Spatial transcriptomic characterization of COVID-19 pneumonitis identifies immune pathways related to tissue injury

Cross AR[1], de Andrea CE[2], Landecho Acha MF[2,3], Cerundolo L[1], Etherington R[4], Denney L[4], Ho LP[4], Roberts I[5], Klenerman P[3], Hester J[1], Melero I[2], Sansom SN[6*], Issa F[1*]

1. Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
2. Department of Immunology and Immunotherapy, Clínica de la Universdad de Navarra, Pamplona, Spain
3. Department of Internal Medicine, Clinica de la Universdad de Navarra, Pamplona, Spain
4. Nuffield Department of Medicine, University of Oxford, Oxford, UK
5. Department of Cellular Pathology, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
6. Kennedy Institute of Rheumatology, University of Oxford, UK

*Joint senior authors. Corresponding author: fadi.issa@nds.ox.ac.uk

ABSTRACT

Severe lung damage in COVID-19 is known to involve complex interactions between diverse populations of immune and stromal cells. In this study, we applied a spatial transcriptomics approach to better delineate the cells, pathways and genes responsible for promoting and perpetuating severe tissue pathology in COVID-19 pneumonitis. Guided by tissue histology and immunohistochemistry we performed a targeted sampling of dozens of regions representing a spectrum of diffuse alveolar damage from the post-mortem lung of three COVID-19 patients. Application of a combination of differential gene expression, weighted gene correlation network, pathway and spatial deconvolution analysis stratified the sampled regions into five distinct groups according to degree of alveolar damage, levels of cytotoxic inflammation and innate activation, epithelial reorganization, and fibrosis. Integrative network analysis of the identified groups revealed the presence of proliferating CD8 T and NK cells in severely damaged areas along with signatures of cytotoxicity, interferon signalling and high expression of immune cell chemoattractants (including *CXCL9/10/11 and CCL2*). Areas of milder damage were marked by innate immune signalling (including *TLR response, IL-1, IL-6*) together with signatures of antigen presentation, and fibrosis. Based on these data we present a cellular model of tissue damage in terminal COVID-19 that confirms previous observations and highlights novel opportunities for therapeutic intervention.

INTRODUCTION

Mortality following severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is largely related to the antiviral response and immune-mediated lung injury[1]. Histopathologically, coronavirus disease 2019 (COVID-19) pneumonitis is associated with diffuse alveolar damage (DAD), fibrosis, leukocytic infiltrates, and microvascular thromboses[2-4]. Features of DAD include alveolar wall thickening, interstitial expansion, hyaline membrane deposition, and pneumocyte hyperplasia. Emerging studies have begun to describe the transcriptomic profiles of these histopathological features, although this has been with a specific emphasis on the analysis of foci of SARS-CoV-2 infection[5-7]. To the best of our knowledge, severe pathology is not consistently associated with active viral replication, and the absence or relatively low levels of SARS-CoV-2 RNA or antigen in lung tissue from prolonged cases of severe COVID-19[5] supports this model of inflammation-induced disease. The immune contributors and biological pathways associated with the severe alveolar injury therefore remain unclear. A greater understanding of the host response to COVID-19 within the lung and correlation to DAD would complement the increasing knowledge of both tissue and blood-based immune profiles[8].

Advanced spatial profiling techniques provide new tools to identify the distribution of proteins and RNAs *in situ* within tissue, allowing a spatial dissection of biological processes in and around specific histological features of interest[9,10]. However, published studies have yet to leverage the technique to map immune spatial heterogeneity to tissue pathology through network analysis. Here, we employ an advanced multiplexed *in situ* hybridization tissue analysis platform to provide detailed transcriptomic descriptions across multiple spatially discrete areas in lung samples from three COVID-19 patients. Uniquely, these tissues were obtained via open sampling at the point of death, which ensured high quality RNA analyses and avoided the caveats associated with late autopsies. We provide a comprehensive visualisation of the processes active within each sampled area, mapping gene modules to degree of damage. To further the analysis, we develop a network analytical approach which may be used as a template for other spatial genomic studies, and which facilitates the unbiased investigation of spatially divergent cellular identities, interactions, and pathways.

RESULTS

Immune cell infiltration is associated with severe local tissue damage in COVID-19

The histological and immune cell landscape within COVID-19 lung tissue from three patients with fatal disease was assessed ensuring coverage of the intra-tissue range of disease (mean specimen area 1.78cm^2; **Supplementary Table 1** for patient information). Each sample featured a spectrum of DAD comprising type II pneumocyte hyperplasia, hyaline membrane formation, and interstitial expansion. There was a non-uniform distribution of immune infiltrates (**Figure 1A-B and Supplementary Table 2**). Histological features beyond DAD included vacuolated macrophages, oedema, vascular thrombi, and squamous metaplasia. Bulk detection for SARS-CoV-2 nucleocapsid RNA in patient lung samples found low or absent viral loads (**Supplementary Table 1**) and SARS-CoV-2 Spike RNA could not be detected with *in situ* hybridisation although weak and sparse foci of intracellular SARS-CoV-2 viral protein were observed within morphologically-defined alveolar macrophages (*data not shown*). Immunofluorescent staining for CD3, CD68 and pan-cytokeratin (panCK) enabled the identification of lymphocytes, myeloid cells, epithelial cells, and general tissue architecture (**Figure 1B**). Based on this staining we identified areas of interest (AOIs) that represented the

spectrum of DAD severity and immune composition (**Figure 1B and 1C**). These were annotated for histopathological DAD severity by two independent pathologists, as well as by cell identity using an automated nucleated cell count analysis for CD3 and CD68 (**Figure 1C**). To profile the pathological pathways active in each AOI, we applied a multiplexed *in situ* hybridisation approach to quantitate the expression of a curated panel of >1800 genes that was enriched in immune targets and augmented with a COVID-19-specific gene set (**Supplementary File 1**). Differential expression analysis comparing mild vs. severe DAD across all sampled AOIs (16 mild and 29 severe) identified n=59 genes with significantly higher expression in severe pathology (Benjamini Hochberg [BH] adjusted p<0.05, |fold change| >1.5) including those encoding chemokines (*CXCL9, CXCL11, and CCL5*), cytotoxic molecules (*GZMA, GZMB, GZMK, PRF1, LYZ, NKG7*), complement factors (*C1QA, C3, C2, C1QB*), and proteins involved in antigen processing and presentation (*CD74, HLA genes, CTSS*) (**Figure 1D**). Genes upregulated in the severe areas showed an a significant over-representation (BH adjusted p<0.05) gene ontology (GO) biological process terms related to T cell activation and differentiation, antigen presentation, cytokine production, cytotoxicity, and response to interferon gamma (**Figure 1E**). By contrast, genes that showed higher expressed in areas of mild damage (n=40, BH adjusted p<0.05, |fold change| >1.5) included those with known roles in immune-mediated repair (*IL1RL1, AREG, CXCL5*) together with a significant over-representation (BH adjusted p<0.05) of pathways associated with wound healing, epithelial re-organization and adhesion, as well as extracellular matrix deposition (**Figure 1D and 1E**).

Network analysis implicates CD8 T cells, mononuclear phagocytes and the TLR, interferon, and IL-1 signaling pathways in lung inflammation in COVID-19

In order to perform an unbiased exploration of the cellular and phenotypic variation present in the set of the profiled AOIs we employed the weighted gene correlation network analysis (WGCNA) method[11]. This analysis identified 17 distinct modules of co-expressed genes (n=27–266 genes per module, median 88) (**Figure 2A**). We began a systematic characterization of these identified gene modules by identifying biological processes and pathways over-represented amongst the gene members of each module (**Figure 2B**). Next we assessed the association of the modules with specific cell types by correlating expression of the module eigengenes (the representative module expression patterns) with separate cell type abundance estimates that were determined for each AOI by automatic cell type deconvolution analysis[12] (**Figure 2C and Supplementary Figure 2**). Finally, we investigated correlation of the module eigengenes with the expression of known cell type marker genes (**Figure 2D**) and with genes involved in immune signaling, the complement system, viral infection and fibrosis (**Figure 2E**). Based on these analyses we named each module according to its cell type associations and biological pathway enrichment.

As expected for lung tissue, we found a set of stromal modules representing (i) "Epithelial cells" (Blue, containing *EPCAM*), (ii) "Type 2 pneumocytes" (Red, containing the surfactant encoding genes *SFTPB, SFTPC and SFTPD*), (iii) fibroblasts ("Fibroblast phenotype", Black, containing *COL1A1, COL3A2, COL5A1* and *THY1*) and (iv) "Vasculature" (Yellow, containing *CDH5, THBD, ENG*), which all showed corresponding pathway and cell type associations (**Figures 2B-D**). Immunofluorescent imaging established the presence of CD68+ macrophages within sampled areas (**Figure 1B** and **1C**) and we found three modules that showed a high correlation with *CD68* expression (**Figure 2D**). These comprised (i) an "Alveolar macrophage" module (Cyan) that also displayed high expression correlation with the macrophage receptor *MARCO* and enrichment for the "phagocytosis, engulfment" GO biological process, (ii) a "Macrophage identity" module (Grey60) that also showed high expression correlation with

the mannose receptor *MRC1* (a marker of alternative "M2" macrophages) and *SIRPA*, and (iii) an "Antigen presentation" module (Greenyellow) that showed a high correlation with predicted *MARCO⁻* macrophage abundance along with high expression correlation with *SIGLEC1* and *C1QA/B* as might be expected for interstitial macrophages. We also noted a "IFITM2/HSP/ECM" module (Pink) that contained *MERTK* and *PDGFRA* and a module with pathway enrichments for "Apelin/mTOR signalling" (Purple).

Severe COVID-19 is associated with massive lung immune cell infiltration, and in line with this we discovered modules of genes with clear associations with lymphocytes and myeloid cells. The "Cytotoxicity and T cells" module (Turquoise) was associated with the GO "Interferon-gamma production" biological process, as well as the KEGG "T cell receptor signalling" and "Natural Killer mediated cytotoxicity" pathways (**Figure 2B**). The expression of this module was positively associated with predicted presence of CD8⁺ cytotoxic T cells, NK cells and activated dendritic cells (DCs) (**Figure 2C**). This module contained known T cell markers including *CD3D/E*, *CD2* and *CD8A* as well as genes associated with cytotoxicity such as *PRF1* and *GNLY* (**Figure 2D**). It also showed a high correlation with the expression of specific chemokines including *CXCL9*, *CXCL10* and *CXCL11* and had the highest correlation with the expression of *IFNG* (**Figure 2E**). The "TLR signaling and monocytes" module (Midnightblue) contained the gene encoding the classical monocyte marker *CD14* along with *CD163* (**Figure 2D**). In addition, we discovered several modules that contained genes associated with active inflammation including an "Interferon responses" module (Brown), a "TLR and IL-1 signaling" module (Salmon) and an "IL-1 response: IL-6/IL-8" module (Lightcyan) that were named according to their pathway enrichments (**Figure 2B**) and correlations with chemokine/cytokine expression (**Figure 2E**). We found three sets of genes associated with more general cellular processes including "Cell cycling" (Lightgreen), "Chromatin remodelling" (Lightyellow) and "Hypoxic response" (Green) which are indicative of local immune cell proliferation and oxygen stress.

