dotted lines highlight the differences in proportions of broad response categories. **f**, UMAP embedding shows good integration (using the scvi-tools package) of IMC lung datasets from the Brazilian, US and Malawian COVID-19 cohorts based on 17 common antibody markers. **g**, Comparison of immune cell frequencies in IMC data from Brazilian, Malawian and US cohorts after integration shown in **f**; some major cell group differences are highlighted by dotted lines. Dashed box highlights apoptotic alveolar macrophages that are present only in the Malawian cohort. **h**, Comparison of stromal cell frequencies in IMC data from Brazilian, Malawian and US cohorts after integration shown in **f**. AM, alveolar macrophage; DAD, diffuse alveolar macrophage; EM, effector memory; IM, interstitial macrophage, mac, macrophage; neut, neutrophil; NK, natural killer; Treg, regulatory T; T2N, type II pneumocyte hyperplasia.

nasal cells better paralleled lung response than blood cells, supporting previous COVID-19 (refs. 40,41) and non-COVID-19 (ref. 42) studies that highlighted the utility of nasal cells to predict respiratory immune responses.

We also assessed whether cytokine responses in plasma or nasal fluid could distinguish the inflammatory or IFN-γ response in patients with COVID-19 versus patients with LRTD. In nasal fluid but not blood, there was a non-significant trend toward several cytokines being higher

in patients with COVID-19 than in patients with LRTD (Supplementary Fig. 5a,b). Using a pseudobulk approach in blood, nasal and lung cells, there was also no difference between *IFNG* or other cytokine genes between patients with COVID-19 and patients with LRTD (Supplementary Fig. 5c–e). Thus, IFN responses identified in single-cell data were not identified by bulk protein or transcriptomic approaches. This may reflect the greater discriminatory power of single-cell methods given small numbers in our study.
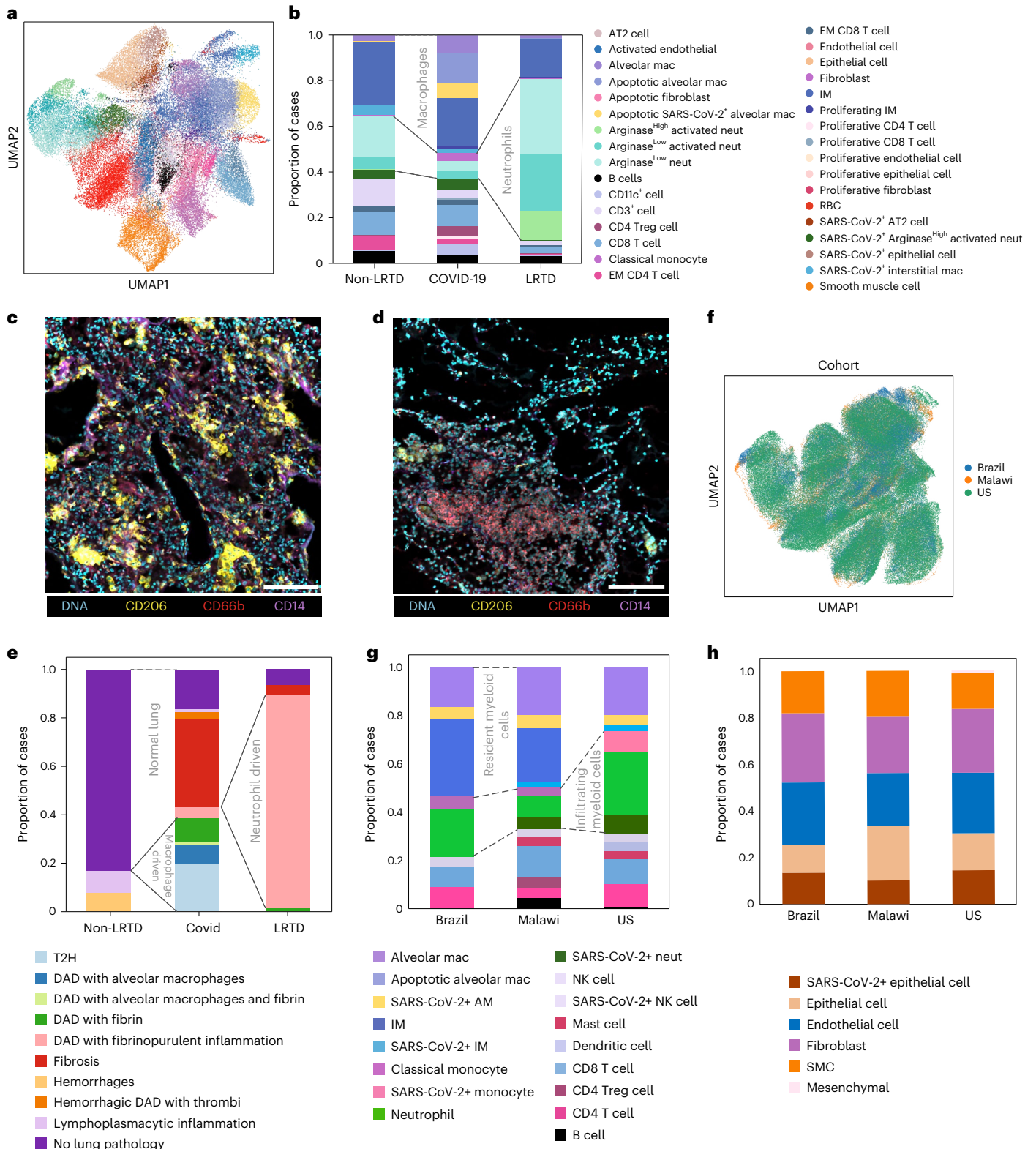
Different myeloid interactions predict alveolar and vascular damage

To assess the role of IFN-γ-responding resident macrophages in lung parenchymal pathology and neutrophil interactions in vascular pathology, and to predict molecular interactions for potential therapeutic targets, we used cell interaction methods. First, we performed unbiased receptor–ligand analysis of lung scRNA-seq data. A high proportion of predicted interactions involved lung-resident macrophages interacting with stromal cells and immune cells (Supplementary Fig. 6), fitting with findings from lesion analysis. We then did a targeted analysis of the interaction between T cells and alveolar macrophages, which identified a specific receptor–ligand interaction between *IFNG* from T helper cells (Th1) and IFN-γ receptor 1 (*IFNGR1*) on alveolar macrophages (Fig. 6a).

To validate these interactions in a spatial context in IMC, we conducted neighborhood enrichment analysis, which models cell proximity to predict likely cellular interactions. In the non-LRTD group, there were no significant interactions, whereas the LRTD group was completely dominated by neutrophil interactions (Extended Data Fig. 6a,b). In the COVID-19 group, several neighborhood enrichments were prominent—principally CD206[high] alveolar macrophages (with and without SARS-CoV-2 and apoptosis) with stromal cells, apoptotic fibroblasts and type II pneumocytes (Fig. 6b, left, and Extended Data Fig. 7a,b). This supports a role in tissue damage of CD206[high] macrophages. In contrast, the most prominent neighborhood enrichment for neutrophils was between SARS-CoV-2[+], Arg[high] neutrophils and activated endothelial cells, implicating neutrophils in vascular pathology (Fig. 6b, left, and Extended Data Fig. 7c).

To spatially resolve this IFN-γ response in Malawian patients, we integrated scRNA-seq and IMC data and mapped gene expression profiles onto IMC cells using a recently developed pipeline[43] (Fig. 6b). The integrated output showed upregulation of IFN-γ response genes, including *HLA-DR*, *IFI30* and *APOE*, and the inducible component of the IFN-γ receptor (*IFNGR2*) in tissue-resident CD206[high] alveolar and interstitial macrophages (Fig. 6c, right). Notably, the IFN-γ response was most prominent in the SARS-CoV-2[+] and apoptotic CD206[high] macrophage populations, predicted to interact with apoptotic fibroblasts and type II pneumocytes in the neighborhood analysis (Fig. 6c, left). Thus, mapping scRNA-seq data onto our IMC data not only validates the IFN-γ response but also implicates these IFN-γ-responding cells in lung stromal cell damage. Additionally, in situ hybridization staining across patients and 138 ROIs also highlighted significantly higher numbers of *IFNGR2*[+] cells in patients with COVID-19 than in non-LRTD controls but not between non-LRTD patients and patients with LRTD (Fig. 6d and Extended Data Fig. 8a–f). *IFNGR2* was predominantly in CD206[high] cells, which could be observed in diffuse alveolar damage lesions (Fig. 6e). In contrast, the number of *IFNG*[+] cells was not significantly increased in patients with COVID-19 (Fig. 6f), validating findings from scRNA-seq (Extended Data Fig. 4c). Thus, multiple orthogonal methods demonstrate an IFN-γ response in CD206[high] lung-resident macrophages, and this is best explained by the responsiveness of these cells rather than increased inflammation and *IFNG* production.

We then looked at validated interactions in COVID-19 in closer detail in scRNA-seq data to predict interactions that might indicate therapeutic targets. Macrophage interactions were frequently with type II pneumocytes (Extended Data Fig. 9a,b), in keeping with type II pneumocytes cells being a principal infected cell[44]. Several of these interactions involved macrophage inhibitory factor (*MIF*) from type II pneumocytes with *CD74*, *CD44* and *CXCR4* on macrophages, a classical response chain in macrophages and a key initiator of proliferation, chemotaxis and activation[45]. *ICAM-1* on type II pneumocytes was predicted to signal to integrins (*ITGB2-ITGAM*) on alveolar macrophages, an interaction involved in cellular attachment during recruitment. Another strong predicted interaction was *IL-34-CSF1R*, involved in triggering macrophage activation and chemotaxis. Reciprocally, there were several interactions between alveolar macrophages and epithelial cells that support their role in alveolar pathology, consistent with IMC data. These included *SPP1* and *TGFB1* with integrin (*ITGB6*) in AT1 cells (Extended Data Fig. 9a), interactions implicated in lung pathology and fibrosis[38,46,47]. We identified multiple neutrophil interactions with endothelial cells, indicating processes involved in neutrophil attachment to the vascular wall (for example, *ITGAL-ICAM-1*) and of activation by neutrophil granule proteins (*GRN-TNFRSF1A*) (Extended Data Fig. 9c,d), providing molecular validation supporting their role in coagulation, endothelial activation and vascular pathology, as suggested by analysis of lesions using IMC.

Discussion

We conducted a postmortem study and characterized pulmonary, blood and nasal immune responses in COVID-19 using histology, scRNA-seq and high-dimensional imaging in a Malawian population. We initially hypothesized that an attenuated immune response to SARS-CoV-2 in SSA populations might lead to high lung viral burden and, thus, to severe disease being the consequence of direct viral effects. This would indicate a need for different treatment approaches from Northern Hemisphere cohorts where hyperinflammation is predominant. Reassuringly, instead, we found a robust immune response, comparatively low levels of virus and many histopathological and immunological similarities to non-African cohorts, even in immunosuppressed patients with HIV. However, there were also differences that may have implications for therapy. We identified a dominant IFN-γ response in lung-resident macrophages, increased in comparison to a large multi-country integrated HLCA dataset. Spatially resolved interaction analysis and scRNA-seq receptor–ligand analysis implicated these IFN-γ-responding resident macrophages in lung damage. In contrast, IL6, TNF and type I/III IFN responses were not as prominent as in other cohorts.

There is crossover among the responses of different IFNs, yet this dominant IFN-γ response is noteworthy in this context as prior infection exposures, including malaria, have been shown to induce augmented IFN-γ response[17], specifically through epigenetic changes, termed trained immunity[18], and IFN-γ, enhanced by prior Bacille Calmette-Guérin (BCG) exposure, has been implicated in clearance of SARS-CoV-2 infection[48]. Such responses may be a double-edged sword in COVID-19, being generally protective (through more rapid viral clearance) but, in a subset of patients, leading to accelerated hyperinflammation and collateral tissue damage. To test the specific hypothesis that malaria, or other infections, is the driver of these different immune responses in COVID-19 in SSA populations would require larger studies of both SSA and non-SSA populations with different levels of exposure to these infections.

Considering the potential for immediate translation, existing therapies for COVID-19 target JAK/STAT (baricitinib), IL6 (toculimazab/

Fig. 3 | Lung single-cell atlas highlights IFN-γ response in alveolar macrophages. a, UMAP visualization of 66,882 lung cells across our cohort, colored by broad cell types cluster. **b**, UMAP visualization of 29,217 lung immune cells reclustered at a higher resolution to characterize the immune landscape, colored by cell type. **c**, UMAP visualization of 37,090 stromal lung cells reclustered at a higher resolution to characterize the stromal landscape, colored by cell type. **d**, Volcano plot showing top DE genes in alveolar macrophages in COVID-19 compared to LRTD with a significant adjusted *P* value (<0.05) and a log fold change

