Published in final edited form as:

Nature. 2021 February ; 590(7847): 635–641. doi:10.1038/s41586-020-03148-w.

Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia

Rogan A. Grant[1,*], **Luisa Morales-Nebreda**[1,*], **Nikolay S. Markov**[1,*], **Suchitra Swaminathan**[2,3], **Melissa Querrey**[4], **Estefany R. Guzman**[3], **Darryl A. Abbott**[3], **Helen K. Donnelly**[1], **Alvaro Donayre**[1], **Isaac A. Goldberg**[1], **Zasu M. Klug**[1], **Nicole Borkowski**[1], **Ziyan**

[**]These corresponding authors jointly supervised this work and are listed in alphabetical order.
[*]These authors contributed equally to this work and will list themselves first on their CVs.
Consortia: The NU SCRIPT Study Investigators:
Rogan A. Grant[1], Luisa Morales-Nebreda[1], Nikolay S. Markov[1], Suchitra Swaminathan[2,3], Melissa Querrey[4], Estefany R. Guzman[3], Darryl A. Abbott[3], Helen K. Donnelly[1], Alvaro Donayre[1], Isaac A. Goldberg[1], Zasu M. Klug[1], Nicole Borkowski[1], Ziyan Lu[1], Hermon Kihshen[1], Yuliya Politanska[1], Lango Sichizya[1], Mengjia Kang[1], Ali Shilatifard[5,6], Chao Qi[7], Jon W. Lomasney[7], A. Christine Argento[1], Jacqueline M. Kruser[1], Elizabeth S. Malsin[1], Chiagozie O. Pickens[1], Sean B. Smith[1], James M. Walter[1], Anna E. Pawlowski[8], Daniel Schneider[8], Prasanth Nannapaneni[8], Hiam Abdala-Valencia[1], Ankit Bharat[1,4], Cara J. Gottardi[1], GR Scott Budinger[1], Alexander V. Misharin[1,6], Benjamin D. Singer[1,5,6], Richard G. Wunderink[1,6], Ajay A. Wagh[1], Alan R. Hauser[9], Alexis Rose Wolfe[1], Anjali Thakrar[2], Anjana V. Yeldandi[7], Ann A. Wang[1], Anne R. Levenson[1], Anthony M. Joudi[1], Betty Tran[1], Catherine A. Gao[1], Chitaru Kurihara[4], Clara J. Schroedl[1], Curt M. Horvath[9,10], Daniel Meza[1], David D. Odell[4], David W. Kamp[1], Deborah R. Winter[2], Egon A. Ozer[9], Elisheva D. Shanes[7], Elizabeth T. Bartom[5,6], Emily J. Rendleman[5], Emily M. Leibenguth[1], Firas Wehbe[8], Gabrielle Y. Liu[1], Gaurav T. Gadhvi[2], Heliodoro Tejedor Navarro[11], Jacob I. Sznajder[1], Jane E. Dematte[1], Jasmine Le[1], Jason M. Arnold[1], Joanne C. Du[1], John Coleman[1], Joseph I. Bailey[1], Joseph S. Deters[1], Justin A. Fiala[1], Justin Starren[8], Karen M. Ridge[1], Katharine Secunda[1], Kathleen Aren[2], Khalilah L. Gates[1], Kristy Todd[1], Lindsey D. Gradone[1], Lindsey N. Textor[1], Lisa F. Wolfe[1], Lorenzo L. Pesce[4], Luís A. Nunes Amaral[11], Madeline L. Rosenbaum[1], Manoj Kandpal[8], Manu Jain[1], Marc A. Sala[1], Mark Saine[4], Mary Carns[2], Michael J. Alexander[1], Michael J. Cuttica[1], Michelle H. Prickett[1], Nabiha H. Khan[5], Navdeep S. Chandel[1,6], Nicholas D. Soulakis[8], Orlyn R. Rivas[1], Patrick C. Seed[12], Paul A. Reyfman[1], Pearl D. Go[8], Peter H. S. Sporn[1], Phillip R. Cooper[1], Rade Tomic[1], Radhika Patel[1], Rafael Garza-Castillon[4], Ravi Kalhan[1], Richard I. Morimoto[10], Ruben J. Mylvaganam[1], Samuel S. Kim[4], Samuel W.M. Gatesy[13], Sanket Thakkar[4], Sarah Ben Maamar[8], SeungHye Han[1], Sharon R. Rosenberg[1], Sophia Nozick[9], Stefan J. Green[14], Susan R. Russell[1], Taylor A. Poor[1], Taylor J. Zak[7], Theresa A. Lombardo[8], Thomas Stoeger[11], Todd Shamaly[8], Ziyou Ren[1].
[9] Department of Microbiology-Immunology, Feinberg School of Medicine, Northwestern University
[10] Department of Molecular Biosciences, Weinberg College of Art and Sciences, Northwestern University
[11] Department of Chemical and Biological Engineering, McCormick School of Engineering, Northwestern University
[12] Division of Pediatric Infectious Diseases, Stanley Manne Research Institute, Ann and Robert H. Lurie Children's Hospital
[13] Division of Infectious Diseases, Department of Medicine, Feinberg School of Medicine, Northwestern University
[14] Genome Research Core, University of Illinois at Chicago.

Authors contribution:

G.R.S.B., A.V.M., B.D.S. and R.G.W. conceived and designed the project and provided funding. R.A.G., L.M.-N., N.S.M., G.R.S.B., A.V.M., B.D.S. and R.G.W. curated and analyzed the data and wrote the manuscript. C.O.P., J.M.K., J.M.W., B.D.S. and R.G.W. performed adjudication of clinical outcomes. A.D., H.K.D., I.A.G., N.B. and Z.M.K. recruited patients into the study and coordinated interactions between clinical and research teams. A.C.A., A.B., B.D.S., C.O.P., E.S.M., G.R.S.B., J.M.K., J.M.W., L.M.-N., R.G.W. and S.B.S. performed BAL procedures and sample collection. A.S. contributed equipment. C.Q. performed clinical microbiological assessment of BAL fluid. J.W.L. performed autopsies on patients who died from COVID-19 and provided biological materials. M.Q. performed RNAscope assay, imaging and analysis. A.E.P., D.S., M.K., P.N., performed programmatic extraction of clinical data and data curation. C.J.G. and E.S.M. coordinated and performed sample collection from healthy volunteers and provided biological materials. D.A.A., E.R.G., A.V.M. and S.S. processed BAL fluid samples and performed flow cytometry and cell sorting. D.A.A., H.K., L.S., R.A.G. and Z.L. performed post-sorting sample processing including RNA isolation for RNA-seq. B.D.S., A.V.M. and S.S. analyzed flow cytometry data. A.V.M. and Z.L. performed single-cell RNA-seq. H.A.-V. and Y.P. performed bulk RNA-seq. All authors provided edits and feedback on the manuscript.

Authors do not declare conflict of interest.

Data availability:

Bulk RNA-seq counts tables and metadata are included as Supplemental Data Files 2 and 3. Single-cell RNA-seq counts tables and integrated objects are available through GEO with accession number GSE155249. Raw data are available through SRA/dbGaP phs002300.v1.p1.

Code availability:

All code used for analysis is available at: https://github.com/NUPulmonary/2020_Grant. High-level bulk and single-cell RNA-seq data can be explored via our in-house data browsers at https://www.nupulmonary.org/covid-19/.

Lu[1], **Hermon Kihshen**[1], **Yuliya Politanska**[1], **Lango Sichizya**[1], **Mengjia Kang**[1], **Ali Shilatifard**[5,6], **Chao Qi**[7], **Jon W. Lomasney**[7], **A. Christine Argento**[1], **Jacqueline M. Kruser**[1], **Elizabeth S. Malsin**[1], **Chiagozie O. Pickens**[1], **Sean B. Smith**[1], **James M. Walter**[1], **Anna E. Pawlowski**[8], **Daniel Schneider**[8], **Prasanth Nannapaneni**[8], **Hiam Abdala-Valencia**[1], **Ankit Bharat**[1,4], **Cara J. Gottardi**[1], **GR Scott Budinger**[1,**], **Alexander V. Misharin**[1,6,**], **Benjamin D. Singer**[1,5,6,**], **Richard G. Wunderink**[1,6,**], **The NU SCRIPT Study Investigators**

[1]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Feinberg School of Medicine, Northwestern University

[2]Division of Rheumatology, Department of Medicine, Feinberg School of Medicine, Northwestern University

[3]Robert H Lurie Comprehensive Cancer Research Center, Feinberg School of Medicine, Northwestern University

[4]Division of Thoracic Surgery, Department of Surgery, Feinberg School of Medicine, Northwestern University

[5]Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University

[6]Simpson Querrey Institute for Epigenetics, Feinberg School of Medicine, Northwestern University

[7]Department of Pathology, Feinberg School of Medicine, Northwestern University

[8]Clinical and Translational Sciences Institute, Feinberg School of Medicine, Northwestern University

Abstract

Some patients infected with Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) develop severe pneumonia and the acute respiratory distress syndrome (ARDS)[1]. Distinct clinical features in these patients have led to speculation that the immune response to virus in the SARS-CoV-2-infected alveolus differs from other types of pneumonia[2]. We collected bronchoalveolar lavage fluid samples from 88 patients with SARS-CoV-2-induced respiratory failure and 211 patients with known or suspected pneumonia from other pathogens and subjected them to flow cytometry and bulk transcriptomic profiling. We performed single-cell RNA-seq on 10 bronchoalveolar lavage fluid samples collected from patients with severe COVID-19 within 48 hours of intubation. In the majority of patients with SARS-CoV-2 infection, the alveolar space was persistently enriched in T cells and monocytes. Bulk and single-cell transcriptomic profiling suggested that SARS-CoV-2 infects alveolar macrophages, which in turn respond by producing T cell chemoattractants. These T cells produce interferon-gamma to induce inflammatory cytokine release from alveolar macrophages and further promote T cell activation. Collectively, our results suggest that SARS-CoV-2 causes a slowly-unfolding, spatially limited alveolitis in which alveolar macrophages harboring SARS-CoV-2 and T cells form a positive feedback loop that drives persistent alveolar inflammation.

A minority of patients infected with Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) develop severe pneumonia requiring mechanical ventilation and these

patients account for almost all of the morbidity and mortality associated with the infection[1,3–5]. The reported 20–40% mortality among patients with severe SARS-CoV-2 pneumonia, combined with a severe systemic inflammatory response in some patients, have led to speculation that the pathobiology of SARS-CoV-2 pneumonia is distinct from pneumonia caused by other respiratory viral and bacterial pathogens[2].

We obtained bronchoalveolar lavage (BAL) samples from 88 patients with respiratory failure secondary to severe SARS-CoV-2 pneumonia and compared them with BAL samples prospectively collected before and during the pandemic from 211 patients with pneumonia secondary to other pathogens and 42 intubated patients without pneumonia. For many patients, we were able to obtain samples within 48 hours of intubation and sequentially over the course of the illness, allowing us to gain insights about the early pathogenesis and progression of COVID-19-induced respiratory failure. We profiled BAL samples using flow cytometry and performed bulk transcriptomic profiling of alveolar macrophages. Additionally, we performed single-cell RNA-sequencing on BAL fluid collected less than 48 hours after intubation from 10 patients with severe SARS-CoV-2 pneumonia. We used these data to develop a testable hypothesis to explain SARS-CoV-2 pathobiology (Fig. 1a).

Demographics of the cohort:

Samples were collected as part of the Successful Clinical Response in Pneumonia Therapy (SCRIPT) Systems Biology Center, an observational study of patients with severe pneumonia, defined as those requiring mechanical ventilation. We prospectively enrolled 88 of the 179 patients with SARS-CoV-2-induced pneumonia and respiratory failure requiring mechanical ventilation in our intensive care unit (ICU) (Fig. 1b, Extended Data Fig. 1a–c; Extended Data Table 1).

We compared patients with COVID-19 with 253 mechanically ventilated patients enrolled in the two years before and during the pandemic. These included patients intubated for reasons other than pneumonia (non-pneumonia control), patients diagnosed with severe non-SARS-CoV-2 viral pneumonia (other viral pneumonia), or patients with severe pneumonia secondary to bacterial or fungal pathogens (other pneumonia). Compared with these patients, patients with SARS-CoV-2 pneumonia were similar in age, race, and sex, but had a significantly higher self-reported Hispanic or Latino ethnicity and body mass index (Fig. 1c–f, Extended Data Table 1).

Severity of illness estimated using the sequential organ failure assessment (SOFA) score and the acute physiology score (APS) was similar in patients with SARS-CoV-2 pneumonia compared with other pneumonia and was comparable to that observed in a recent study of ARDS[6] (Fig. 1g,h, Extended Data Table 1). Patients with severe SARS-CoV-2 pneumonia had longer lengths of stay in the ICU and required longer periods of ventilation compared with all pneumonia and non-pneumonia controls (Fig. 1i,j). On the day of their first BAL, patients with severe SARS-CoV-2 pneumonia had increased levels of C-reactive protein compared with patients with other pneumonias, while other biomarkers of inflammation were similar (Extended Data Fig. 1d–i). The BAL sampling rate per day among patients with COVID-19 was not higher than the sampling rate among patients with other pneumonias

(see Extended Data Fig. 1j–l). Mortality was not different between patients with SARS-CoV-2 pneumonia compared with the entire cohort (Fig. 1k).

BAL fluid composition in SARS-CoV-2 pneumonia.

We began by performing flow cytometry on BAL samples collected within 48 hours of intubation (Fig. 2a,b, Extended Data Fig. 2a–e). We found that, despite severe pneumonia requiring mechanical ventilation, only 31% of patients with severe COVID-19 exhibited neutrophilia in BAL fluid (Fig. 2a,b). Instead, we found that in patients with severe SARS-CoV-2 pneumonia, the alveolar space was significantly enriched for CD4+ and CD8+ T cells and monocytes (Fig. 2a,b, Extended Data Fig. 3a).

Distinct macrophage response in COVID-19.

As macrophages respond to alterations in their micro-environment[7], we reasoned that changes in the alveolar macrophage transcriptome may reflect unique features of the SARS-CoV-2-infected alveolus. We isolated alveolar macrophages from BAL fluid collected within the first 48 hours after intubation from patients with severe SARS-CoV-2 pneumonia and compared them with alveolar macrophages from patients with pneumonia secondary to other pathogens, non-pneumonia controls, and healthy volunteers. k-means clustering of the 1,194 significantly variable genes ($q < 0.05$, likelihood-ratio test) across diagnosis identified 5 clusters (Fig. 2c, Supplemental Data Files 1–3). Notably, the majority of patients with COVID-19 clustered together. Cluster 1 contained genes specifically upregulated in patients with COVID-19 and was characterized by genes involved in the response to interferon (IFN). Cluster 1 also included genes encoding the chemokines *CCL7*, *CCL8*, and *CCL13,* which drive recruitment of monocytes and T cells.

To detect viruses in alveolar macrophages we aligned RNA-seq reads to a hybrid genome including the human, SARS-CoV-2, and Influenza A/California/07/2009 reference genomes. An additional negative-strand SARS-CoV-2 transcript, which is transiently formed during its replication, was added to the SARS-CoV-2 genome to detect replicating virus[8]. We detected SARS-CoV-2 transcripts in alveolar macrophage transcriptomes from 67% of samples with PCR-confirmed SARS-CoV-2 infection. In 38% of these samples, we detected both positive- and negative-strand SARS-CoV-2 transcripts (Extended Data Fig. 3b,c).

SARS-CoV-2 pneumonia persists over time.

Unlike other types of pneumonia, SARS-CoV-2 pneumonia is characterized by a long duration between symptom onset and the development of respiratory failure (6–12 days) and a prolonged course of mechanical ventilation (Fig. 1b,i,j)[1,4]. To determine whether the unique cellular composition of the BAL fluid and the interferon response signature in alveolar macrophages persists over the course of the disease, we analyzed samples obtained early after intubation (<48 hours) and all samples obtained >48 hours after intubation. Hierarchical clustering of BAL samples from all time points from serially sampled patients as well as direct pairwise comparison of early (<48 hours after intubation) versus late (>48 hours after intubation) samples between groups demonstrated that, in comparison with BAL

samples from other pneumonia, samples from patients with COVID-19 were persistently enriched for T cells (Fig. 3a,b; Extended Data Fig. 4a,b). These findings persisted in the 40% of patients with severe SARS-CoV-2 pneumonia who had a superinfecting pathogen detected over the course of mechanical ventilation (Extended Data Table 2). Nevertheless, BAL fluid neutrophils were increased in samples collected after 48 hours of mechanical ventilation in patients with COVID-19, likely attributable to the duration of mechanical ventilation and the attendant risk of bacterial superinfection (Extended Data Fig. 4c).