Severe alveolar damage in COVID-19 is linked with myeloid cell antigen presentation, T cell cytotoxicity and expression of the CXCL9/10/11 interferon response genes

Within the AOIs from the each of three patients we found that severe DAD was consistently associated with the expression of the "Interferon responses", "Cytotoxicity and T cells", "Chromatin remodelling", "Antigen presentation", and TLR signaling ("TLR and IL-1 signalling" or "TLR signaling and monocytes") modules. The "IL-1 response: IL-6/IL-8" module did not show an obvious correlation with DAD severity while higher expression of gene modules associated with "Epithelial cells" and "Vasculature" was associated with the lowest DAD scores (**Figure 3A).** We next sought to better understand the variation in cellular and immune phenotype that was present in the sampled AOIs. To do so we hierarchically clustered the AOIs by expression of the WGCNA module eigengenes. This analysis revealed 5 groups of AOIs with different transcriptional signatures and associations with severity (**Figure 3B-C**). We numbered these spatial groups of AOI according to their association with severe damage: spatial group 1 contained the lowest proportion of severe AOIs while spatial group 5 was comprised exclusively of AOIs from regions with severe DAD. The breakdown of the identified AOI groups by severity, mean presence of CD3+ cells and by patient of origin is shown in Figure **4A**.

Spatial group 1 included AOIs from patients B and C that were marked with high expression of the stromal eigengenes (with expression of the vascular genes *ENG, PECAM1, CDH5*; the surfactant genes *SFTPA1, STFPB*, and *SFTBC*; and the epithelial genes *KRT19, EPCAM* and

CDH1), as well as the "IL-1 response: IL-6/IL-8" eigengene, and that were mostly of mild severity. The AOIs in spatial group 2 all came from patient A, showed expression of macrophage and "Hypoxic response" eigengenes in addition to that of many of the stromal signatures in spatial group 1, but also a number of macrophage and antigen-presenting genes (*HLA-DR* genes, *MARCO, MRC1* and *CD68*), and were associated with both mild and severe DAD. The remaining three spatial groups (3, 4 and 5) lacked enrichment of the epithelial eigengenes, shared higher expression of the "Interferon responses" module and were predominantly composed of AOIs with severe DAD and higher fractions of CD3+ T cells. AOIs from spatial groups 3 and 5 showed higher expression of the "Cytotoxicity and T cells" and "cell cycling" eigengenes, while the AOIs from spatial group 4 (which were all from patient B) showed higher expression of the "TLR signalling and monocytes" eigengene. Spatial groups 3 and 5 shared expression of NK, CD8, and cytotoxicity genes (e.g. *NKG7, CD8A, GNLY, GZMK*, and *PRF1*), yet differed in their expression of interferon genes (spatial group 3) vs. macrophage genes (spatial group 5). Spatial group 4 however had a strong association with both interferon and TLR signalling genes (e.g. *IFNA1, IRF1, TLR2, TLR4, TLR5, TLR6*), but also had a weak overlap with both cytotoxicity and macrophage identity genes. Similar to spatial group 2, the AOIs in spatial group 5 also showed higher expression of the "Hypoxic response" eigengene and were also mostly from patient A. Importantly, despite variability in immune signatures across AOIs, there was a consistency in lymphocyte characteristics across all patients. AOIs with a lymphocytic infiltrate demonstrated a consistent phenotype enriched principally with the "Cytotoxicity and T cells" gene module, the expression of which showed little spatial heterogeneity across AOIs of similar cellular makeup.

Finally, we constructed correlation networks to investigate the associations between the expression of the WGCNA eigengenes, genes encoding immune signaling factors, and the predicted cell type abundances in each of the 5 spatial groups. Inspection of the networks from the three spatial groups associated with the highest levels of damage (spatial groups 3,4,5; **Figure 4A**) identified the following correlates: (i) expression of the WGCNA modules "Cytotoxicity and T cells", "Antigen presentation", "Interferon response", "Cell cycling" and "Chromatin remodelling", (ii) predicted abundances of "MARCO- macrophages", "Activated DC", "NK", "CD8 T cells", "Dividing T cells", "Dividing macrophages" and (iii) expression of *CCL2/3/4/5/13/18/19, CXCL9/10/11, FASLG, IFNA1, IFNA17, IL16/18/32, LTB, TNF, TNFSF10, TNFSF13B and XCL2* as core features of severe DAD in all three patients. The "TLR and IL-1 signalling" module eigengene was implicated in the spatial group 3 and 4 networks from patients B and C while the "IL-1 response: IL-6/IL-8" module was observed in the spatial group 1 and 2 networks from patients A, B and C suggesting that IL-1 signaling is centrally involved in the COVID-19-induced lung tissue damage. In agreement with the computational predictions, immunofluorescent confocal microscopy analysis confirmed the presence of a large number of T cells in the AOIs with the highest levels of tissue damage (see representative inset images, **Figure 4B-F, Figure 3B, and Figure 4A**). Unexpectedly, the expression of IL-6, IL-6 receptors, viral receptors/co-factors, genes involved in fibrosis, and complement regulatory genes were largely associated with alveolar epithelium, endothelium, and alveolar macrophages (**Figure 4G**). These areas, although associated with innate IL-1 signalling, correlated with the less inflamed spatial groups (1 and 2, **Figure 4B-C**).

In summary these analyses suggest that severe tissue damage in COVID-19 involves an ensemble of interacting and proliferating immune cells in which myeloid cells such as MARCO-macrophages and DCs are activated by TLR-mediated signaling, express IL-1 and interferon alpha, and present antigen to cytotoxic lymphocytes driving the production of *IFNG* and a specific set of chemokines and cytokines. This includes high expression of *CXCL9, CXCL10* and *CXCL11*, factors which are known to act via CXCR3 to promote immune cell chemotaxis,

extravasation and activation[13], as well as IL-32 which is known to stimulate TNFa and IL-6 secretion from macrophages[14], and CCL19 which acts via CCR7 to promote DC and central memory T cell migration[15].

DISCUSSION

Analytically, the correlation network-based analysis approach presented here provides a powerful means for the interpretation of spatial transcriptomic data in a novel and evolving field. Differential gene expression analysis between histological categories has been the most commonly employed approach in the analysis of spatial datasets but this method can reduce complex heterogeneity into binary categories and limit the potential for novel insight into spatial variation. In this study, we therefore integrated WGCNA and cell type deconvolution to perform an unbiased analysis of spatial variation in gene expression and cell type abundance. Together with pathological and immunohistological cell count analysis this approach permitted elucidation of the biological circuits active in the lung tissue of patients with critical COVID-19, demonstrating the potential of spatial genomics and correlation-based approaches to help discover the cellular basis of disease pathology.

The early innate response to COVID-19 is dominated by antiviral cytokine secretion, principally type I/III interferons, that have been reported to inhibit SARS-CoV-2 replication in vitro[16-18], and which mediate downstream protective effects via IFN-stimulated gene (ISG) products[19]. While impaired peripheral blood IFN responses have been reported[20], it is not clear if this also reflects local responses, particularly since IFN genes have been detected in bronchio-alveolar lavage (BAL) fluid[21] and lung tissue[5]. More recent work has identified an upregulation of type I interferons in peripheral blood mononuclear cells from patients with severe COVID-19 which is not seen in other viral pneumonias such as severe influenza[22]. We found IFN responses in the lung to be largely associated with DAD severity, although there was variability in associated modules and genes even within the most severe spatial AOI groups. A reduced expression of IFN, TLR and IL-1 signalling modules was accompanied with a loss in chemokine expression. Specifically, there was spatial variability in the association with TLR signalling, specific chemokines (e.g. *CXCL9*), and T cell cytotoxicity, with the more moderate spatial groups associating with TLR/IL-1 signalling and the most severe groups displaying an IFN signalling signature without obvious TLR/IL-1 response. This may represent a switch from TLR dependent type I (*IFNA1*/*IFNA17*) IFN production by myeloid cells in the less severe AOIs to secondary T and NK cell driven type III (*IFNG*) responses in areas of severe tissue damage[23].

Myeloid cell dysregulation is a key finding of severe or progressive COVID-19 infection[24-27]. A single cell RNA sequencing (scRNAseq) study of BAL fluid from severe COVID-19 cases identified an increase in the ratio of CD14+HLA-DR^{lo} inflammatory monocytes to tissue-resident alveolar macrophages[28]. This myeloid-neutrophil crosstalk may be associated with neutrophil extracellular trap formation (NETosis) and ongoing inflammation[29]. However, we did not detect a strong neutrophil signature in our tissue samples, which is in keeping with reports of low neutrophil abundance in the lung in late stage disease[6] but may also relate to our sampling strategy, probe panel or to technical problems associated with the detection of low abundance neutrophil mRNA or its degradation by neutrophil RNAse[30]. Our spatial analysis could separate alveolar from infiltrating monocytes, and while both populations expressed antigen presenting genes, only interstitial monocytes associated with lymphocytes and were correlated with severe damage.

IL-6 is a key inflammatory cytokine in viral infection and is produced by a wide range of cell types including macrophages, fibroblasts, and endothelial cells[31]. Its importance in COVID-19 stems from its association with disease severity[32-34] and the success of tocilizumab (anti-IL-6R) therapy[35]. Unexpectedly, our data revealed IL-6 response genes to most prominently associate with the less damaged areas of spatial groups 3 and 4 that had the highest abundance of macrophages, fibroblasts, and vascular cells. It is possible that IL-6-targeting therapies act to protect areas which have not yet progressed to severe irreversible damage. This may in turn explain the divergent innate signatures seen in Patient A who received tocilizumab, where there was a weaker correlation between fibroblasts and macrophages with severity together with reduced TLR/IL-1 signalling.