of more than 0.5 using MAST followed by Bonferroni multiple test correction. **e**, Dotplot showing the average gene module score of IFN response pathways across alveolar macrophages in COVID-19 and LRTD. **f**, Violin plots showing the gene module score across alveolar macrophages in gene sets associated with the gamma, alpha, beta, lambda and IL6 response in COVID-19 compared to LRTD. Black lines indicate the mean value across all cells, with the log fold change between means across conditions annotated above the plots. gdT cell, gamma-delta T cell; NK, natural killer; NS, not significant; Treg, regulatory T.
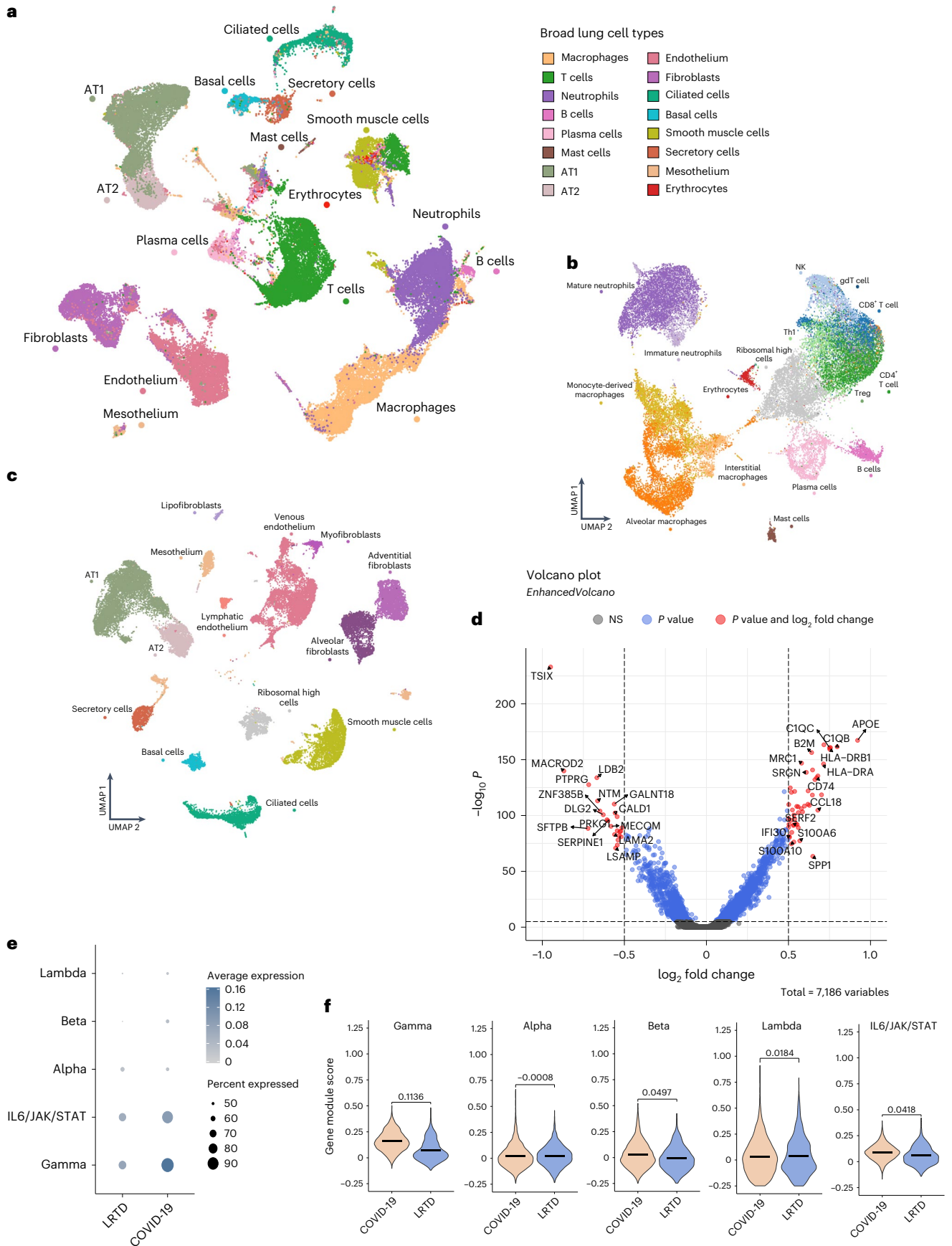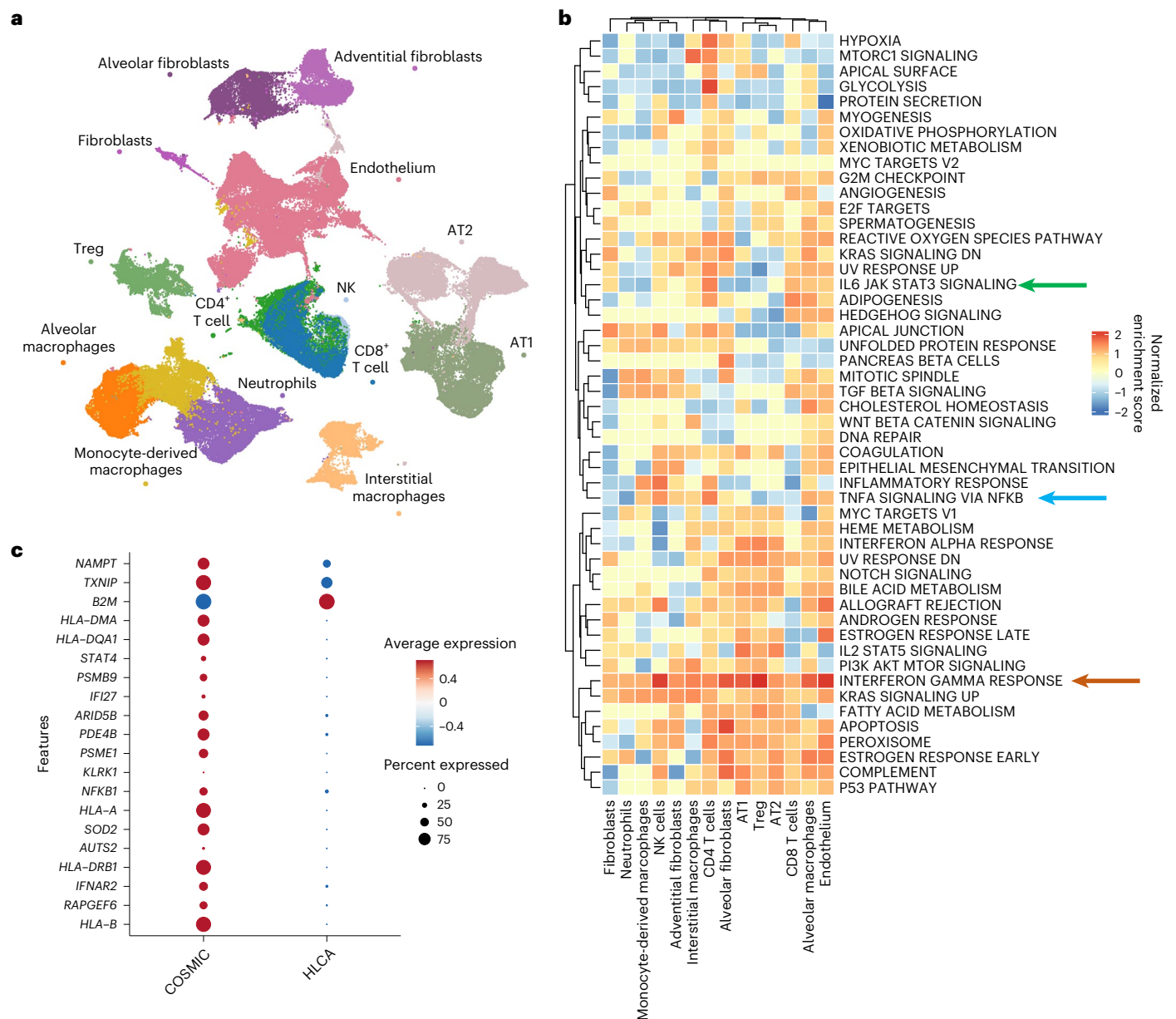
Fig. 4 | Integration with HLCA COVID-19 cohorts highlights dominant T cell macrophage IFN-γ axis in Malawian patients with COVID-19. a, UMAP visualization of 147,935 lung cells deriving from integrating cells from patients with COVID-19, patients with LRTD and non-LRTD patients from our cohort with cells from the HLCA from COVID-19, patients with LRTD and non-LRTD patients. Clusters are colored by cell type. **b**, Heatmap showing pathway analysis for DE genes in our COVID-19 cohort compared to the HLCA COVID-19 cohort. Shown are the 50 canonical hallmark gene sets (for list, see Supplementary Information) colored by the normalized enrichment score for each cell type. Gene Ontology pathways of interest are indicated by arrows (IL6 JAK STAT3 SIGNALING, green; TNFA SIGNALING VIA NFKB, blue; INTERFERON GAMMA RESPONSE, orange). **c**, Dotplot showing the average expression of top DE genes in the lung alveolar macrophages that contribute the highest in the hallmark gene set 'INTERFERON GAMMA RESPONSE' pathway in our COVID-19 cohort compared to the HLCA COVID-19 cohort. NK, natural killer; Treg, regulatory T.

sarilumab) or TNF (infliximab)[15,16]. JAK/STAT signaling is a conserved pathway for IFN responses, including IFN-γ[49]. Thus, our data, if corroborated, support potential efficacy of baricitinib over other treatments. Baricitinib is a small molecule (tablet) and, thus, highly suited to wide distribution[15].

Our data have several limitations. Our cohort is small and in a single center. We could not fully control for all variables, leaving the cause of different immune responses uncertain, including a potential impact of different viral variants. Although single-cell methods have a higher capacity to resolve complex data in small sample sizes, many analyses in our study are underpowered. It is, thus, unclear how representative our data are of the wider Malawian or other SSA populations. Studies

in other settings, and, ideally, large multi-center studies, are needed. Although this would be a complex undertaking, we have demonstrated that single-cell methods are feasible in an SSA setting, and our study provides useful templates. Although lung samples cannot readily be obtained in live patients, postmortem studies have limitations: cells may change or degrade, and pathological processes present early in disease are likely missed. However, postmortem studies in Northern Hemisphere settings with longer postmortem intervals identified validated targets[7]. Although minimally invasive autopsy is more feasible and acceptable than traditional open autopsy, blinded sampling may attenuate the identification and sampling of areas of pathology. However, except for large airway pathology, which was not sampled,
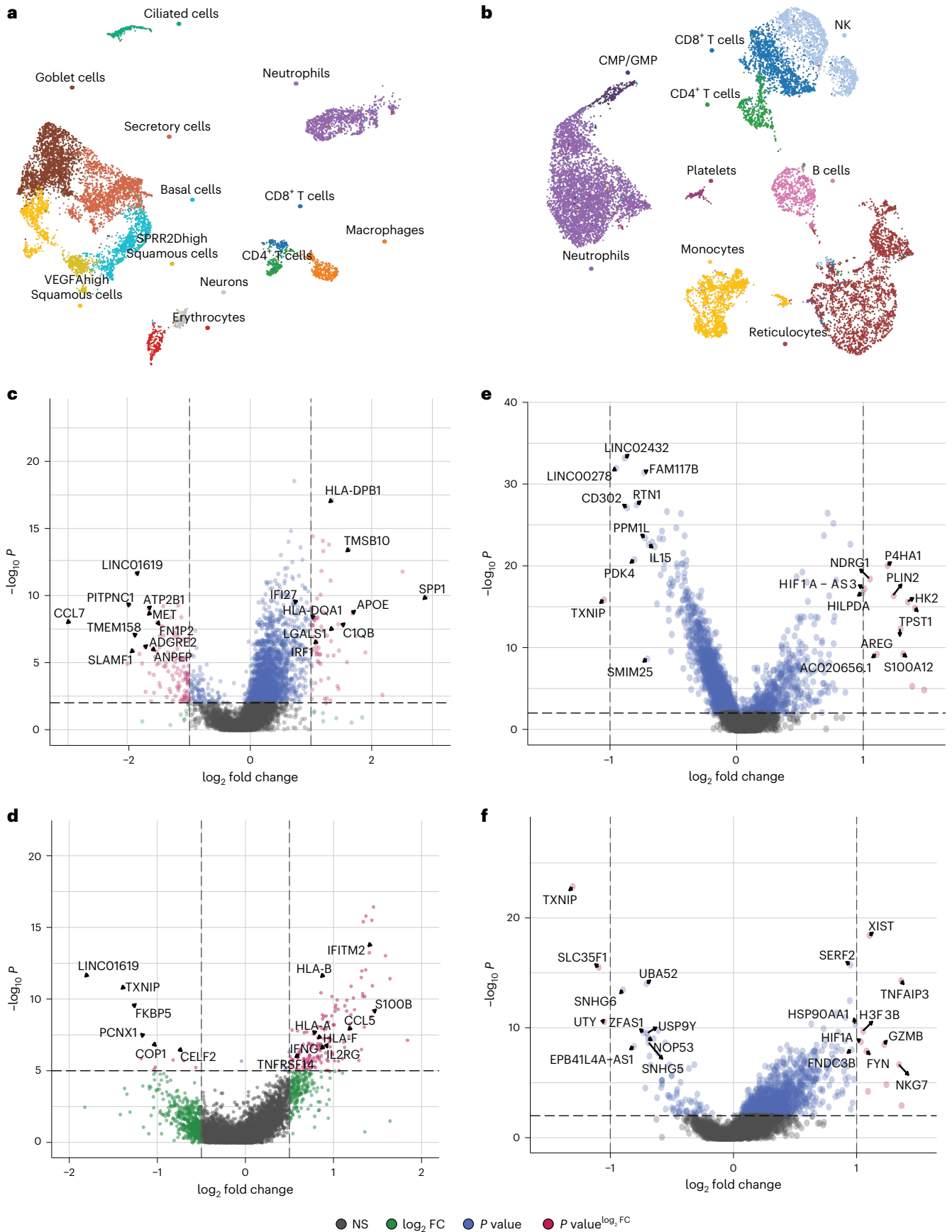
Fig. 5 | scRNA-seq of nasal and blood cells: nasal but not blood cells parallel lung IFN-γ response. a, UMAP visualization of 8,098 nasal cells across our cohort, colored by broad cell types. **b**, UMAP visualization of 13,350 peripheral blood cells across our cohort, colored by broad cell types. **c,d**, Volcano plots showing top DE genes in nasal macrophages (**c**) and T cells (**d**) in COVID-19 compared to LRTD with a significant adjusted P value (<0.05) and a log fold change of more than 0.5. **e,f**, Volcano plots showing top DE genes in peripheral blood monocytes (**e**) and CD4+ T cells (**f**) in COVID-19 compared to LRTD with a significant adjusted P value (<0.05) and a log fold change of more than 0.5. FC, fold change; NK, natural killer.
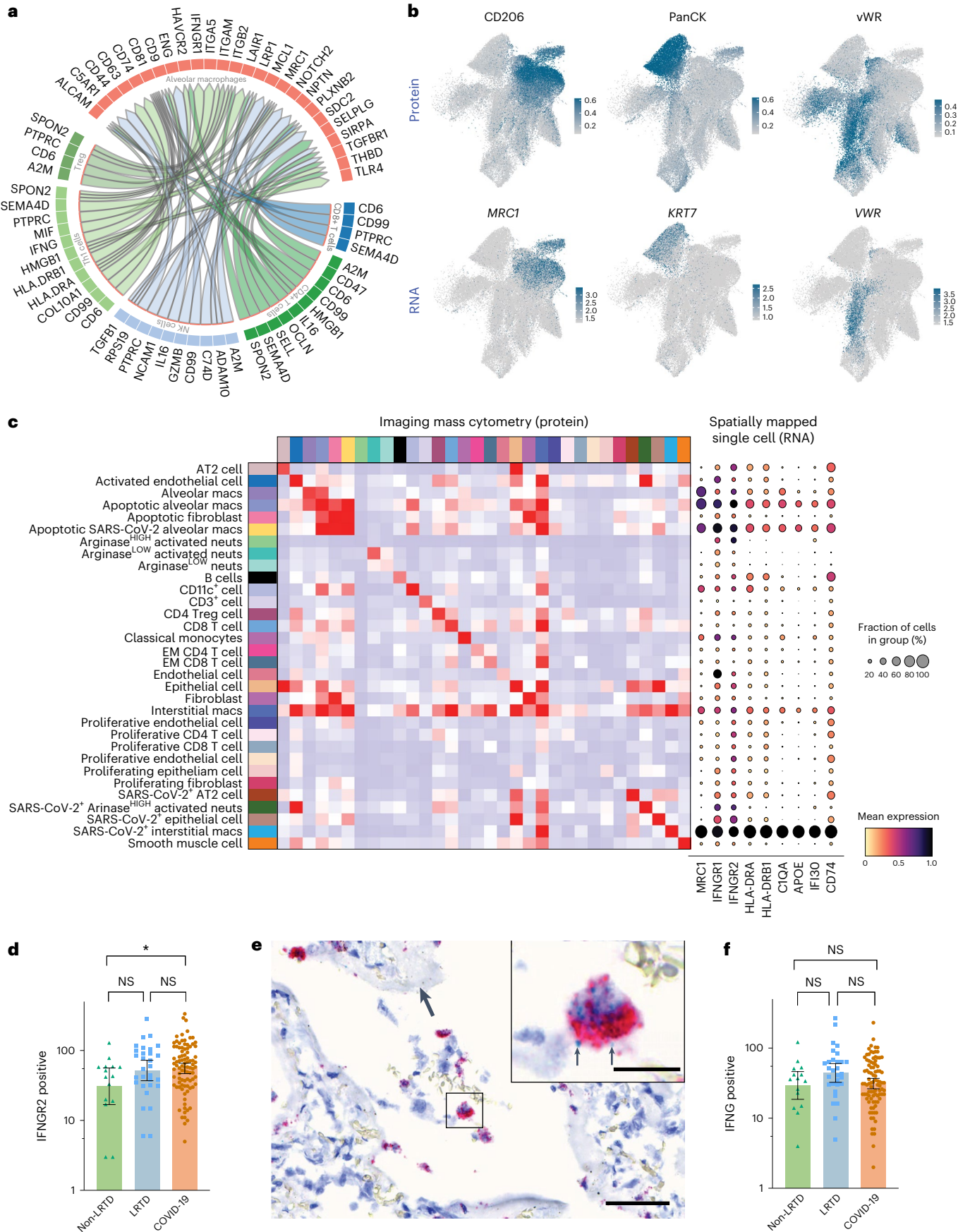
Fig. 6 | Spatially resolved cell interaction analysis predicts molecular mechanisms of alveolar and endothelial pathology. a, Circos plot showing the top cell–cell interactions from immune cells to alveolar macrophages in Malawian patients with COVID-19 versus Malawian patients with LRTD. Segments are colored by cell type with ligands and receptors labeled on the outside. Direction of the arrows shows the senders of communications that are expressing a given ligand to the receiver cell type expressing its cognate receptor. Inner tracks on sender segments are colored by the receiving cell type for ease of interpretation. **b**, UMAP plots to show expression levels of different hallmark proteins in different clusters by IMC and then below RNA levels from scRNA-seq data imputed by MaxFuse. **c**, Heatmaps showing co-localized cell types from IMC data, providing insight into potentially interacting cell types in the lung in patients with COVID-19; comparator LRTD and non-LRTD are in Extended Data Fig. 9. **d**–**f**, Quantification of mRNA in situ staining for IFNGR2 and IFNG in tissue. In total, 138 ROIs were taken based on multiple sampled areas from the left and right lung in nine patients with COVID-19, in three patients with LRTD and in two non-LRTD patients. Separate TMA sections were dual stained for either IFNGR2 and MRC1 (CD206) or IFNG and CD3E mRNA by in situ hybridization, and then the number of cells positive for each stain within respective cells

of interest IFNGR2 in CD206+ cells and IFNG in CD3E+ cells was analyzed by automatic quantification. Each dot represents the quantities of positive cells in an independent tissue core that were used as replicates for analysis in **d** and **f**. These data were log transformed and analyzed using one-way ANOVA and Tukey's multiple comparison test to adjust for multiple comparisons and a pre-defined alpha level of 0.05. Colored bars show the geometric mean, and error bars show the 95% confidence interval. **d**, Compared to non-LRTD patients, there were significantly higher numbers of IFNGR2+ cells in patients with COVID-19 but not in patients with LRTD (*$P = 0.0441$). **e**, Co-staining of IFNγR2 (red) and CD206 gene (green) using mRNA probes in lungs of patients infected with SARS-CoV-2. Lung of patients with COVID-19 shows, in the periphery of the damaged alveolar space fibrin (empty arrows) and in the lumen of the alveoli, cells with macrophage morphology expressing IFNGR2 (red signal, rectangle). The insert shows a higher magnification of the rectangle with a macrophage expressing CD206 in green (black arrows) and abundant IFNGR2 in red. Scale bars, 60 μm and 15 μm, respectively. Hematoxylin counterstaining. **f**, No significant difference was observed in quantities of IFNG+ cells among the different groups (NS, not significant; $P = 0.111$). EM, effector memory; macs, macrophages; neuts, neutrophils; Treg, regulatory T.

most COVID-19 features were identified. The studies that we used for comparisons had considerable variation in methods and demographics from ours, which may induce noise and bias. We used data integration methods that reduce, but do not eliminate, these. Reassuringly, findings were validated both by comparison to other cohorts and by orthogonal IMC data and targeted in situ staining.