We then performed RNA-seq on alveolar macrophages isolated early (<48 hours) and later over the course of mechanical ventilation. k-means clustering of the 2,323 significantly variable genes ($q < 0.05$, likelihood-ratio test) across diagnoses identified five clusters (Fig. 3c; Extended Data Fig. 5a; Supplemental Data Files 1–3). Of note, genes and related GO biological processes associated with these clusters were remarkably similar to those identified in samples collected within 48 hours after intubation (Fig. 2c). Specifically, samples from patients with COVID-19 continued to cluster together and were differentiated by increased expression of genes in Cluster 1, which was enriched for interferon-response genes and T cell chemokines (Fig. 3c, Extended Data Fig. 5a,b).

In order to identify gene modules in alveolar macrophages that distinguish pneumonia type and outcome, we performed weighted gene co-expression network analysis (WGCNA) (Extended Data Fig. 5c, Supplemental Data File 1). Module 15 was enriched for interferon-responsive genes and correlated with the detection of SARS-CoV-2 transcripts, levels of C-reactive protein, CD8+ T cell abundance, and COVID-19 diagnosis ($R = 0.28, 0.30, 0.28, 0.25$, respectively). Notably, all SARS-CoV-2 genes included in this analysis were assigned to this module, further underscoring the relevance of the disease diagnosis category. Consistent with the results of k-means clustering, module 15 was enriched for type I and type II interferon response genes (GO:0060337, GO:0060333). We observed a significant negative correlation between interferon signaling and the duration of mechanical ventilation (Extended Data Fig. 5d–f). In addition, we identified modules related to macrophage maturation, including Module 4, which exhibited a positive correlation with the percentage of $CD206^{hi}$ alveolar macrophages ($R = 0.61$), and Module 12, which exhibited a negative correlation with percentage of $CD206^{hi}$ macrophages ($R = -0.62$) (Extended Data Fig. 5c). The unique features of SARS-CoV-2 pneumonia are further illustrated by a UMAP projection of all bulk RNA-seq samples, which separated largely by diagnosis, module 15 gene expression, and abundance of T cells (Extended Data Fig. 5g).

We took advantage of the serial samples collected from the same patients with COVID-19 and explored the relationship between expression of canonical interferon-response genes and patient outcomes. Analysis of the serial samples from patients with COVID-19 who were discharged home (5 patients) or to an inpatient facility (2 patients) demonstrated decreased expression of interferon-response genes as the disease progressed, while in patients who were discharged to a long-term acute care facility (LTAC, 5 patients) or died (2 patients) the expression of interferon-response genes was largely unchanged (Extended Data Fig. 5h). Consistent with these findings, we observed a significant negative correlation between abundance of SARS-CoV-2 transcripts in patients with confirmed COVID-19 and the time since intubation ($\rho = -0.49$, Spearman correlation; Extended Data Fig. 5i).

Immune cell circuits in SARS-CoV-2 pneumonia.

We performed single-cell RNA-seq on 10 patients with COVID-19 from whom BAL samples were collected within 48 hours of intubation (Fig. 4a, Extended Data Fig. 6a–d). We included two additional patients, one with bacterial pneumonia and one non-pneumonia control (Extended Data Fig. 6e–h). Analysis of an integrated object resolved multiple clusters corresponding to macrophages and other cell types (Supplemental Data File 4). Alveolar macrophages contained six clusters. Four clusters of monocyte-derived alveolar macrophages (MoAM) were characterized by *CCL2* expression and gradually increasing expression of genes associated with alveolar macrophage maturation. The other two clusters expressed markers of tissue-resident alveolar macrophages (TRAM) (Fig. 4b, Extended Data Fig. 6a). We did not detect expression of type I interferons in our single-cell dataset or in other publicly available single-cell RNA-seq datasets[9,10] obtained from BAL fluid later in the clinical course of patients with COVID-19 (data not shown). In contrast, expression of type II interferon (*IFNG*) was detected in T cells from all 10 patients with COVID-19 (Fig. 4c, Extended Data Fig. 6b). These results suggest that the interferon-response gene signature we observed in alveolar macrophages may be in response to IFNγ released from activated T cells.

As expected, positive- and negative-strand SARS-CoV-2 transcripts were detected in epithelial cells. Surprisingly, we also detected SARS-CoV-2 transcripts in migratory *CCR7*+ dendritic cells, MoAM2 and TRAM2, which do not express *ACE2* (Fig. 4d, Extended Data Fig. 6c). Coronaviruses generate large numbers of positive-strand transcripts from a single negative-strand template[8,11]. Consistent with this known biology, we detected more transcripts for positive compared with negative strands in both our single-cell and bulk RNA-seq data (Extended Data Fig. 3c, Extended Data Fig. 6d and Fig. 4d). These results suggest that alveolar macrophages harbor SARS-CoV-2 and that they may support viral replication as has been reported for SARS-CoV and MERS-CoV[12–14].

TRAM2 harboring SARS-CoV-2 exhibited a distinct transcriptional program compared with uninfected TRAM1 (Fig. 4b,h, Extended Data Fig. 6g,h, Supplemental Data File 5). In patients with COVID-19, genes distinguishing infected TRAM2 from non-infected TRAM1 included several chemokines and cytokines important for T cell and monocyte recruitment, such as *CCL4, CCL20*, *CXCL10* and *CXCL11.* TRAM2 also expressed genes *IL1B, TNFSF10*, and *DEFB1* (Fig. 4b,e–h, Supplemental Data File 6). Finally, infected TRAM2 were marked by increased expression of interferon-response genes compared with non-infected TRAM1 (Fig. 4b,h).

IL-6 induces the transcription of clotting factors in the liver and tissue factor in the endothelium to promote thrombosis, and elevated levels of IL-6 predict mortality in patients with COVID-19[15–17]. Because single-cell RNA-seq analyses of peripheral blood from patients with COVID-19 failed to identify circulating cells producing IL-6[18,19], some have suggested that IL-6 is produced by inflammatory cells in the alveolus[17]. In our dataset, the overall expression of *IL6* was low and was primarily restricted to very immature MoAM1 originating from two patients (Extended Data Fig. 7a). Furthermore, the expression of *IL6* in BAL fluid cells from patients with COVID-19 was comparable to previously reported *IL6*

expression in alveolar macrophages from patients with later-stage COVID-19[9] and lower than *IL6* expression observed in stromal and endothelial cells from a published single-cell RNA-seq dataset describing the healthy or fibrotic lung[20]. Consistently, in our bulk RNA-seq data from flow cytometry-sorted alveolar macrophages, *IL6* expression was not different between patients with pneumonia secondary to SARS-CoV-2 compared with other respiratory pathogens, although it was higher than in healthy controls, in whom *IL6* transcripts were never detected (Extended Data Fig. 5b). While we did not sample neutrophils, Liao et al. and Chua et al. did not detect increased neutrophil *IL6* expression in their single-cell RNA-seq datasets of BAL fluid from patients with SARS-CoV-2 pneumonia[9,10].

In murine models of influenza A viral pneumonia, IFNγ drives apoptosis of tissue-resident alveolar macrophages[21]. We therefore wondered whether the CD206^{hi} alveolar macrophages in patients with severe SARS-CoV-2 pneumonia were truly tissue-resident alveolar macrophages or maturing monocyte-derived alveolar macrophages. We performed cell-type deconvolution of our bulk RNA-seq data and found that in patients with SARS-CoV-2 pneumonia only a small fraction of alveolar macrophages were tissue-resident alveolar macrophages. Instead, the majority of alveolar macrophages were monocyte-derived alveolar macrophages (MoAM3) (Extended Data Fig. 7b,c).

The BAL procedure samples cells from a lung segment, which includes many thousands of alveoli. Therefore, the detection of infected and non-infected alveolar macrophages in the same BAL fluid sample in our single cell RNA-seq data suggests infected and uninfected alveoli will co-exist within a single lung segment. To test this prediction, we examined lung tissue from a patient who died from SARS-CoV-2 pneumonia while forgoing treatments, save comfort measures. Using smFISH we detected positive- and negative-strand SARS-CoV-2 transcripts in the lung epithelium and alveolar macrophages (Fig. 4i and Supplemental Data File 7, Supplemental Data Video 1). Positive-strand SARS-CoV-2 transcripts exhibited diffuse cytoplasmic staining, and the number of punctae exceeded that of negative transcripts. As predicted from our model of lung infection, cells containing SARS-CoV-2 transcripts were also detected in a spatially-restricted manner, as regions of the lung containing the virus were adjacent to the regions with nearly-normal architecture where viral particles were absent.

Discussion:

A minority of patients with severe SARS-CoV-2 pneumonia develop respiratory failure, but it is these patients who account for almost all of the morbidity, mortality, and socioeconomic cost associated with the COVID-19 pandemic. We systematically sampled the alveolar space in patients with new-onset respiratory failure secondary to SARS-CoV-2 pneumonia and compared these samples with those from a large comparison cohort of patients with pneumonia from other respiratory pathogens collected before and during the pandemic. We used these data to develop a model for the unique pathobiology of SARS-CoV-2 pneumonia (Fig. 1a). Our model proposes that SARS-CoV-2 initially infects and replicates in epithelial cells in the nasopharynx, which express relatively high levels of *ACE2* in comparison with epithelial cells in lower airways or the distal lung[22,23]. Whether by progressive movement

distally in the tracheobronchial tree or via aspiration of nasopharyngeal contents, some virus gains access to the distal alveolar space. In the alveolar space, SARS-CoV-2 infects alveolar epithelial cells and tissue-resident alveolar macrophages[24]. A transcriptional program is activated within infected tissue-resident alveolar macrophages that promotes the recruitment of memory T cells to the alveolar space. There, memory T cells become activated, releasing IFNγ, which activates tissue-resident alveolar macrophages harboring SARS-CoV-2. Eventually, these tissue-resident alveolar macrophages die and monocyte-derived alveolar macrophages are recruited, which in turn become infected with SARS-CoV-2 to sustain the inflammatory signaling loop with T cells. These infected alveolar macrophages may act as a Trojan horse, transferring the virus to adjacent lung regions, slowly propagating SARS-CoV-2 infection across the lung.

Our model is informed by a wealth of causal data generated in cell and animal models of SARS-CoV-2 and the related coronaviruses SARS-CoV and MERS-CoV. SARS-CoV-2 proteins have been shown to suppress type I interferon responses[24–26], and we did not detect ongoing expression of type I interferons in our bulk or single-cell RNA-seq data[25,27]. Our data suggesting that SARS-CoV-2 replicates in alveolar macrophages are consistent with reports that SARS-CoV, MERS-CoV, and SARS-CoV-2 can infect macrophages *in vitro*[12–14,24,28], and the detection of SARS-CoV-2 in alveolar macrophages in autopsy studies of patients with SARS-CoV-2 pneumonia[29].

Alveolar macrophage infection might result from phagocytosis of infected alveolar epithelial cells followed by viral escape from the lysosome. Alternatively, alveolar macrophages might be directly infected, as was shown for SARS-CoV and MERS-CoV[12,14]. Finally, antibody-dependent enhancement has been suggested from cell and animal models of SARS-CoV and MERS-CoV infection[30,31].

Activation of memory T cells leads to IFNγ production, local proliferation of activated memory T cells, mounting of inflammatory responses, and recruitment of monocytes and T cells. We observed enrichment of CD4+ and CD8+ T cells in the alveolar space of patients with SARS-CoV-2 pneumonia relative to pneumonia secondary to other pathogens. Furthermore, our single-cell RNA-seq data confirmed production of IFNγ by both CD4+ and CD8+ T cells. These results raise the question of how T cells in the alveolar space become activated following SARS-CoV-2 infection. While tissue-resident alveolar macrophages are poor antigen-presenting cells and do not convert naïve T cells into effector T cells[32], a low level of antigen presentation by alveolar macrophages might be sufficient to activate pre-existing memory T cells that cross-react with SARS-CoV-2. Existence of such cross-reactive memory T cells has been reported for SARS-CoV[33] and for SARS-CoV-2[34–38]. A larger number of SARS-CoV-2 cross-reactive memory T cells was observed independently in elderly people and in patients with severe COVID-19, compared with those with mild COVID-19[38]. These cross-reactive T cells exhibited lower avidity and reduced antiviral responses in response to stimulation with SARS-CoV-2 peptides compared with T cells from patients who recovered from COVID-19. This mechanism might explain the epidemiology of severe SARS-CoV-2 pneumonia, which disproportionately affects elderly individuals.

Single-cell RNA-seq atlases and smFISH studies of the normal human lung show that only a small number of alveolar epithelial cells express *ACE2*, the gene encoding the receptor for SARS-CoV-2 entry[22,23]. In contrast, the sialic acid residues that serve as receptors for influenza A virus are abundantly expressed in alveolar type 2 cells[22]. Thus, while influenza A infects large numbers of cells leading to rapid viral replication, widespread injury, robust antiviral responses, and death of infected epithelial cells, infection by SARS-CoV-2 is likely to lead to spatially localized areas of infection. This could explain the localized areas of ground glass infiltrates observed in chest computed tomography in minimally symptomatic patients with COVID-19[39]. We speculate that alveolar macrophages harboring SARS-CoV-2 might spread virus between alveoli. For example, tissue-resident alveolar macrophages, once thought to be sessile, were recently shown to travel between alveoli through pores of Kohn, particularly during viral infection[40]. In each new area of infection, positive feedback loops between alveolar macrophages harboring the virus and activated T cells could promote persistent injury and inflammation.

Our model of SARS-CoV-2 pneumonia as a slowly-progressive, spatially-restricted infection explains some of the unusual clinical features of COVID-19. Most notably, the clinical course of severe SARS-CoV-2 pneumonia is much longer than that of other respiratory viruses. The time from the onset of symptoms to respiratory failure in patients with SARS-CoV-2 infection is 6–12 days, compared with 1–3 days or even less in patients with influenza A virus infection[4,41]. Furthermore, in our cohort the duration of mechanical ventilation and ICU stay were much longer in patients with severe SARS-CoV-2 pneumonia, compared with other respiratory pathogens causing pneumonia, despite a similar severity of illness and mortality. This longer clinical course might also explain some of the systemic complications of the disease. Although sparsely sampled, the levels of inflammatory biomarkers in the blood were similar in patients with COVID-19 and those with other etiologies of pneumonia in our cohort, confirming more recent systematic reports[42,43]. This observation raises the intriguing possibility that the increased number of systemic complications observed in patients with SARS-CoV-2 pneumonia are related to prolonged, rather than higher level, exposure to circulating inflammatory cytokines.

In summary, we present a dataset that supports a testable model, in which alveolar macrophages harboring SARS-CoV-2 form positive feedback loops with IFNγ-secreting T cells to promote alveolitis in patients with severe COVID-19.

Methods:

Human subjects:

All human subjects research was approved by the Northwestern University Institutional Review Board. Samples from patients with COVID-19, viral pneumonia, other pneumonia, and non-pneumonia controls were collected from participants enrolled in the Successful Clinical Response In Pneumonia Therapy (SCRIPT) study STU00204868. Alveolar macrophages from healthy volunteers were obtained under study STU00206783. Autopsy tissues were obtained per usual protocols and analyzed under study STU00079445. All study participants or their surrogates provided informed consent.

Patients ≥ 18 years of age with suspicion of pneumonia based on clinical criteria (including but not limited to fever, radiographic infiltrate, and respiratory secretions) were screened for enrollment into the SCRIPT study. Inability to safely perform BAL or NBBAL were considered exclusion criteria[44]. In our center, patients with respiratory failure are intubated based on the judgement of bedside clinicians for worsening hypoxemia, hypercapnia, or work of breathing refractory to high-flow oxygen or non-invasive ventilation modes. Extubation occurs based on the judgement of bedside clinicians following a protocol-driven trial of spontaneous breathing in patients demonstrating physiologic improvement in their cardiorespiratory status during their period of mechanical ventilation.