Chemokine ligands were strongly represented in our dataset, with the key leukocyte-recruiting ligands *CCL5*, *CCL18* and *CCL19* clearly associating with severe pathology. In addition, we found a very strong association of *CXCL9*, *CXCL10* and *CXCL11* with severe histopathological damage suggesting a role for the CXCL9, CXCL10, CXCL11/CXCR3 axis in driving lung damage in COVID-19. These chemokines may be produced by macrophages, fibroblasts, or endothelial cells on stimulation with IFNg[13,36] and are crucial in the recruitment of NK cells as well as cytotoxic T cells, therefore presenting a potential avenue for therapeutic intervention. In a scRNAseq study of BAL, *CXCL9/10* as well as granzyme-expressing cytotoxic CD8 T cells have been detected, with differences between moderate and severe patient groups principally in degree of proliferation and TCR clonality[28]. A reduction in peripheral blood NK cells is a signature of COVID-19 severity, yet data from BAL scRNAseq show that NK transcripts are present at equivalent levels to normal lung[28]. It is likely therefore that the phenotype of NK cells and their activation profile is of relevance to disease severity[37]. We identified a cytotoxic gene module within areas of lymphocytic infiltrate with overlapping CD8 and NK cell signatures. These could not be independently resolved, suggesting either co-localisation or a converging phenotype. Despite contrasting findings in the peripheral blood, activated and SARS-CoV-2 reactive CD8+ T cells have been found to be associated with severe disease, indicating active T cell immunity[38]. CD4 transcripts correlated with myeloid rather than T cell modules in our dataset, which may reflect a relative lack of CD4+ T cells in areas of severe damage. Regulatory mechanisms were largely undetectable apart from *IL-10R* expression in severe areas (associating with T cell modules) and complement regulatory genes (*CD46, CD55, CD59*) in milder areas. Together with our data, these studies of lung immune pathology contrast with peripheral blood data where CD8 lymphopaenia and exhaustion is a feature of severe COVID-19 infection[39] and may highlight dissimilarities in tissue and systemic responses[8].

Fibrosis is of particular interest given the emerging data linking its development with the long-term respiratory complications of COVID-19[40,41]. While DAD is a defining feature of COVID-19[42], post-mortem studies have highlighted fibrotic changes to be commonplace in patients with a longer duration of illness[43], similar to the histopathological findings in our series. This fibrosis is widely reported and characterized by lung parenchymal remodelling, fibroblast proliferation, and airspace obliteration[2,42-44]. Interestingly, our data highlight the association of fibrotic genes and pathways (in particular TGF-B signalling) with epithelial and endothelial cells as well as *MARCO+* macrophages in areas of lower DAD severity. These areas were distinct in their expression of prostaglandins such as *PTGS2*, *PLA2G4F*, and *PTGER4*, supporting a potential role for prostaglandin inhibition in COVID-19 independent of its anti-thrombotic effects[45]. While immune modulatory agents have received attention as therapies in the treatment of COVID-19, we suggest that further work should focus on understanding how and when to intervene on the pathways promoting fibrosis.

Of note, complement activating genes were expressed in areas that were over-represented in pathways involved in cytotoxicity, T cells, and antigen presentation and were largely associated with areas of severe DAD. By contrast, complement inhibitory genes (*CD46, CD55, CD59*) were associated with epithelial and endothelial cells in milder areas of damage. Several studies are now emerging implicating complement or its regulation in COVID-19. In one study, patients with severe COVID-19 were found to have increased levels of Factor B and Ba, a zymogen which downregulates complement inhibitory protein expression[46]. In another, C3 expression was found to correlate with viral load, with an increase in its expression in AT2 cells from lung biopsy and BAL scRNAseq datasets, and its processing into its active form in SARS-CoV-2 infected lung epithelial cell lines in vitro[47]. There is therefore good evidence to support exploration of complement inhibition in severe COVID-19.

Within the lung, COVID-19 manifests in a wide spectrum of DAD and fibroproliferation, making the dissection of pathways specific to areas of damage a challenge using bulk assessment techniques. However, the histopathological features are non-specific and there are no clear findings that differentiate SARS-CoV-2 from a number of other respiratory viral infections, particularly those developing following infection with other betacoronaviridae such as SARS-CoV and MERS-CoV[48-53]. Our data therefore provide a clearer picture of the immune landscape within lung tissue after severe COVID-19 infection through the assessment of dozens of histopathological regions and using a network-based analytical workflow. Overall, our data suggest a model (**Figure 5**) in which terminal COVID-19 lung damage is initiated by TLR-mediated myeloid cell activation which involves type I IFN signalling, antigen presentation, complement activation and IL-1 expression. This leads to T and NK cell activation, IFNG and IL-32 driven immune cell proliferation and immune cell recruitment including via the CXCL9, CXCL10, CXCL11/CXCR3 axis. In this model, tissue damage is linked to CD8 T and NK cell cytotoxicity as well as to complement activation. In parallel our data indicate that elevated levels of IL-1, IL-6 and TNF (the so-called "cytokine storm") also drive abnormal stromal cell responses including fibrosis (and, as shown by others, endothelitis[54]). This model highlights a series of partially interlinked processes including innate immune activation, leukocyte activation, immune cell recruitment, cytokine storm, and stromal dysregulation as possible targets for therapeutic intervention in severe COVID-19 (red targets, **Figure 5**). Currently deployed therapeutics such as dexamethasone and tocilizumab tend to target cytokine storm and while effective are not curative in all cases. The partial independence of the processes identified, and multiple feedback loops present, suggest that combination approaches targeting different aspects of immune and tissue dysregulation may be needed. While therapeutics are available or in development to target the virus (such as remdesivir and recombinant soluble ACE2[55]), innate immune activation (such as intranasal IFNB1a/b[56]) and the complement system (via complement inhibition[57]) less attention has been paid to immune cell recruitment. Our data suggest that manipulation of the CXCL9, CXCL10, CXCL11/CXCR3 axis may be a valuable strategy in severe COVID-19 as this pathway is implicated upstream of both tissue damage and cytokine storm (**Figure 5**). Finally, the limited quantities of SARS-CoV-2 detected in the lung tissue from our patients suggests that identified processes of immune activation and tissue dysregulation may become (at least partially) uncoupled from viral infection in terminal cases. This observation highlights a likely need for stage-specific therapeutic interventions and suggests that specific therapeutics may promote different outcomes in early and late disease.

MATERIALS AND METHODS

Ethics statement

This study was approved by the ethics committee of the University of Navarra, Spain (15/05/2020) and the Medical Sciences Interdivisional Research Ethics Committee of the University of Oxford (Approval R76045/RE001). Tissues were stored at the John Radcliffe Hospital according to Human Tissue Authority regulations (Licence 12433).

Patients and tissue processing

Medical records of the patients were retrospectively reviewed[58] and three COVID-19 patients were selected for in-depth analysis based on: their clinical manifestation of ARDS (**Supplementary Table 1**), typical COVID-19 histology (**Supplementary Table 2**) with a 4-5 score on the Brescia-COVID Respiratory Severity Scale, and a lung-restricted presence of SARS-COV-2 (absence in heart, liver and kidney biopsies). Post-mortem lung tissues were obtained through open biopsy at the point of death and processed as described[58]. In brief, tissues were immediately fixed in neutral buffered formalin for <24 hours and then paraffin embedded. Sections (5μm each) were cut for H&E staining, ISH and DSP analysis. Six pathologists reviewed the histology in and agreed on the gross histological characteristics (**Supplementary Table 2**). RNA was extracted from 4-8 x 5μm sections for quantification of SARS-CoV-2 nucleocapsid (N) and envelope protein (E) transcripts (**Supplementary Table 2**).

NanoString GeoMx Digital Spatial Profiling

This technique was carried out according to manufacturer's recommendations for GeoMx-NGS RNA BOND RX slide preparation (manual: MAN-10131-02). Deparaffinization, rehydration, heat-induced epitope retrieval (ER2 for 20 minutes at 100°C) and enzymatic digestion (1μg/ml proteinase K for 15 minutes at 37°C) were carried on the Leica BOND-RX. Tissues were incubated with 10% neutral buffered formalin for 5 minutes and 5 minutes with NBF Stop buffer. The tissue sections hybridized with the oligonucleotide probe mix (Cancer Transcriptome Atlas and COVID-19 spike-in panel) overnight, then blocked and incubated with PanCK-532 (AE1+AE3, Novus), CD3-647 (UMAB54, Origene), CD68-594 (KP1, Santa Cruz) and DNA dye (Syto13-488; Invitrogen) for 1 hour. Tissue sections were then loaded into the GeoMx platform and scanned for immunofluorescent signal.

At least 14 AOIs from each COVID-19 tissue were selected for analysis, spanning on average 0.2mm^2 (range: 0.05 to 0.33mm^2) with exclusion of empty space. Areas were selected to represent the spectrum of alveolar injury within each tissue covering regions of a) mild/moderate injury with some conservation of alveolar architecture and b) severe injury with a loss of alveolar structure and significant inflammation. Selected AOIs occupy alveolar and interstitial spaces, except from A_16 and A_17 containing bronchiolar epithelium. The severity grade of each AOI was confirmed post-hoc by 2 pathologists.

After selection of areas of interest (AOIs), UV light directed at each AOI released oligonucleotides that were collected and prepared for sequencing. Illumina i5 and i7 dual indexing primers were added during PCR (4 μL of collected oligonucleotide/AOI) to uniquely index each AOI. AMPure XP beads (Beckman Coulter) were used for PCR purification. Library concentration as measured using a Qubit fluorometer (Thermo Fisher Scientific) and quality was assessed using a Bioanalyzer (Agilent). Sequencing was performed on an Illumina NextSeq

2000 and FASTQ files were processed by the NanoString DND pipeline, resulting in count data for each target probe in each AOI.

Analysis of immunofluorescent images for cell counts

The number of nuclei, CD3$^+$ and CD68$^+$ cell counts were determined using CellProfiler software. A pipeline was designed to quantify circular objects within RBG files for each AOI[59]. Global manual intensity thresholds were set for object identification of nuclei and CD3$^+$ cells, whilst CD68$^+$ cells were identified by adapted intensity thresholds. The efficacy of object identification for each AOI was visually confirmed.

Quality control and pre-processing of GeoMx transcript expression data

Quality control and initial data exploration was conducted using the GeoMx DSP Analysis Suite. Sequencing quality per AOI was examined and an under-sequenced area (B_04) with zero deduplicated reads was excluded. Expression of each transcript was measured by 5 or more probes; outlier probes (geomean probe in all AOI/geomean of all probes for a transcript = <0.1; or probe failure in the Grubbs outlier test in over 20% AOIs) were excluded and the remaining probes combined to generate a single (post "biological probe QC") expression value per gene target per AOI.