Our data highlight the value of a combined scRNA-seq and high-dimensional imaging approach. They provide spatial and receptor–ligand validation for a role of IFN-γ-responding tissue-resident macrophages in alveolar damage and for neutrophils in endothelial activation. The data highlight specific molecular interactions involved in these processes. If validated by further work, some of these interactions may highlight additional plausible targets for intervention—for example, *MIF*, for which several small molecules are in clinical development for therapy in inflammatory disorders[45]. Our de-identified data, provided open access and through visualization tools, make an important resource for furthering the global understanding of COVID-19 pathogenesis and immune responses in SSA populations, as part of the HCA.

Online content

Any methods, additional references, Nature Portfolio 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/s41591-024-03354-3.

References

1. Regev, A. et al. The Human Cell Atlas. *eLife* **6**, e27041 (2017).
2. Lindeboom, R. G. H., Regev, A. & Teichmann, S. A. Towards a Human Cell Atlas: taking notes from the past. *Trends Genet.* **37**, 625–630 (2021).
3. Rood, J. E., Maartens, A., Hupalowska, A., Teichmann, S. A. & Regev, A. Impact of the Human Cell Atlas on medicine. *Nat. Med.* **28**, 2486–2496 (2022).
4. Majumder, P. P., Mhlanga, M. M. & Shalek, A. K. The Human Cell Atlas and equity: lessons learned. *Nat. Med.* **26**, 1509–1511 (2020).
5. Divangahi, M. et al. Trained immunity, tolerance, priming and differentiation: distinct immunological processes. *Nat. Immunol.* **22**, 2–6 (2021).
6. Mangino, M., Roederer, M., Beddall, M. H., Nestle, F. O. & Spector, T. D. Innate and adaptive immune traits are differentially affected by genetic and environmental factors. *Nat. Commun.* **8**, 13850 (2017).
7. Delorey, T. M. et al. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* **595**, 107–113 (2021).
8. Melms, J. C. et al. A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).
9. Ilieva, M., Tschaikowski, M., Vandin, A. & Uchida, S. The current status of gene expression profilings in COVID-19 patients. *Clin. Transl. Discov.* **2**, e104 (2022).
10. Sikkema, L. et al. An integrated cell atlas of the lung in health and disease. *Nat. Med.* **29**, 1563–1577 (2023).
11. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
12. COVID Tissue Atlas Consortium et al. Single-nuclei characterization of pervasive transcriptional signatures across organs in response to COVID-19. *eLife* **12**, e81090 (2023).
13. Da Silva Filho, J. et al. A spatially resolved single-cell lung atlas integrated with clinical and blood signatures distinguishes COVID-19 disease trajectories. *Sci. Transl. Med.* **16**, eadk9149 (2024).
14. van de Veerdonk, F. L. et al. A guide to immunotherapy for COVID-19. *Nat. Med.* **28**, 39–50 (2022).
15. Hall, M. W., Joshi, I., Leal, L. & Ooi, E. E. Immune immunomodulation in coronavirus disease 2019 (COVID-19): strategic considerations for personalized therapeutic intervention. *Clin. Infect. Dis.* **74**, 144–148 (2022).
16. Alijotas-Reig, J. et al. Immunomodulatory therapy for the management of severe COVID-19. Beyond the anti-viral therapy: a comprehensive review. *Autoimmun. Rev.* **19**, 102569 (2020).
17. de Jong, S. E. et al. Systems analysis and controlled malaria infection in Europeans and Africans elucidate naturally acquired immunity. *Nat. Immunol.* **22**, 654–665 (2021).
18. Franklin, B. S. et al. Malaria primes the innate immune response due to interferon-γ induced enhancement of Toll-like receptor expression and function. *Proc. Natl Acad. Sci. USA* **106**, 5789–5794 (2009).
19. Afolabi, M. O. et al. Prevalence and distribution pattern of malaria and soil-transmitted helminth co-endemicity in sub-Saharan Africa, 2000–2018: a geospatial analysis. *PLoS Negl. Trop. Dis.* **16**, e0010321 (2022).
20. Cohee, L. M. et al. Understanding the intransigence of malaria in Malawi. *Am. J. Trop. Med. Hyg.* **107**, 40–48 (2022).
21. Mategula, D. et al. Malaria burden stratification in Malawi-a report of a consultative workshop to inform the 2023-2030 Malawi Malaria Strategic Plan. *Wellcome Open Res.* **8**, 178 (2023).
22. Menendez, C. et al. Avoiding another lost decade in reducing malaria burden in African infants and young children. *Lancet Glob. Health* **10**, e1385–e1386 (2022).
23. Breiman, R. F. et al. Postmortem investigations and identification of multiple causes of child deaths: an analysis of findings from

the Child Health and Mortality Prevention Surveillance (CHAMPS) network. *PLoS Med.* **18**, e1003814 (2021).

24. Chawana, R. et al. Potential of minimally invasive tissue sampling for attributing specific causes of childhood deaths in South Africa: a pilot, epidemiological study. *Clin. Infect. Dis.* **69**, S361–S373 (2019).

25. Taylor, A. W. et al. Initial findings from a novel population-based child mortality surveillance approach: a descriptive study. *Lancet Glob. Health* **8**, e909–e919 (2020).

26. Ackermann, M. et al. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. *N. Engl. J. Med.* **383**, 120–128 (2020).

27. Borczuk, A. C. et al. COVID-19 pulmonary pathology: a multi-institutional autopsy cohort from Italy and New York City. *Mod. Pathol.* **33**, 2156–2168 (2020).

28. Duarte-Neto, A. N. et al. Ultrasound-guided minimally invasive tissue sampling: a minimally invasive autopsy strategy during the COVID-19 pandemic in Brazil, 2020. *Clin. Infect. Dis.* **73**, S442–S453 (2021).

29. Himwaze, C. M. et al. Post-mortem examination of hospital inpatient COVID-19 deaths in Lusaka, Zambia—a descriptive whole-body autopsy series. *Int J. Infect. Dis.* **108**, 363–369 (2021).

30. Menter, T. et al. Postmortem examination of COVID-19 patients reveals diffuse alveolar damage with severe capillary congestion and variegated findings in lungs and other organs suggesting vascular dysfunction. *Histopathology* **77**, 198–209 (2020).

31. Mudenda, V. et al. Histopathological evaluation of deceased persons in Lusaka, Zambia with or without coronavirus disease 2019 (COVID-19) infection: results obtained from minimally invasive tissue sampling. *Clin. Infect. Dis.* **73**, S465–S471 (2021).

32. Xu, Z. et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir. Med.* **8**, 420–422 (2020).

33. Nienhold, R. et al. Two distinct immunopathological profiles in autopsy lungs of COVID-19. *Nat. Commun.* **11**, 5086 (2020).

34. Rendeiro, A. F. et al. The spatial landscape of lung pathology during COVID-19 progression. *Nature* **593**, 564–569 (2021).

35. Mulder, K. et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. *Immunity* **54**, 1883–1900 (2021).

36. Prasse, A. et al. A vicious circle of alveolar macrophages and fibroblasts perpetuates pulmonary fibrosis via CCL18. *Am. J. Respir. Crit. Care Med.* **173**, 781–792 (2006).

37. Leclerc, E., Fritz, G., Weibel, M., Heizmann, C. W. & Galichet, A. S100B and S100A6 differentially modulate cell survival by interacting with distinct RAGE (receptor for advanced glycation end products) immunoglobulin domains. *J. Biol. Chem.* **282**, 31317–31331 (2007).

38. MacDonald, L. et al. COVID-19 and RA share an SPP1 myeloid pathway that drives PD-L1+ neutrophils and CD14+ monocytes. *JCI Insight* **6**, e147413 (2021).

39. Sposito, B. et al. The interferon landscape along the respiratory tract impacts the severity of COVID-19. *Cell* **184**, 4953–4968 (2021).

40. Roukens, A. H. E. et al. Prolonged activation of nasal immune cell populations and development of tissue-resident SARS-CoV-2-specific CD8+ T cell responses following COVID-19. *Nat. Immunol.* **23**, 23–32 (2022).

41. Morton, B. et al. Distinct clinical and immunological profiles of patients with evidence of SARS-CoV-2 infection in sub-Saharan Africa. *Nat. Commun.* **12**, 3554 (2021).

42. Carniel, B. F. et al. Pneumococcal colonization impairs mucosal immune responses to live attenuated influenza vaccine. *JCI Insight* **6**, e141088 (2021).

43. Chen, S. et al. Integration of spatial and single-cell data across modalities with weakly linked features. *Nat. Biotechnol.* **42**, 1096–1106 (2024).

44. Bosmuller, H., Matter, M., Fend, F. & Tzankov, A. The pulmonary pathology of COVID-19. *Virchows Arch.* **478**, 137–150 (2021).

45. Bilsborrow, J. B., Doherty, E., Tilstam, P. V. & Bucala, R. Macrophage migration inhibitory factor (MIF) as a therapeutic target for rheumatoid arthritis and systemic lupus erythematosus. *Expert Opin. Ther. Targets* **23**, 733–744 (2019).

46. Shan, M. et al. Cigarette smoke induction of osteopontin (SPP1) mediates T_H17 inflammation in human and experimental emphysema. *Sci. Transl. Med.* **4**, 117ra119 (2012).

47. Das, A. et al. Molecular and immune signatures, and pathological trajectories of fatal COVID-19 lungs defined by in situ spatial single-cell transcriptome analysis. *J. Med Virol.* **95**, e29009 (2023).

48. Hilligan, K. L. et al. Bacterial-induced or passively administered interferon gamma conditions the lung for early control of SARS-CoV-2. *Nat. Commun.* **14**, 8229 (2023).

49. Taks, E. J. M., Moorlag, S., Netea, M. G. & van der Meer, J. W. M. Shifting the immune memory paradigm: trained immunity in viral infections. *Annu. Rev. Virol.* **9**, 469–489 (2022).

James Nyirenda[1,2,3,26], Olympia M. Hardy [1,26], João Da Silva Filho [1,4,26], Vanessa Herder [5], Charalampos Attipa[1,3,6,7], Charles Ndovi[1,2,3], Memory Siwombo[8,9], Takondwa Rex Namalima[8], Leticia Suwedi[3,10], Georgios Ilia[5], Watipenge Nyasulu[1,2,3], Thokozile Ngulube[1,3], Deborah Nyirenda[3,10], Leonard Mvaya[3], Joseph Phiri[3], Dennis Chasweka[9], Chisomo Eneya[9], Chikondi Makwinja[9], Chisomo Phiri[9], Frank Ziwoya[9], Abel Tembo[9], Kingsley Makwangwala[9], Stanley Khoswe[9], Peter Banda[8,9], Ben Morton [3,11], Orla Hilton [12], Sarah Lawrence[13], Monique Freire dos Reis[14,15], Gisely Cardoso Melo[16], Marcus Vinicius Guimaraes de Lacerda[16,17,18], Fabio Trindade Maranhão Costa[19], Wuelton Marcelo Monteiro[15,16], Luiz Carlos de Lima Ferreira[15,16], Carla Johnson [1], Dagmara McGuinness [1], Kondwani Jambo [3,11], Michael Haley[20], Benjamin Kumwenda[9], Massimo Palmarini [5], Donna M. Denno[13], Wieger Voskuijl[8,9,21], Steve Bvuobvuo Kamiza[9,27], Kayla G. Barnes[3,22,23,24,25,27], Kevin Couper[20,27], Matthias Marti [1,4,27 ✉], Thomas D. Otto [1,27 ✉] & Christopher A. Moxon [1,2,3,11,22,27 ✉]

[1]School of Infection and Immunity, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow, UK. [2]Blantyre Malaria Project, Kamuzu University of Health Sciences, Blantyre, Malawi. [3]Malawi-Liverpool-Wellcome Programme, Kamuzu University of Health Sciences, Blantyre, Malawi. [4]Universität Zürich, Institut für Parasitologie, Zurich, Switzerland. [5]MRC-University of Glasgow Centre for Virus Research, Glasgow, UK. [6]The Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK. [7]The Roslin Institute, University of Edinburgh, Edinburgh, UK. [8]Queen Elizabeth Central Hospital, Blantyre, Malawi. [9]Kamuzu University of Science of Health Sciences, Blantyre, Malawi. [10]International Public Health, Liverpool School of Tropical Medicine, Liverpool, UK. [11]Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK. [12]Department of Infectious Diseases, Imperial College London, London, UK. [13]Department of Global Health and Pediatrics, University of Washington, Seattle, WA, USA. [14]Department of Education and Research, Oncology Control Centre of Amazonas State (FCECON), Manaus, Brazil. [15]Postgraduate Program in Tropical Medicine, University of Amazonas State, Manaus, Brazil. [16]Tropical Medicine Foundation Dr. Heitor Vieira Dourado, Manaus, Brazil. [17]Institute Leônidas & Maria Deane, Fiocruz, Manaus, Brazil. [18]The University of Texas Medical Branch, Galveston, TX, USA. [19]Department of Genetics, Evolution, Microbiology and Immunology, University of Campinas, Campinas, Brazil. [20]Division of Immunology, Immunity to Infection & Respiratory Medicine, Faculty of Biology, University of Manchester, Manchester, UK. [21]Department of Global Health, Amsterdam University Medical Centers, Amsterdam, The Netherlands. [22]Department of Tropical Disease Biology, Liverpool School of Tropical Medicine, Liverpool, UK. [23]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [24]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [25]Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK. [26]These authors contributed equally: James Nyirenda, Olympia M. Hardy, João Da Silva Filho. [27]These authors jointly supervised this work: Steve Bvuobvuo Kamiza, Kayla G. Barnes, Kevin Couper, Matthias Marti, Thomas D. Otto, Christopher A. Moxon. ✉e-mail: matthias.marti@uzh.ch; thomasdan.otto@glasgow.ac.uk; christopher.moxon@glasgow.ac.uk

Methods

Ethics

This study complies with all relevant ethical regulations. The protocol for the Malawian study was approved by the National Health Scientific Research Committee in Malawi (protocol number 07/09/1913) and by the Medical Veterinary Life Sciences ethics committee in Glasgow (protocol number 200190041). The study protocol for the Brazilan study was approved by the local research ethics committee at Tropical Medicine Foundation Dr. Heitor Vieira Dourado, Manaus, Western Brazilian Amazon (protocol numbers CAAE:30152620.1.0000.0005 and CAAE:32077020.6.0000.0005). Additional studies on this cohort were published separately[13,50]. We also used open-access de-identified IMC data from a published US-based autopsy study conducted at New York Presbyterian/Weill Cornell Medicine Hospital, for which the study protocol was approved by the institutional review board at Weill Cornell Medical College[34]. Informed consent was taken from the families of deceased patients for all patients at all sites.