We collected data and samples from patients enrolled in SCRIPT from June 15, 2018 to July 6, 2020 in our intensive care unit (ICU) at Northwestern Memorial Hospital in Chicago. We prospectively enrolled 88 of the 179 patients with SARS-CoV-2-induced pneumonia and respiratory failure requiring mechanical ventilation in our intensive care unit (Fig. 1b, Extended Data Fig. 1a, b; Extended Data Table 1), all but one of whom had been discharged at the time of our submission. Management of patients with COVID-19 was guided by protocols published and updated on the Northwestern Medicine intranet as new information became available over the pandemic. Clinical laboratory testing including studies ordered on BAL fluid was at the discretion of the care team; however, quantitative cultures, multiplex PCR (BioFire Film Array Respiratory 2 panel), and automated cell count and differential were recommended by local ICU protocols. Most patients also underwent urinary antigen testing for *Streptococcus pneumoniae* and *Legionella pneumophilia* serogroup 1 on admission. Clinicians were encouraged to manage all patients, including those with COVID-19, according to ARDSNetwork protocols including the use of a higher PEEP/lower FiO_2 strategy for those with severe hypoxemia[45, 46]. Prone positioning (16 hours per day) was performed in all patients with a PaO_2/FiO_2 <150 who did not have contraindications[47]. In those who had a response to prone positioning evidenced by improved oxygenation, prone positioning was repeated. Esophageal balloon catheters (Cooper Surgical) were placed at the discretion of the care team to estimate transpulmonary pressure and optimize PEEP, particularly in patients with a higher-than-normal BMI.

Autopsy specimen used for sm-FISH (RNAscope) (Fig. 4i): An 81-year old woman with end-stage renal disease and cirrhosis was admitted for a fever of 38.5°C. She had a positive nasopharyngeal swab for SARS-CoV-2. The patient developed increased O_2 requirements and was subsequently transferred to the COVID ICU. In the ICU, the decedent developed hypotension to 60s/40s and after discussion with the clinical team, the decedent's family elected to focus on comfort care. The patient died 8 days after admission.

NBBAL and BAL procedures:

Consent was obtained from patients or their surrogates for bronchoscopic procedures. Bronchoscopic BAL was performed in intubated ICU patients with flexible, single-use Ambu aScope (Ambu) devices. Patients were given sedation and topical anesthetic at the physician proceduralist's discretion. Vital signs were monitored continuously throughout the procedure. The bronchoscope was wedged in the segment of interest based on available chest imaging or intra-procedure observations, aliquots of 30 ml of normal saline at a time,

generally 90–120 ml total, were instilled and aspirated back. The fluid returned following the first aliquot was routinely discarded. Samples were split (if sufficient return volume was available) and sent for clinical studies and an aliquot reserved for research. A similar procedure was applied to non-bronchoscopic BAL (NBBAL); however, NBBAL was performed with directional (lateral) but not visual guidance, and as usual procedural care by a respiratory therapist rather than a pulmonologist[44].

For bronchoscopies performed in COVID-19 patients, additional precautions were taken to minimize the risk to healthcare workers including only having essential providers present in the room, clamping of the endotracheal tube, transient disconnection of the inspiratory limb from the ventilator, and preloading of the bronchoscope through the adapter[48]. Sedation and neuromuscular blockade to prevent cough, was administered for these procedures at the physician's discretion. In most cases of early bronchoscopy, the procedure was performed immediately after intubation, taking advantage of neuromuscular blockade administered for the intubation procedure.

For all patients with COVID-19, samples were collected from regions of greatest chest radiograph abnormality by a critical care physician using a disposable bronchoscope. The majority of samples prior to the pandemic were collected by respiratory therapists using a non-bronchoscopic bronchoalveolar lavage (NBBAL) catheter that is the same diameter as a standard bronchoscope with the catheter directed to the most radiographically affected lung. For both bronchoscopic and NBBAL, the recommended instillate volume was 120 ml and the initial aliquot was discarded if adequate return was obtained[44].

Pneumonia adjudication:

Five critical care physicians (JMK, COP, BDS, JMW, RGW) retrospectively adjudicated patients as COVID-19 pneumonia, non-COVID-19 viral pneumonia, pneumonia secondary to other pathogens, or non-pneumonia controls (intubated for reasons other than pneumonia), according to a standardized adjudication procedure (the adjudication protocol can be found in Supplemental Data File 8). Non-pneumonia controls were defined as patients who underwent BAL to exclude pneumonia but had negative quantitative cultures, a negative multiplex PCR for viral and bacterial pathogens (when available) and negative urine antigens for *Streptococcus pneumoniae* and *Legionella pneumophilia* serogroup 1, as well as an alternative diagnosis. The treating clinician's impression was considered, but concordance was not required. Subsequent course and the entirety of the clinical record was used for adjudication. Some of the patients adjudicated as non-pneumonia controls developed ventilator-associated pneumonia later in their clinical course. Viral pneumonia was diagnosed based on detection of a respiratory viral pathogen from either a nasopharyngeal swab or BAL fluid in the appropriate clinical setting. Bacterial pneumonia was defined as positive quantitative cultures with > 10^2 colony forming units/ml, detection of a bacterial pathogen by PCR analysis of BAL fluid or a positive urine antigen. Over the course of the study, BAL fluid was analyzed using an MRSA PCR (MRSA/SA SSTI) and the BioFire FilmArray Respiratory 2 (RP2) Panel and Pneumonia panels. SARS-CoV-2 was detected with a variety of assay platforms including the Cepheid Gene Expert, Abbott ID NOW, Becton-Dickinson, and a locally-developed and validated PCR. For some patients

without COVID-19, the diagnosis of pneumonia was made based on clinical suspicion, radiographic findings, and response to antimicrobial therapy in the absence of an identified pathogen. Ventilator-associated pneumonia was diagnosed as detection of a new respiratory pathogen using quantitative culture or PCR more than 48 hours after intubation, the detection of a new respiratory pathogen on serial BAL samples, or the reappearance of a respiratory pathogen after a negative BAL on a subsequent study. Clinical laboratory data were obtained from the Northwestern Medicine Enterprise Data Warehouse using Structured Query Language (SQL). APS and SOFA scores were generated from the Electronic Health Record using previously validated programming.

Flow cytometry and cell sorting:

NBBAL and BAL samples were filtered through a 70-μm cell strainer, pelleted by centrifugation at 300 rcf for 10 min at 4°C, followed by hypotonic lysis of red blood cells with 2 ml of BD PharmLyse reagent for 2 minutes. Lysis was stopped by adding 13 ml of MACS buffer (Miltenyi Biotech). Cells were pelleted again and resuspended in 100 μl of 1:10 Fc-Block (Human TruStain FcX, Biolegend) in MACS buffer, and a 10-μl aliquot was taken for counting using K2 Cellometer (Nexcelom) with AO/PI reagent. The volume of Fc-Block was adjusted so the concentration of cells was always less than 5×10^7 cells/ml and the fluorophore-conjugated antibody cocktail was added in 1:1 ratio (Extended Data Table 3). After incubation at 4°C for 30 minutes, cells were washed with 5 ml of MACS buffer, pelleted by centrifugation, and resuspended in 500 μl of MACS buffer with 2 μl of SYTOX Green viability dye (ThermoFisher). Cells were sorted on a FACS Aria III SORP instrument using a 100-μm nozzle at 20 psi. Cells were sorted into 300 μl of MACS buffer for bulk RNA-seq or 300 μl of 2% BSA in PBS for single-cell RNA-seq. Sample processing was performed in BSL-2 facility using BSL-3 practices. Analysis of the flow cytometry data was performed using FlowJo 10.6.2. using uniform sequential gating strategy (Extended Data Fig. 2) reviewed by three investigators (SS, BDS, AVM). Immune populations were defined using canonical markers as shown in Extended Data Fig. 2[44, 49]. Alveolar macrophages were defined by the expression of CD206 (mannose receptor). We further subdivided alveolar macrophages into $CD206^{lo}$ alveolar macrophages, which represent differentiating monocyte-derived alveolar macrophages, and $CD206^{hi}$ alveolar macrophages, which include both tissue-resident alveolar macrophages, present in the lung prior to the onset of pneumonia, and mature monocyte-derived alveolar macrophages. Since $CD206^{hi}$ alveolar macrophages can be found in BAL fluid from both healthy volunteers and patients across different types of pneumonia, and the presence of $CD206^{lo}$ alveolar macrophages varies across the patients with pneumonia[49–51], we focused our transcriptomic analysis on $CD206^{hi}$ alveolar macrophages. Abundance of specific populations in individual BAL fluid samples can be found in Supplemental Data File 9.

Bulk RNA-seq of flow cytometry-sorted alveolar macrophages:

Immediately after sorting, cells were pelleted by centrifugation and lysed in 350 μl of RLT Plus lysis buffer (Qiagen) supplemented with 2-mercaptoethanol. Lysates were stored at −80°C until RNA isolation using the AllPrep DNA/RNA Micro kit according to the manufacturer's protocol (Qiagen). RNA quality and quantity were assessed using TapeStation 4200 High Sensitivity RNA tapes (Agilent), and RNA-seq libraries were

prepared from 250 pg of total RNA using SMARTer Stranded Total RNA-seq Kit v2 (Takara Bio). Libraries were pooled using dual indexing and sequenced on a NextSeq 500 instrument (Illumina), 75 cycles, single-end, to an average sequencing depth of 19.55M reads.

FASTQ files were generated using bcl2fastq (Illumina). To enable detection of viral RNA, a custom hybrid genome was prepared by joining FASTA, GFF, and GTF files for GRCh37.87, SARS-CoV-2 (NC_045512.2), and Influenza A/California/07/2009 (GCF_001343785.1), which was the dominant strain of influenza throughout BAL fluid collection at our hospital[52]. An additional negative strand transcript spanning the entirety of the SARS-CoV-2 genome was then added to the GTF and GFF files to enable detection of SARS-CoV-2 replication. Normalized counts tables later revealed high enrichment of SARS-CoV-2 transcripts in diagnosed COVID-19 patients, and enrichment of IAV genes in patients marked as other viral pneumonia. Of note, since our alveolar macrophage sorting strategy for bulk RNA-seq (Extended Data Fig. 2a,b) only focused on $CD206^{hi}$ cells, our bulk RNA-seq data likely underestimate infection of alveolar macrophages infected with SARS-CoV-2 (Fig. 2d).

To facilitate reproducible analysis, samples were processed using the publicly available nf-core/RNA-seq pipeline version 1.4.2 implemented in Nextflow 19.10.0 using Singularity 3.2.1–1 with the minimal command nextflow run nf-core/rnaseq -r 1.4.2 –singleEnd -profile singularity –reverseStranded --three_prime_clip_r2 3[53–55]. Briefly, lane-level reads were trimmed using trimGalore! 0.6.4 and aligned to the hybrid genome described above using STAR 2.6.1d[56]. Gene-level assignment was then performed using featureCounts 1.6.4[57]. Putative sample swaps were identified first by comparing known patient sex with sex determined by levels of *XIST* and *RPS4Y1* expression, followed by single nucleotide polymorphism analysis with NGSCheckMate version 1.0.0 in FASTQ mode using default settings[58]. Samples exhibiting unexpected correlation were excluded from analysis.

Bulk differential expression analysis:

All analysis was performed using custom scripts in R version 3.6.3 using the DESeq2 version 1.26.0 framework[59]. Correspondence between lanes was first confirmed by principal component analysis before merging counts using the command collapseReplicates(). One outlier sample from the Other Pneumonia group with low RIN score and exhibiting extreme deviation on PCA and poor alignment and assignment metrics was excluded from downstream analysis. For differential expression analysis (DEA), both proportion of alveolar macrophage from flow cytometry data and diagnosis were used as explanatory factors. A “local” model of gene dispersion was employed as this better fit dispersion trends without obvious overfitting, and gene outlier replacement was disabled; otherwise default settings were used. See code for details.

k-means clustering of bulk samples.

A custom-built function was used (available at https://github.com/NUPulmonary/utils/blob/master/R/k_means_Figure.R) for k-means clustering. Briefly, variable genes were identified using a likelihood-ratio test (LRT) with local estimates of gene dispersion in DESeq2 with diagnosis as the full model as well as a reduced model corresponding to intercept alone.

Genes with q ≥ 0.05 were discarded. Extant genes were then clustered using the Hartigan-Wong method with 25 random sets and a maximum of 1000 iterations using the kmeans function in R stats 3.6.3. Samples were then clustered using Ward's method and plotted using pheatmap version 1.0.12. GO term enrichment was then determined using Fisher's exact test in topGO version 2.38.1, with org.Hs.eg.db version 3.10.0 and GO.db version 3.10.0 as references.

Weighted gene coexpression network analysis (WGCNA).

WGCNA was performed manually using WGCNA version 1.69 with default settings unless otherwise noted[60]. Genes with counts > 5 and detection in at least 10% of samples were included in the analysis. To best capture patterns of co-regulation, a signed network was used. Using the pickSoftThreshold function, we empirically determined a soft threshold of 7 to best fit the network structure. A minimum module size of 30 was chosen to isolate relatively large gene modules. Module eigengenes were then related back to patient and sample metadata using biweight midcorrelation. Module GO enrichment was then determined as above using Fisher's exact test in topGO version 2.38.1, with org.Hs.eg.db version 3.10.0 as a reference. UMAP plotting was performed using uwot version 0.1.8 using the first 20 principal components of the same genes used in WGCNA analysis after Z-scaling and centering, with a minimum distance of 0.2[61]. Default parameters were otherwise used.

Single-cell RNA-seq of flow cytometry-sorted BAL cells:

For patients with COVID-19 we limited our analysis to samples in which flow cytometry identified distinct populations of $CD206^{hi}$ and $CD206^{lo}$ macrophages (Patient 1, 2, 3, 4, 5, 7, 8, 9, A and B). We included two additional patients, one with bacterial pneumonia secondary to infection with *Pseudomonas aeruginosa* and *Acinetobacter baumannii* (assigned as "other pneumonia", Patient 6) and one intubated for airway protection to facilitate endoscopy for severe gastrointestinal bleeding without pneumonia (assigned as a "non-pneumonia control", Patient C) to examine nonspecific effects of inflammation and mechanical ventilation, respectively, on transcriptomic signatures.

Samples were enriched via flow cytometry sorting for live cells, excluding granulocytes. Cells were sorted into 2% BSA in DPBS, pelleted by centrifugation at 400 rcf for 5 min at 4°C, resuspended in 0.1% BSA in DPBS to ~1000 cells/μl concentration. Concentration was confirmed using K2 Cellometer (Nexcelom) with AO/PI reagent and cells were loaded on 10x Genomics Chip A with Chromium Single Cell 5' gel beads and reagents (10x Genomics) aiming to capture ~5,000–10,000 cells per library. Libraries were prepared according to the manufacturer's protocol (10x Genomics, CG000086_RevM). After quality checks, single-cell RNA-seq libraries were pooled and sequenced on a NovaSeq 6000 instrument.

Data was processed using the Cell Ranger 3.1.0 pipeline (10x Genomics). To enable detection of viral RNA, reads were aligned to a custom hybrid genome containing GRCh38.93 and SARS-CoV-2 (NC_045512.2). An additional negative-strand transcript spanning the entirety of the SARS-CoV-2 genome was then added to the GTF and GFF files

to enable detection of SARS-CoV-2 replication. Data were processed using Scanpy v1.5.1[62], doublets were detected with scrublet v0.2.1[63] and removed, ambient RNA was corrected with FastCAR (https://github.com/LungCellAtlas/FastCAR), and multisample integration was performed with BBKNN v1.3.12[64]. Only human transcripts were used during integration, selection of highly variable genes and clustering, SARS-CoV-2 transcripts did not influence clustering. Gene set enrichment analysis was performed with signatures retrieved from the gsea-msigdb.org website[65] using following terms: HALLMARK_INTERFERON_GAMMA_RESPONSE M5913, HALLMARK_INTERFERON_ALPHA_RESPONSE M5911. Computations were automated with snakemake v5.5.4[66].

With exclusion of patients A and B, single-cell RNA-seq was performed without multiplexing, using cells from a single patient per single 10x Genomics chip channel. Cells from patients A and B were split into three 10x Genomics chip channels: sample 14 contained cells from patient A, sample 15 contained cells from patient B and sample 16 contained cells from patients A and B multiplexed together. To assign cells from this sample to patients, we used souporcell v2.0[67] (commit 34eade2ad3a361f045a31f53fee58c2e0c49423f) with the list of common variants for GRCh38 genome, provided on the souporcell page. We ran souporcell for samples 14, 15 and 16 with the number of clusters k=2. We computed Pearson correlation between integer-coded single-nucleotide polymorphisms in genotypic clusters in sample pairs 14–16 and 15–16 to determine which genotypic clusters come from the same patients. Genotypic doublets and unassigned cells were discarded. See code for details.

Deconvolution of bulk RNA-seq alveolar macrophage signatures

Deconvolution of bulk RNA-seq alveolar macrophage signatures was performed using AutoGeneS v1.0.3[68] and signatures derived from the integrated single-cell RNA-seq object. We used an integrated single-cell RNA-seq object containing the first six subjects included into analysis (patients 1–6) to train the AutoGeneS model. Signatures were automatically identified from 4500 highly variable genes with function optimize(ngen=200, seed=0, nfeatures=2000, mode="fixed") (see code at https://github.com/NUPulmonary/2020_Grant). The model was then applied to bulk RNA-seq data to estimate the proportion of specific cell types using regression. We used bulk RNA-seq samples from healthy volunteers (which contain only tissue-resident alveolar macrophages and do not contain inflammatory monocyte-derived alveolar macrophages) to validate and optimize selection of the cluster-specific genes and deconvolution results.