We evaluated the performance of two normalization strategies, upper quartile and quantile normalization, by investigating their ability to standardise the expression distributions of housekeeping genes, negative control probes as well as the full expression distribution between the AOIs. The expression values for negative control probes that were not reported as global outliers were appended to the gene expression matrix. Based on this assessment (and the results of subsequent PCA analyses) we proceeded with the quantile normalized expression values. We investigated the influence of known technical and biological factors on the variance structure of the dataset by performing PCA of the log2(n+1) transformed quantile-normalized expression values. This analysis revealed that the first component in the data was associated with the aligned read depth statistic. We therefore corrected the quantile normalized expression values for this technical factor using Limma "removeBatchEffect" function. Finally, to distinguish gene expression from background noise, we modelled the expression distribution of the set of negative probes, defining the median of the negative probe expression values + 2*median absolute deviations as a robust detection threshold. In total we retained 1631 genes that passed this detection threshold in at least 2 AOIs. The normalized expression distributions and sample PCA plots obtained following normalization, aligned read depth correction and expression level filtering are shown in Supplementary Figure 1.

Differential gene expression and over-representation analysis

Differential gene expression was calculated for each gene between areas of mild/moderate and severe alveolar damage using linear mixed models (fixed effect: severity, random variable: patient identity). Significance was estimated using Satterthwaite's degrees of freedom method and *p* values adjusted using Benjamini-Hochberg methods (R libraries: lme4, lmerTest). Over-representation analysis for Gene Ontology Biological Processes (GO.BP) terms was performed on genes with >1.5 fold change and FDR > 0.05 (R library: clusterProfiler[60]) with the parameters: pvalueCutoff = 0.05 and qvalueCutoff = 0.10. Redundant and duplicated pathways were removed for **Figure 1D**; for full list of pathways see **Supplementary data**.

Weighted gene correlation network analysis (WGCNA)

WGCNA was applied to the log2(n+1) transformed, quantile normalized, aligned read corrected and filtered expression values. The parameter values were set as follows: minimum fraction of non-missing samples for a gene to be considered good=0.5, minimum number of non-missing samples for a gene to be considered good=4, minimum number of good genes=4, cut height for removing outlying samples=100 (no samples removed), minimum number of objects on branch to be considered a cluster=2, network type=signed-hybrid, soft power=4, adjacency correlation function=bicor, adjacency distance function=dist, TOM type=signed, minimum module size (number of genes)=10, dissimilarity threshold used for merging modules: 0.25. The analysis identified 17 distinct modules labelled with colours, plus a module of 3 unassigned genes (grey module). The WGCNA analysis was performed using pipeline_wgcna.py (https://github.com/sansomlab/cornet).

The expression patterns of the modules were summarized by calculation of module eigengenes using the WGCNA package (which defines a modules' eigengene as the first principal component of the expression of the modules' gene members). AOIs were hierarchically clustered by expression of the module eigengenes (Pearson correlation distance; optimized leaf ordering) to identify n=5 distinct groups (R "cutree" function). The over-representation of Gene Ontology (GO) categories, KEGG pathways and Reactome pathways in module gene members was tested using one-sided Fishers exact tests (https://github.com/sansomlab/gsfisher) using the union of gene members from all the modules as the background geneset. Representative genesets and pathways that showed significant (BH adjusted $p < 0.05$) over-representations are displayed in the figures (with full results are provided in supplementary table X). WGCNA modules were further characterized by assessing the correlation of the module eigengenes with a) histological severity (Pearson correlation) b) predicted cell abundances (as estimated using Spatial Decon; Spearman correlation) and c) that of selected genes (Spearman correlation, R library Hmisc).

Cell deconvolution of GeoMx transcript expression data

Deconvolution of cell types from the gene expression data was performed using the SpatialDecon R library[12]. For this analysis the full matrix of quantile normalized gene counts (without correction), mean negative probe counts and the cell profile matrix 'Lung_plus_neut' were applied as input. The cell profile matrix retrieved from the package was generated from the Human Cell Atlas lung scRNAseq dataset and appended with neutrophil profiles as described in Desai et al. (2020)[5]. For correlation analyses, the abundance of cell types is reported for cell types present with > 2 relative abundance in > 2 AOIs. Complete cell type output in **Supplementary figure.**

Construction of AOI group correlation networks

We constructed correlation networks to investigate the relationship between the WGCNA modules, the estimated cell abundances (from SpatialDecon) and the expression levels of immune signalling genes for each of the n=5 AOI spatial groups (R igraph library; layout = layout_with_dh). Nodes were scaled in size and/or colour according to cell abundance, normalized gene expression or WGCNA module eigengene expression. Drawn edges represent significant correlations ($p < 0.05$) and were weighted according to the value of the correlation coefficient (Spearman's Rho). WGCNA modules were included in the network if they showed a

positive module eigengene expression in >70% of the AOI of the relevant group. Immune signalling genes from the KEGG 'cytokine-cytokine receptor interaction' pathway (hsa04060) (human) were included if they i) correlated with the positively expressed gene modules for a given spatial group and ii) showed a median expression above the expression detection threshold (as defined in the pre-processing section) in a given AOI group. Cell type abundance estimates were included if they i) correlated with the positively expressed gene modules for a given spatial group and ii) if the cell types were present with > 2 relative abundance in > 2 of the 46 AOIs.

FIGURE AND SUPPLEMENTARY DATA LEGENDS

<u>Figure 1:</u> **A spectrum of diffuse alveolar damage and inflammation was observed within and across COVID-19 lung biopsies.** Selection and annotation of areas of diffuse alveolar damage in COVID-19 lung tissues for transcriptomic analysis. (**A**) Haematoxylin and eosin-stained sections of the lung samples from patients A, B and C (left; scale bars: 5mm). Merged immunofluorescence (IF) images of the lung samples from patients A, B and C (right; scale bars: 5mm, pan cytokeratin-green, DNA-blue, CD3-red, CD68-yellow). IF sections are within 15μm of the H&E-stained sections. 47 areas of interest (AOI) selected for transcript profiling are highlighted, of which 1 AOI was excluded from further analysis for sequencing failure. (**B**) Representative IF images of AOIs demonstrating the morphology and immune infiltrate observed within areas of severe and mild/moderate diffuse alveolar damage (pan cytokeratin-green, DNA-blue, CD3-red, CD68-yellow). AOIs spanned on average $0.2mm^2$ (range: 0.05 to $0.33mm^2$) with exclusion of empty space. (**C**) The proportion of $CD3^+$ and $CD68^+$ cells of total nucleated cells were derived from the immunofluorescent imaging, plotted for each AOI and coloured by histological severity (mild/moderate-light blue, severe-pink) or inclusion of bronchiolar epithelium (grey). AOI are annotated with their origin from patient A (circle), patient B (triangle), and patient C (square). (**D**) Differential gene expression between areas of severe and mild/moderate damage. Coloured and annotated genes have a fold change expression > 1.5 and a Benjamini-Hochberg (BH) adjusted *p* value <0.05 in a linear mixed model. (**E**) Selected GO biological processes significantly over-represented in genes differentially expressed between mild and severe areas of damage (BH corrected p < 0.05, one-sided Fisher's exact test); see also **Supplementary Table 3**.

<u>Figure 2:</u> **Identification and characterization of gene modules with spatially heterogenous expression in COVID-19 lung tissue.** Application of weighted gene correlation network analysis (WGCNA) to spatial transcriptomic data identified 17 modules of co-expressed genes (each designated a colour name). (**A**) The number of genes assigned to each module. (**B**) Selected GO Biological processes, KEGG and Reactome pathways that were significantly over-represented in the detected modules (BH adjusted *p* values < 0.05; one-side Fisher's exact test; see also **Supplementary Table 3**). (**C**) Correlation between estimated cell type abundance (as determined by cell deconvolution) and WGCNA module eigengene expression (all AOIs; only positive Spearman correlation values are shown). (**D**) Association of cell type marker and immune function genes with the WGCNA module eigengenes (only positive Spearman correlation values are shown; asterisks indicate high module membership > 0.7, *p* value < 0.05). Simplified module aliases (right) describe each module based on their association with cell lineages, function-associated genes and defining biological pathways. (**E**) Genes were selected for impact on pathology: *chemokines/receptors, cytokines/receptors, viral receptors and cofactors, IL-6 signalling, complement, fibrosis* and *prostaglandins*. Genes were correlated with the eigengene modules (Spearman's correlation; asterisks indicate high WGCNA membership > 0.7, *p* value < 0.05).

<u>Figure 3:</u> **Combination of WGCNA modules reveals five groups of alveolar damage with distinct transcriptional profiles.** (**A**) The severity of tissue damage was correlated to module eigengene expression separately for the AOIs of each patient (Pearson correlation). (**B**)WGCNA module eigengene expression is shown for each AOI. Sampled areas (AOI) are annotated with patient identity, the severity of damage and the percentage of CD3$^+$ and CD68$^+$ cells of total nucleated cells. Hierarchical clustering of the 46 areas (AOI) by expression of the WGCNA module eigengenes identified 5 "spatial groups" with distinct patterns of module expression. (**C**) The expression of selected genes that exemplify key differences between the 5 "spatial groups" of AOIs.

<u>Figure 4:</u> **Transcriptomic subgroups of alveolar damage are associated different soluble factor profiles.** (**A**) Key characteristics of the five spatial AOI groups: the average percentage of CD3$^+$ lymphocytes of nucleated cells (x axis), the percentage of AOI with severe alveolar damage (y axis), and patient of origin (pie charts). (**B-F**) Correlation networks for each spatial group. Networks were constructed with selected WGCNA eigengene modules (grey nodes), cytokines and chemokines red gradient nodes) and cell abundance estimates from cell deconvolution (blue gradient nodes) (see methods for node inclusion criteria). Nodes are scaled in size and/or colour to the relative module eigengene expression (grey nodes), the normalized gene counts (red gradient nodes) and the cell abundance (blue gradient nodes). Edges represent positive correlations (Spearman; p<0.05) between a gene/cell type and eigengene modules across the whole dataset and are weighted to the Spearman Rho coefficient (+0.1 to +1). Three immunofluorescent images of representative AOI from each spatial group are shown to illustrate major morphological features and cell composition (pan cytokeratin-green, DNA-blue, CD3-red, CD68-yellow). (**G**) Positive correlations between spatial groups and the expression of genes selected for impact on pathology is shown: *chemokines/receptors, cytokines/receptors, viral receptors and cofactors, IL-6 signalling, complement, fibrosis* and *prostaglandins* (Spearman's rank correlation, white 0 to red +1).