Patients

We recruited patients aged 45–75 years who were admitted to QECH, Blantyre, between October 2020 and July 2021, during which there were two epidemiological waves driven by different SARS-CoV-2 variants: Beta (December 2020–February 2021) and Delta (May–July 2021)[41]. Patients admitted with respiratory signs were routinely tested for SARS-CoV-2 at QECH. We recruited patients into three groups based on clinical criteria: (1) a COVID-19 group ($n = 9$) with clinical features suggesting acute respiratory distress (ARDS, oxygen requirement and respiratory signs on either clinical examination or chest X-ray changes or both) and who had at least one nasal swab positive for SARS-CoV-2 on admission; (2) a non-COVID-19 LRTD group ($n = 5$) with clinical signs of ARDS but negative for SARS-CoV-2 on admission and during hospitalization; and (3) a no-LRTD, COVID-19-negative group ($n = 2$) with no oxygen requirement and no clinical signs of LRTD and for which the admission and any subsequent nasal swabs were negative for SARS-CoV-2 on polymerase chain reaction (PCR) (Fig. 1b and Extended Data Table 1). Clinical, premortem and postmortem laboratory data were entered into REDCap; double entry was used and checked by a third investigator, with discrepant results resolved by consulting the original source. The study only recruited patients who died between 24:00 and 12:00 to minimize the postmortem interval and to avoid doing any autopsies at night. None of the patients included had received any SARS-CoV-2 vaccine; only approximately 2% of the Malawian population had received a first dose by study completion.

Minimally invasive autopsy

We used minimally invasive tissue sampling (MITS) to conduct autopsies with large-bore needle biopsies of organ samples rather than full autopsy[23]. Being more culturally acceptable, MITS is widely used to determine cause of death in pediatric studies[23–25], showing good concordance with full autopsy[24]. From our ongoing pediatric MITS studies in Malawi, we adapted protocols for adult patients with COVID-19 to obtain tissue suitable for scRNA-seq and IMC, based on the protocol from the Child Health and Mortality Prevention Surveillance (CHAMPS) network but with adaptations. A larger-caliber needle (11 gauge) was used for biopsies to obtain larger tissue samples. Samples were taken from the brain through supraorbital sampling from both left and right sides. From each lung, samples were taken from lower-middle and upper zones from a single entry point, angling the needle to sample different areas. Nasal cells were collected from the nasal inferior turbinate using curettes (ASL Rhino-Pro, Arlington Scientific). Two curettes were collected from each nostril, and the cells were placed immediately into ice-cold HypoThermosol (STEMCELL Technologies). Cells were transported on ice in a cold box immediately to the laboratory and were spun at 300g for 5 min for either immediate processing for scRNA-seq or storage in a CryoStor 10 (see below). Nasal fluid was collected using matrix strips (Nasosorption, Hunt Developments). One strip was used per nostril. Personal protective equipment (PPE) was worn by all staff involved in the autopsies and for all work in the laboratory. Laboratory work on samples was performed in vented laminar flow hoods.

Processing and storage of samples

Biopsies from each organ were collected in three different ways for different downstream workflows: (1) for paraffin embedding for histology and IMC, put in 10% neutral buffered formalin; (2) for viable cells, put in ice-cold HypoThermosol (STEMCELL Technologies) for transport to the laboratory and then slow freeze in a CryoStor 10 (STEMCELL Technologies); and (3) for snap-frozen cells, put in cryovials and then seal and immediately submerge in liquid nitrogen.

Biopsies were fixed in 10% neutral buffered formalin for 4–8 h, rinsed in water and then embedded in paraffin blocks. Samples for viable cells were rinsed and cut into pieces of approximately 20–50 mm and then put into ice-cold CryoStor for 15–30 min before transfer to a −80 °C freezer in a chilled cryogenic storage container (CoolCell, Corning).

Blood cells collected into sodium heparin tubes were separated from plasma by spinning at 400g for 10 min. Plasma was then removed and spun for an additional 10 min at 1,500g, and plasma was frozen in aliquots at −80 °C. Cells were resuspended in 10% FBS in PBS, and PBMCs were separated using Ficoll-Paque with a 27-min spin at 450g and either used immediately for scRNA-seq or pelleted and resuspended in ice-cold CryoStor 10 and then moved to a −80 °C freezer in a chilled cryogenic storage container (CoolCell, Corning). The next day, samples were moved from the −80 °C freezer to liquid nitrogen for long-term storage. Snap-frozen samples were transferred in a liquid nitrogen dewar and then moved to liquid nitrogen storage tanks for long-term storage.

Pathology and organ-specific scoring

Formalin-fixed tissues were paraffin embedded (FFPE) for lung, bone marrow, brain, spleen and liver to make blocks. FFPE blocks were sectioned at 2–4-μm thickness, mounted on glass slides and stained with H&E. A medical pathologist (S.K.) reviewed tissue slides, alongside patient histories and antemortem laboratory results per standard clinical practice, and completed an organ-specific scoring proforma that included COVID-19 features (Supplementary Table 3). Then, for a non-biased assessment, two additional pathologists, blinded to diagnosis, scored the lung pathology in all patients using systematic scoring criteria. Lung tissue was scored independently by two additional pathologists (C.A. and V.H.) who were blinded to patient history and previous diagnoses. After individual scoring, any discrepancies were discussed by joint review of the slides until a consensus was reached. The lung scoring was semi-quantitative for the parameters indicated in Extended Data Fig. 1a–c. Subsequently, we characterized each sample with a dominant histological characteristic—for example, fibrinopurulent inflammation/pneumonia in case the neutrophil infiltration with fibrin extravasation was marked next to a mild infiltrate of lymphocytes, plasma cells and macrophages. Whole-tissue slides from lung samples in our nine patients with COVID-19 can be accessed in their entirety and visualized at various magnifications, as if they were observed under a microscope, using our virtual microscope tool: https://covid-atlas.cvr.gla.ac.uk (de-identified slides will be uploaded and publicly viewable upon publication).

After scoring, in each lung biopsy, the most representative areas were manually selected based on the scoring performed on the H&E-stained section to create the TMAs with cores of 1 mm in diameter using the TMA Grand Master (3DHISTECH) and CaseViewer software (version 2.4.0119028). At least eight ROIs were taken from each case (four left, four right). From the newly created TMA-FFPE blocks, 4-mm-thick sections were cut and used for downstream IMC, in situ hybridization or bright-field immunohistochemistry.

Cause of death attribution

A panel consisting of the pathologist who reviewed the patients, respiratory physician, intensive care physician, infectious disease physician and two trainee doctors reviewed all the patients to assign a cause of death. Codes assigning death were given according to International Classification of Diseases (ICD) codes and using the standard coding system used for death certification. The review consisted of a review of the clinical notes, premortem and postmortem laboratory results and the pathology report. Each member reviewed the documents independently and reached an individual verdict. When there were discrepancies, a consensus was reached through discussion.

Multiparameter cytokine assay

Cytokine levels were measured in plasma and nasal fluid samples using Luminex with the Inflammation 20-Plex Human ProcartaPlex panel (Thermo Fisher Scientific, EPX200-12185-901) according to the manufacturer's protocol and levels measured with a Luminex MagPix device. Data were transformed with a $\log_2$ and for the visualization with ComplexHeatmap in R with a z-score by cytokine.

IMC

Sections from TMAs underwent deparaffinization, followed by antigen retrieval at 96 °C for 30 min in Tris-EDTA at pH 8.5. Non-specific binding was blocked with 3% BSA for 45 min, followed by incubation with lanthanide-conjugated primary antibodies (overnight at 4 °C), which were diluted in PBS with 0.5% BSA (Supplementary Information). Antibodies were conjugated with metals using Maxpar Antibody Labeling Kits (Standard BioTools) and were validated with positive control tissue (tonsil and spleen for immune-targeted antibodies). Slides were then washed with 0.1% Triton X-100 in PBS, followed by nuclear staining with iridium (1:400; Intercalator-Ir, Standard Bio Tools) for 30 min at room temperature and, finally, briefly (10 s) washed with ultrapure water and air dried. Images were acquired on a Hyperion imaging mass cytometer as per the manufacturer's instructions (Standard BioTools). Each TMA core was imaged in a separate ROI.

IMC analysis

Pre-processing, imaging denoise, cell segmentation and extraction of single-cell features were performed using a combination of Python and R packages, including ImcSegmentationPipeline, IMC-Denoise[51] and DeepCell[13,34,52]. For the single-cell analysis, the annotated data object was generated, and protein expression raw measurements were normalized at the 99th percentile to remove outliers. In Scanpy (version 1.9.1), principal component analysis (PCA), batch correction and Harmony data integration were performed to compute and plot the uniform manifold approximation and projection (UMAP) embeddings (umap-learn Python package, version 0.5.3). Next, automated cell type assignment using the Python package Astir (version 0.1.4) was applied to identify the major cell types expected to be found in the lung tissue according to the antibody panel used. For cell assignment with Astir, the following information to label cells based on a broad ontogeny (metaclusters and major cell types) and the proteins (lineage markers) to be most expressed in each expected cell type were used: (1) macrophage: CD163, CD206, CD14, CD16, CD68, CD11c, Iba1; (2) neutrophil: CD66b, Arginase1; (3) CD8 T cells: CD3, CD8; (4) CD4 T cells: CD3, CD4; (5) B cells: CD20; (6) endothelium: CD31; (7) fibroblast: Collagen1; (8) SMC: smooth muscle actin; epithelial: PanCK; RBCs: CD235ab.

After cell assignment, cells labeled as 'other' or 'unknown' were filtered out from downstream analysis, and the annotated data object was subset into the major cell types identified—that is, macrophages, neutrophils, lymphoid, vascular, epithelial and stromal—and Phenograph Louvain clustering (with 200 nearest neighbors) was performed for each cell population separately using a small set of specific lineage marker and functional proteins. The finer cell type annotation was used to evaluate the frequency and absolute counts of cell types across clinical groups, histopathological lesions and HIV status. Differential abundance analysis was also performed using the scanpro and scCODA Python packages[53] and the miloR R package (version 1.4.0)[54]. Spatial statistics analysis based on the coordinates of the cells in the ROIs was performed using the Python package Squidpy (version 1.2.2)[55]. These coordinates were used to plot spatial graphs and to calculate and plot neighborhood enrichment scores[13].

Integration of Malawian IMC data with other available IMC COVID-19 lung data

IMC COVID-19 data from postmortem lung samples from published Brazilian[13] and US[34] fatal cohorts were integrated with the Malawian IMC dataset. First, datasets were concatenated in Scanpy taking the 'inner' (intersection) of all common protein markers in the panels across the three IMC datasets. Then, with scvi-tools[56], we applied different integration methods, such as Harmony and variational autoencoder (VAE)-based methods, such as scVI and scANVI. Analysis of the UMAP embedding of the integrated versus non-integrated data showed that Harmony and scANVI performed better, and, in downstream analysis, we used Harmony-integrated output. Next, cell identities were standardized (label harmonization), which refers to a process of checking that labels are consistent across the datasets that are being integrated. Finally, cell frequencies in the postmortem lung across all three cohorts were plotted, and differential abundance analysis was performed using scanpro (https://github.com/loosolab/scanpro) and scCODA Python packages[57] and the miloR R package (version 1.4.0)[58].

Dissociation of lung cells from frozen samples and single-nuclei preparation

Lung samples were dissociated both from fresh samples and from slow-frozen samples that had been stored in liquid nitrogen. Slow-frozen cells were defrosted in a water bath at 37 °C, and then pieces of tissue were transferred to RPMI 1640 medium with 25 mM HEPES and L-glutamine (Thermo Fisher Scientific) and 40% heat-inactivated FBS (Thermo Fisher Scientfic). Fresh or defrosted frozen cells were then dissociated, adapting a previously published protocol for lung dissociation[57]. Samples were dissociated in a buffer containing 400 mg ml^{-1} Liberase DL (Sigma-Aldrich), 32 U ml^{-1} DNAse I (Roche) and 1.5% BSA in PBS (without calcium and magnesium). The tissue was put in buffer (four times weight:volume) in a GentleMACS C-tube (Miltenyi Biotec, 130-096-334), minced using scissors and then run on a GentleMACS dissociator (Miltenyi Biotec, 130-093-235) on the manufacturer's program 'C-lung 01_02'. Dissociation was achieved by warming the tissue on an orbital shaker in a chamber at 37 °C for 30 min and running 'C-lung 01_02' twice more: once at 15 min and once at 30 min. The enzyme was neutralized by diluting with 10 ml of ice-cold 20% FBS, containing 32 U ml^{-1} DNase. The sample was then filtered through a 100-μm strainer (Corning, 352360), and samples were subsequently kept on ice with all centrifuge and antibody incubation steps at 4 °C. Cells were pelleted by spinning at 300g for 5 min at 4 °C. RBCs were removed by incubating with ACK lysing buffer (Thermo Fisher Scientific, A1049201) for 5 min at room temperature. For frozen cells, debris and dead cells were removed using a debris removal solution (Miltenyi Biotec, 130-109-398) and a dead cell removal kit (Miltenyi Biotec, 130-090-101), respectively, according to the manufacturer's protocol.

Single nuclei were isolated from snap-frozen lung tissue samples using a previously published method[7]. Tissue was kept on dry ice/liquid nitrogen until processing was started. Tissue was placed into a GentleMACS C-tube containing 2 ml of freshly prepared nuclei extraction buffer that contained RNAse inhibitors: 0.2 U μl^{-1} RNaseIN Plus RNAse inhibitor (Promega) and 0.1 U μl^{-1} SUPERasin RNAse inhibitor (Thermo Fisher Scientific). Dissociation was achieved by running the C-tube on the GentleMACS dissociator on program 'm_spleen_01' for 1 min. The sample was filtered using a 40-μm strainer and spun at 500g for 10 min at 4 °C. Pellet was then resuspended in 500 μl of 1× ST without RNAse

inhibitor and filtered again using a 35-µm strainer. A 10-µl volume was loaded on a hemocytometer for counting.

Single-cell and single-nuclei partitioning and library preparation

10x 3′ v3 chemistry was used for all samples. For fresh lung samples, we loaded 10,000 cells into one channel of a 10x chip (1000120). For fresh nasal and blood samples, we labeled the nasal and blood samples with different hashtags and pooled them at a 1:1 ratio and loaded 10,000–20,000 cells. For frozen nuclei and single-cell samples, we pooled samples from 3–6 different patients aiming for equal ratios and loaded 20,000–40,000 cells per nuclei. Libraries were prepared according to the manufacturer's protocol and sequenced with an Illumina NextSeq 2000. To make these data available for analysis by others, reads were submitted to ArrayExpress (E-MTAB-13544).

Single-cell data processing

For all tissue compartments, the data were analyzed through the following steps. (1) Processing of the raw reads. 5′ scRNA-seq data along with the 3′ snRNA-seq runs were demultiplexed using Cell Ranger 'mkfastq'. Reads were mapped to a concatenated human GRCh38, SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, GenBank MN908947.3) and HIV (human immunodeficiency virus 1, GenBank AF033819.3) reference genome to generate count matrices using Cell Ranger 'cellranger count' (version 7.0). (2) Ambient RNA removal. To reduce potential noise driven from empty droplets or ambient RNA captured in our samples, we used the tool SoupX (version 1.6.2)[58] and used corrected expression matrices in subsequent analyses. (3) Quality control and filtering. Data were analyzed using the Seurat package (version 4.3)[59] in R (version 4.2) with mitochondrial gene expression thresholding applied on individual samples. In addition, cells that were expressing more than 150 genes were retained to maximize discovery of cell types. (4) Normalization and variance stabilization. Samples were merged and normalized using the SCTransform() function, selecting the top 3,000 variable genes to drive the downstream clustering. Additionally, effects of mitochondrial gene expression, ribosomal gene expression and cell cycle were regressed out. (5) Integration. PCA was run on all merged data objects. The embeddings were then fed into the standard Harmony (version 0.1.1)[56] integration pipeline. (6) Clustering and dimensionality reduction. An appropriate number of principal components (PCs) were selected to generate the UMAP. PCs were used to determine the k-nearest neighbors for each cell for the shared nearest neighbor (SNN) graph construction, followed by clustering at resolution 0.3. (7) Cell type annotation. Identification of cluster markers for the lung and nasal datasets were calculated by running FindAllMarkers() using MAST, followed by Bonferroni multiple test correction. We specified that genes must be expressed in at least 25% of cells (min.pct = 0.25) with a log fold change of 0.25. Cell types were manually annotated, leveraging canonical cell type markers reported from existing literature and curated datasets. Peripheral blood clusters were annotated using the consensus label transfer algorithm SingleR (version 2.0.0)[60] using the Azimuth Reference PBMC atlas (https://zenodo.org/records/4546839). Cells with low mapping scores were reanalyzed and manually annotated as above. (8) Gene Ontology (GO) and pathway analysis. DE genes across conditions were calculated using the FindMarkers() function using MAST. Genes were defined as DE with a significance threshold of less than 0.05 and a log fold change threshold of 0.25, followed by Bonferroni correction. Gene set enrichment analysis (GSEA) was done using the fgsea package (1.3.0)[61] using 50 canonical hallmark gene sets as described in the Molecular Signatures Database (MSigDB) (version 7.5.1)[62]. (9) Module scoring. Gene module scoring was calculated using the AddModuleScore() function of gene sets taken from MSigDB and AmiGO 2 (ref. 63) that related to IFN responses (lambda (GO:0034342), alpha (GO:0035455), IFN-β (GO:0035456), IFN-γ (GO:0034341), IL6/JAK/STAT (*HALLMARK_IL6_JAK_STAT3_SIGNALING*) and TNF (*HALLMARK_TNFA_SIGNALING_VIA_NFKB*)).

log fold changes in module scores were calculated using the log2 + 1 of the differential means across a cell type. (10) Cell–cell communication analysis. Inference of cellular communications was computed using the multinichenetR (version 1.0.3) package[64] with a log fold change cutoff of 0.5 being expressed in at least 10% of cells across conditions.