RNAscope of paraffin lung slices:

RNAscope Multiplex V2 manual assay from ACDbio was performed on paraffin-embedded 5 μm slices of lung tissue using mild digest times according to manufacturer instructions as we have described. Probes used were RNAscope Probe-V-nCoV2019-S-C3 (catalog number 848561) with Akoya Bio Opal Dye 520 using the 488 nm laser line and RNAscope Probe-nCoV2019-orf1ab-sense-C2 (catalog number 859151) with Opal Dye 690 using 640 nm laser line. After RNAscope assay was complete, slides were washed in TBST (1X TBS pH 7.6 with 0.1% Tween-20) for 2 minutes with agitation twice. Slides were incubated in the

dark at room temperature for 30 minutes with 10% normal goat serum in 1X TBS with 1% bovine serum albumin (BSA). The blocking solution was removed from slides via gentle flicking. Slides were then incubated in primary antibody solution containing anti-CD206 antibody (clone C-10) conjugated to AF546 (Santa Cruz Biotechnology Cat# sc-376232, RRID:AB_10989352) at 1:100 dilutions in TBS 1% BSA for 1 h at room temperature in the dark. Slides were rinsed using TBST for 5 minutes with agitation twice. Slides were rinsed in TBST buffer for 5 minutes twice. Slides were then mounted and dried overnight. Images were acquired at the Center for Advanced Microscopy at Northwestern University Feinberg School of Medicine using the Nikon W1-Spinning Disk Confocal microscope. Nucleus was added to the images using machine-based learning network trained on one patient using DAPI and brightfield images in Nikon Elements. Final images were rendered using Fiji.

Statistical analysis:

Statistical analysis was performed using base R version 3.6.3 with tidyverse version 1.3.0[69] and Python 3.6. For all comparisons, normality was first assessed using a Shapiro-Wilk test and manual examination of distributions. For parameters exhibiting a clear lack of normality, nonparametric tests were employed. In cases of multiple testing, P-values were corrected using false-discovery rate (FDR) correction. In Python, we used the mannwhitneyu function from scipy package version 1.3.1[70] for nonparametric tests, and corrected for multiple testing with the statsmodels package version 0.10.1[71]. Adjusted P-values < 0.05 were considered significant. Two-sided statistical tests were performed in all cases.

Visualization:

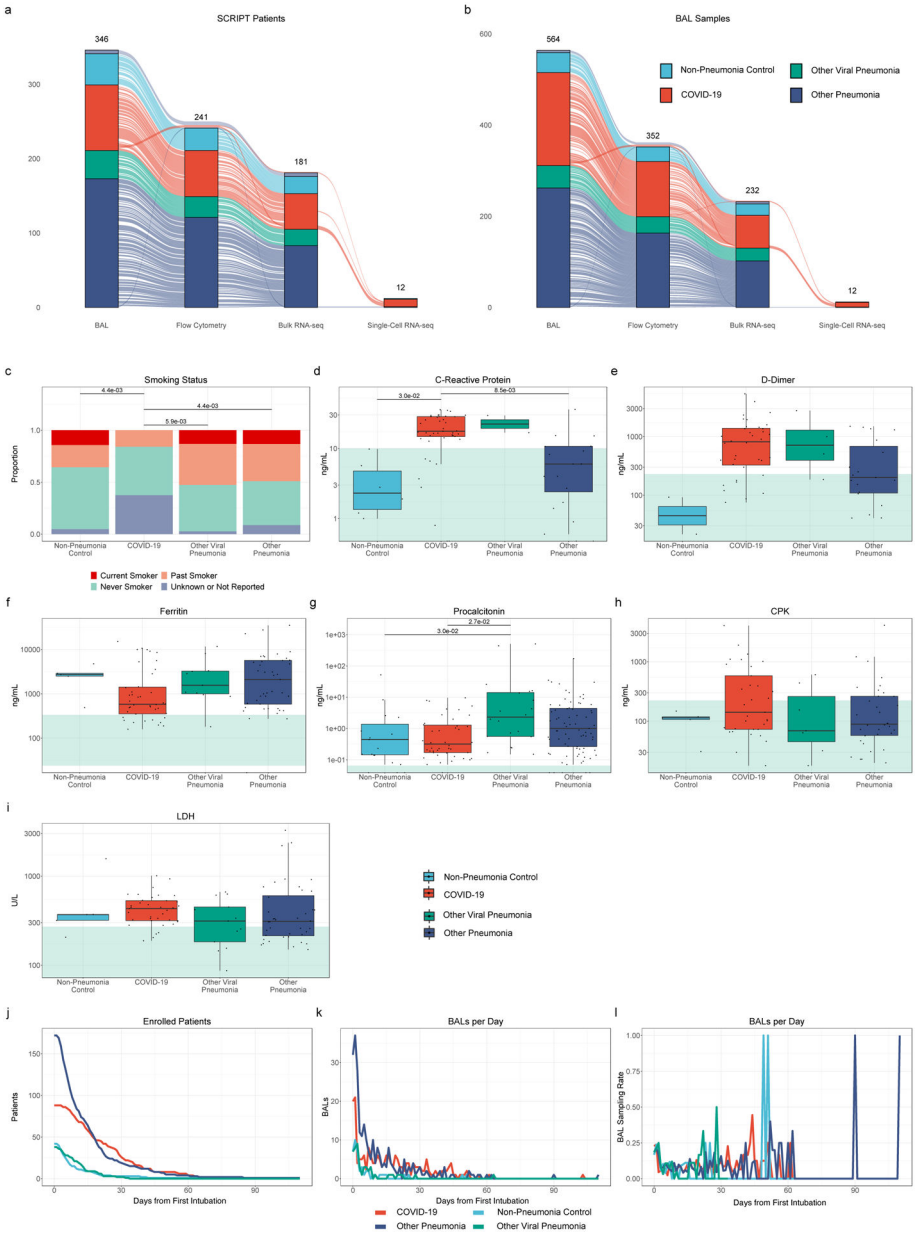
Plotting was performed in figures 1, 2, 3, extended data figures 1, 2, and 3 using ggplot2 version 3.3.1 unless otherwise noted. Comparisons for these figures were added using ggsignif version 0.6.0. Heatmaps on figures 2 and 3 were generated using pheatmap version 1.0.12. Sankey/Alluvial plots on Extended Data Fig. 1 were generated using ggalluvial version 0.12.0[72]. Figure layouts for Figures 1, 2, 3, 4, and Extended Data Figures 1, and 3 were generated using patchwork version 1.01 and edited in Adobe Illustrator 2021. Figure 4 and extended data figures 4 and 5 were generated with matplotlib v3.2.1[73]. For all boxplots shown, box limitsq represent the interquartile range (IQR) with a centerline at the median. Whiskers represent the largest point ≤ 1.5 * IQR. All points are overlaid.

Study limitations:

First, this is an observational study, and in the absence of a specific intervention targeting a necessary component of our model, our data are hypothesis-generating. Moreover, our observational cohorts are heterogeneous with respect to treatments received and other processes of care. Second, while we made every attempt to standardize BAL fluid volumes, the number of alveoli sampled and the return volume during both bronchoscopic and non-bronchoscopic BAL is variable. Hence, our observations are relative rather than quantitative. Third, our flow cytometry panels, while rigorous, were limited by the number of antibodies that could be used for high volume characterization of clinical samples. Fourth, our BAL samples were collected as part of clinical care; therefore, sicker patients were more likely to be sampled. Finally, while we made every effort to standardize care for patients with SARS-CoV-2 pneumonia in our ICU, some of our patients were enrolled in clinical trials of

remdesivir or sarilumab, many patients received unproven therapies off label including hydroxychloroquine and tocilizumab, and our study largely preceded reports on the use of steroids in patients with severe SARS-CoV-2 pneumonia (Extended Data Table 1). Despite these limitations, our systems approach to understand SARS-CoV-2 pathobiology provides a model with testable predictions that can serve as a template for the design of targeted interventions in patients with severe disease.

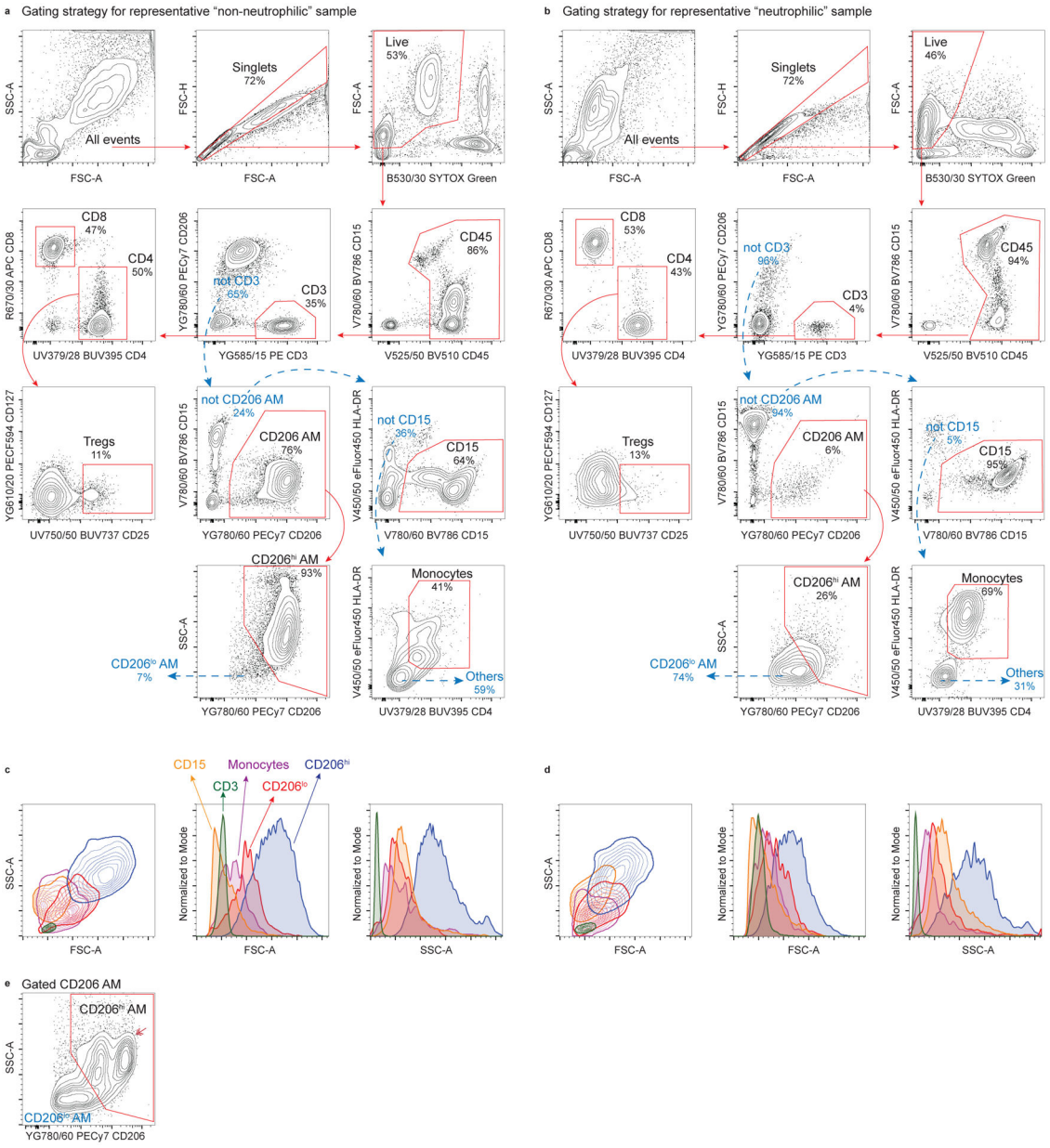
Extended Data



Extended Data Fig. 1: Overview of the study and biomarkers.

We compared BAL fluid obtained sequentially from 88 patients with severe SARS-CoV-2 pneumonia requiring mechanical ventilation with 38 patients with confirmed pneumonia

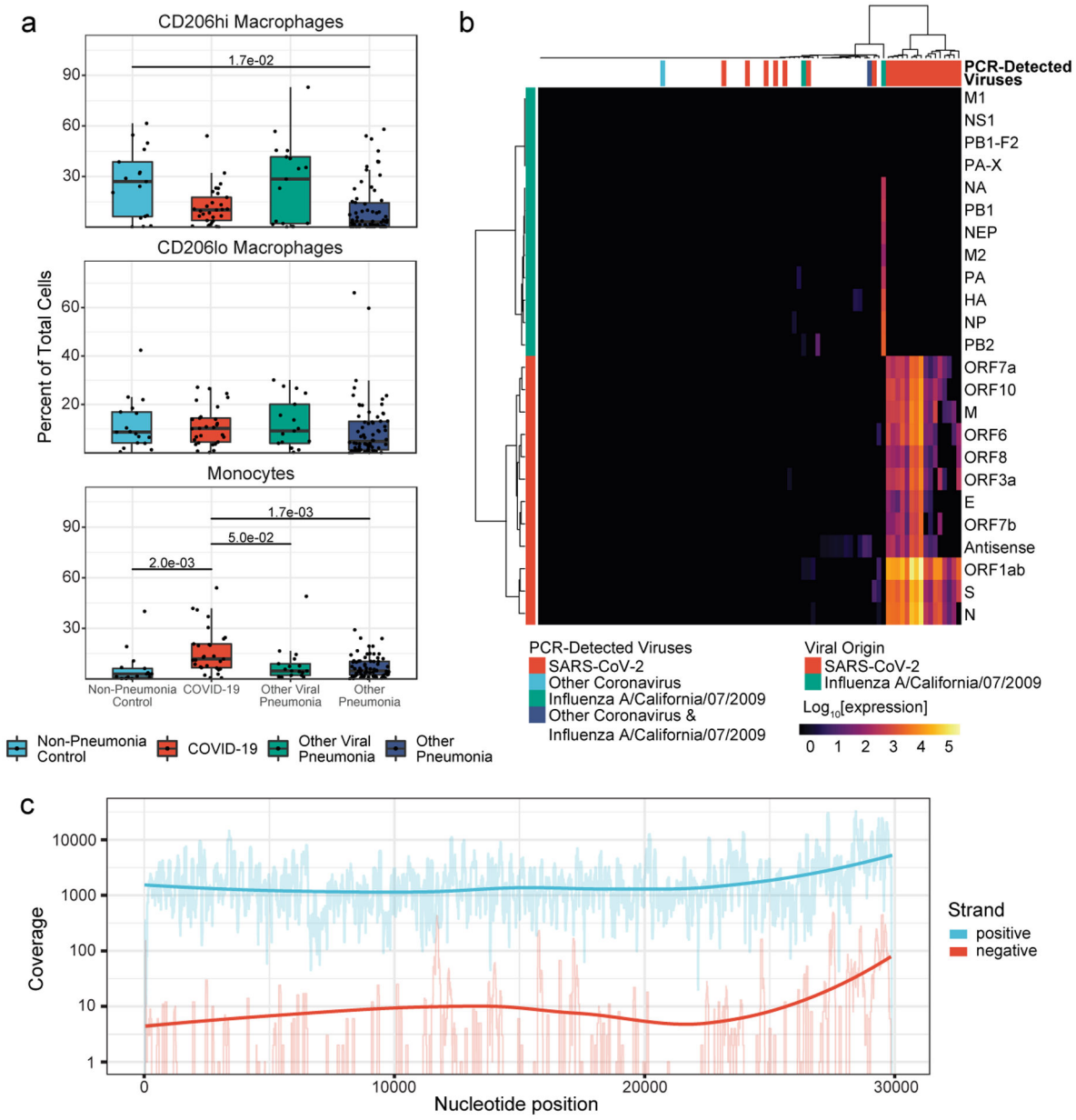
secondary to other respiratory viruses (other viral pneumonia), 173 patients with non-viral pneumonia (other pneumonia) and 42 mechanically ventilated patients without pneumonia undergoing BAL (non-pneumonia controls). **a**. Sankey diagram illustrating steps in analysis performed for at least one BAL sample for participants with COVID-19, other viral pneumonia, non-viral pneumonia (other pneumonia), non-pneumonia controls, and healthy controls. This includes samples from patients 1) enrolled in the SCRIPT study (346 patients), 2) analyzed via flow cytometry (241 patients), 3) for whom bulk RNA-seq was performed on flow cytometry-sorted alveolar macrophages (181 patients) and 4) for whom single-cell RNA-seq was performed on cells from BAL fluid (10 patients with SARS-CoV-2 pneumonia, 1 patient with bacterial pneumonia and 1 patient intubated for reasons other than pneumonia (gastrointestinal bleeding requiring endoscopy, a non-pneumonia control)). Some samples were cryopreserved and sorted post-cryorecovery. Because cryopreservation affects the number of neutrophils, these samples were not included in flow cytometric analysis but were used for bulk RNA-seq profiling of flow cytometry-sorted alveolar macrophages. Samples for which flow or bulk RNA-seq analysis was skipped are represented by alluvia flowing over the grouping bars. **b**. Sankey diagram illustrating steps in analysis performed for all BAL samples from participants with COVID-19, other viral pneumonia, non-viral pneumonia (other pneumonia) and non-pneumonia controls. This includes samples from patients 1) enrolled in the SCRIPT study (564 samples), 2) analyzed via flow cytometry (352 samples), 3) for whom bulk RNA-seq was performed on flow cytometry-sorted alveolar macrophages (232 samples) and 4) for whom single-cell RNA-seq was performed on cells from BAL fluid (12 samples). **c**. Self-reported smoking status. Significantly fewer active smokers were observed in the COVID-19 cohort as compared with all control groups ($q < 0.05$, pairwise Chi-square tests of proportions with continuity and FDR correction). **d–i.** Biomarkers: C-reactive protein (CRP), D-dimer, ferritin, procalcitonin, creatine phosphokinase (CPK), lactate dehydrogenase (LDH). The green-shaded area indicates the normal range. **j**. Number of patients remaining on mechanical ventilation. **k**. Number of BAL per day of mechanical ventilation. **l**. The BAL sampling rate per day among patients with COVID-19 was not higher than the sampling rate among patients with other pneumonias.



Extended Data Fig. 2: Representative gating strategy to identify immune cell subsets in BAL samples.

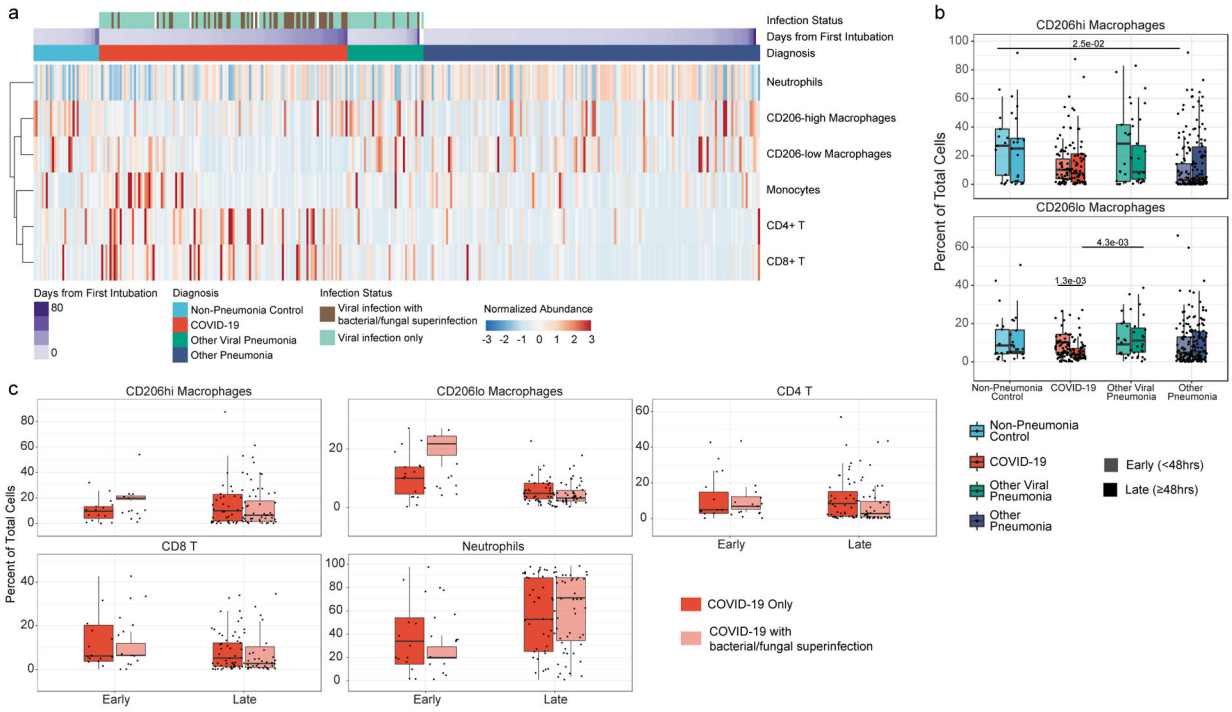
a. We developed a gating strategy that allowed us to quantify immune cell populations including monocytes, alveolar macrophage subsets, and T cell subsets. Importantly, we defined alveolar macrophages by their expression of CD206, subdividing them into early monocyte-derived alveolar macrophages ($CD206^{lo}$) and more mature ($CD206^{hi}$) alveolar macrophages. T cells were identified as CD3-positive and further subdivided into CD4+ and CD8+ T cells. Tregs were identified as CD3+CD4+CD25+CD127−. Neutrophils were identified as CD15+ cells. Monocytes were identified as HLA-DR+CD4+CD206− cells. Of note, only $CD206^{hi}$ alveolar macrophages were flow cytometry-sorted for bulk RNA-seq analysis (Figs. 2 and 3); hence, early monocyte-derived alveolar macrophages (MoAM1 and MoAM2 in our single-cell RNA-seq data, Fig. 4a–h) were not captured in bulk RNA-seq.

Representative sample from a patient without neutrophilia is shown. Solid red arrows indicate direct sequential gating, dashed blue arrows indicate Boolean “not” gates. Numbers on plots indicate the percentage of the parent population. Axis labels indicate laser line (UV – 355 nm, V – 405 nm, B – 488 nm, YG – 552 nm, R – 640 nm), bandpass filter, fluorochrome and antigen/dye. **b**. Representative sample from a patient with neutrophilia illustrates loss of $CD206^{hi}$ alveolar macrophages and influx of monocyte-derived $CD206^{lo}$ alveolar macrophages. **c–d**. Contour plot and histogram overlays illustrating forward (FSC) and side scatter (SSC) properties of the CD3+ T cells (CD3), CD15+ neutrophils (CD15), monocytes, $CD206^{lo}$ alveolar macrophages ($CD206^{lo}$), and $CD206^{hi}$ alveolar macrophages ($CD206^{hi}$) in the representative sample from a patient without neutrophilia (**c.**) and with neutrophilia (**d.**). Note that neutrophils have higher side scatter than monocytes. **e**. Representative contour plots illustrating a sample with two distinct populations of $CD206^{hi}$ alveolar macrophages. Single-cell RNA-seq analysis (Fig. 4a–h) suggests that $CD206^{hi}$ alveolar macrophages (double-head arrow) are *bona fide* tissue-resident alveolar macrophages.



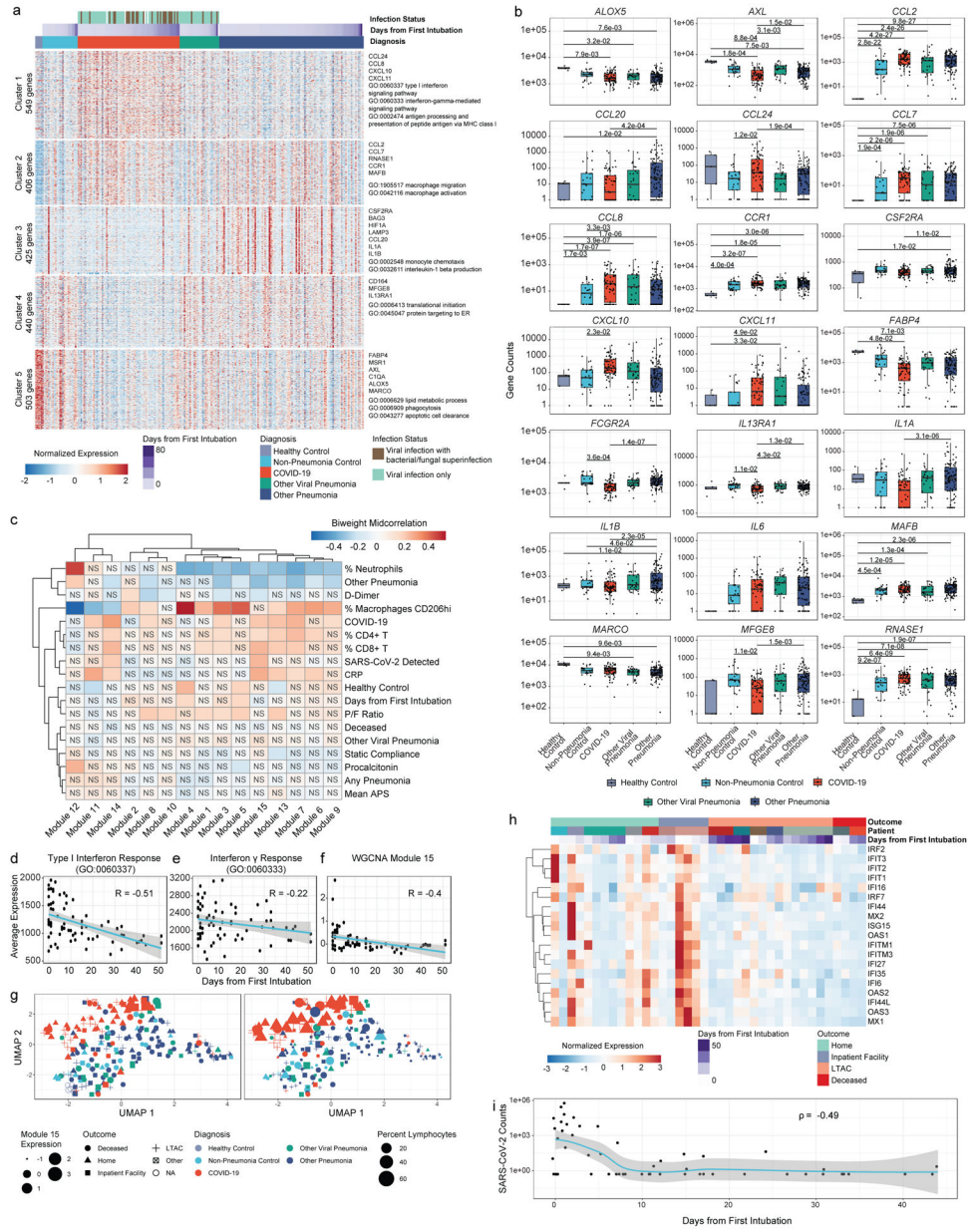
Extended Data Fig. 3. At the time of intubation, the alveolar space in patients with severe SARS-CoV-2 pneumonia is enriched for T cells and monocytes and contains alveolar macrophages harboring SARS-CoV-2 RNA.

a. Proportions of cells detected within 48 hours of intubation ($q < 0.05$, pairwise Wilcoxon rank-sum tests with FDR correction). Comparisons are not significant unless otherwise noted. **b**. Hierarchical clustering of viral reads for SARS-CoV-2 and influenza A/California/07/2009 virus using Ward's method. Log_{10}(DESeq2-normalized counts) are shown. **c**. Cumulative coverage plot of RNA-seq reads from flow cytometry-sorted alveolar macrophages aligned to the SARS-CoV-2 genome.



Extended Data Fig. 4. The BAL fluid from patients with SARS-CoV-2 pneumonia is persistently enriched for T cells irrespective of superinfection status.

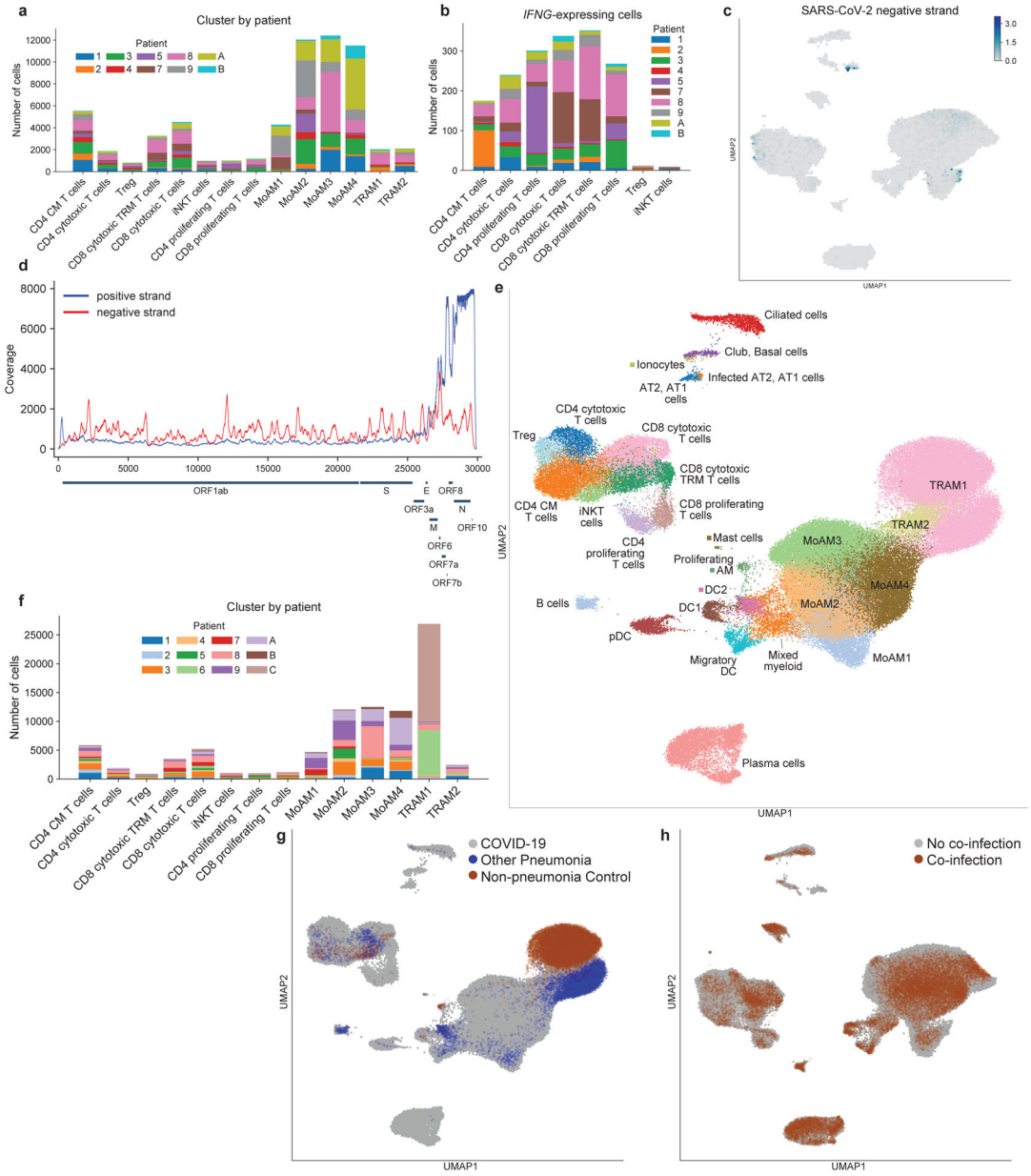
a. Heatmap of flow cytometry data demonstrating composition of BAL samples from all time points, grouped by diagnosis and ordered by the duration of mechanical ventilation. Column headers are color-coded by the diagnosis, duration of mechanical ventilation (white color indicates chronically ventilated patients), and presence or absence of superinfection ("Infection Status"). "Infection Status" refers only to the "COVID-19" and "Other Viral Pneumonia" groups; blanks in these groups refer to samples for which microbiology data were incomplete and infectious status could not be determined. "Viral infection only" refers to viral pathogens as the only detected pathogen in a sample and "Viral infection with bacterial/fungal superinfection" refers to detection of a viral pathogen with one or more bacterial or fungal co-pathogens. **b.** Comparison of percentage of $CD206^{lo}$ and $CD206^{hi}$ alveolar macrophages between early (<48 hours after intubation) and late (>48 hours of mechanical ventilation) samples ($q < 0.05$, pairwise Wilcoxon rank-sum tests with FDR correction). **c**. Comparison of CD4+ and CD8+ T cell and neutrophil abundance in the COVID-19 group, with and without superinfection in early and late sampling. Superinfection is represented by lighter bars. Differences between groups are not significant after FDR correction.



Extended Data Fig. 5. SARS-CoV-2 pneumonia is characterized by a persistent interferon-response signature in alveolar macrophages.

a. k-means clustering of the 2,323 significantly variable genes ($q < 0.05$, likelihood-ratio test) across diagnosis, columns represent each individual patient, grouped by diagnosis and ordered by day from first intubation. “Infection Status” refers only to the “COVID-19” and “Other Viral Pneumonia” groups; blanks in these groups refer to samples for which microbiology data were incomplete and infectious status could not be determined. “Viral infection only” refers to viral pathogens as the only detected pathogen in a sample and “Viral infection with bacterial/fungal superinfection” refers to detection of a viral pathogen with one or more bacterial or fungal co-pathogens. Representative genes and GO biological processes are shown for each cluster. Column headers are color-coded by diagnosis and duration of mechanical ventilation (white color indicates chronically ventilated patients). **b**.

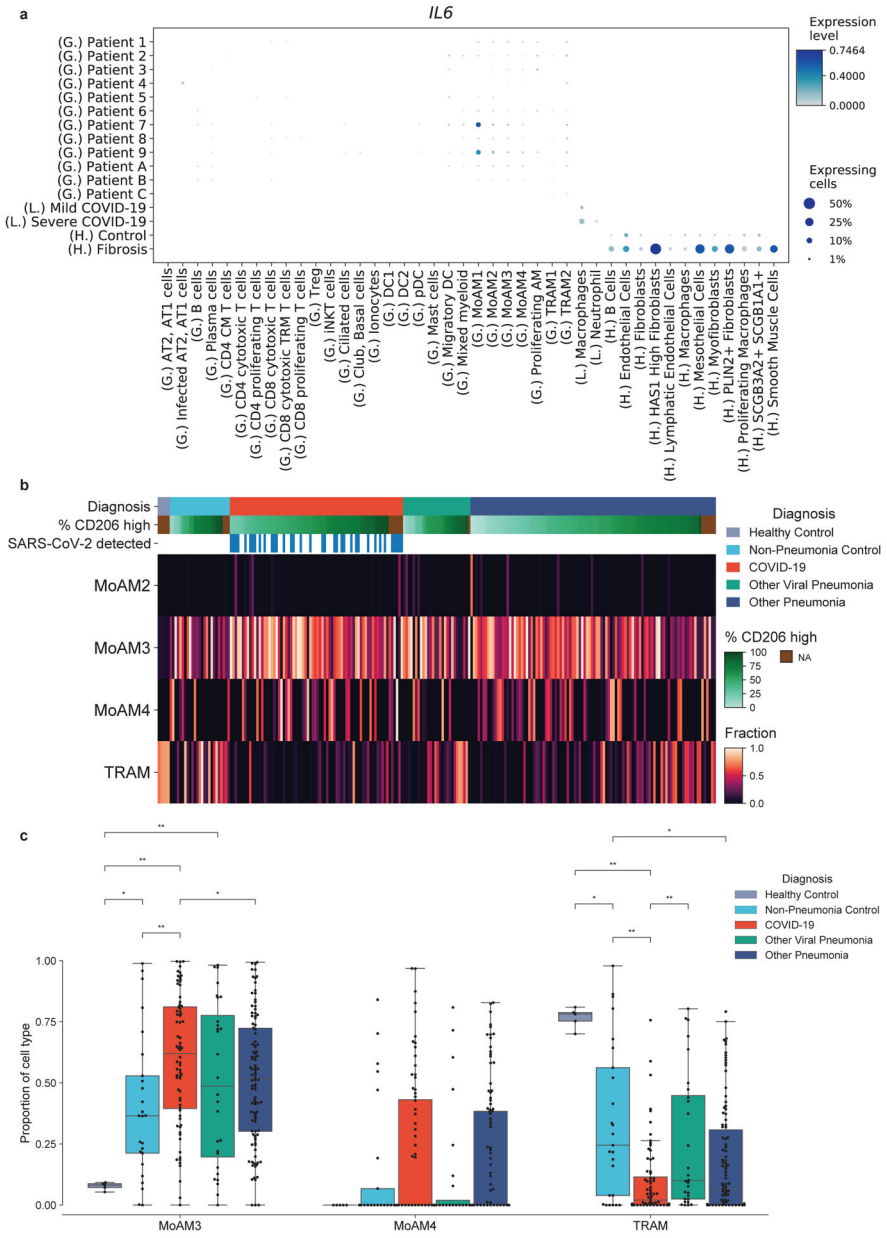
Expression of selected genes between the groups. Note that expression of *IL6* is not increased in any group. All significant comparisons are shown ($q < 0.05$, Wald test with FDR correction in DESeq2). **c**. Weighted gene co-expression network analysis (WGCNA). **d–f**. Interferon-response signatures in alveolar macrophages from patients with COVID-19 gradually decrease over the course of disease. Correlation between average expression of genes from GO:0060337 “type I interferon signaling pathway” ($R = -0.51$, $P = 5.7\times10^{-6}$, Pearson correlation; panel **d**.), GO:0060333 “interferon-gamma-mediated signaling pathway” ($R = -0.22$, $P = 0.06$, Pearson correlation; panel **e**.), module 15 of WGCNA ($R = -0.40$, $P = 5.5\times10^{-4}$, Pearson correlation; panel **f.**) and time on mechanical ventilation. Gray boundaries represent 95% confidence intervals. **g**. UMAP projections of all bulk RNA-seq samples. Average expression of WGCNA module 15 (left) and percent of CD3+ T cells in BAL (right) are shown by point area. **h**. Heatmap demonstrating time-dependent changes in gene expression of the canonical interferon-response genes from module 15 from patients with positive outcomes (discharged home or inpatient facility) or poor outcomes (discharged to a long-term acute care facility (LTAC) or deceased. **i**. Correlation between detection of SARS-CoV-2 reads and disease progression ($\rho = -0.49$, $p = 8.3\times10^{-4}$, Spearman correlation).



Extended Data Fig. 6. Single-cell RNA-seq identifies a positive feedback loop between IFNγ-producing T cells and SARS-CoV-2-infected alveolar macrophages.

a. Subsets of alveolar macrophages and T cells are represented by the cells from all 10 patients with COVID-19. **b**. *IFNG* expression is detected in T cells from all 10 patients with COVID-19, T cells with at least one count of *IFNG* were used for analysis. **c**. Detection of the SARS-CoV-2 negative strand. Density projection plot, with expression averaged within hexagonal areas on UMAP. **d**. Coverage plot of single-cell RNA-seq reads aligned to the SARS-CoV-2 genome. Cumulative data from 10 patients. Reads were aligned to genes on the positive strand or to the entire negative strand. **e**. UMAP plot showing integrative analysis of 105,715 cells isolated from 10 patients with severe COVID-19 within 48 hours after intubation (see Fig. 4a–h), one intubated patient with bacterial pneumonia and one intubated non-pneumonia control patient. AT1 – alveolar epithelial type 1 cells; AT2 –

alveolar epithelial type 2 cells; DC1 – conventional dendritic cells type 1, *CLEC9A*+; DC2 – conventional dendritic cells type 2, *CD1C*+; Migratory DC – migratory dendritic cells, *CCR7*+; pDC – plasmacytoid dendritic cells, *CLEC4C*+; Mixed myeloid – mixed cluster containing transitory dendritic cells and very immature monocyte-derived macrophages; TRAM – tissue-resident alveolar macrophages; MoAM – monocyte-derived alveolar macrophages; iNKT cells – invariant natural killer T cells; Treg – regulatory T cells, *FOXP3*+. **f**. Cells from non-pneumonia control (patient 6) and a patient with bacterial pneumonia (Patient C) primarily contribute to the TRAM1 cluster and have limited contribution to MoAM clusters. **g**. UMAP plot showing cells from non-pneumonia control (patient 6) and a patient with bacterial pneumonia (Patient C) from the integrative analysis on Extended Data Fig. 5e. h. Presence of co-infection does not affect clustering. UMAP plot showing cells from patients with and without co-infection from the integrative analysis in Fig. 4.



Extended Data Fig. 7. Deconvolution of bulk RNA-seq demonstrates loss of tissue-resident alveolar macrophages and persistence of mature monocyte-derived alveolar macrophages in patients with severe COVID-19.

a. Dot plot showing *IL6* expression across cell types. Dot size is proportional to the number of cells expressing *IL6* in the corresponding cluster. Data from this study (G.) are presented per patient, data from Liao et al. 2020 (L.) and Haberman et al. 2020 (H.) are averaged by condition. **b**. Heatmap demonstrating proportion of alveolar macrophage subsets predicted from deconvolution analysis. Data are grouped by condition and ordered by proportion of $CD206^{hi}$ alveolar macrophages. **c**. Proportion of alveolar macrophage cell types in patients with COVID-19 in comparison to other types of pneumonia obtained from deconvolution analysis (pairwise Wilcoxon rank-sum tests with FDR correction, * $q < 0.05$, ** $q < 0.01$, *** $q < 0.001$).

Extended Data Table 1:

Demographics of the SCRIPT cohort, grouped by diagnosis. All patients were admitted to Northwestern Memorial Hospital in Chicago between June 15, 2018 and July 6, 2020. Bronchoscopy was most commonly performed as part of routine clinical care to guide antimicrobial therapy. Patients with ARDS were managed using a high PEEP, low tidal volume ventilation strategy and were ventilated prone when suggested by ARDSnet guidelines. Patients received steroids and off label IL-6 receptor antagonists (tocilizumab or sarilumab) at the discretion of the clinical team. Any hydroxychloroquine use is reported, but hydroxychloroquine was routinely stopped upon ICU admission. Some patients were enrolled in multicenter placebo-controlled trials of remdesivir or sarilumab, treatment assignment is blinded.

	COVID-19 (N = 88)	Non-Pneumonia Control (N = 42)	Other Pneumonia (N = 173)	Other Viral Pneumonia (N = 38)
Ethnicity				
Hispanic or Latino	38/ 88 (43.2%)	6/ 42 (14.3%)	13/ 173 (7.5%)	8/ 38 (21.1%)
Not Hispanic or Latino	39/ 88 (44.3%)	33/ 42 (78.6%)	148/ 173 (85.5%)	27/ 38 (71.1%)
Unknown or Not Reported	11/ 88 (12.5%)	3/ 42 (7.1%)	12/ 173 (6.9%)	3/ 38 (7.9%)
Race				
American Indian/ Alaska Native	0/ 88 (0.0%)	0/ 42 (0.0%)	1/ 173 (0.6%)	1/ 38 (2.6%)
Asian	1/ 88 (1.1%)	2/ 42 (4.8%)	7/ 173 (4.0%)	1/ 38 (2.6%)
Black/ African American	20/ 88 (22.7%)	7/ 42 (16.7%)	38/ 173 (22.0%)	7/ 38 (18.4%)
Unknown or Not Reported	33/ 88 (37.5%)	5/ 42 (11.9%)	13/ 173 (7.5%)	4/ 38 (10.5%)
White	34/ 88 (38.6%)	28/ 42 (66.7%)	114/ 173 (65.9%)	25/ 38 (65.8%)
Age				
Minimum	21	19	22	34
Median (IQR)	59.50 (46.00, 68.25)	62.00 (46.75, 71.50)	65.00 (51.00, 73.00)	59.50 (52.25, 69.00)
Mean (SD)	57.09 ± 14.64	59.14 ± 17.89	61.44 ± 16.06	60.66 ± 13.64
Maximum	86	90	99	88
Sex				
Female	30/ 88 (34.1%)	19/ 42 (45.2%)	69/ 173 (39.9%)	17/ 38 (44.7%)
Male	58/ 88 (65.9%)	23/ 42 (54.8%)	104/ 173 (60.1%)	21/ 38 (55.3%)
Discharge Status				
Deceased	22/ 88 (25.0%)	15/ 42 (35.7%)	60/ 173 (34.7%)	14/ 38 (36.8%)
Home	38/ 88 (43.2%)	18/ 42 (42.9%)	37/ 173 (21.4%)	8/ 38 (21.1%)
Inpatient Facility	16/ 88 (18.2%)	7/ 42 (16.7%)	53/ 173 (30.6%)	14/ 38 (36.8%)
LTAC	11/ 88 (12.5%)	2/ 42 (4.8%)	22/ 173 (12.7%)	2/ 38 (5.3%)
Other	1/ 88 (1.1%)	0/ 42 (0.0%)	1/ 173 (0.6%)	0/ 38 (0.0%)
Smoking Status				
Current Smoker	0/ 88 (0.0%)	6/ 42 (14.3%)	23/ 173 (13.3%)	5/ 38 (13.2%)

	COVID-19 (N = 88)	Non-Pneumonia Control (N = 42)	Other Pneumonia (N = 173)	Other Viral Pneumonia (N = 38)
Never Smoker	41/ 88 (46.6%)	25/ 42 (59.5%)	73/ 173 (42.2%)	17/ 38 (44.7%)
Past Smoker	14/ 88 (15.9%)	9/ 42 (21.4%)	62/ 173 (35.8%)	15/ 38 (39.5%)
Unknown or Not Reported	33/ 88 (37.5%)	2/ 42 (4.8%)	15/ 173 (8.7%)	1/ 38 (2.6%)
Days of Intubation				
Minimum	1	0	0	1
Median (IQR)	18.00 (10.00, 33.00)	4.00 (2.00, 8.75)	7.00 (3.00, 13.00)	7.00 (3.00, 13.00)
Mean (SD)	25.48 ± 25.06	7.93 ± 11.48	10.62 ± 10.98	9.79 ± 8.91
Maximum	153	58	68	34
Unknown/ Missing	0 (0.00%)	0 (0.00%)	1 (0.58%)	0 (0.00%)
Length of ICU Stay				
Minimum	3	1	1	1
Median (IQR)	21.00 (12.75, 31.50)	5.50 (3.00, 11.50)	9.00 (5.00, 18.00)	10.50 (4.50, 13.75)
Mean (SD)	24.49 ± 16.96	10.17 ± 12.87	13.87 ± 13.41	11.39 ± 8.10
Maximum	85	68	84	33
Unknown/ Missing	0 (0.00%)	0 (0.00%)	2 (1.16%)	0 (0.00%)
BMI				
Minimum	20.62	16.56	12.46	17.34
Median (IQR)	31.21 (26.76, 37.09)	27.41 (24.59, 33.83)	25.89 (21.90, 32.32)	26.78 (23.19, 32.75)
Mean (SD)	33.17 ± 9.14	30.77 ± 9.78	27.48 ± 8.15	28.73 ± 8.37
Maximum	81.81	70.6	70.25	55.76
Unknown/ Missing	0 (0.00%)	1 (2.38%)	1 (0.58%)	0 (0.00%)
SOFA				
Minimum	1.83	1.75	1	2
Median (IQR)	7.02 (5.49, 9.33)	6.58 (4.87, 10.49)	6.54 (4.47, 9.66)	6.04 (4.03, 8.88)
Mean (SD)	7.44 ± 2.65	7.38 ± 3.31	7.35 ± 3.62	6.54 ± 3.02
Maximum	15.84	13.75	18.5	13.17
Unknown/ Missing	0 (0.00%)	0 (0.00%)	2 (1.16%)	0 (0.00%)
APS				
Minimum	33	28.25	21.2	26.33
Median (IQR)	70.27 (57.75, 82.11)	60.67 (46.31, 80.88)	64.22 (51.14, 80.55)	70.60 (51.20, 78.77)
Mean (SD)	69.54 ± 16.98	65.08 ± 21.89	66.85 ± 23.08	66.68 ± 19.89
Maximum	103.36	115.5	161	106.5
Unknown/ Missing	13 (14.77%)	1 (2.38%)	19 (10.98%)	1 (2.63%)
Remdesivir Study				
Enrolled	17/ 88 (19.3%)	0/ 42 (0.0%)	0/ 173 (0.0%)	0/ 38 (0.0%)
Not Enrolled	71/ 88 (80.7%)	42/ 42 (100.0%)	173/ 173 (100.0%)	38/ 38 (100.0%)
Sarilumab Study				
Enrolled	30/ 88 (34.1%)	0/ 42 (0.0%)	0/ 173 (0.0%)	0/ 38 (0.0%)

	COVID-19 (N = 88)	Non-Pneumonia Control (N = 42)	Other Pneumonia (N = 173)	Other Viral Pneumonia (N = 38)
Not Enrolled	58/ 88 (65.9%)	42/ 42 (100.0%)	173/ 173 (100.0%)	38/ 38 (100.0%)
Steroids				
Treated	30/ 88 (34.1%)	0/ 42 (0.0%)	0/ 173 (0.0%)	0/ 38 (0.0%)
Unreated	58/ 88 (65.9%)	42/ 42 (100.0%)	173/ 173 (100.0%)	38/ 38 (100.0%)
Hydroxychloroquine				
Treated	17/ 88 (19.3%)	0/ 42 (0.0%)	0/ 173 (0.0%)	0/ 38 (0.0%)
Unreated	71/ 88 (80.7%)	42/ 42 (100.0%)	173/ 173 (100.0%)	38/ 38 (100.0%)

Extended Data Table 2:

Pneumonia-causing pathogens detected in the SCRIPT cohort.

	Pathogen
1	Achromobacter species
2	Acinetobacter baumannii
3	Acinetobacter ursingii
4	Elizabethkingia meningoseptica
5	Enterobacter aerogenes
6	Enterobacter cloacae complex
7	Enterococcus faecalis
8	Enterococcus faecium
9	Escherichia coli
10	Klebsiella oxytoca
11	Klebsiella pneumoniae
12	Staphylococcus aureus
13	Proteus mirabilis
14	Providencia stuartii
15	Pseudomonas aeruginosa
16	Serratia marcescens
17	Stenotrophomonas maltophilia
18	Stomatococcus species
19	Streptococcus agalactiae (Group B)
20	Streptococcus pneumoniae
21	Streptococcus pseudopneumoniae
22	Viridans streptococcus

Extended Data Table 3:

Flow cytometry panels, reagents and instruments configuration.

Panel 1: Used in original SCRIPT study prior to March 13, 2020							
	Laser	Filter	Dye	Antigen	Clone	Final dilution	Catalog #/Identifier

1	355 nm	379/28	BUV395	CD4	RPA-T4	1:40	BD Biosciences Cat# 564724, RRID:AB_2738917
2		750/50	BUV737	CD25	2A3	1:20	BD Biosciences Cat# 564385, RRID:AB_2744342
3	405 nm	450/50	eFluor450	HLA-DR	L243	1:40	Thermo Fisher Scientific Cat# 48-9952-42, RRID:AB_1603291
4		780/60	BV786	CD15	HI98	1:20	BD Biosciences Cat# 563838, RRID:AB_2738444
5	488 nm	530/30	SYTOX Green	L/D	na	1:250	Thermo Fisher Scientific Cat# S34860
6	561 nm	585/15	PE	CD3	SK7	1:40	Thermo Fisher Scientific Cat# 12-0036-42, RRID:AB_10805512
7		610/20	PECF594	CD127	HIL-7R	1:20	BD Biosciences Cat# 562397, RRID:AB_11154212
8		780/60	PECy7	CD206	19.2	1:40	Thermo Fisher Scientific Cat# 25-2069-42, RRID:AB_2573426
9	640 nm	670/30	APC	CD8	SK1	1:20	BioLegend Cat# 344721, RRID:AB_2075390

Panel 2: Used on all samples starting March 13, 2020

	Laser	Filter	Dye	Antigen	Clone	Final dilution	Catalog #/Identifier
1	355 nm	379/28	BUV395	CD4	RPA-T4	1:40	BD Biosciences Cat# 564724, RRID:AB_2738917
2		750/50	BUV737	CD25	2A3	1:20	BD Biosciences Cat# 564385, RRID:AB_2744342
3	405 nm	450/50	eFluor450	HLA-DR	L243	1:40	Thermo Fisher Scientific Cat# 48-9952-42, RRID:AB_1603291
4		520/50	BV510	CD45	HI30	1:20	BioLegend Cat# 304036, RRID:AB_2561940
5		780/60	BV786	CD15	HI98	1:20	BD Biosciences Cat# 563838, RRID:AB_2738444
6	488 nm	530/30	SYTOX Green	L/D	na	1:250	Thermo Fisher Scientific Cat# S34860
7	561 nm	585/15	PE	CD3	SK7	1:40	Thermo Fisher Scientific Cat# 12-0036-42, RRID:AB_10805512
8		610/20	PECF594	CD127	HIL-7R	1:20	BD Biosciences Cat# 562397, RRID:AB_11154212
9		780/60	PECy7	CD206	19.2	1:40	Thermo Fisher Scientific Cat# 25-2069-42, RRID:AB_2573426
10	640 nm	670/30	APC	CD8	SK1	1:20	BioLegend Cat# 344721, RRID:AB_2075390
11		670/30	APC	EpCAM	9C4	1:40	BioLegend Cat# 324208, RRID:AB_756082

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments and funding:

Northwestern University Flow Cytometry Core Facility is supported by NCI Cancer Center Support Grant P30 CA060553 awarded to the Robert H. Lurie Comprehensive Cancer Center. Cell sorting was performed on BD FACSAria SORP cell sorter purchased through the support of NIH 1S10OD011996-01. Imaging work was performed at the Northwestern University Center for Advanced Microscopy supported by NCI Cancer Center Support Grant P30 CA060553 awarded to the Robert H Lurie Comprehensive Cancer Center. This research was supported in part through the computational resources and staff contributions provided by the Genomics Compute Cluster which is jointly supported by the Feinberg School of Medicine, the Center for Genetic Medicine, and Feinberg's Department of Biochemistry and Molecular Genetics, the Office of the Provost, the Office for Research, and Northwestern Information Technology. The Genomics Compute Cluster is part of Quest, Northwestern University's high-performance computing facility, with the purpose to advance research in genomics. We would like to thank the Office for Research Safety and Institutional Biosafety Committee at Northwestern University for their help and guidance in setting up protocols and workflows for safe processing of samples from patients with COVID-19. Single-cell RNA sequencing was performed with support from Simpsons Querrey Institute for Epigenetics.

Rogan Grant was supported by NIH grant T32AG020506-18. Luisa Morales-Nebreda was supported by T32HL076139 and F32HL151127. Ankit Bharat was supported by NIH grants HL145478, HL147290, and HL147575. Cara J. Gottardi was supported by NIH grants GM129312 and HL134800. GR Scott Budinger was supported by NIH grants U19AI135964, P01AG049665, P01AG04966506S1, R01HL147575 and Veterans Affairs grant I01CX001777. Alexander V. Misharin was supported by NIH grants U19AI135964, P01AG049665, R56HL135124, R01HL153312 and NUCATS COVID-19 Rapid Response Grant. Benjamin D. Singer was supported by NIH awards K08HL128867, U19AI135964, R01HL149883, and P01AG049665. Richard G. Wunderink was supported by NIH grant U19AI135964 and a GlaxoSmithKline Distinguished Scholar in Respiratory Health grant from the CHEST Foundation.

References

1. Yang X et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med 8, 475–481 (2020). [PubMed: 32105632]
2. Gattinoni L et al. COVID-19 Does Not Lead to a "Typical" Acute Respiratory Distress Syndrome. American Journal of Respiratory and Critical Care Medicine vol. 201 1299–1300 (2020). [PubMed: 32228035]
3. Baud D et al. Real estimates of mortality following COVID-19 infection. Lancet Infect. Dis 20, 773 (2020).
4. Zhou F et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 395, 1054–1062 (2020). [PubMed: 32171076]
5. RECOVERY Collaborative Group et al. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report. N. Engl. J. Med (2020) doi:10.1056/NEJMoa2021436.
6. National Heart, Lung, and Blood Institute PETAL Clinical Trials Network et al. Early Neuromuscular Blockade in the Acute Respiratory Distress Syndrome. N. Engl. J. Med 380, 1997–2008 (2019). [PubMed: 31112383]
7. Watanabe S, Alexander M, Misharin AV, Budinger GRS & Others. The role of macrophages in the resolution of inflammation. J. Clin. Invest 129, 2619–2628 (2019). [PubMed: 31107246]
8. Kim D et al. The Architecture of SARS-CoV-2 Transcriptome. Cell 181, 914–921.e10 (2020). [PubMed: 32330414]
9. Liao M et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat. Med 26, 842–844 (2020). [PubMed: 32398875]
10. Chua RL et al. COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. Nat. Biotechnol (2020) doi:10.1038/s41587-020-0602-4.
11. Sola I, Almazán F, Zúñiga S & Enjuanes L Continuous and Discontinuous RNA Synthesis in Coronaviruses. Annu Rev Virol 2, 265–288 (2015). [PubMed: 26958916]
12. Zhou J et al. Active replication of Middle East respiratory syndrome coronavirus and aberrant induction of inflammatory cytokines and chemokines in human macrophages: implications for pathogenesis. J. Infect. Dis 209, 1331–1342 (2014). [PubMed: 24065148]

13. Yip MS et al. Antibody-dependent infection of human macrophages by severe acute respiratory syndrome coronavirus. Virol. J 11, 82 (2014). [PubMed: 24885320]
14. Cheung CY et al. Cytokine responses in severe acute respiratory syndrome coronavirus-infected macrophages in vitro: possible relevance to pathogenesis. J. Virol 79, 7819–7826 (2005). [PubMed: 15919935]
15. Soberanes S et al. Metformin Targets Mitochondrial Electron Transport to Reduce Air-Pollution-Induced Thrombosis. Cell Metab. 29, 503 (2019). [PubMed: 30726761]
16. Price LC, McCabe C, Garfield B & Wort SJ Thrombosis and COVID-19 pneumonia: the clot thickens! The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology vol. 56 (2020).
17. Vabret N et al. Immunology of COVID-19: Current State of the Science. Immunity 52, 910–941 (2020). [PubMed: 32505227]
18. Wilk AJ et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat. Med (2020) doi:10.1038/s41591-020-0944-y.
19. Schulte-Schrepping J et al. Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. Cell 182, 1419–1440.e23 (2020). [PubMed: 32810438]
20. Habermann AC et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Science Advances 6, eaba1972 (2020). [PubMed: 32832598]
21. Califano D, Furuya Y & Metzger DW Effects of Influenza on Alveolar Macrophage Viability Are Dependent on Mouse Genetic Strain. J. Immunol 201, 134–144 (2018). [PubMed: 29760191]
22. Sungnak W et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat. Med 26, 681–687 (2020). [PubMed: 32327758]
23. Hou YJ et al. SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. Cell (2020) doi:10.1016/j.cell.2020.05.042.
24. Chu H et al. Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. Clin. Infect. Dis (2020) doi:10.1093/cid/ciaa410.
25. Blanco-Melo D et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. Cell 181, 1036–1045.e9 (2020). [PubMed: 32416070]
26. Konno Y et al. SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is Increased by a Naturally Occurring Elongation Variant. Cell Rep. 32, 108185 (2020). [PubMed: 32941788]
27. Hadjadj J et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science (2020) doi:10.1126/science.abc6027.
28. Jaume M et al. Anti-severe acute respiratory syndrome coronavirus spike antibodies trigger infection of human immune cells via a pH- and cysteine protease-independent FcγR pathway. J. Virol 85, 10582–10597 (2011). [PubMed: 21775467]
29. Martines RB et al. Pathology and Pathogenesis of SARS-CoV-2 Associated with Fatal Coronavirus Disease, United States. Emerg. Infect. Dis 26, 2005–2015 (2020). [PubMed: 32437316]
30. Sariol A & Perlman S Lessons for COVID-19 Immunity from Other Coronavirus Infections. Immunity 53, 248–263 (2020). [PubMed: 32717182]
31. Liu L et al. Anti-spike IgG causes severe acute lung injury by skewing macrophage responses during acute SARS-CoV infection. JCI Insight 4, (2019).
32. Guilliams M, Lambrecht BN & Hammad H Division of labor between lung dendritic cells and macrophages in the defense against pulmonary infections. Mucosal Immunol. 6, 464–473 (2013). [PubMed: 23549447]
33. Gioia C et al. T-Cell response profiling to biological threat agents including the SARS coronavirus. Int. J. Immunopathol. Pharmacol 18, 525–530 (2005). [PubMed: 16164833]
34. Grifoni A et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell 181, 1489–1501.e15 (2020). [PubMed: 32473127]
35. Le Bert N et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature (2020) doi:10.1038/s41586-020-2550-z.

36. Mateus J et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science (2020) doi:10.1126/science.abd3871.

37. Braun J et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. Nature 587, 270–274 (2020). [PubMed: 32726801]

38. Bacher P et al. Low avidity CD4+ T cell responses to SARS-CoV-2 in unexposed individuals and humans with severe COVID-19. Immunity 0, (2020).

39. Oran DP & Topol EJ Prevalence of Asymptomatic SARS-CoV-2 Infection: A Narrative Review. Ann. Intern. Med (2020) doi:10.7326/M20-3012.

40. Neupane AS et al. Patrolling Alveolar Macrophages Conceal Bacteria from the Immune System to Maintain Homeostasis. Cell 183, 110–125.e11 (2020). [PubMed: 32888431]

41. Louie JK et al. Factors associated with death or hospitalization due to pandemic 2009 influenza A(H1N1) infection in California. JAMA 302, 1896–1902 (2009). [PubMed: 19887665]

42. Leisman DE et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. Lancet Respir Med (2020) doi:10.1016/S2213-2600(20)30404-5.

43. Kox M, Waalders NJB, Kooistra EJ, Gerretsen J & Pickkers P Cytokine Levels in Critically Ill Patients With COVID-19 and Other Conditions. JAMA (2020) doi:10.1001/jama.2020.17052.

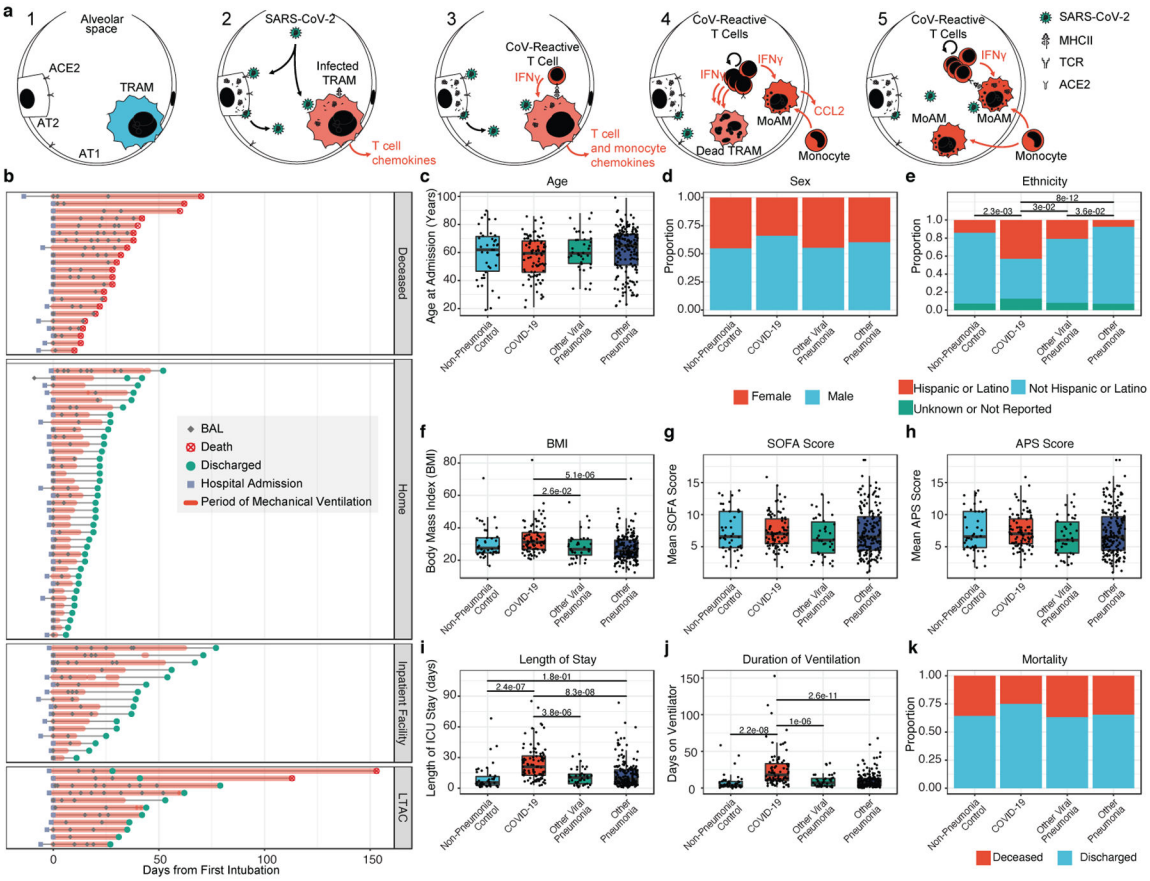


Fig. 1. Schematic and demographics of the SCRIPT cohort.

a. Schematic illustrating the interpretation of the main findings. 1. The normal alveolus contains *ACE2*-expressing alveolar type 1 and type 2 cells (AT1 and AT2, respectively) and tissue-resident alveolar macrophages (TRAM). 2. SARS-CoV-2 infects AT1 and AT2 cells and TRAM. Infected TRAM express T cell chemokines. 3. Cross-reactive or *de novo*-generated effector-memory T cells recognize SARS-CoV-2 antigens presented by TRAM and produce IFN**γ**, further activating TRAM to produce cytokines and chemokines. 4. Activated T cells proliferate and continue to produce IFN**γ**, eventually leading to death of infected TRAM and recruitment of monocytes, which rapidly differentiate into monocyte-derived alveolar macrophages (MoAM). 5. Recruited MoAM become infected with SARS-CoV-2, continuing to present antigens to T cells and maintaining the feedback loop until viral clearance is achieved. **b.** Timing of hospital admission (square), BAL fluid collection (diamonds), length of mechanical ventilation (bold red line), and hospital stay (thin grey line) in patients with severe COVID-19 grouped by outcomes (crossed open red circles – death; green circles – discharged). Day 0 is defined as the day of the first intubation. **c**. Distribution of patient age. Differences not significant by pairwise t-test with FDR correction. **d**. Proportions of females (red) and males (blue) (pairwise Chi-square tests of proportions with continuity and FDR correction). **e**. Self-reported ethnicity (pairwise Chi-square tests of proportions with continuity and FDR correction). **f**. Body mass index (BMI) (t-test with FDR correction). **g**. The Sequential Organ Failure Assessment (SOFA) score (pairwise Wilcoxon rank-sum tests with FDR correction). **h**. The Acute Physiology Score

(APS) (pairwise Wilcoxon rank-sum tests with FDR correction). **i**. ICU length of stay (pairwise t-tests with FDR correction). **j**. The duration of mechanical ventilation (pairwise t-tests with FDR correction). **k**. Mortality in patients with COVID-19 was similar to patients in other groups (25% vs. 35%, P = 0.10, $\chi^2 = 2.63$, Chi-square tests of proportions).

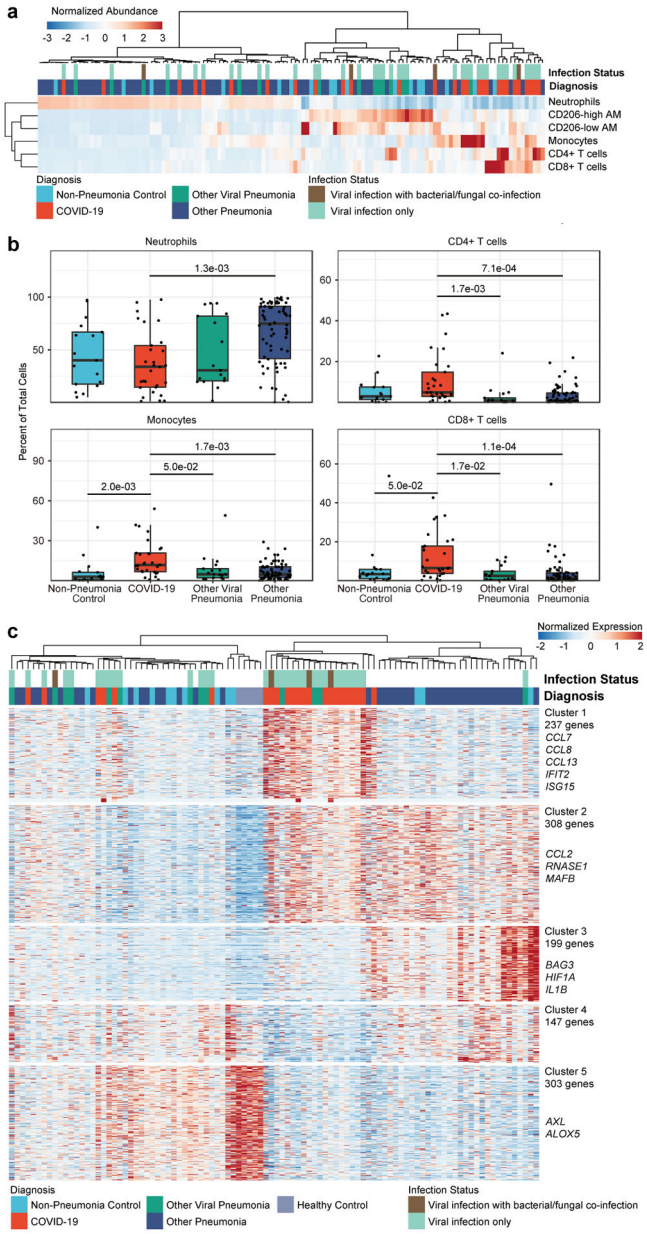


Fig. 2. At the time of intubation, the alveolar space in patients with severe SARS-CoV-2 pneumonia is enriched for T cells and monocytes and contains alveolar macrophages harboring SARS-CoV-2 RNA and expressing interferon response genes.

a. Hierarchical clustering of flow cytometry data from BAL samples collected within 48 hours of intubation. Column headers are color-coded by the diagnosis and presence or absence of co-infection (“Infection Status”). “Infection Status” refers only to the “COVID-19” and “Other Viral Pneumonia” groups. “Viral infection only” refers to viral pathogens as the only detected pathogen in a sample and “Viral infection with bacterial/fungal co-infection” refers to detection of a viral pathogen with one or more bacterial or fungal co-pathogens. Samples were clustered by Euclidean distance using Ward’s method. **b.** Proportions of cells detected within 48 hours of intubation ($q < 0.05$, pairwise Wilcoxon rank-sum tests with FDR correction). Comparisons are not significant unless otherwise

noted. **c**. k-means clustering of the 1,194 significantly variable genes ($q < 0.05$, likelihood-ratio test) across diagnosis. Columns represent each individual sample and are clustered using Ward's method. Column headers are color-coded by the diagnosis and presence or absence of a co-infecting pathogen ("Infection Status"). "Infection Status" refers only to the "COVID-19" and "Other Viral Pneumonia" groups; blanks in these groups refer to samples for which microbiology data were incomplete and infectious status could not be determined. "Viral infection only" refers to viral pathogens as the only detected pathogen in a sample and "Viral infection with bacterial/fungal co-infection" refers to detection of a viral pathogen with one or more bacterial or fungal co-pathogens. Representative genes and GO biological processes are shown for each cluster.

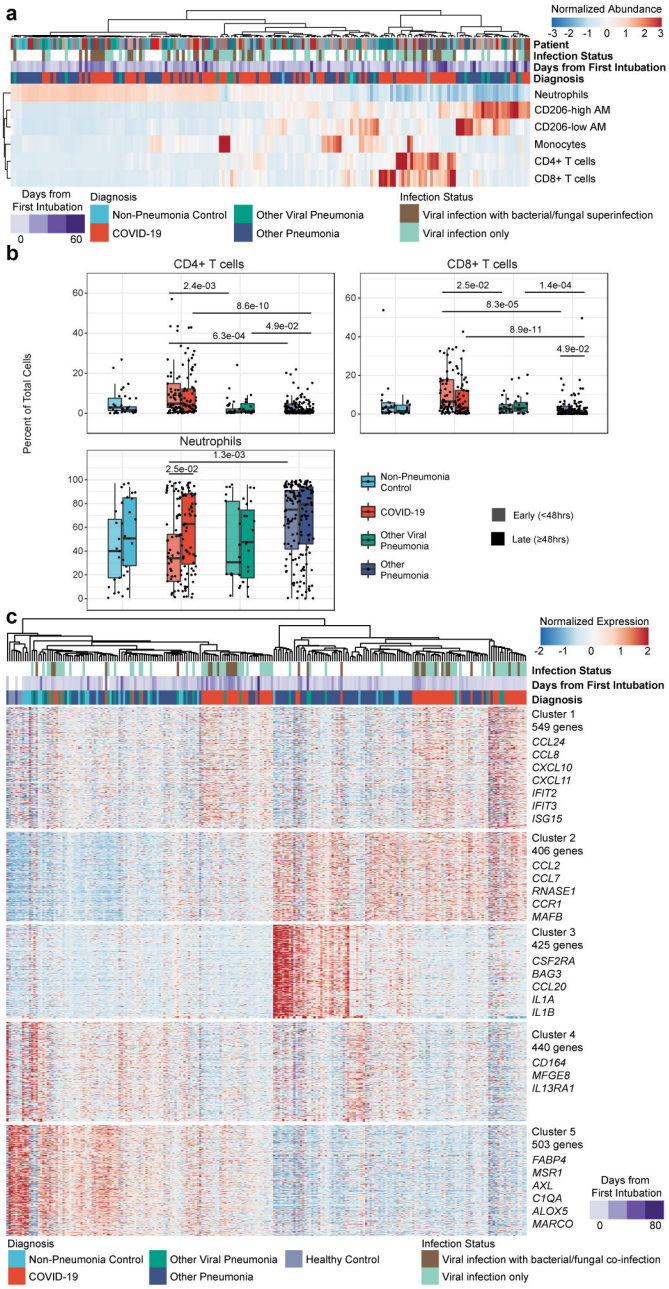


Fig. 3. The BAL fluid from patients with SARS-CoV-2 pneumonia is persistently enriched for T cells and characterized by an interferon-response signature in alveolar macrophages.

a. Hierarchical clustering of flow cytometry analysis of BAL samples from all time points from patients with serial sampling (n > 1) based on their composition. Column headers are color-coded by the diagnosis, patient, duration of mechanical ventilation, and presence or absence of superinfection (“Infection Status”). “Infection Status” refers only to the “COVID-19” and “Other Viral Pneumonia” groups; blanks in these groups refer to samples for which microbiology data were incomplete and infectious status could not be determined. “Viral infection only” refers to viral pathogens as the only detected pathogen in a sample and “Viral infection with bacterial/fungal superinfection” refers to detection of a viral

pathogen with one or more bacterial or fungal co-pathogens. Samples were clustered by Euclidean distance using Ward's method. **b.** Comparison of percentage of CD4+ and CD8+ T cells and neutrophils between early (<48 hours after intubation) and late (>48 hours of mechanical ventilation) samples ($q < 0.05$, pairwise Wilcoxon rank-sum tests with FDR correction). **c.** k-means clustering of the 2,323 significantly variable genes ($q < 0.05$, likelihood-ratio test) across diagnosis, columns represent each individual patient and clustered using Ward's method. Representative genes and GO biological processes are shown for each cluster. Column headers are color-coded by the diagnosis, duration of mechanical ventilation (white color indicates chronically ventilated patients), and presence or absence of superinfection ("Infection Status"). "Infection Status", "Viral infection only" and "Viral infection with bacterial/fungal superinfection" are as in panel **a.**

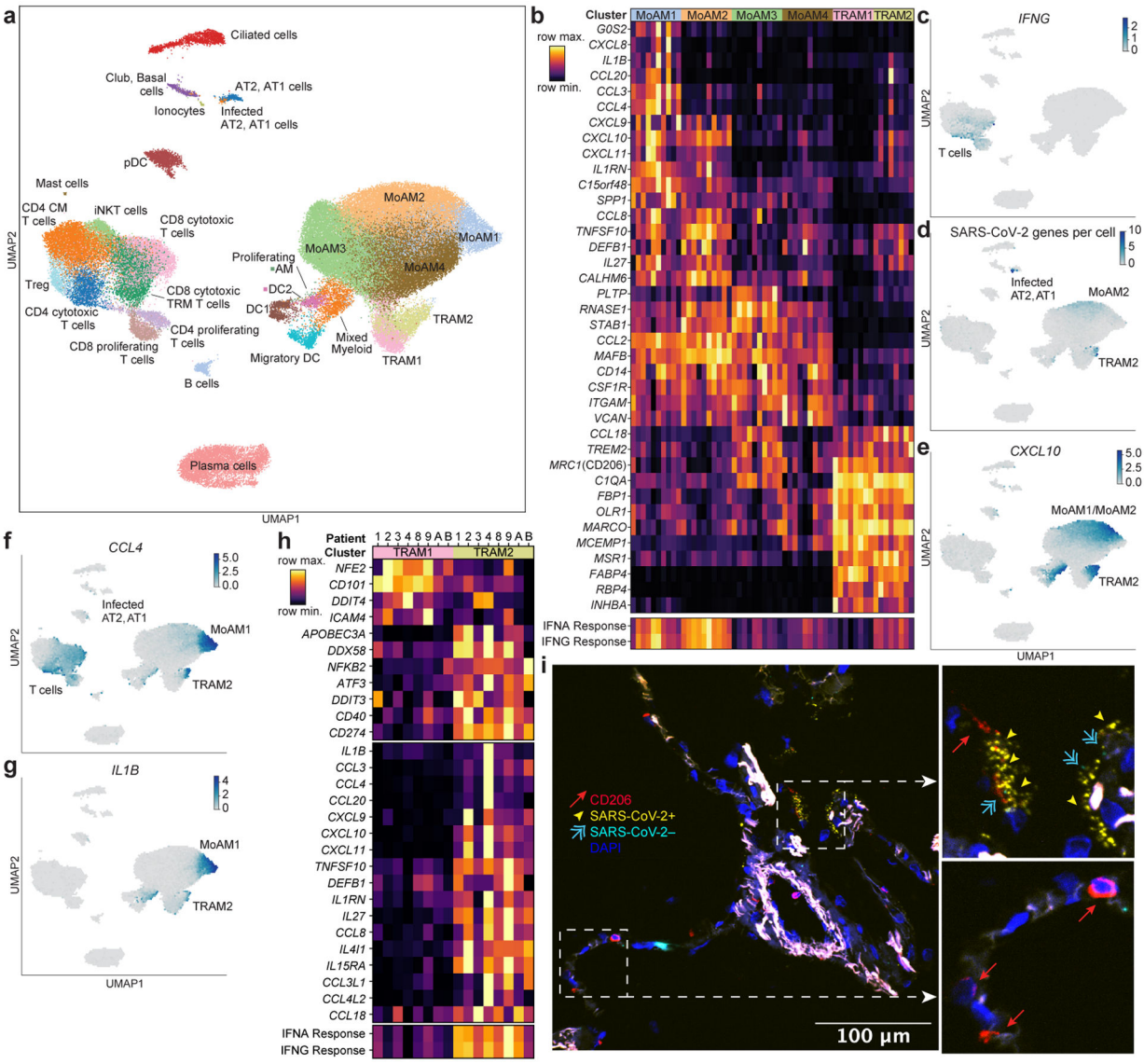


Fig. 4. Single-cell RNA-seq identifies a positive feedback loop between IFNγ-producing T cells and SARS-CoV-2-infected alveolar macrophages.

a. UMAP plot showing integrated analysis of 77,146 cells isolated from 10 patients with severe COVID-19 within 48 hours after intubation. AT1 – alveolar epithelial type 1 cells; AT2 – alveolar epithelial type 2 cells; DC1 – conventional dendritic cells type 1, *CLEC9A*+; DC2 – conventional dendritic cells type 2, *CD1C*+; Migratory DC – migratory dendritic cells, *CCR7*+; pDC – plasmacytoid dendritic cells, *CLEC4C*+; Mixed myeloid – mixed cluster containing transitory dendritic cells and very immature monocyte-derived macrophages; TRAM – tissue-resident alveolar macrophages; MoAM – monocyte-derived alveolar macrophages; iNKT cells – invariant natural killer T cells; Treg – regulatory T cells, *FOXP3*+. **b.** Heatmap demonstrating expression of the selected genes of interest in two subsets of tissue-resident (TRAM1 and TRAM2) and four subsets of monocyte-derived (MoAM1, MoAM2, MoAM3, MoAM4) alveolar macrophages. Each column represents a single patient with COVID-19. **c**. Expression of *IFNG* is restricted to T cells. **d**. Detection of

SARS-CoV-2 transcripts. Plot shows cumulative number of SARS-CoV-2 genes plus negative strand. **e-g**. Specific upregulation of selected cytokines and chemokines in TRAM2: *CXCL10* (**e.**), *CCL4* (**f.**), *IL1B* (**g.**). Density projection plots, expression averaged within hexagonal areas on UMAP. **h**. Heatmap demonstrating selected differentially expressed genes between two subsets of tissue-resident alveolar macrophages (TRAM1 and TRAM2). **i**. SARS-CoV-2 infection is spatially restricted. Combined immunofluorescence microscopy for CD206, a marker of mature macrophages (red arrows), and smFISH (RNAscope) for positive- (yellow arrowheads) and negative-strand (cyan doublehead arrows) SARS-CoV-2 transcripts. Inserts show infected and non-infected CD206-positive alveolar macrophages in the adjacent alveoli.