<u>Figure 5</u>: **A cellular model of late-stage lung tissue damage in terminal COVID-19.** Our data indicate that dysregulation of lung tissue homeostasis in late-stage terminal COVID-19 involves (1) SARS-CoV-2 mediated TLR activation of myeloid cells such as alveolar macrophages and dendritic cells that leads to (2) complement system activation, (3) Type I interferon (*IFNA1/A17*) expression and (4) antigen presentation. The type I interferon signaling and antigen presentation are hypothesized to lead to (5) activation of NK and T cells resulting in their expression *IFNG* and *IL32* driving (6) proliferation and activation of immune cells in the tissue, including that of *MARCO-* macrophages. These activated immune cells are thought to (7) recruit peripheral monocytes, T and NK cells into the tissue via the CXCL9, CXCL10, CXCL11/CXCR3 axis. Tissue damage (8) is understood to be a consequence of NK and CD8 cell cytotoxicity, complement activation as well as the actions of infiltrating myeloid cells. Cytokine storm (9) involves IL-1 expression from activated myeloid cells as well as TNF and IL6 expression from activated leukocytes and infiltrating immune cells. High levels of these pro-inflammatory cytokines lead to detectable (10) IL-1/IL-6 responses in epithelial cells and is suggested to cause (11) fibrosis. In addition, (12) endothelitis has been reported in the literature[54]. Established and potential therapeutic targets are indicated with red target symbols. Only genes and cell types (with the exception of endothelitis) for which there is evidence of involvement from this study are shown in the figure.

<u>Supplementary Figure 1:</u> **Normalization and variance of transcriptomic data across areas of interest (AOIs).** (**A**) The distribution of all gene and (**B**) selected housekeeping gene expression levels after the application of quantile normalization across the AOIs. (**C**) Principal

component analysis of the variance between the sampled AOIs (PCA performed with the pre-processed quantile normalised, log2 transformed, aligned read corrected and expression-level filtered gene expression values). Here PCs 1-4 are shown, annotated by the severity of alveolar damage (severe: pink/square, mild/moderate: light blue/circle, respiratory epithelium: grey triangle), patient identity (A: blue/circle, B: green/triangle, C: pink/square), percentage of CD3$^+$ (red gradient) and CD68$^+$ cells (blue gradient) of total nucleated cells.

Supplementary Figure 2: Deconvolution of cell lineages across sampled areas in COVID-19 lung tissue. (A) The abundance of cell lineages was estimated from the RNA expression data obtained from each sampled area (AOI) across the post-mortem COVID-19 lung tissues (R package SpatialDecon using cell profile matrix derived of healthy lung scRNAseq appended with neutrophil populations 'Lung_plus_neut'). The abundance is hierarchically clustered and annotated with patient identity, the percentage of CD3 or CD68 cells of total nuclei, and the alveolar damage severity. **(B)** The estimated cell abundance is plotted with the percentage of CD3+ (top panel) or CD68+ (bottom panel) cells of total nuclei, calculated from analysis of the immunofluorescent images (linear model; confidence interval = 0.95). **(C)** The correlation coefficients were calculated between deconvoluted cell types (> 2 abundance in > 2 AOI) and genes selected for potential impact on pathology: *chemokines/receptors, cytokines/receptors, viral receptors and cofactors, IL-6 signalling, complement, fibrosis* and *prostaglandins* (Spearman's rank correlation, white 0 to red +1).

DATA AVAILABILITY

Phase contrast images of H&E-stained slides and RBG composite images taken of immunofluorescent stained slides will be deposited in GEO together with NanoString GeoMx DSP data and metadata. Annotated scripts for the computational analysis are available from https://github.com/sansomlab/covidlung.

AUTHOR CONTRIBUTIONS

C.E.DA., M.F.LA. and I.M. obtained consent, clinical data and the patient samples. I.R. and C.E.DA. reported on tissue histology. A.R.C, L.C. and F.I. designed and performed the experiments. A.R.C., F.I. and S.S. analysed data and generated the figures. A.R.C, S.S. and F.I. acquired funding and oversaw the project. A.R.C., J.H., P.K., I.M. and F.I. conceived the work and all authors contributed to writing the manuscript.

ACKNOWLEDGEMENTS

FUNDING STATEMENT

KEY WORDS

REFERENCES

1   Speranza, E. *et al.* Single-cell RNA sequencing reveals SARS-CoV-2 infection dynamics in lungs of African green monkeys. *Sci Transl Med* **13**, doi:10.1126/scitranslmed.abe8146 (2021).

2   Wichmann, D. *et al.* Autopsy Findings and Venous Thromboembolism in Patients With COVID-19: A Prospective Cohort Study. *Ann Intern Med* **173**, 268-277, doi:10.7326/M20-2003 (2020).

3   Grasselli, G. *et al.* Pathophysiology of COVID-19-associated acute respiratory distress syndrome: a multicentre prospective observational study. *Lancet Respir Med* **8**, 1201-1208, doi:10.1016/S2213-2600(20)30370-2 (2020).

4   Short, K. R., Kroeze, E., Fouchier, R. A. M. & Kuiken, T. Pathogenesis of influenza-induced acute respiratory distress syndrome. *Lancet Infect Dis* **14**, 57-69, doi:10.1016/S1473-3099(13)70286-X (2014).

5   Desai, N. *et al.* Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. *Nat Commun* **11**, 6319, doi:10.1038/s41467-020-20139-7 (2020).

6   Rendeiro, A. F. *et al.* The spatial landscape of lung pathology during COVID-19 progression. *Nature*, doi:10.1038/s41586-021-03475-6 (2021).

7   Delorey, T. M. *et al.* COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature*, doi:10.1038/s41586-021-03570-8 (2021).

8   Ahern, D. J. *et al.* A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *medRxiv*, 2021.2005.2011.21256877, doi:10.1101/2021.05.11.21256877 (2021).

9   Marx, V. Method of the Year: spatially resolved transcriptomics. *Nat Methods* **18**, 9-14, doi:10.1038/s41592-020-01033-y (2021).

10   Merritt, C. R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* **38**, 586-599, doi:10.1038/s41587-020-0472-9 (2020).

11   Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).

12   Danaher, P. *et al.* Advances in mixed cell deconvolution enable quantification of cell types in spatially-resolved gene expression data. *bioRxiv*, doi:https://doi.org/10.1101/2020.08.04.235168 (2020).

13   Tokunaga, R. *et al.* CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. *Cancer Treat Rev* **63**, 40-47, doi:10.1016/j.ctrv.2017.11.007 (2018).

14   Ribeiro-Dias, F., Saar Gomes, R., de Lima Silva, L. L., Dos Santos, J. C. & Joosten, L. A. Interleukin 32: a novel player in the control of infectious diseases. *J Leukoc Biol* **101**, 39-52, doi:10.1189/jlb.4RU0416-175RR (2017).

15   Yan, Y. *et al.* CCL19 and CCR7 Expression, Signaling Pathways, and Adjuvant Functions in Viral Infection and Prevention. *Front Cell Dev Biol* **7**, 212, doi:10.3389/fcell.2019.00212 (2019).

16   Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036-1045 e1039, doi:10.1016/j.cell.2020.04.026 (2020).

17   Lokugamage, K. G. *et al.* Type I Interferon Susceptibility Distinguishes SARS-CoV-2 from SARS-CoV. *J Virol* **94**, doi:10.1128/JVI.01410-20 (2020).

18   Stanifer, M. L. *et al.* Critical Role of Type III Interferon in Controlling SARS-CoV-2 Infection in Human Intestinal Epithelial Cells. *Cell Rep* **32**, 107863, doi:10.1016/j.celrep.2020.107863 (2020).

19   Pfaender, S. *et al.* LY6E impairs coronavirus fusion and confers immune control of viral disease. *Nat Microbiol* **5**, 1330-1339, doi:10.1038/s41564-020-0769-y (2020).

20   Hadjadj, J. *et al.* Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718-724, doi:10.1126/science.abc6027 (2020).

21   Zhou, Z. *et al.* Heightened Innate Immune Responses in the Respiratory Tract of COVID-19 Patients. *Cell Host Microbe* **27**, 883-890 e882, doi:10.1016/j.chom.2020.04.017 (2020).

22   Lee, J. S. *et al.* Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. *Sci Immunol* **5**, doi:10.1126/sciimmunol.abd1554 (2020).

23   Perkins, D. J. & Vogel, S. N. Space and time: New considerations about the relationship between Toll-like receptors (TLRs) and type I interferons (IFNs). *Cytokine* **74**, 171-174, doi:10.1016/j.cyto.2015.03.001 (2015).

24   Mathew, D. *et al.* Deep immune profiling of COVID-19 patients reveals patient heterogeneity and distinct immunotypes with implications for therapeutic interventions. *bioRxiv*, doi:10.1101/2020.05.20.106401 (2020).

25   Vabret, N. *et al.* Immunology of COVID-19: Current State of the Science. *Immunity* **52**, 910-941, doi:10.1016/j.immuni.2020.05.002 (2020).

26   Giamarellos-Bourboulis, E. J. *et al.* Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. *Cell Host Microbe* **27**, 992-1000 e1003, doi:10.1016/j.chom.2020.04.009 (2020).

27   Mann, E. R. *et al.* Longitudinal immune profiling reveals key myeloid signatures associated with COVID-19. *Sci Immunol* **5**, doi:10.1126/sciimmunol.abd6197 (2020).

28   Liao, M. *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med* **26**, 842-844, doi:10.1038/s41591-020-0901-9 (2020).

29   Barnes, B. J. *et al.* Targeting potential drivers of COVID-19: Neutrophil extracellular traps. *J Exp Med* **217**, doi:10.1084/jem.20200652 (2020).

30   He, J. *et al.* Single-cell analysis reveals bronchoalveolar epithelial dysfunction in COVID-19 patients. *Protein Cell* **11**, 680-687, doi:10.1007/s13238-020-00752-4 (2020).

31   Mauer, J., Denson, J. L. & Bruning, J. C. Versatile functions for IL-6 in metabolism and cancer. *Trends Immunol* **36**, 92-101, doi:10.1016/j.it.2014.12.008 (2015).

32   Mehta, P. *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033-1034, doi:10.1016/S0140-6736(20)30628-0 (2020).

33   Thwaites, R. S. *et al.* Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. *Sci Immunol* **6**, doi:10.1126/sciimmunol.abg9873 (2021).

34   Mudd, P. A. *et al.* Distinct inflammatory profiles distinguish COVID-19 from influenza with limited contributions from cytokine storm. *Sci Adv* **6**, doi:10.1126/sciadv.abe3024 (2020).

35   Group, R. C. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* **397**, 1637-1645, doi:10.1016/S0140-6736(21)00676-0 (2021).

36   House, I. G. *et al.* Macrophage-Derived CXCL9 and CXCL10 Are Required for Antitumor Immune Responses Following Immune Checkpoint Blockade. *Clin Cancer Res* **26**, 487-504, doi:10.1158/1078-0432.CCR-19-1868 (2020).

37   Maucourant, C. *et al.* Natural killer cell immunotypes related to COVID-19 disease severity. *Sci Immunol* **5**, doi:10.1126/sciimmunol.abd6832 (2020).