Hashtag demultiplexing

Hashtag reads were quantified using CITE-seq-Count (version 1.4.4)[65] and demultiplexed using cellHashR (version 1.0.1)[66]. The following methods were tested: BFF$_{cluster}$, BFF$_{raw}$ (10), GMM-Demux[67], Seurat HTODemux[59] and DropletUtils hashedDrops[68], with HTODemux resulting in the highest number of singlets that were used for analysis.

Single-nucleotide polymorphism splitting of multiplexed runs

Demultiplexing of runs was carried out using the single-nucleotide polymorphism (SNP) clustering algorithm Souporcell[69] to identify distinct genotypes and assign cells to different individuals. For each run, we set the number of clusters (k) to the expected number of genotypes in the run ($k = 2–6$), and cell barcodes were assigned to each cluster. Cluster barcodes were then used to subset the input BAM file across human leukocyte antigen (HLA) loci of the multiplexed runs, under the assumption that these would be distinct regions of the genome for each individual. Using Integrative Genomics Viewer (IGV), we visualized SNP distributions at a set allele frequency of 0.2 and compared the subset BAM files to BAM files from individual runs. Iteratively, Souporcell clusters were assigned to samples through the following rationale: (1) matching SNP distributions to independent sequencing runs, (2) through mapping to sex chromosomes or (3) through the process of elimination where an independent sequencing run genotype was not available. In scenarios where Souporcell failed to identify the expected number of genomes, we assigned cluster barcodes to matching genotypes from independent sample runs regardless of expected k. After successful demultiplexing, we identified which cells derived from which patient and were able to proceed with downstream single-cell analyses as outlined above (see 'Single-cell data processing' subsection).

HLCA integration

The HLCA[10] was filtered down, retaining cells that were taken from the lung and lung parenchyma. These included studies originating from the Northern Hemisphere, with lung cell data in COVID-19, pneumonia and healthy controls. Cell type annotations harmonized with our analyses (*AT1, AT2, EC arterial, EC capillary, EC venous, Fibroblasts, Innate lymphoid cell, NK, Macrophages, Monocytes, T cell lineage*) were selected. To have sufficient power for downstream analyses with our cohort, we randomly subsampled each cell type within each disease condition to create a normalized atlas of 100,000 cells to integrate with our lung atlas. Processing and integration steps were followed as described previously for the Malawian cohort using 38 PCs and a clustering resolution of 0.2. Manual cluster annotation was performed by running FindAllMarkers(), leveraging canonical cell type markers.

Pseudobulking single-cell nasal and blood

To make our nasal and blood scRNA-seq comparable with Luminex cytokine data, we assigned all cells to a unified identifier ('pseudo_cluster') to pool cells belonging from different cell type clusters together. Then, the average expression of the different cytokines on the Luminex panel were visualized using ComplexHeatmap[70] and a z-score of the counts (Supplementary Fig. 5). For the statistical tests of genes associated with the IFN-γ pathway, we used a Welch two-sample t-test.

Exploring viral reads in samples

To identify SARS-CoV-2-infected cells in our lung dataset, we quantified the number of unique molecular identifiers (UMIs) that were detected after mapping with Cell Ranger across our single-cell datasets. A given

cell was deemed to be infected if it expressed at least two UMIs of genes mapping to the SARS-CoV-2 genome.

Integration of Malawian COVID-19 lung IMC data with Malawian COVID-19 lung snRNA-seq data

Lung IMC and snRNA-seq data, exclusively from Malawian patients with COVID-19, were integrated with the recently developed integration tool MaxFuse, which integrates data across weakly linked modalities, such as protein and RNA expression, through cross-modality matching and iterative smoothed embedding[43]. Highly variable features (s.d. > 0.3 for the RNA expression and s.d. > 0.1 for the protein expression) shared between both datasets were retrieved based on a protein-to-gene correspondence list, produced by the MaxFuse authors and edited to include specific protein markers in our IMC panel (Supplementary Information). Cell counts used for each modality included IMC (53,762 cells) and snRNA-seq (36,616 cells). Previously normalized and batch-corrected IMC protein expression and snRNA-seq RNA expression were used as MaxFuse input. All values were capped between 5% and 95% quantiles for visualization purposes. With the resulting integration, expression levels of IFN-γ response-related genes (*IFNGR1*, *IFNGR2*, *HLA-DRA*, *HLA-DRB1*, *C1QA*, *APOE*, *IFI30* and *CD74*) and IFN-γ signature score were determined and plotted in the lung cells derived from the IMC data.

In situ hybridization co-staining for *CD3* and *IFNG* and *CD206* (*MRC1*) and *IFNGR2*

In situ staining was performed on TMAs with 138 ROIs using the same TMAs and patients used for IMC, covering multiple lung regions from left and right lungs in nine patients with COVID-19, three patients with LRTD and two non-LRTD patients. Consecutive slides were used for two dual staining panels: one for *IFNG* and *CD3E* and the other for *IFNGR2* and *MRC1* (CD206). Slides were stained according to the manufacturer's instructions (product codes: 322452 and 322500, ACD, Bio Techne) using the probes Hs IFNG-C1, Hs IFNGR2-C1, Hs-MRC1-C2 and Hs CD3E-C2 (product codes: 310501, 553971-C2, 1269501-C1 and 583921-C2, ACD, Bio Techne) and positive and negative control probes PPIB/POLR2A and DapB (product codes: 321641 and 320751, ACD Bio Techne). Slides were digitized and scanned with standard settings at ×80 magnification using the Motic EasyScan Infinity 60 digital slide scanner (I. Miller Microscopes). For quantification of positive cells, we used HALO software (version 3.6.4134.362) with the AI module (3.6.4134) and the FISH module (version 3.2.3) for cell detection after deconvolution.

Immunohistochemistry

Immunohistochemistry was performed in an autostainer using the Envision kit and DAB chromogen (product codes: K4003 and K4001, Agilent Technologies) with anti-CD206/MRC1 (E2L9N) or anti-CD3 antibodies (product codes: 91992, Cell Signaling Technologies, and A0452, Agilent Technologies). Slides were digitized and scanned at ×20 magnification using an Aperio VERSA 8 slide scanner (Leica Biosystems) and Aperio VERSA 1.0.4.125 software (Leica Biosystems).

Statistics and reproducibility

No statistical method was used to predetermine sample size. We excluded nine single-cell sequencing runs that had few to no cells and that did not pass standard quality control metrics. Within our lung atlas, a population of cells (*n* = 1,348) was excluded that we deemed to be low-quality cells that almost exclusively derived from one multiplexed single-nuclei sequencing run that exhibited extremely low UMI counts. Two non-COVID-19 patients with LRTD were excluded from IMC runs as they had evidence of active TB lung disease because of theoretical safety concerns, as IMC can generate aerosol. Pathologists were blinded to patient groups for systematic scoring of the lung, and investigators conducting the in situ validation experiments undertook staining and automated scoring on the TMAs blinded to which samples were from which case or group. For other experiments and analyses, investigators were not blinded to case groups. Samples were sequenced as multiplex, including patients from different groups, and IMC was run on TMAs as a single run, in both instances to reduce batch effect.

Ethics and inclusion statement

Malawian researchers with clinical, laboratory, analysis and medical ethics expertise were involved throughout the research process from conception to manuscript preparation. The main research questions were determined by Malawian clinical and laboratory researchers alongside international researchers who were living and working in Malawi. Before conducting the study, we undertook a full sensitization process for the study with all staff on the recruiting wards in our hospital to discuss the study and consider the best way of sensitively conducting recruitment and informed consent. This work was led by two social scientists (L.S. and D.N.), one specialized in bioethics (D.N.). Details of our approach and considerations for recruitment are published as a chapter in a casebook separately[71]. Extensive research and laboratory infrastructure already exists in Malawi through a medical university (Kamuzu University of Health Sciences) and several internationally funded research programs. Building on this, as part of this project, local research capacity was enhanced by establishing a single-cell platform in Malawi and training local scientists and by additional training of local scientists in tissue processing. As a result, all tissue processing and cell partitioning and library preparation for single-cell and single-nuclei sequencing were done in Malawi. The research protocol was approved by a Malawian research ethics review committee (National Health Service Research Ethics Committee) and, in the United Kingdom, by the University of Glasgow Medicine Veterinary and Life Sciences Research Ethics Committee. Safety of staff was ensured by conducting renovations to create a dedicated autopsy room for COVID-19 autopsies and by providing PPE and cleaning solutions and training all staff that were patient facing or involved in sample collection and handling in their appropriate use. Laboratory work was conducted in a laminar flow hood using PPE. Local and regional research, including autopsy studies and investigative work, were considered throughout the study and are appropriately cited.

Reporting summary

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

Data availability

scRNA-seq: Raw data and processed count matrices are deposited at the EBI ArrayExpress (accession number E-MTAB-13544). Fully processed RDS objects of the scRNA-seq analysis and IMC can be found through the GitHub repository (https://github.com/olympiahardy/COSMIC_Malawi_Covid_Atlas) and through the following Zenodo records: https://doi.org/10.5281/zenodo.13898422 (ref. 72) and https://doi.org/10.5281/zenodo.13899297 (ref. 73).

The atlases are browsable using the Cellxgene VIP platform hosted by the University of Glasgow at the following URLs:

Lung Atlas: https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Lung_Atlas.h5ad/

Lung Immune Atlas: https://cellatlas-cxg.mvls.gla.ac.uk/view/COSMIC_Lung_Immune_Atlas.h5ad/

Lung Stromal Atlas: https://cellatlas-cxg.mvls.gla.ac.uk/view/COSMIC_Lung_Stromal_Atlas.h5ad/

Nasal Atlas: https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Nasal_Atlas.h5ad/

Blood Atlas: https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Blood_Atlas.h5ad/

Histopathology slides on virtual microscope: https://covid-atlas.cvr.gla.ac.uk

Metadata for the patients (without identifying information) are provided in Extended Data Table 1 and Supplementary Tables 1 and 2.

IMC: https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COS-MIC_IMC_Lung.h5ad/

Code availability

All R scripts for the scRNA-seq analysis and figure generation can found at https://github.com/olympiahardy/COSMIC_Malawi_Covid_Atlas. Python scripts to process the imaging mass cytometry and figure generation can be found at https://github.com/joaolsf/Spatial_Single_Cell_Lung_Atlas_Malawi_COVID.

References

50. Santana, M. F. et al. Hemorrhagic and thrombotic manifestations in the central nervous system in COVID-19: a large observational study in the Brazilian Amazon with a complete autopsy series. *PLoS ONE* **16**, e0255950 (2021).
51. Lu, P. et al. IMC-Denoise: a content aware denoising pipeline to enhance imaging mass cytometry. *Nat. Commun.* **14**, 1601 (2023).
52. Geuenich, M. J. et al. Automated assignment of cell identity from single-cell multiplexed imaging and proteomic data. *Cell Syst.* **12**, 1173–1186 (2021).
53. Buttner, M., Ostner, J., Muller, C. L., Theis, F. J. & Schubert, B. scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12**, 6876 (2021).
54. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using *k*-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
55. Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
56. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
57. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
58. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
59. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
60. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
61. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* https://doi.org/10.1101/060012 (2021).
62. Gene Ontology Consortium et al. The Gene Ontology knowledgebase in 2023. *Genetics* **224**, iyad031 (2023).
63. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
64. Browaeys, R. et al. MultiNicheNet: a flexible framework for differential cell–cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. Preprint at *bioRxiv* https://doi.org/10.1101/2023.06.13.544751 (2023).
65. Roelli, P. B., Flynn, B. S. & Gui, G. Hoohm/CITE-seq-Count: 1.4.2 (1.4.2). *Zenodo* https://doi.org/10.5281/zenodo.2590196 (2019).
66. Boggy, G. J. et al. BFF and cellhashR: analysis tools for accurate demultiplexing of cell hashing data. *Bioinformatics* **38**, 2791–2801 (2022).
67. Xin, H. et al. GMM-Demux: sample demultiplexing, multiplet detection, experiment planning, and novel cell-type verification in single cell sequencing. *Genome Biol.* **21**, 188 (2020).
68. Lun, A. T. L. et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
69. Heaton, H. et al. Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020).
70. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
71. Mormina, M., Suwalowska, H., Schneiders, M. L. in *Research Ethics in Epidemics and Pandemics: A Casebook* (eds Bull, S. et al.) 173–192 (Springer Nature, 2024).
72. Hardy, O. M. Spatially resolved single-cell atlas unveils a distinct cellular signature of fatal lung COVID-19 in a Malawian population. https://doi.org/10.5281/zenodo.13898422 (2024).
73. Joao Luiz, D. S. F. Spatially resolved single-cell atlas unveils a distinct cellular signature of fatal lung COVID-19 in a Malawian population. https://doi.org/10.5281/zenodo.13899297 (2024).

Acknowledgements

Author contributions

Research protocol, Malawi (C.A.M., W.V., S.L. and D.C.). Sensitization, Malawi (L.S., T.R.N., M.S. and D.N.). Recruitment of patients, Malawi (M.S., T.R.N., L.S., D.C., C.E., A.T. and F.Z.). Data entry, Malawi (C.M.). Autopsies, Malawi (S.B.K. and C.A.). Pathology reports, Malawi (S.K.). Histopathology scoring and identification of regions of interest (C.A. and V.H.). Clinical characterization and cause of death attribution, Malawi (T.R.N., M.S., S.K., P.B., B.M., S.L., W.V. and D.M.D.). Laboratory protocols and processing of tissues, Malawi (J.N., C.N., K.M., W.N., T.N., S.K., C.P. and C.A.M.). Setup of single-cell protocols and optimization of single-cell methods, Malawi (J.N. and C.A.M.). Single-cell experiments (J.N., C.A.M., C.N., L.M., K.J. and J.P.). Design and conduct