38   Saini, S. K. *et al.* SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8(+) T cell activation in COVID-19 patients. *Sci Immunol* **6**, doi:10.1126/sciimmunol.abf7550 (2021).

39   Kusnadi, A. *et al.* Severely ill COVID-19 patients display impaired exhaustion features in SARS-CoV-2-reactive CD8(+) T cells. *Sci Immunol* **6**, doi:10.1126/sciimmunol.abe4782 (2021).

40   Fraser, E. Long term respiratory complications of covid-19. *Bmj* **370**, m3001, doi:10.1136/bmj.m3001 (2020).

41   Spagnolo, P. *et al.* Pulmonary fibrosis secondary to COVID-19: a call to arms? *Lancet Respir Med* **8**, 750-752, doi:10.1016/S2213-2600(20)30222-8 (2020).

42   Carsana, L. *et al.* Pulmonary post-mortem findings in a series of COVID-19 cases from northern Italy: a two-centre descriptive study. *Lancet Infect Dis* **20**, 1135-1140, doi:10.1016/S1473-3099(20)30434-5 (2020).

43   Grillo, F., Barisione, E., Ball, L., Mastracci, L. & Fiocca, R. Lung fibrosis: an undervalued finding in COVID-19 pathological series. *Lancet Infect Dis* **21**, e72, doi:10.1016/S1473-3099(20)30582-X (2021).

44   Fox, S. E. *et al.* Pulmonary and cardiac pathology in African American patients with COVID-19: an autopsy series from New Orleans. *Lancet Respir Med* **8**, 681-686, doi:10.1016/S2213-2600(20)30243-5 (2020).

45   Chow, J. H. *et al.* Aspirin Use Is Associated With Decreased Mechanical Ventilation, Intensive Care Unit Admission, and In-Hospital Mortality in Hospitalized Patients With Coronavirus Disease 2019. *Anesthesia and analgesia* **132**, 930-941, doi:10.1213/ANE.0000000000005292 (2021).

46   Yan, B. *et al.* SARS-CoV-2 drives JAK1/2-dependent local complement hyperactivation. *Sci Immunol* **6**, doi:10.1126/sciimmunol.abg0833 (2021).

47   Ma, L. *et al.* Increased complement activation is a distinctive feature of severe SARS-CoV-2 infection. *bioRxiv*, doi:10.1101/2021.02.22.432177 (2021).

48   Xu, Z. *et al.* Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir Med* **8**, 420-422, doi:10.1016/S2213-2600(20)30076-X (2020).

49   Barton, L. M., Duval, E. J., Stroberg, E., Ghosh, S. & Mukhopadhyay, S. COVID-19 Autopsies, Oklahoma, USA. *Am J Clin Pathol* **153**, 725-733, doi:10.1093/ajcp/aqaa062 (2020).

50   Menter, T. *et al.* Postmortem examination of COVID-19 patients reveals diffuse alveolar damage with severe capillary congestion and variegated findings in lungs and other organs suggesting vascular dysfunction. *Histopathology* **77**, 198-209, doi:10.1111/his.14134 (2020).

51   Franks, T. J. *et al.* Lung pathology of severe acute respiratory syndrome (SARS): a study of 8 autopsy cases from Singapore. *Hum Pathol* **34**, 743-748, doi:10.1016/s0046-8177(03)00367-8 (2003).

52   Martines, R. B. *et al.* Pathology and Pathogenesis of SARS-CoV-2 Associated with Fatal Coronavirus Disease, United States. *Emerg Infect Dis* **26**, 2005-2015, doi:10.3201/eid2609.202095 (2020).

53   Ng, D. L. *et al.* Clinicopathologic, Immunohistochemical, and Ultrastructural Findings of a Fatal Case of Middle East Respiratory Syndrome Coronavirus Infection in the United Arab Emirates, April 2014. *Am J Pathol* **186**, 652-658, doi:10.1016/j.ajpath.2015.10.024 (2016).

54   Varga, Z. *et al.* Endothelial cell infection and endotheliitis in COVID-19. *Lancet* **395**, 1417-1418, doi:10.1016/S0140-6736(20)30937-5 (2020).

55   Zoufaly, A. *et al.* Human recombinant soluble ACE2 in severe COVID-19. *Lancet Respir Med* **8**, 1154-1158, doi:10.1016/S2213-2600(20)30418-5 (2020).

56   Sallard, E., Lescure, F. X., Yazdanpanah, Y., Mentre, F. & Peiffer-Smadja, N. Type 1 interferons as a potential treatment against COVID-19. *Antiviral Res* **178**, 104791, doi:10.1016/j.antiviral.2020.104791 (2020).

57   Zelek, W. M. *et al.* Complement Inhibition with the C5 Blocker LFG316 in Severe COVID-19. *Am J Respir Crit Care Med* **202**, 1304-1308, doi:10.1164/rccm.202007-2778LE (2020).

58   Recalde-Zamacona, B. *et al.* Histopathological findings in fatal COVID-19 severe acute respiratory syndrome: preliminary experience from a series of 10 Spanish patients. *Thorax* **75**, 1116-1118, doi:10.1136/thoraxjnl-2020-215577 (2020).

59   Lamprecht, M. R., Sabatini, D. M. & Carpenter, A. E. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* **42**, 71-75, doi:10.2144/000112257 (2007).

60   Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

SUPPLEMENTARY TABLES

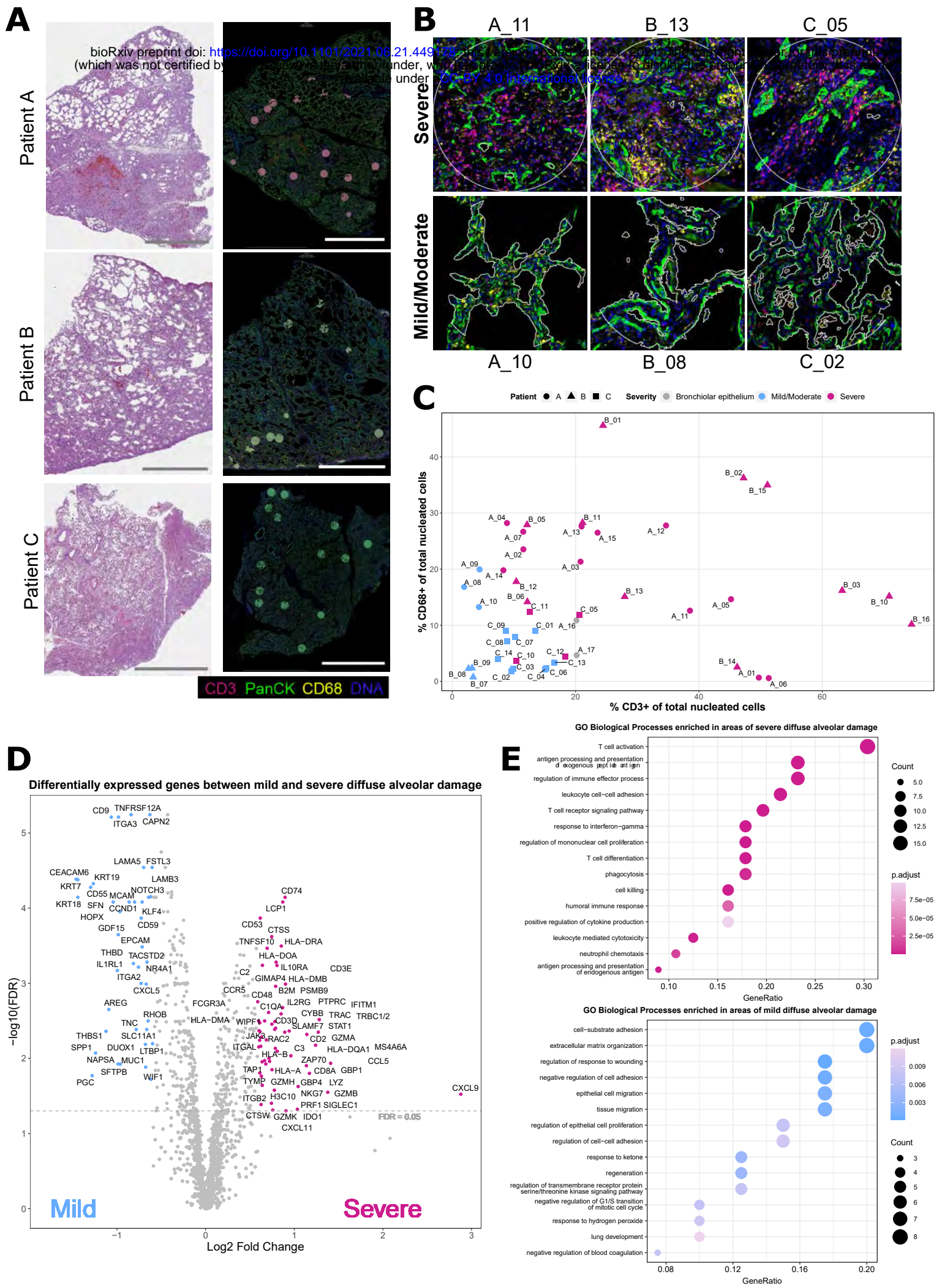### Supplementary Table 1: Patient characteristics

Patient	Age/ Sex	Medical history	CT findings	Lymphocytes (10⁹/L) NR:1.2-4	Neutrophils (10⁹/L) NR: 1.5-6.5	CRP (mg/dl) NR: ≤0.5	Ferritin (ng/ml) NR:15-150	D-dimer (ng/ml) NR:150-500	Platelets (10⁹/L) NR: 150-450	Troponin T (ng/L) NR: ≤14	Pro-BNP (pg/ml) NR: ≤222	Treatment	DCC (ICU)
A	64/F	HTN, heart failure, atrial fibrillation, lymphoma	GGO, consolidation	0.91	2.19	9.17	n/a	1880	12	n/a	n/a	HCQ, AZM, CCS, Tocilizumab Remdesivir	46 (19)
B	80/M	HTN, COPD	GGO, RETp (Gen), pleural effusion	0.24	10.99	0.43	768	2120	133	27.3	2854	HCQ, AZM, CRO, CCS, LPV/r	17 (0)
C	60/F	HTN, type II diabetes, obesity, dyslipidaemia, heart failure, ischaemic cardiomyopathy	GGO, consolidation	0.85	3.58	18.94	998	7010	310	495.2	14858	HCQ, AZM, CRO, CCS LPV/r	29 (24)

Clinical, CT and laboratory findings. Hypertension (HTN); Ground-glass opacities (GGO); C-Reactive protein (CRP); Hydroxychloroquine (HCQ); Azithromycin (AZM); Lopinavir/ritonavir (LPV/r); Corticosteroids (CCS); Ceftriaxone (CRO); Intensive Unit Care (ICU). Duration of clinical course (DCC).

Supplementary Table 2: Histological characteristics

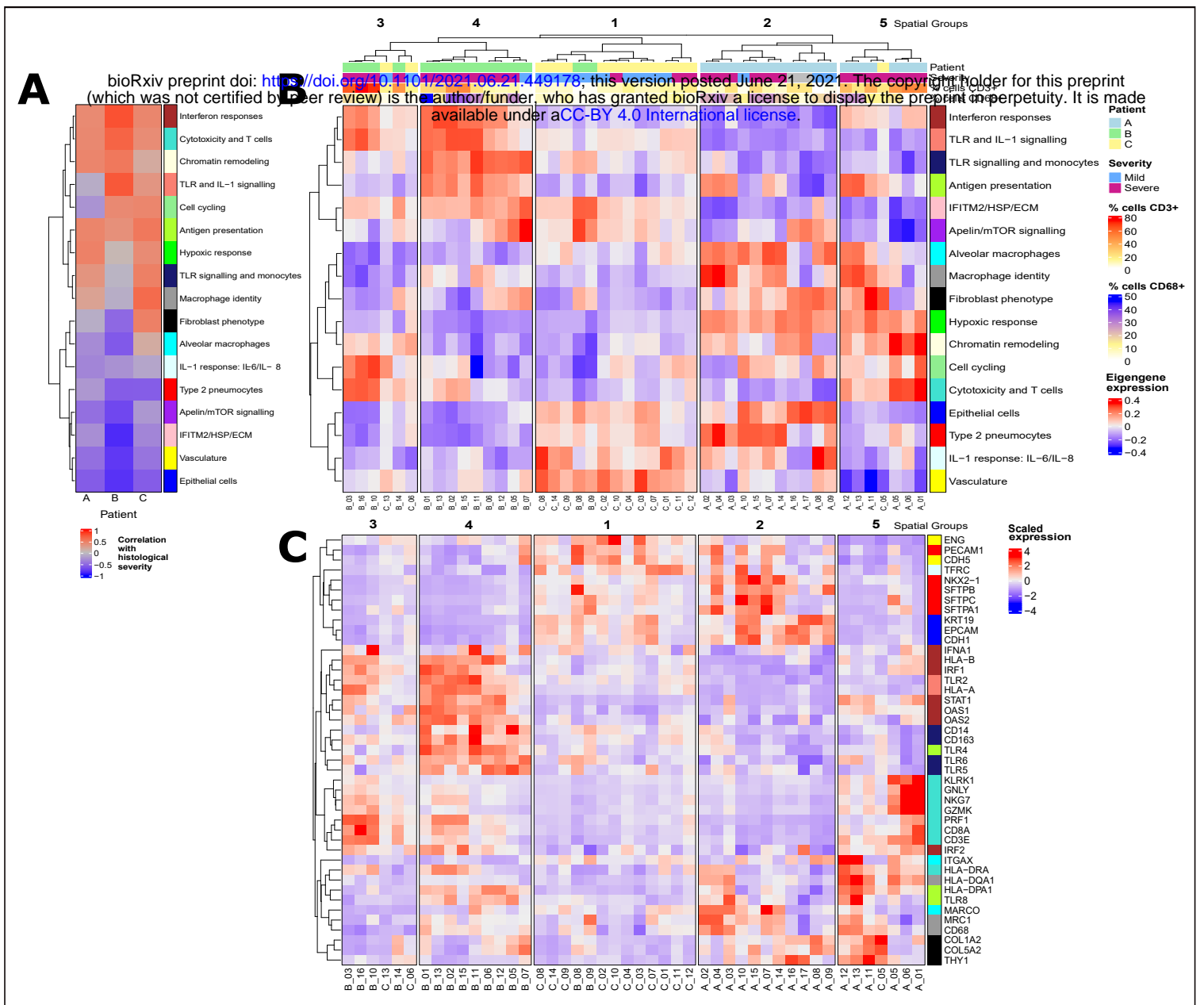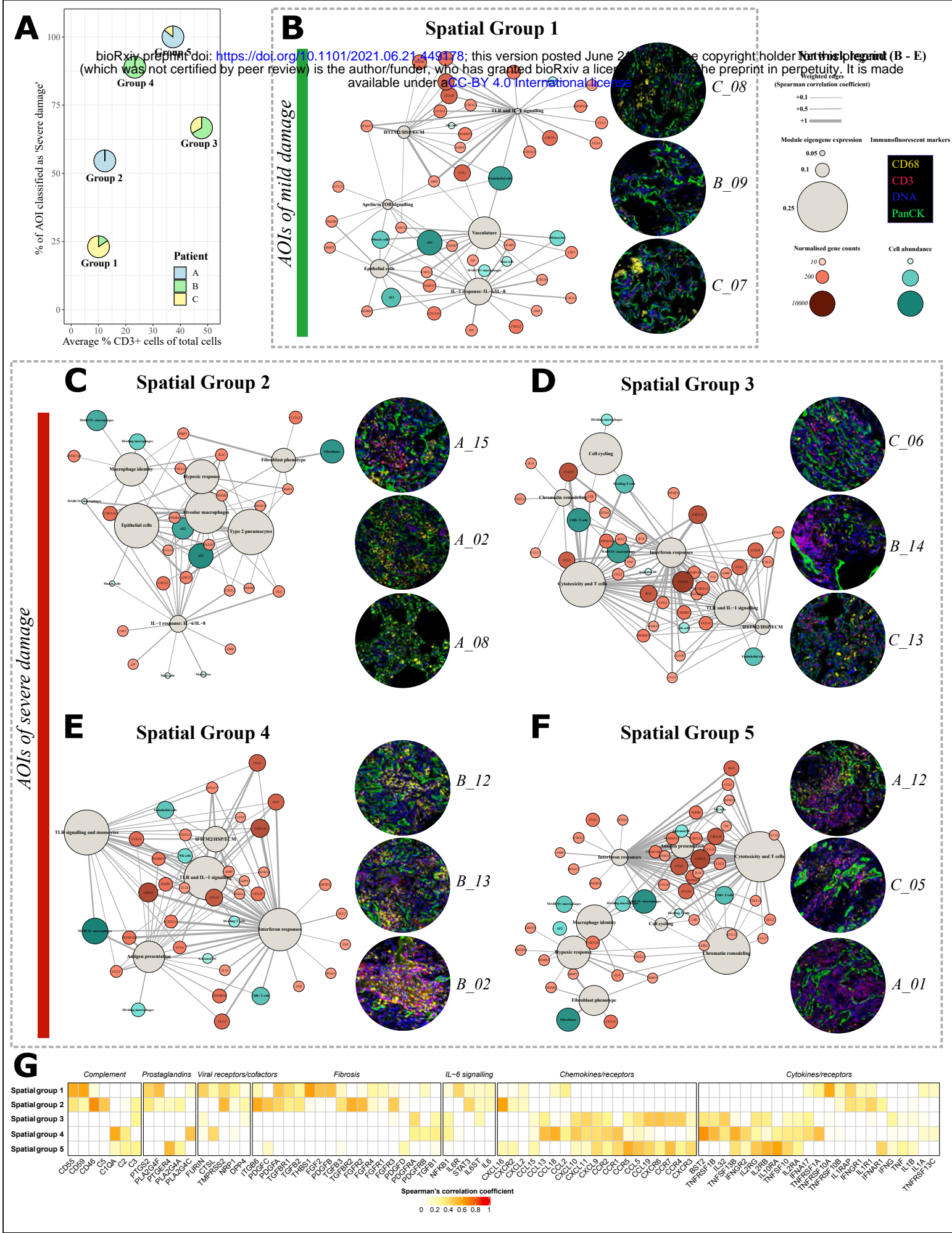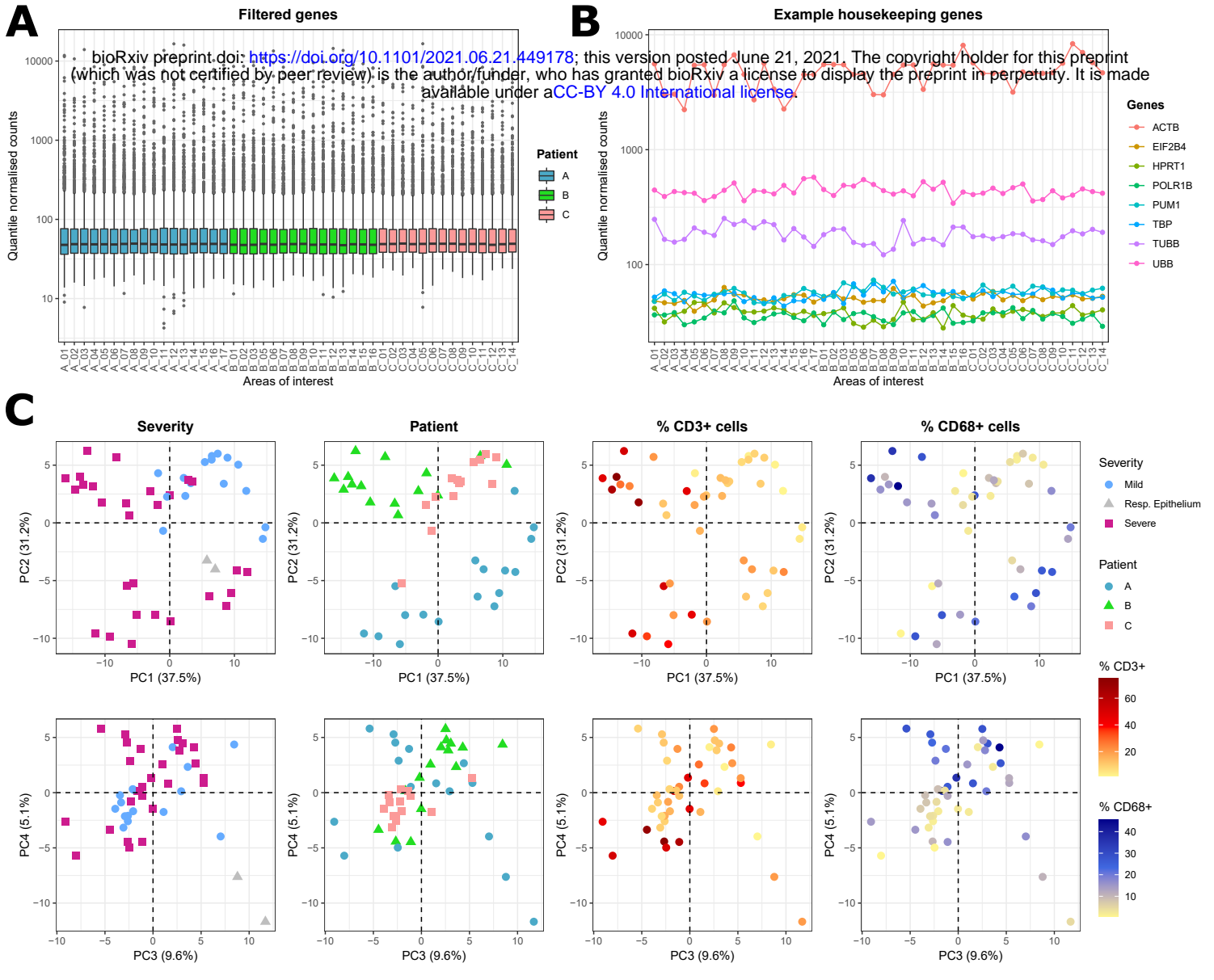| Patient	Histopathological Description	Cellular findings	Lung tissue SARS-CoV-2 qPCR
A	This displays DAD consistent with a later organising or proliferative phase of the pathology, with fewer hyaline membranes, continued type II pneumocyte hyperplasia, and organized pneumonia. The tissue is split into two segments: one segment features severe DAD, interstitial inflammation and haemorrhage and one segment features mild/moderate DAD with conserved alveolar structure (airspace). The immune infiltrate was diffusely and patchily distributed across the section. This tissue exhibits endarteritis and medium sized-vessel thrombi.	Type II pneumocyte hyperplasia and hypertrophy with reactive atypia; Vacuolated macrophages; Hyaline membranes; Normal vascular growth; Medium size-vessel partial thrombi; Interstitial pneumonia with mild inflammation; Squamous metaplasia.	Gen N: 34.2 Gen E: 35.8
B	This tissue features widespread diffuse alveolar damage (DAD) with evident hyaline membranes and type II pneumocyte hyperplasia. These features are present across the tissue, yet the severity ranges from mild/moderate to severe. The interstitium is expanded in many areas, demonstrating a fibroproliferative response. There is evidence of diffuse and patchy interstitial inflammation across the tissue with infiltrate morphology consistent with lymphocytes, macrophages and plasma cells.	Type II pneumocyte hyperplasia and hypertrophy with reactive atypia; Vacuolated macrophages; Hyaline membranes; Vascular wall hyperplasia; Medium size-vessel partial thrombi; Squamous metaplasia.	Gen N: 27.3 Gen E: 25.5
C	This features DAD consistent with an organising or proliferative phase of the pathology, with fewer hyaline membranes, continued type II pneumocyte hyperplasia and organising pneumonia. Interstitial inflammation is present and diffusely distributed across the section. The vasculature exhibits arterial thrombi and endarteritis with intimal expansion.	Type II pneumocyte hyperplasia and hypertrophy with reactive atypia; Vacuolated macrophages; Hyaline membranes; Moderate inflammation; Vascular wall hyperplasia; Medium size-vessel partial thrombi; Focal edema.	Gen N: Neg Gen E: Neg

Figure 1

Figure 2
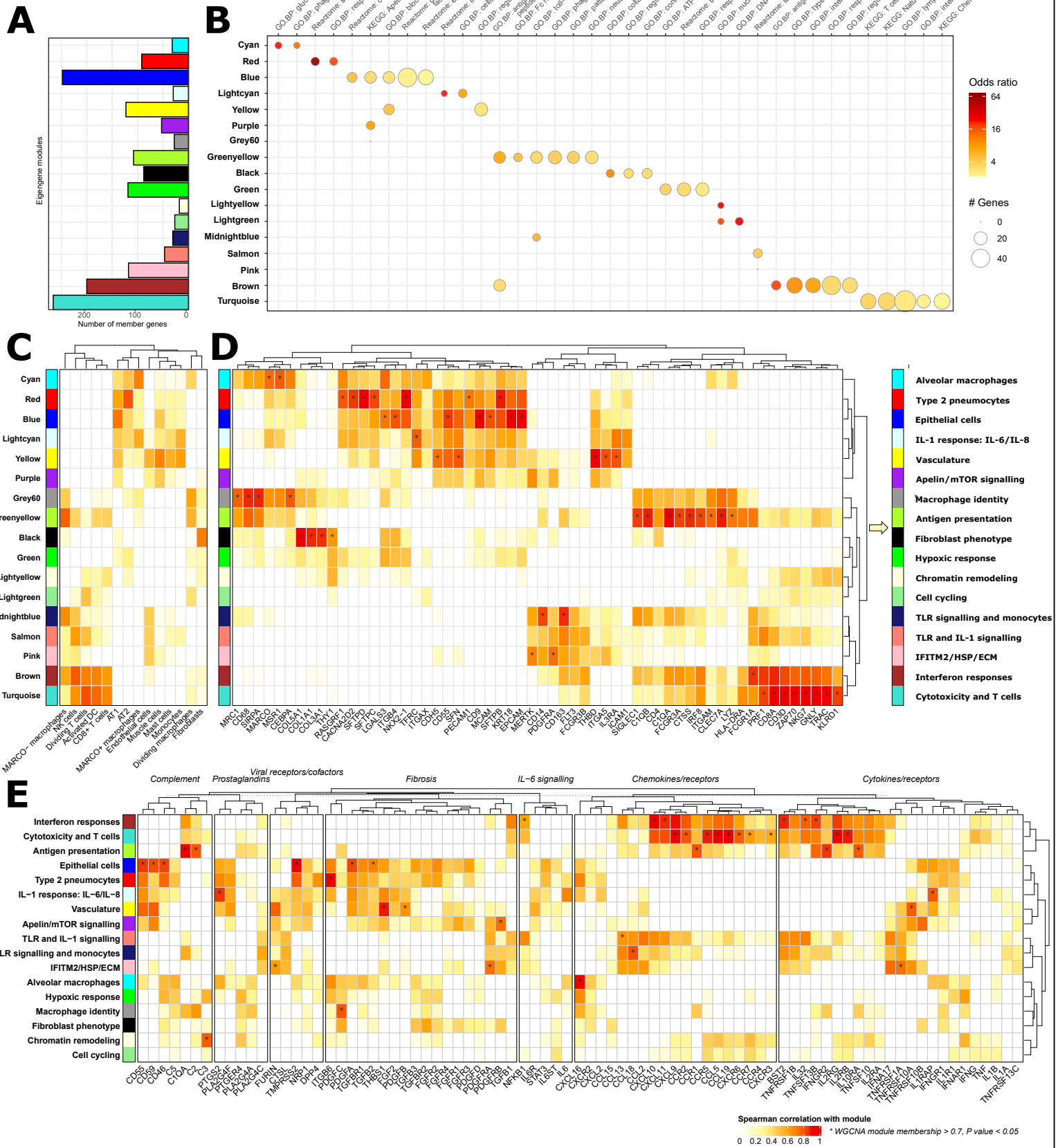
Figure 3

Figure 4

Figure 5
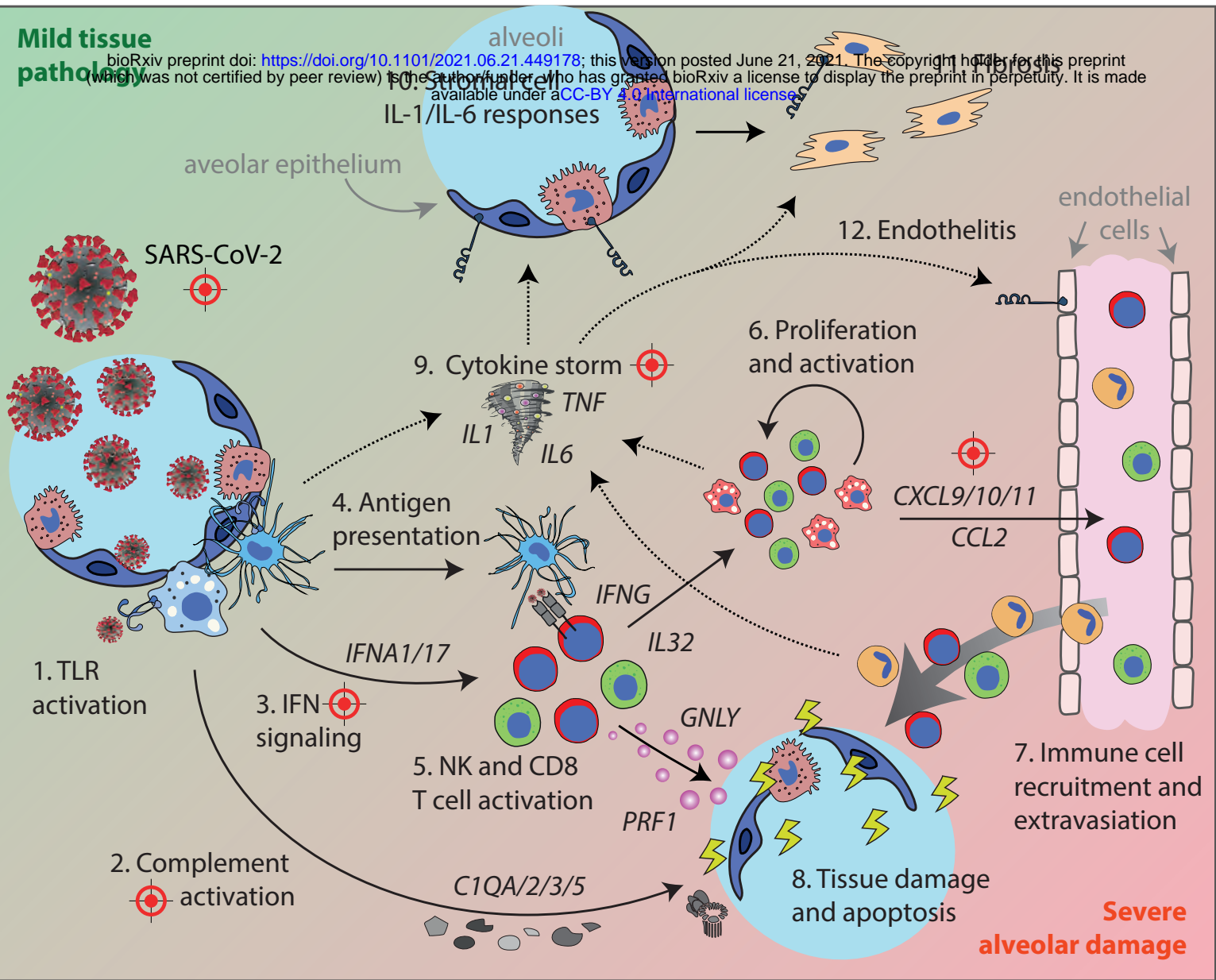
Supplementary figure 1

Supplementary figure 2