Competing interests

Additional information

Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | The histopathology of fatal Covid19 versus fatal non-Covid19 LRTD and non-LRTD in Malawian cases. Histopathology in the left and right lungs of the 16 cases was scored systematically using pre-defined criteria by two pathologists who were blinded to clinical information. We used identical scoring to a Brazil cohort that we have published on separately. A – C are violin plots of the distribution of scores to highlight comparisons between different group, central thick bars highlight the median and outer bars the interquartile range. In all three graphs a two-sided unpaired t-test was used to compare lesion frequencies with no correction for multiple comparisons * denotes $p < 0.05$ with specific p values given below. (**a**) Comparison of histological features between COVID-19 (n = 9) and non-COVID-19 fatal lower respiratory tract disease (LRTD) cases (n = 5). p values for significant individual comparisons: syncytia p = 0.008; type II hyperplasia p = 0.016; vascular congestion p = 0.031; lymphocytes p = 0.0032; granulation p = 0.034. (**b**) Comparison of histological features between HIV + COVID-19 cases (n = 5) and HIV- COVID-19 cases (n = 4). No comparisons had a p < 0.05 (**c**) Comparison of COVID-19 cases from Malawi cohort (n = 9) with cases from Brazil cohort (n = 20). p values for significant individual comparisons; vascular congestion, alveolar oedema DAD p = <0.00001; Alveolar thickening p = 0.0004; haemorrhage p = 0.0008; type II hyperplasia p = 0.0013; alveolar emphysema p = 0.0015; syncytia p = 0.0039. (**d**) PCA of cases split by groups. (**e**) UMAP of same data, including HIV status.
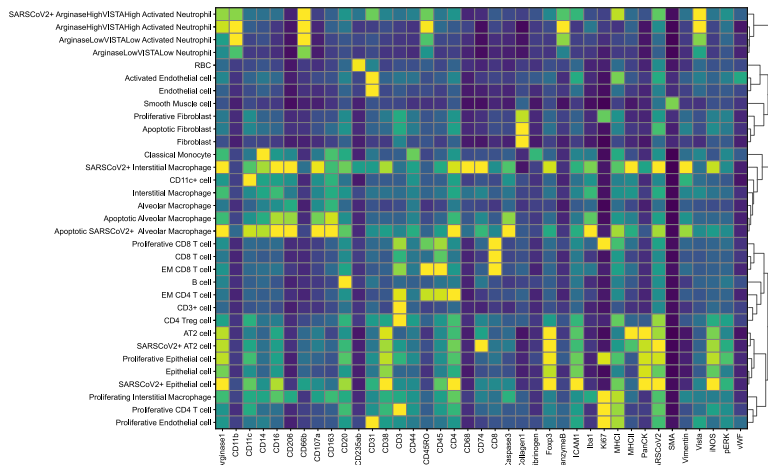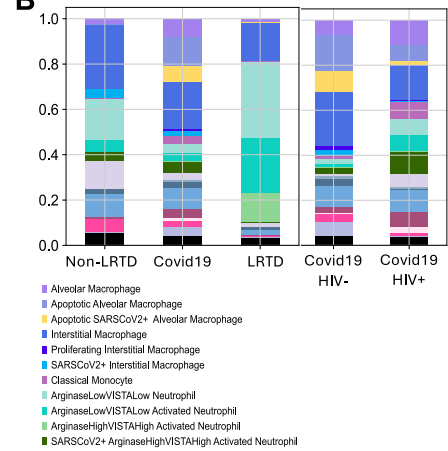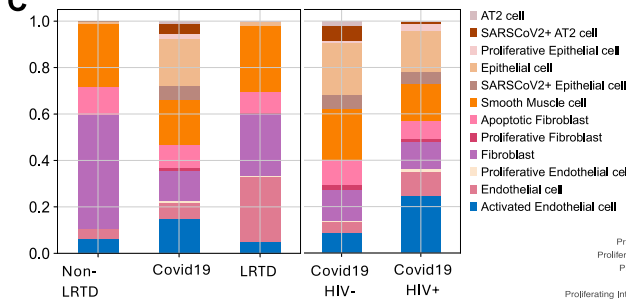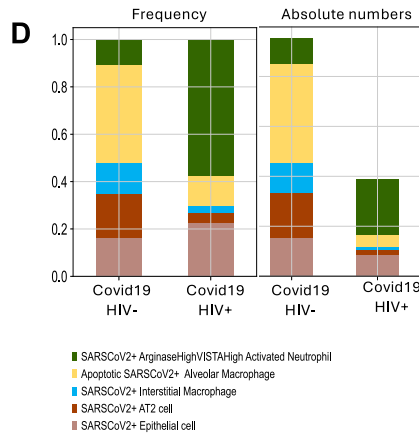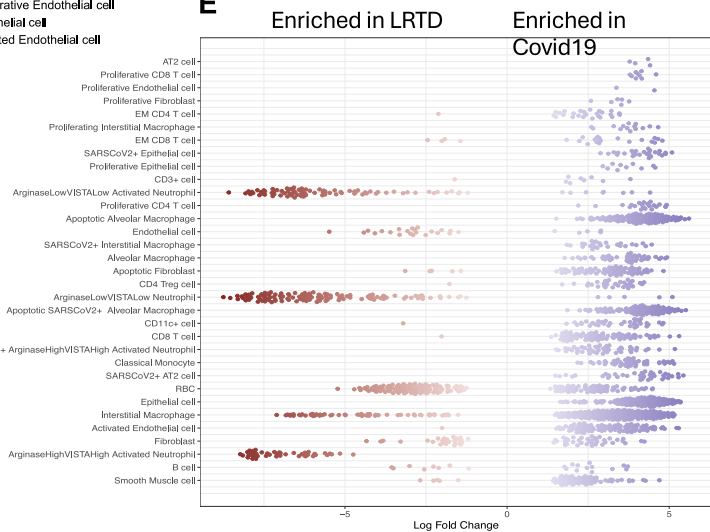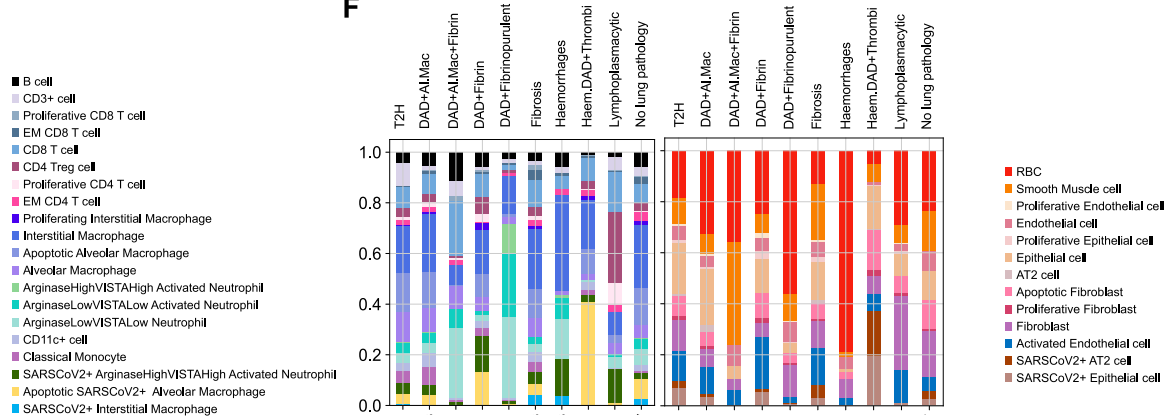
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Cell atlas and phenotype of cell types identified in the post-mortem lung tissue determined by Imaging Mass Cytometry (IMC).
a) Phenotype representation of each cell type identified in the lung samples. The heatmap shows the mean expression of each protein marker in the IMC panel in each cell type identified in the post-mortem lung tissue. (**b**) Frequency of the immune cell types identified in the post-mortem lung samples by IMC according to clinical groups and according to HIV status within the COVID-19 group. (**c**) Frequency of the stromal cell types identified in the post-mortem lung samples by IMC according to clinical groups and according to HIV status within the COVID-19 group. (**d**) Frequency and absolute numbers of SARS-CoV-2 Ag+ cells in the myeloid and epithelial compartments, determined by IMC, in the post-mortem lung samples according to HIV status within the COVID-19 group. (**e**) Cell type enrichment analysis of the cell populations identified in Malawi lung IMC data. The comparison shown is between COVID-19 versus LRTD cases. To correct for multiple testing, the spatial false discovery rate (FDR) was calculated and only dots with spatial FDR < 0.05 are shown. (**f**) Cellular landscape of histopathological lesions based on matched H&E and IMC analysis of post-mortem lung samples from the different clinical groups. The lesions were pooled, and the graph shows the average proportion of each cell type in each lesion type.
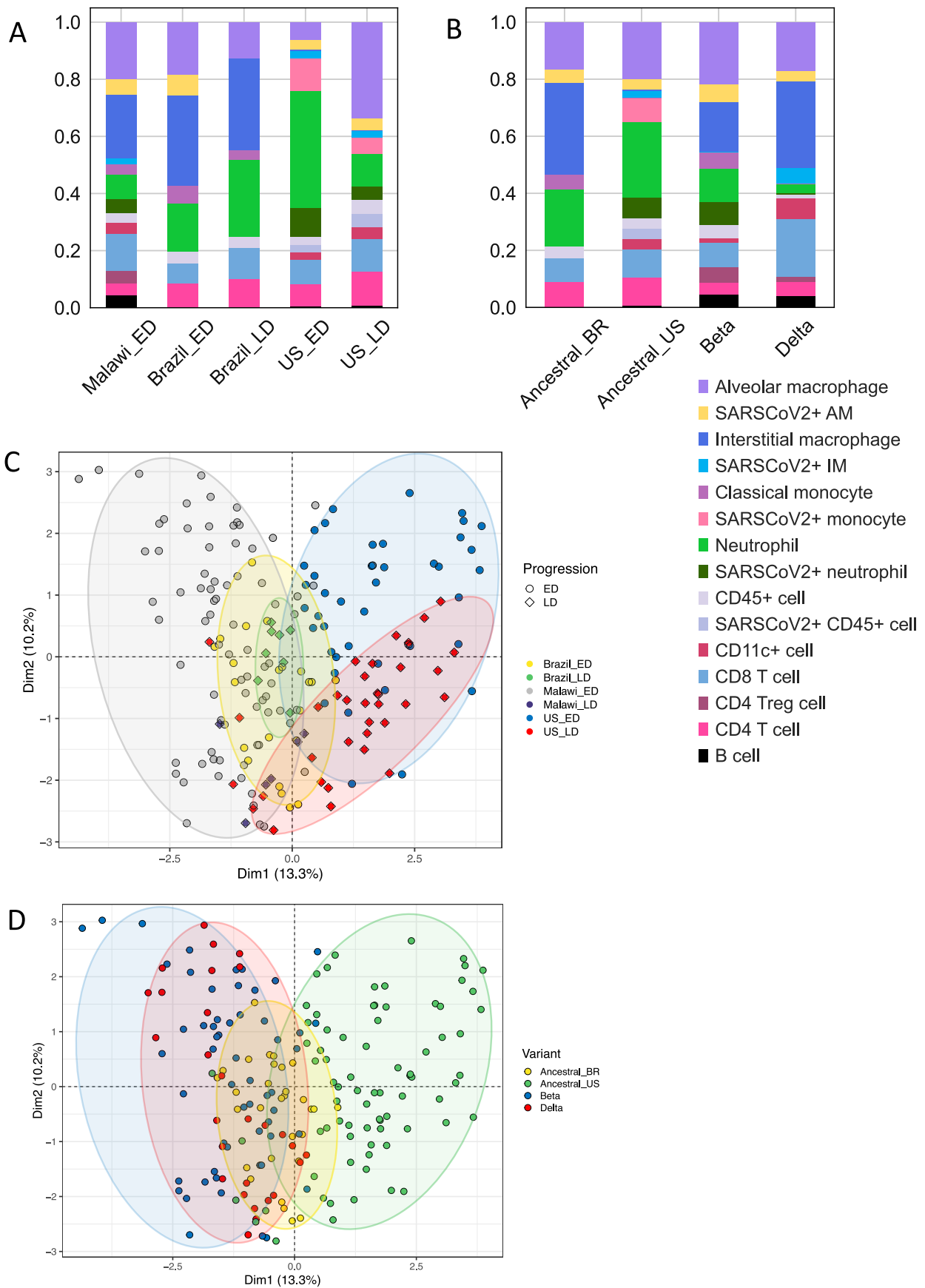
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Effect of disease stage and viral variant on lung immune composition. (a,b) Immune cell proportions from lung imaging mass cytometry data in: **(a)** the Malawi cohort versus Brazil and USA early death (ED; that is cases who died within two weeks of illness onset) and late death (LD; that is cases who died after 2 weeks of illness). When compared with Malawi cases, USA early death (ED) have an even higher proportion of neutrophils and a lower proportion of macrophages than late death (LD). **(b)** Lung proportions in Brazil and USA cohorts who had the ancestral variant compared with Malawi cases with Beta and Malawi cases with Delta variant. The low proportion of neutrophils and high proportion of CD206[high] macrophages in the Malawi cases are present regardless of variant. **(c, d)** Principal component analysis of of IMC lung immune cell composition data where cases are coloured by geographical location of the cohort (Malawi, Brazil, USA) and either: **(c)** disease stage (early versus late death) or **(d)** viral variant. Each dot is a separate lung sample from a different lung region from tissue microarrays (Malawi and USA) or lung blocks (Brazil). Coloured oval areas indicate where the majority of samples from each group cluster drawn by visual estimation using the same colours as those for the dots as indicated in the legend. For both graphs samples cluster principally by geographical location of the cohort and not by disease stage or viral variant.

A

Lung Immune Compartment

- Alveolar macrophages
- Interstitial macrophages
- Monocyte-derived macrophages
- Mature neutrophils
- Immature neutrophils
- CD8+ T cell
- NK
- gdT cell
- CD4+ T cells
- Th1
- Treg
- Plasma cells
- B cells
- Mast cells
- Ribosomal high cells
- Cycling cells
- Erythrocytes

Lung Stromal Compartment

- AT1
- AT2
- Venous endothelium
- Lymphatic endothelium
- Adventitial fibroblasts
- Alveolar fibroblasts
- Myofibroblasts
- Lipofibroblasts
- Ciliated cells
- Basal cells
- Mesothelium
- Smooth muscle cells
- Secretory cells
- Soup
- Ribosomal high cells

B

C — IFNG

D

E

Signature.type:
- Alpha Response
- Beta response
- Gamma Response
- Lambda response
- TNF response

Group:
- COVID-19
- LRTD

Cluster:
- Alveolar macrophages
- Interstitial macrophages
- Monocyte-derived macrophages
- CD16+ Neutrophils
- CD16- Neutrophils
- CD8+ T cells
- Naive CD4+ T cells
- NK cells
- gdT cells
- Th1
- T reg
- Plasma cells
- B cells
- Mast cells
- AT1
- AT2

Gene Module Score

Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Lung cell proportions and gene module scores.
(**a-b**) Cell type proportion bar plots of lung cell types in (**a**) Immune cells and
(**b**) Stromal cells corresponding with Fig. 3b and c, grouped by disease group.
(**c**) Violin plot to show a comparison between COVID-19 and LRTD cases of
expression of IFNG (the IFN-γ gene) in a pseudo bulk analysis of lung scRNA-seq
including all T-cells. (**d**) Plot shows expression levels of different IFNG module
genes in lung alveolar macrophages between COVID-19 and LRTD cases. Line is at
1:1 ratio, hence dots to the left of the line indicate genes with higher expression in
COVID-19 cases and to the right of the line indicates genes with higher expression
in LRTD. Dots for genes for the IFN-γ receptors IFNGR1 and IFNGR2 and for IFNG
(the IFN-γ gene) are indicated. (**e**) Heatmap showing the mean gene module score
across cells in gene sets associated with the *alpha, beta, gamma, lambda and
TNF response*. Cell types have been grouped by COVID-19 and LRTD to show the
difference in response and module score values have been scaled between −1 and 1.

A

COVID–19$_{HLCA}$ vs No–LRTD$_{HLCA}$ COVID–19$_{Malawi}$ vs No–LRTD$_{HLCA}$ COVID–19$_{HLCA}$ vs COVID–19$_{Malawi}$

Alveolar Macs   CD4   CD8   Alveolar Macs   CD4   CD8
Mo-derived-Macs   Mo-derived-Macs

Lung

Cluster
- Alveolar macrophages
- Interstitial macrophages
- Monocyte derived marcophages
- Neutrophils
- CD4 T cells
- T reg
- NK cells
- CD8 T cells
- AT1
- AT2
- Endothelium
- Alveolar Fibroblasts
- Adventitial Fibroblasts
- Fibroblasts

B

Alpha   Beta   Gamma   Lambda   IL6/JAK/STAT

Early   Late

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

Extended Data Fig. 5 | Heatmap of interferon response genes in lung.
(a) Heatmaps showing the log fold change of up/down-regulated interferon response genes taken from immunologic gene sets involved in the immune response. Comparisons include the change in interferon response in cells from the HLCA COVID-19 cohort compared to HLCA control cases (left), the Malawi COVID-19 cohort compared to control cases from the HLCA (middle) and interferon responses from our COVID-19 cohort compared to the HLCA

COVID-19 cohort (right). **(b)** Violin plots of IFN and IL6 response modules in lung tissue macrophages in USA cases (from Delorey et al) comparing early death cases that died within 2 weeks of illness onset (early) and late death cases that died after two weeks of illness (late). All p-values were calculated using a two-sample Wilcoxon test with Bonferroni multiple test correction (Alpha $p < 2.2e\text{-}16$ ****, Beta $p < 2.2e\text{-}16$ ****, Gamma $p = 0.7141$ ns, Lambda $p = 0.09$ **, IL6 $p = 0.8939$ ns).
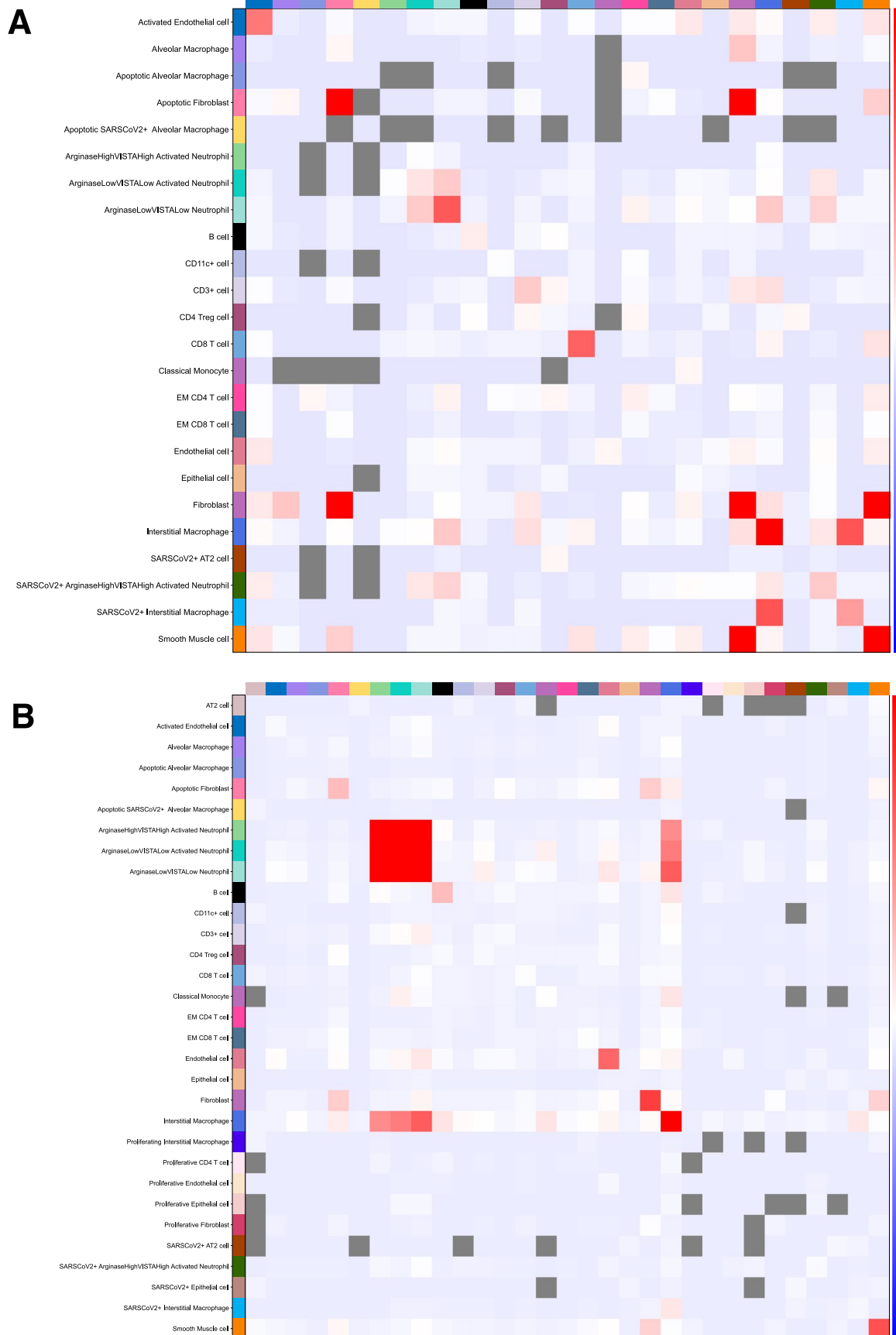
Extended Data Fig. 6 | Predicted cell–cell interactions in the IMC datasets. Heatmaps for the non-LRTD group (**a**) and LRTD (**b**) showing co-localised cell types as shown by the IMC providing insight into potentially interacting cell types in the lung, shown for comparison with the same data from COVID-19 cases Fig. 6b (main figures).
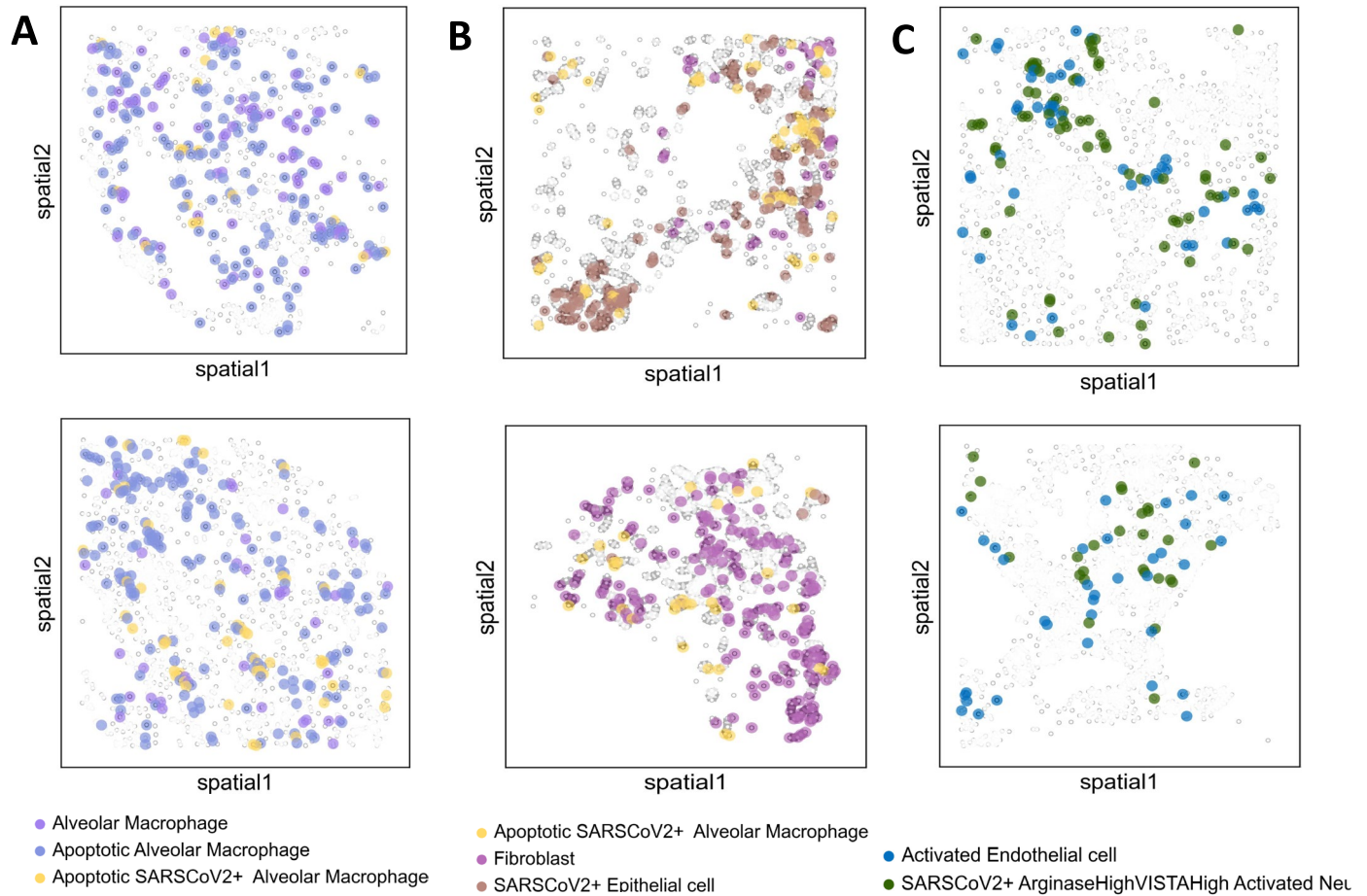
Extended Data Fig. 7 | Visualisation of cell-cell interactions in the IMC datasets. Cellular maps showing the spatial location of specific immune cells – highlighting spatially enriched macrophage and neutrophil interactions identified in Fig. 6b (main figures). (**a**) shows interactions between alveolar macrophages (purple), apoptotic alveolar macrophages (blue) and apoptotic alveolar macrophages (yellow). (**b**) shows interactions between apoptotic alveolar macrophages (yellow) fibroblast (lilach) and SARSCoV2+ Epithelial cells (purple) (**c**) shows interactions between activated endothelial cells (blude) and SARSCoV2+ Neutrophils (green).
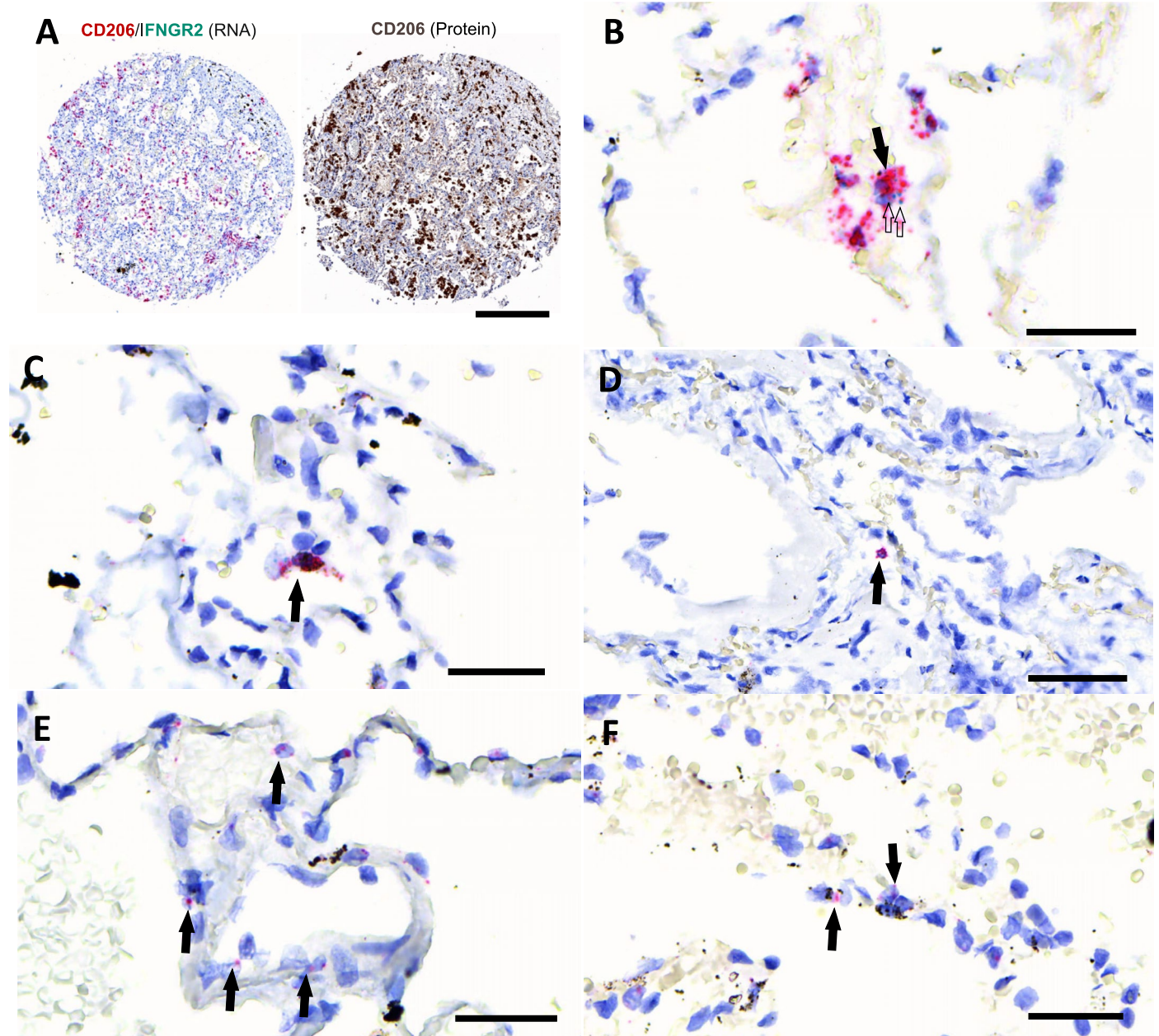
Extended Data Fig. 8 | Dual in situ staining for CD206 and IFNGR2 (inducible IFN-γ receptor) and CD3 and IFNG (IFN-γ gene) for validation of IFN-γ response. 138 regions of interest were taken based on multiple sampled areas from the left and right lung in: 9 COVID-19 cases in 3 LRTD cases and 2 non-LRTD cases **a**) Shows adjacent sections of cores from a COVID-19 (Cos009) case to demonstrate concordance of CD206 RNA and protein staining. The left has had in situ staining (MRC1/CD206 in red and IFNGR2 in green) and the right image shows immunohistochemistry using an anti-CD206 antibody (staining in brown), bar 300μm. **b**) In situ staining for CD206 and IFNGR2 in LRTD case (Cos004); few cells are present with co-staining of IFNGR2 (green, arrow) and CD206 (red, empty arrow), bar 30μm. **c**) In situ staining for CD206 and IFNGR2 for non-LRTD case (Cos 016) only single positive cells are detected in general expressing IFNGR2 (green, arrow) as well as CD206, bar 30μm. **d**) In situ staining for IFNG and CD3 in a LRTD case (Cos003), CD3 red signal arrows, no green signal for IFNG has been detected, bar 60μm. **e**) LRTD case (Cos011), CD3, red signal, single positive cells, no co-staining with IFNG, arrows, bar 30μm. **f**) Non-LRTD cases (Cos016), CD3 in red, black pigment is interpreted as anthracosis, bar 30μm.
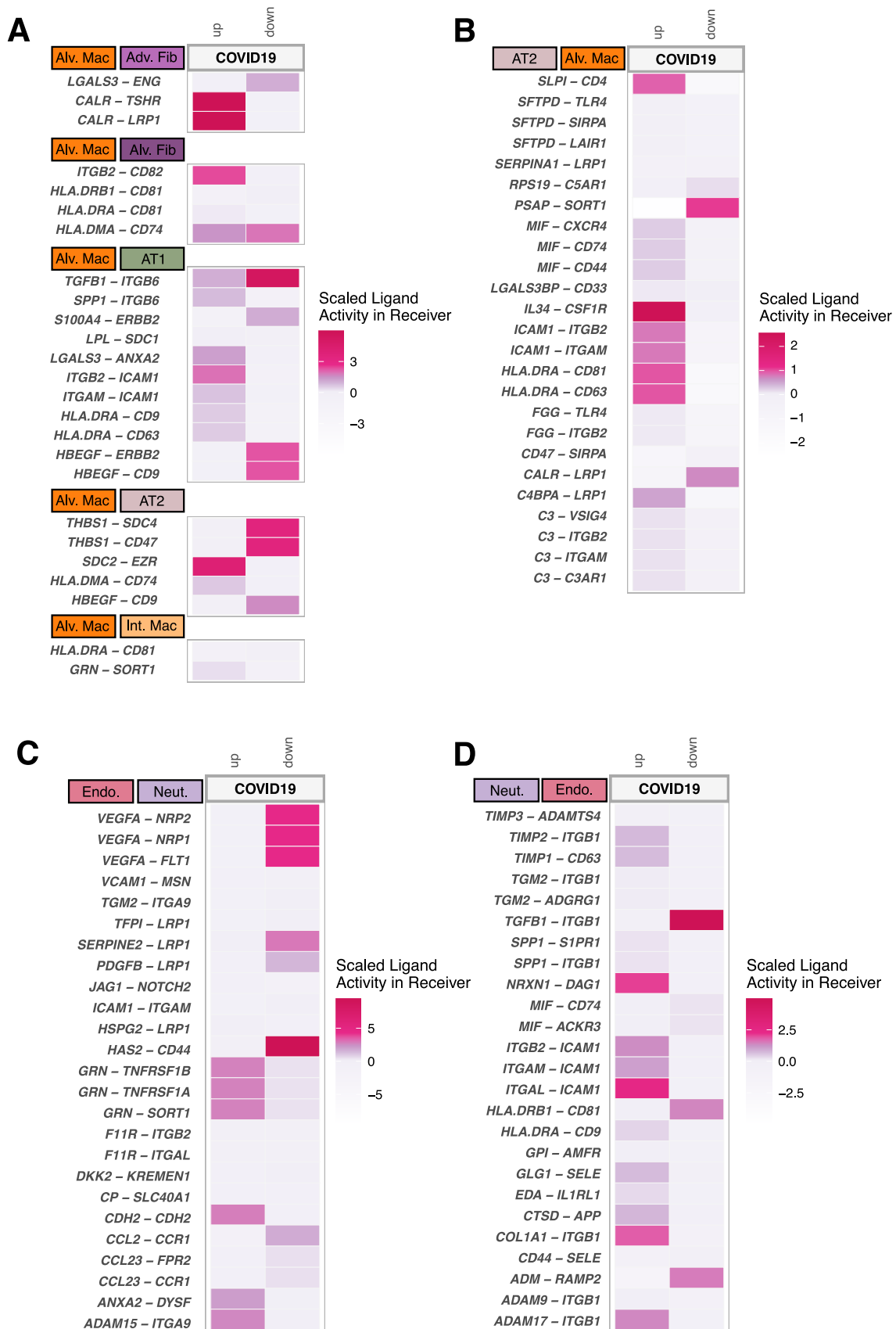
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Predicted receptor ligand interactions in single cell data. a) Heatmap showing up/down-regulated interactions in COVID-19 compared to LRTD driven by AT2 pneumonocytes to alveolar macrophages. Coloured boxes indicate cell type with the ligand-expressing cell type followed by the receptor-expressing cell type. **b**) Heatmap showing up/down-regulated interactions in COVID-19 compared to LRTD driven by lung alveolar macrophages to lung epithelial cells and interstitial macrophages. Coloured boxes indicate cell type with the ligand-expressing cell type followed by the receptor-expressing cell type. **c**) Heatmap showing up/down-regulated interactions in COVID-19 compared to LRTD driven by lung endothelium to neutrophils. Coloured boxes indicate cell type with the ligand-expressing cell type followed by the receptor-expressing cell type. **d**) Heatmap showing up/down-regulated interactions in COVID-19 compared to LRTD driven by neutrophils to lung endothelium. Coloured boxes indicate cell type with the ligand-expressing cell type followed by the receptor-expressing cell type.

Extended Data Table 1 | Characteristics of the patients

Group	Case	Diagnosis	HIV	CD4	ART	Sex	Age (yr)	Ster'ds	Sats	PMI (hr)	Obese/Wasted	Pre-morbidity	S.S to death	Lung sc/sn	Nasal sc	Blood sc	Lung IMC	Nasal Lx	Blood Lx
Covid-19	3	C19	1	126	N	M	55-60	1	87	9	↑	DM2, HT	7		●	●	●	●	●
Covid-19	5	C19	1	142	N	M	55-61	1	82	5	↑↑	DM2, HT	6				●		●
Covid-19	6	C19	1	234	N	M	50-55	1	40	4	↑↑	DM2, HT	7	●		●	●	●	●
Covid-19	7	C19	1	290	Y	F	50-55	1	92*	5	↑↑↑	Cancer	4		●		●	●	●
Covid-19	8	C19	1	301	Y	F	45-50	1	77*	6	↑↑↑	HT, A	8				●	●	●
Covid-19	9	C19	0	n/a	n/a	F	50-55	1	82	5.5	↑↑	None	8				●	●	●
Covid-19	12	Sepsis + C19	0	n/a	n/a	F	60-65	0	98	9	↓	None	29 (2)*	●	●	●	●	●	●
Covid-19	13	C19	0	n/a	n/a	F	70-75	1	76	5.5	↑↑	DM2, HT	20	●	●		●	●	●
Covid-19	15	C19	0	n/a	n/a	M	55-60	1	28	10.5	→	None	5	●	●		●	●	●
No LRTD	2	Sepsis	1	missing	Y	F	45-50	0	88	9	↓↓	None	13	●		●			●
No LRTD	16	Stroke	0	n/a	n/a	F	60-65	0	98	2.5	→	HT	9				●	●	
LRTD	1	TB	1	89	Y	M	50-55	1	79	3	→	None	17					●	●
LRTD	4	Bac. Pneum.	0	n/a	n/a	F	60-65	0	92	3	↓↓↓	None	10	●	●		●	●	●
LRTD	10	TB	1	116	Y	F	60-65	0	96	9.5	↓	HT	5	●	●	●	●	●	●
LRTD	11	Lung Cancer	1	n/a	Y	F	50-55	0	99*	10.5	↓↓	None	4				●	●	
LRTD	14	Bac. Pneum.	0	n/a	n/a	M	50-55	0	99	9	→	None	2	●	●	●	●	●	●

Summary table of cases recruited into our study. PMI, postmortem interval in hours. Obese/Underweight indicates nutritional status, determined by a combination of abdominal circumference measurements and mid-arm circumference measurements and based on reference data for men and women in African populations: ↑, overweight; ↑↑, obese; ↑↑↑, morbidly obese; ↓, underweight; ↓↓, severely underweight. Pre-morbidity: DM2, type 2 diabetes mellitus; HT, hypertension. S.S. to death, symptom start to death, indicating the number of days between the first symptoms consistent with COVID-19 (fever, cough, headache, etc.) and death. Lung IMC, imaging mass cytometry; Lung sc, lung cell single-cell RNA-seq; Nasal sc, nasal cell single-cell RNA-seq; Blood sc, blood cell single-cell RNA-seq; Nasal Lx, nasal Luminex, for multiplexed cytokine array on nasal fluid. A dot for each of these parameters indicates that data are available for that assay for that case.

nature portfolio

Corresponding author(s):	Christopher Moxon, Matthias Marti, Thomas Otto
Last updated by author(s):	Oct 6, 2024

Reporting Summary

Nature Portfolio 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 Portfolio 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	Data were collected using clinical reporting forms and entered into REDCap
Data analysis	umap-learn Python package, v 0.5.3; ASsignmenT of sIngle-cell pRoteomics v 0.1.4; miloR R package (v 1.4.0); Spatial Quantification of Molecular Data in Python v. 1.2.2;Seurat; multinichenetR; MaxFuse v0.0.2 All code and software used in the analysis are detailed in our GitHub repositories: https://github.com/olympiahardy/ COSMIC_Malawi_Covid_Atlas and https://github.com/joaolsf/Spatial_Single_Cell_Lung_Atlas_Malawi_COVID

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

scRNA-Seq: Raw data and processed count matrices are deposited at the EBI ArrayExpress (Accession number E-MTAB-13544 (private until publication). Fully processed objects are deposited on Zenodo 10.5281/zenodo.13898422 and here 10.5281/zenodo.13899297 for the scRNA-seq and IMC respectively.
Lung Atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Lung_Atlas.h5ad/
Lung Immune Atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Lung_Immune_Atlas.h5ad/
Lung Stromal Atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Lung_Stromal_Atlas.h5ad/
Nasal Atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Nasal_Atlas.h5ad/
Blood Atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_Blood_Atlas.h5ad/
Histopathology slides on virtual microscope: https://covid-atlas.cvr.gla.ac.uk
Metadata for the cases (without identifying information) is provided in Extended Table 1.
IMC lung atlas - https://cellatlas-cxg.mvls.gla.ac.uk/COSMIC/view/COSMIC_IMC_Lung.h5ad/

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	A priori we aimed to recruit fatal cases with an even sex ratio. The sex of cases and proportions of cases in each group that were male and female sex are indicated in the summary table and the sex of cases is indicated in metadata linked to single cell and imaging mass cytometry data.
Reporting on race, ethnicity, or other socially relevant groupings	All the cases in our study were of African ethnic background and residents in the Southern region of Malawi
Population characteristics	These are summarized in Extended data table 1 and supplementary table 1
Recruitment	We recruited patients aged 45-75 admitted to Queen Elizabeth Central Hospital, Blantyre between October 2020 and July 2021 during which there were two epidemiological waves driven by different SARS-CoV2 variants: Beta (Dec 2020-Feb 2021) and Delta (May-July 2021). Patients admitted with respiratory signs were routinely tested for SARS-CoV2 at QECH. We recruited cases into three groups based on clinical criteria: 1) a Covid19 group (n=9) with clinical features suggesting acute respiratory distress (ARDS, oxygen requirement and either respiratory signs on clinical examination or chest x-ray changes or both) and who had at least one nasal swab positive for SARS-CoV-2 on admission; 2) A non-Covid19 LRTD (lower respiratory tract disease) group (n=5) who had clinical signs of ARDS but were negative for SARS-CoV-2 on admission and during hospitalisation; 3) a no LRTD, COVID-19 negative group (n=2) who had no oxygen requirement and no clinical signs of LRTD and for whom the admission and any subsequent nasal swabs were negative for SARS-CoV2 on PCR (Fig.1b, Extended Data Table 1). The study only recruited cases who died between 12 midnight and 12 noon to minimize the postmortem interval and to avoid doing any autopsies at night. We also aimed to balance sex and HIV status so had to decline some eligible male and some HIV positive cases. Otherwise our cases broadly reflect the demographics of the severe cases on our ward during this period. We do not think this selection introduced systematic bias, but cannot exclude this possibility in this small cohort.
Ethics oversight	The study protocol was approved by the National Health Scientific Research Committee (NHSRC) in Malawi (Protocol number 07/09/1913) and and by the Medical Veterinary Life Sciences ethics committee in Glasgow (protocol number 200190041)

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

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 method was used to predetermine sample size. 16 cases in total were recruited across the three clinical groups with this number determined by logistical feasibility and by funding levels. These numbers were considered to be sufficient for these planned single cells and high-dimensional analysis based on sample sizes in other published studies using similar methods.

Data exclusions	We excluded 9 single cell sequencing runs that had few to no cells and that did not pass standard QC metrics. Within our lung atlas a population of cells (n=1348) were excluded that we deemed to be low quality cells that almost exclusively derived from one multiplexed single nuclei sequencing run that exhibited extremely low UMI counts. 2 non-Covid19 LRTD cases were excluded from IMC runs as they had evidence of active TB lung disease because of theoretical safety concerns as IMC can generate aerosol.
Replication	Not applicable as data were based on human samples, however key findings were validated using orthogonal methods.
Randomization	The experiments were not randomized.
Blinding	Scoring by pathologists and validation experiments were done with investigators blinded to patient group.

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
☐	☒ Antibodies
☐	☐ Eukaryotic cell lines
☐	☐ Palaeontology and archaeology
☐	☐ Animals and other organisms
☐	☒ Clinical data
☐	☐ Dual use research of concern
☐	☐ Plants

Methods

n/a	Involved in the study
☐	☐ ChIP-seq
☐	☐ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Smooth muscle actin (Bio-Rad, Clone: 1A4, Catalogue#: MCA5781GA, Stock conc: 0.66 mg/ml, Dilution: 50) Metal channel: 89
	Cd68 (Thermo, Clone: KP1, Catalogue#: MA5-13324, Stock conc: 0.48 mg/ml, Dilution: 50) Metal channel: 113
	Cd235ab (BioLegend, Clone: HIR2, Catalogue#: 306615, Stock conc: 0.5 mg/ml, Dilution: 200) Metal channel: 115
	Pan-cytokeratin (Biolegend, Clone: AE-1/AE-3, Catalogue#: 914204, Stock conc: 0.63 mg/ml, Dilution: 50) Metal channel: 139
	Cd38 (Standard Biotools, Clone: EPR4106, Catalogue#: 3141018D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 141
	MHC-1 (Abcam, Clone: EMR8-5, Catalogue#: ab70328, Stock conc: 0.27 mg/ml, Dilution: 50) Metal channel: 142
	Vimentin (Standard Biotools, Clone: RV202, Catalogue#: 3143029D, Stock conc: 0.5 mg/ml, Dilution: 100) Metal channel: 143
	Cd14 (Cell Signalling, Clone: D7A2T, Catalogue#: 56082BF, Stock conc: 0.65 mg/ml, Dilution: 50) Metal channel: 144
	ICAM1 (Abcam, Clone: EP1442Y, Catalogue#: ab271852, Stock conc: 0.47 mg/ml, Dilution: 50) Metal channel: 145
	Cd16 (Abcam, Clone: SP175, Catalogue#: ab243925-100ug, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 146
	iNOS (Abcam, Clone: SP126, Catalogue#: ab239990, Stock conc: 0.4 mg/ml, Dilution: 50) Metal channel: 147
	Cd66b (Novus, Clone: G10F5, Catalogue#: g10f5_nb100-77808, Stock conc: 0.41 mg/ml, Dilution: 50) Metal channel: 148
	Cd11b (Abcam, Clone: EP1345Y, Catalogue#: ab52478, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 149
	Cd44 (BioLegend, Clone: IM7, Catalogue#: 103001, Stock conc: 0.49 mg/ml, Dilution: 100) Metal channel: 150
	Cd107a (Standard Biotools, Clone: H4A3, Catalogue#: 3151021D, Stock conc: 0.5 mg/ml, Dilution: 100) Metal channel: 151
	Cd45 (eBioscience, Clone: CD45-2B11, Catalogue#: 14-9457-82, Stock conc: 0.65 mg/ml, Dilution: 50) Metal channel: 152
	Cd31 (Novus, Clone: JC/70A, Catalogue#: jc-70a_nb600-562, Stock conc: 0.53 mg/ml, Dilution: 50) Metal channel: 153
	Cd11c (Standard Biotools, Clone: Polyclonal, Catalogue#: 3154025D, Stock conc: 0.5 mg/ml, Dilution: 100) Metal channel: 154
	Foxp3 (Abcam, Clone: 236A/E7, Catalogue#: ab20034, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 155
	Cd4 (Standard Biotools, Clone: EPR6855, Catalogue#: 3156033D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 156
	SARS-Cov2 (Novus, Clone: Polyclonal, Catalogue#: NB100-56576, Stock conc: 0.5 mg/ml, Dilution: 25) Metal channel: 158
	Von Willebrand Factor (Dako, Clone: Polyclonal, Catalogue#: A0082, Stock conc: 1.27 mg/ml, Dilution: 100) Metal channel: 159
	Vista (Standard Biotools, Clone: D1L2G, Catalogue#: 3160025D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 160
	Cd20 (Standard Biotools, Clone: H1, Catalogue#: 3161029D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 161
	Cd8 (eBioscience, Clone: CD8/144B, Catalogue#: 14-0085-82, Stock conc: 0.55 mg/ml, Dilution: 50) Metal channel: 162
	iba1 (WAKO, Clone: Polyclonal, Catalogue#: 019-19741, Stock conc: 0.42 mg/ml, Dilution: 100) Metal channel: 163
	Arginase-1 (Standard Biotools, Clone: D4E3M, Catalogue#: 3164027D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 164
	Fibrinogen (Abcam, Clone: EPR18145-84, Catalogue#: ab227063, Stock conc: 2.93 mg/ml, Dilution: 100) Metal channel: 165
	Cd74 (Standard Biotools, Clone: LN2, Catalogue#: 3166025D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 166
	Granzyme B (Standard Biotools, Clone: EPR20129-217, Catalogue#: 3167021D, Stock conc: 0.5 mg/ml, Dilution: 100) Metal channel: 167
	Ki-67 (Abcam, Clone: B56, Catalogue#: ab279657, Stock conc: 0.56 mg/ml, Dilution: 100) Metal channel: 168
	Collagen Type I (Standard Biotools, Clone: Polyclonal, Catalogue#: 3169023D, Stock conc: 0.5 mg/ml, Dilution: 100) Metal channel: 169
	Cd3 (Cell Signalling, Clone: D7A6E, Catalogue#: 85061BF, Stock conc: 0.53 mg/ml, Dilution: 50) Metal channel: 170
	pERK1/2 [T202/Y204] (Standard Biotools, Clone: D13.14.4E, Catalogue#: 3171021D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 171
	Cleaved Caspase 3 (Standard Biotools, Clone: 5A1E, Catalogue#: 3172023A, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 172
	CD45RO (Standard Biotools, Clone: UCHL1, Catalogue#: 3173016D, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 173

MHC2 (Abcam, Clone: TAL1B5, Catalogue#: ab176408, Stock conc: 0.5 mg/ml, Dilution: 50) Metal channel: 174
Cd206 (Cell Signalling, Clone: E2L9N, Catalogue#: 91992, Stock conc: 0.43 mg/ml, Dilution: 50) Metal channel: 175
Cd163 (Bio-Rad, Clone: EDHu-1, Catalogue#: MCA1853, Stock conc: 0.9 mg/ml, Dilution: 50) Metal channel: 196

Validation | Antibodies for immune markers were optimized and validated on formalin-fixed paraffin embedded human lymph nodes. SARS-Cov2 staining was previously validated in COVID lungs from Brazilian cohort, using lungs from pneumonia patients and control lungs as negative controls.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s) | *State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.*

Authentication | *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination | *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*

Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance | *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.*

Specimen deposition | *Indicate where the specimens have been deposited to permit free access by other researchers.*

Dating methods | *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

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

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals | *For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.*

Wild animals | *Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*

Reporting on sex | *Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*

Field-collected samples | *For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*

Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No | Yes
☒ ☐ Public health
☒ ☐ National security
☒ ☐ Crops and/or livestock
☒ ☐ Ecosystems
☒ ☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No | Yes
☒ ☐ Demonstrate how to render a vaccine ineffective
☒ ☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
☒ ☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
☒ ☐ Increase transmissibility of a pathogen
☒ ☐ Alter the host range of a pathogen
☒ ☐ Enable evasion of diagnostic/detection modalities
☒ ☐ Enable the weaponization of a biological agent or toxin
☒ ☐ Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*

Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*
Effect(s) tested	*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

(See Eklund et al. 2016)

Correction	*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis