

# Integrating spatial and mononuclear transcriptome data to elucidate pulmonary microenvironment and cell communication during COVID-19 infection

Ning Zhang<sup>1,2,a,\*</sup>, Luyu Yang<sup>2,b</sup>, Yayu Li<sup>2,c</sup>

<sup>1</sup>School of Pharmacy, Xi'an Medical University, Xi'an, Shaanxi, China

<sup>2</sup>College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China

<sup>a</sup>zhangtyning@163.com, <sup>b</sup>yangluyu225@nwfau.edu.cn, <sup>c</sup>qingliwork@nwfau.edu.cn

\*Corresponding authors

**Abstract:** Although it has been established that the principal cause of death in COVID-19-infected patients is pneumonia and respiratory failure, little is currently known about the effects of COVID-19 on the lungs. Herein, we performed a single-nucleus RNA sequencing analysis of COVID-19 cases and controls using the human lung tissue data. Spatial transcriptomics was combined with single-nucleus RNA sequencing of human lung tissue to identify the specific cell subpopulations in COVID-19 cases and controls. We mapped ligand-receptor networks to specific cell types by combining single-cell and spatial data. Mapping the results revealed that fibroblasts are at the center of intercellular communication. These findings characterize the lung subpopulations, including fibroblasts and epithelial cells, the spatial niches in which they interact and the COVID-19 gene networks involved.

**Keywords:** COVID-19, Lung, Fibroblast cell, Spatial transcriptomics, Microenvironment, Cell communication

## 1. Introduction

According to the Real-time Statistics website, the cumulative number of coronavirus disease 2019 (COVID-19) deaths worldwide has exceeded 6.3 million, including more than 1 million in the United States [1]. The coronavirus disease pandemic has resulted in a sizable increase in individuals with acute respiratory distress syndrome (ARDS), associated with high morbidity and death [2-5].

Before, investigations of healthy persons' single nuclei RNA sequences (snRNA-seq) have shown the tissue effect of host receptors necessary for SARS-CoV-2 entrance [6, 7]. In a study by Blanco and Wilk et al., the bronchoalveolar lavage fluid and blood of patients with varying degrees of COVID-19 infection were analyzed to determine the impact of SARS-CoV-2 infection on immunological responses and cytokine dysregulation [8, 9]. Although snRNA-seq allows for a high-throughput investigation of cell transcriptomes, the geographical context of these transcriptomes is lost during tissue processing. In contrast, spatial transcriptomics (ST) enables profiling transcriptome-wide expression data and retains data on the spatial tissue context, which is hence well suited to studying cell interactions and spatial characteristics of lung gene expression.

Accordingly, we obtained snRNA-seq datasets of lung samples from patients that died from the virus and those with no history of COVID 19 infection. Using annotated snRNA-seq data, we categorized the spots and constructed a cell-type-specific atlas for cell types present in the spatial transcriptomics data. Subsequently, based on the spatial distribution atlas of COVID-19 specific cell types, we constructed the spatial cell communication atlas of COVID-19 specific receptor-ligand pairs in the lungs, which provided the foothold for further research.

## 2. Materials and Methods

### 2.1. snRNA-seq datasets, preparation and analysis of data

The snRNA-seq datasets were obtained from three samples of COVID-19 deaths and three non-COVID-19 infected cases. We profiled the human lung snRNA-seq dataset (GEO dataset GSE171524), which consisted of 26 490 cells using the 10× Genomics platform [10].

Data filtering and quality assurance were the first steps in data preparation. Low-quality cells and cells with outliers (in terms of library complexity) were excluded; we excluded cells from each sequencing library with fewer than 50 gene quantiles based on the gene distribution detected in each cell. Cells with more than 20% transcripts from mitochondrial genes were also eliminated. Then, genes were ordered by standard deviation, and the top 1500 genes with the highest variance were used for the research. A scale factor of 100,000 and log transformation was used to reduce dimensionality in the snRNA-Seq dataset; the total expression data were normalized by the log-transformed total expression value.

For lung clustering recognition, “FindNeighbors” and “FindClusters” were used with 30 PCA dimensions and the resolution was set to 0.8. Cluster cell identity was assigned by manual annotation using the released human lung snRNA-seq dataset. The R package clusterProfiler [11] was used to analyze significantly enriched Gene Ontology (GO) terms. The hallmark signature was scored using Seurat's “AddModuleScore” function with the default settings. At the same time, SCENIC analysis was conducted using tutorial-recommended parameters [12].

## **2.2. Analysis of spatial transcriptomics data**

The human lung 10× Genomics Visium ST datasets were obtained from GSE178361 [13]. Using Seurat, we normalized the data, extracted the 2D coordinates, and plotted the spatial pictures from the slices of 1,045 spots. Next, we used the Seurat package in R and custom scripts to analyze the gene point matrix generated by processing ST data samples [14]. Spots were screened to obtain the minimum detectable gene count of 200 genes for each patient's data, while genes with fewer than 10 read counts or expressed in fewer than two spots were eliminated. By analyzing the top 30 independent components (ICs), dimensions were reduced, and clustering was accomplished.

## **2.3. Spatial cell-cell interaction analysis**

stLearn a unique Python toolbox, was also used to analyze ST data from human lung spatial cell communication [15]. It is well-established that the stLearn leverages the morphological similarity of adjacent spots to normalize gene expression and eliminate the “dropout” noise inherent in technical shortcomings in RNA-seq technology. Using the human lung ST-seq data, we eliminated genes expressed in less than three locations. The counts per million approach then normalized the filtered gene count matrix followed by log transformation and scaling. We used CellPhoneDB's default settings and a curated database version for L-R prediction (3.0.0) [16]. Similar to the snRNA-seq study for Cell-Cell communication through L-R pairings, we applied NicheNet L-R prediction to 10× Genomics Visium ST data. Finally, tissue morphological data was utilized to normalize the gene count matrix by executing the stSME function. Cell-cell interaction was discovered in the human lung, STRING database (<https://string-db.org>) was used to identify ACE2 interaction genes, and stLearn was used to identify the spatial distribution and intensity of ACE2 interaction genes.

# **3. Results**

## **3.1. The cell atlas of COVID-19-infected and control lungs**

To examine the associations between patients with COVID-19 disease and patient characteristics, we first acquired a snRNA-Seq dataset on human lung samples from three COVID-19 decedents and three controls (unexpected death) from a separate study. After strictly filtering the original snRNA-seq expression matrix data, 26490 cells remained for the subsequent analysis. Following quality control, we used Seurat to determine main cell clusters in lungs from a separate study, which included epithelial cells, myeloid cells, fibroblasts, APC-like cells, T cells, endothelial cells, mast cells, neural cells, and B cell clusters, based on their tissue of origin (Figure 1A–C). We conducted differential gene expression (DGE) analysis between cell clusters (Figure 1D; Table S1) and gene set enrichment analysis of each cluster's elevated genes (Figure 1E). Gene ontology (GO) terms associated with “Sphingolipid signaling” and “Focal adhesion” were significantly enriched in Epithelial clusters, while “Phagosome” was enriched in compact Myeloid cells. The fibroblast clusters were involved in Focal adhesion with significant enrichment in ECM receptor interaction (Figure 1E). In addition, fibroblasts exhibited differential expression of LAMA2 (Figure 1D), which promotes lung cell metastasis. Furthermore, the proportion of cells inside each cluster was constant between repeats (Figure 1F).

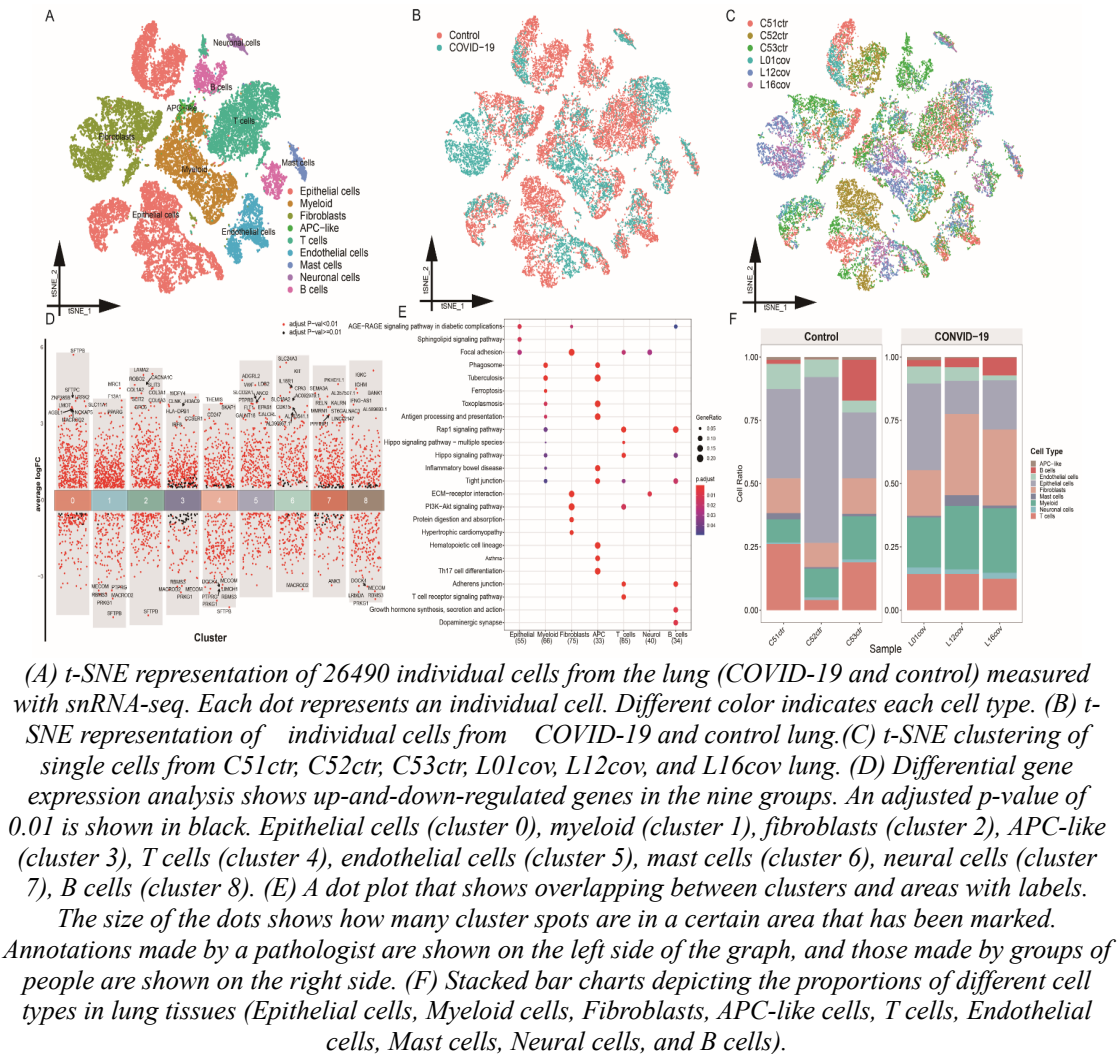


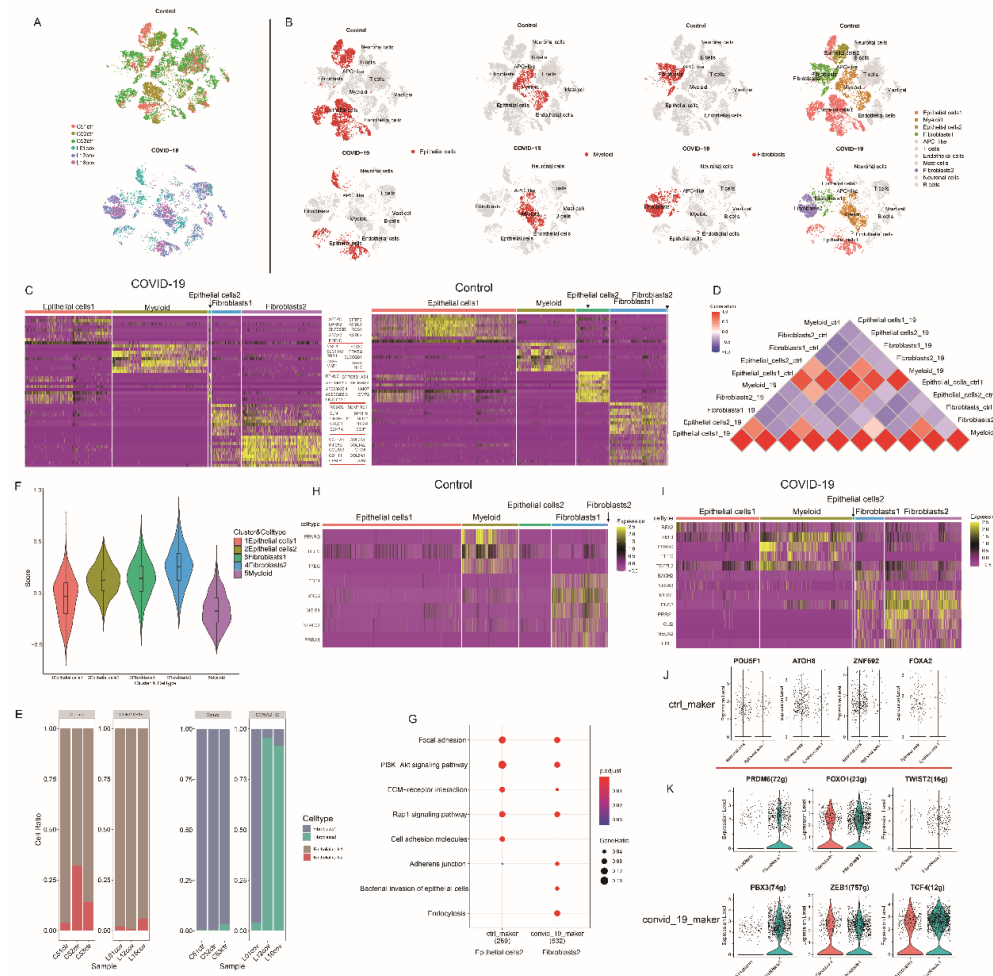
Figure 1: snRNA-seq atlas of the lung (COVID-19 and control).

### 3.2. The analysis of specific cell subsets

Interestingly, we found that COVID-19-infected lungs and the control lungs showed three main types of cells (Figure 2A, B). The proportion of cells was significantly different, including epithelial cells (20%), fibroblasts (14%), and myeloid cells (11%), which may be related to the recognized roles of epithelial cells and fibroblasts and their existence in COVID-19 and control lungs should be further investigated. Canonical correlation analysis [17] demonstrated that COVID-19-infected lungs and control lungs exhibited comparable (epithelial cells and fibroblasts) subpopulations and many common marker genes (Figure 2A–2C). Interestingly, COVID-19 and control cell clustering yielded a difference in subpopulation or cell type and a correlation between normal COVID-19 and control lungs (Figure 2D).

Twenty of the top 50 genes that were differentially expressed between COVID-19 and control lung cells exhibited significantly higher expression in COVID-19 tissue (log fold change >2.7 and p 0.001) (Figure 2F). The epithelial cell subpopulation comprised 22.2% to 42.3% of COVID-19 cells and overlapped with control lung cells (Figure 2E). The Fibroblast\_1 cell subpopulation exhibited more of COVID-19 and control lung cells across patients (Figure 2E). Many signaling pathways were significantly enriched in epithelial cells and fibroblasts, including the Focal adhesion PI3K-Akt signaling pathway, Rap1 signaling pathway, and Focal adhesion in epithelial cell and fibroblast subpopulations in COVID-19-infected and control lungs (Figure 2G). Epithelial cells\_2 was specifically enriched in the interaction between cell adhesion molecules and ECM-receptor. Fibroblast subtype\_2 cells were significantly enriched in adherens junction, Bacterial invasion, and Endocytosis of epithelial cells. Thus, COVID-19 and control lungs both contain Fibroblast cells that largely mimic the key cell states in normal lung Fibroblast and a normal-specific Epithelial subpopulation with no counterpart in normal control lungs.

The association of COVID-19 signature genes with cellular migration and extracellular matrix deconstruction in fibroblasts suggested an invasive behavior (Figure 2G). Moreover, Fibroblasts<sub>2</sub> exhibited the greatest expression of the Hallmark COVID-19 gene signature (n=50 genes) (Figure 2F). A prior study similarly analyzed fibroblasts and epithelial cells, but the expression of typical COVID-19 transcription factors (TFs) was not investigated (Figure 2H, J). Therefore, we conducted single-cell regulatory network inference and clustering (SCENIC), which identified key transcription factors that govern fibroblast transcription and gene expression (Figure 2 G and K). Our findings indicated dysregulated differentiation in fibroblasts and epithelial cells, which may be attributed to the following reasons: (1) Epithelial cells exhibit diminished proliferation capacity in COVID-19 lungs. (2) COVID-19 virus invasion of epithelial cells is related to fibroblasts. (3) Fibroblasts and epithelial subsets expressing COVID-19-related genes are formed.



(A) T-SNE of normal lungs and COVID-19 lungs; (B) t-SNE classifying normal and COVID-19 lungs into nine-cell types. Normal lungs and COVID-19 lungs contain normal fibroblasts and epithelial cell subpopulations. (C) The expression of the top 10 shared genes between fibroblasts and epithelial cells, fibroblast subpopulation, epithelial cell subpopulation and myeloid cell marker genes. (D) Gene overlaps correlation matrix of COVID-19 and normal lungs across all cell subpopulations in snRNA-seq clusters. (E) Stacked bar charts depicting the proportions of different cell subpopulations in lung tissues (Epithelial cells, Fibroblasts). (F) Violin plots of the normal lungs and COVID-19 lungs subpopulations with characteristic marker gene signature scores. (G) Gene ontology (GO) keywords for the top 50 genes that are up-and down-regulated in the subpopulation with differential expression between normal and COVID-19 lungs are shown as a dot plot. The tissues from COVID-19 and normal lungs were contrasted with a subset of differentiated cells. (H, I) Heatmap of the expression of different cell subpopulation-related transcription factors in normal and COVID-19 lungs. (J, K) Violin plots of single-cell regulon scores deduced by SCENIC (g, genes).

Figure 2: Differential expression between normal lungs and COVID-19 lungs.

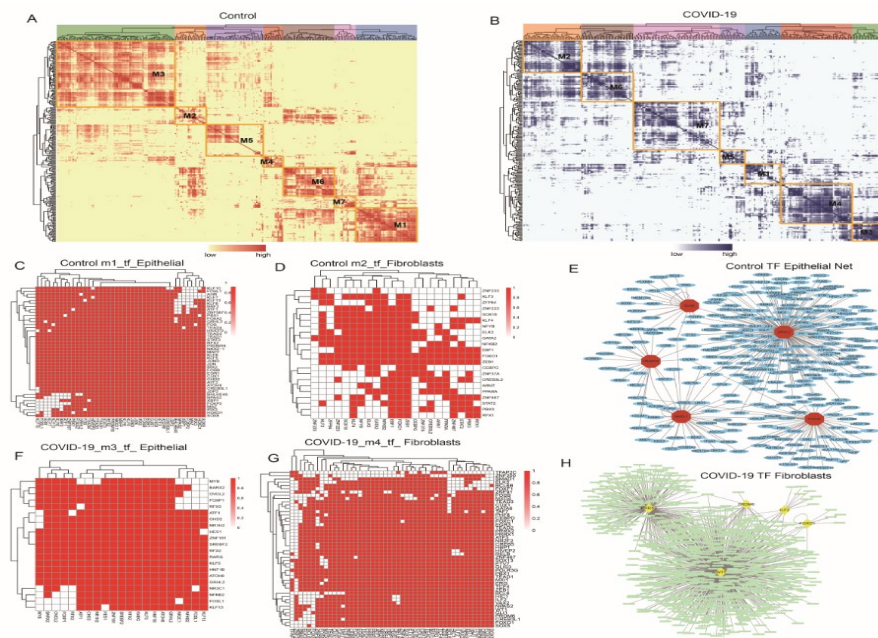
### 3.3. The co-expression patterns of COVID-19 and control modules

It is well-established that transcription factors interact to regulate gene expression levels. To comprehensively define the unique combinatorial patterns, we assessed the atlas-wide similarity of RAS scores for each pair of regulons using the Connection Specificity Index. Transcription factors frequently work together to control gene expression levels. The Connection Specificity Index (CSI) was used to analyze the atlas-wide similarity of RAS scores for each pair of regulons to identify the combinatorial patterns fully. Under normal conditions, 202 regulatory factors were identified in the lung, divided into seven main components. (Figure 3A). About 243 regulons of COVID-19 lungs were arranged into seven primary modules (Figure 3B).

We picked several representative regulators and cell types for each module based on their average activity ratings (Figure 1A). When we projected each module's average activity score onto the tSNE map, we noticed that each module was in its zone, and all highlighted parts had complementary patterns, as shown in Supplementary Figure 1A.

We then concentrated on the normal status modules M1 and M2, which comprised 43 and 22 regulons, respectively (Figure 3A). The M1 module was significantly associated with epithelial cell types, and the epithelial cell co-expressed transcription factors were mainly enriched in RNA polymerase II transcription factor activity and sequence-specific DNA binding (Figure 3C). The M2 module was significantly associated with fibroblasts (Figure 3D). Given that transcription factors achieve their corresponding biochemical functions through target genes, we further analyzed the top 5 TFs of epithelial cells targeting the regulatory network in the normal lungs (Figure 3E).

Next, we focused on modules M3 and M4 in COVID-19 patients, which contained 21 and 48 regulons (Figure 3B). The M3 module was significantly associated with epithelial cells (Figure 3F), while the M4 module was significantly associated with fibroblasts. The epithelial cell co-expressed transcription factors were mainly enriched in RNA polymerase II transcription factor activity and sequence-specific DNA binding (Figure 3G). Given that transcription factors achieve relevant biochemical functions through target genes, we further analyzed the top 5 TFs of fibroblasts targeting the regulatory network in COVID-19 (Figure 3H). In conclusion, key transcription factors in fibroblasts and epithelial cells have important roles in regulating the microenvironment of healthy humans and COVID-19-infected lungs.



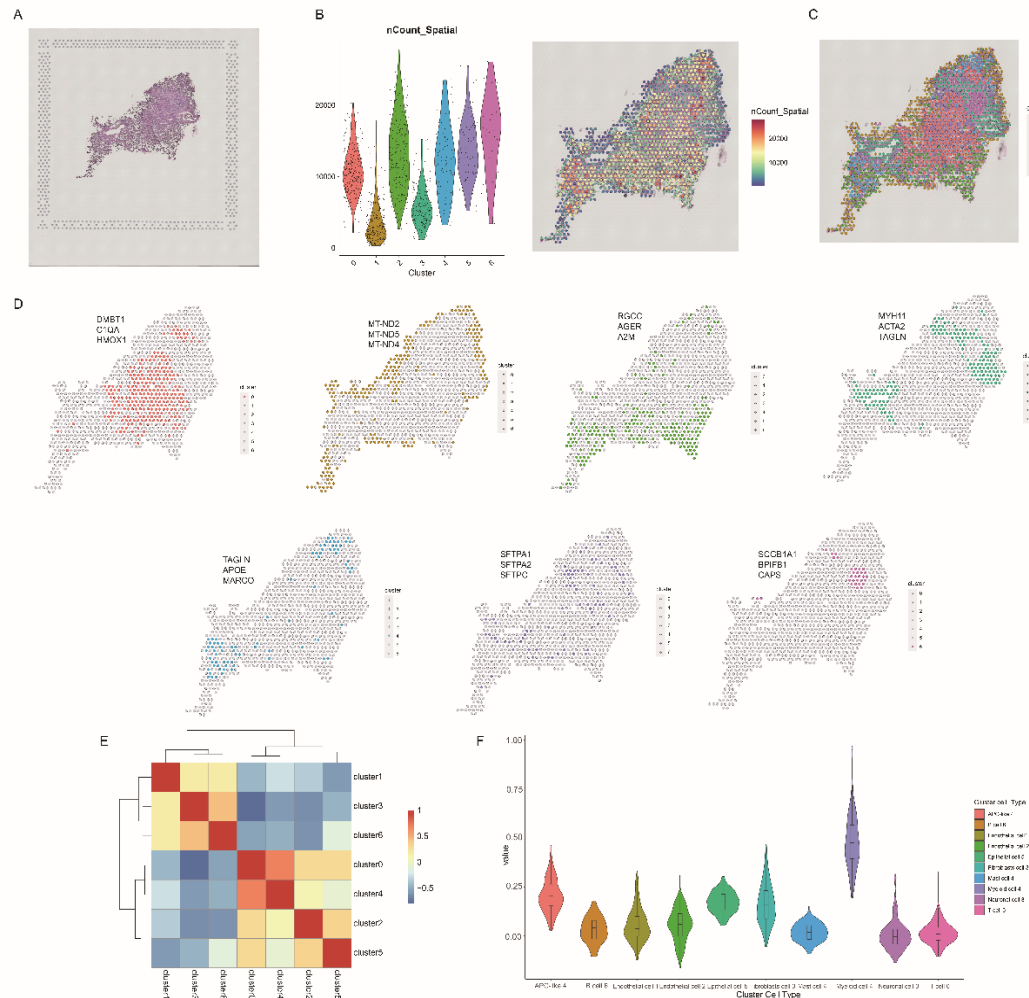
(A, B) Regulon modules were identified using the regulon connection specificity index (CSI) matrix and sample transcription factors, their associated binding motifs, and cell types. (C, D) A zoomed image of modules M1 and M2 reveals the structure of their sub-modules. (E, F) A zoomed image of modules M3 and M4 reveals the structure of their sub-modules. (G) Top 5 Transcription factor-target gene regulatory network in fibroblasts from COVID-19 tissues. (H) Top 5 Transcription factor-target gene regulatory network in epithelial cells from control tissues.

Figure 3: Identification of combinatorial regulon modules.



### 3.4. Atlas of lung spatial transcriptomic cells

To evaluate the spatial position of epithelial cell and fibroblast populations, we conducted ST on selected distal lung sections (Figure 4A). After data quality filter processing, the transcriptome of 1,045 spots was collected (Figure 4B). Seven clusters were obtained by clustering the ST data, and each cluster's marker top 3 genes were obtained for annotation (Figure 4C and 4D). The location of each region was scored using snRNA-seq signature genes for epithelial cells and fibroblasts. These high clusters of epithelial cells and fibroblasts were comparable in different individuals (Figure 4E and 4F).



(A) Hematoxylin and eosin staining of tissue sections (H&E). (B) Violin plots of genes per point across duplicate tissue segment replicate. (C, D) Clusters of adnexal spots projected independently with the top differentially expressed genes found. Clusters of lesions connected with the lungs (clusters covering annotated areas in sections). (E) and (F) Matrix of overlapping genes that are differentially expressed top maker gene in ST clusters across the lung (E). ST Cluster Signature ( $n = 10$  genes) was constructed using highlighted equivalent spatial clusters and violin plots of the ST cluster top marker gene score by cell type in snRNA-seq data (F).

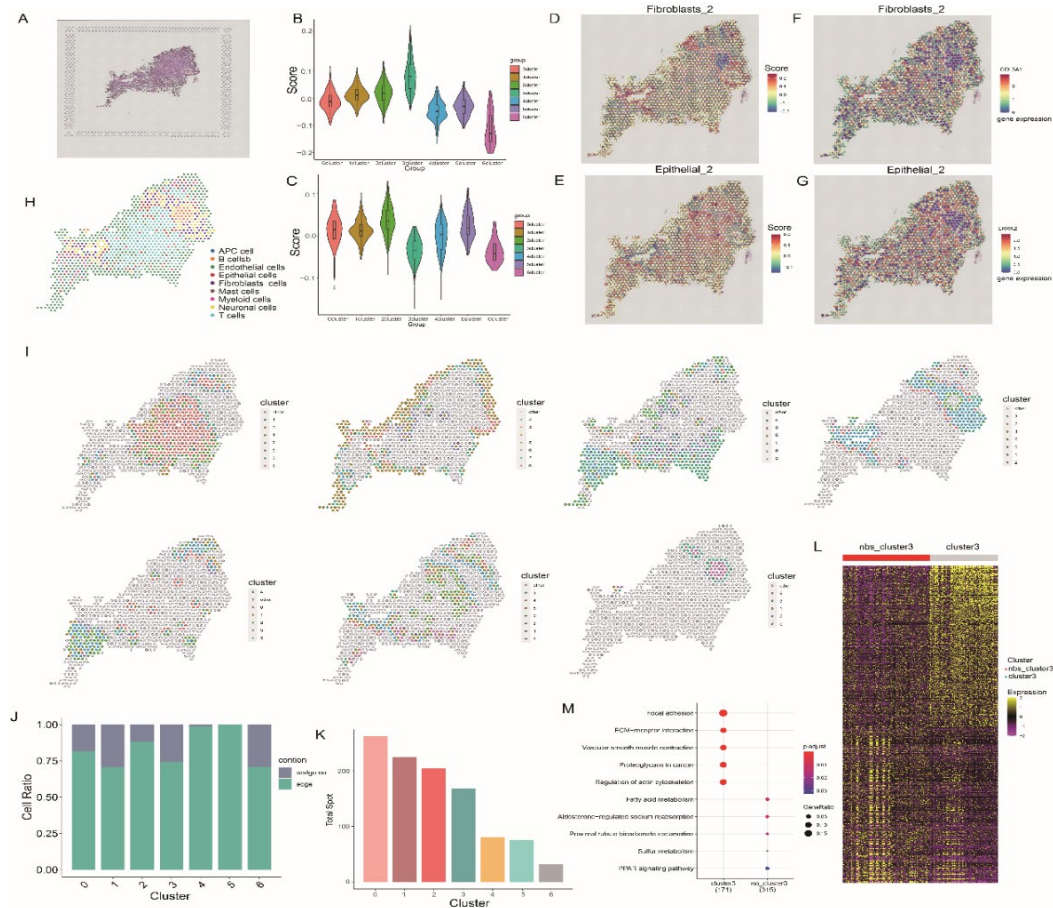
Figure 4: Spatial transcriptomics atlas of lung distal lung sections.

### 3.5. Spatial transcriptomics reveals significant heterogeneity

snRNA-seq demonstrated a clear cluster of epithelial cells (Figure 5A, 5B and 5D) and fibroblasts (Figure 5A, 5C and 5E), including the Fibroblast\_2 cell and Epithelial\_2 cells marker COL3A1 and LRRK2, in each patient (Figure 5F and 5G). In clearly separated lung clusters, the surrounding stroma was enriched in COVID-19-associated fibroblasts and endothelial transcripts, indicating the presence of a fibrovascular niche (Figure 5H).

More than 70% of fibroblast high cluster spots were found at cluster 3 (Figure 5J, 5K), despite the

cluster points being few at the compartment; the remaining neighborhood points were enriched in Fibroblast\_2 top marker genes (Figure 5I). At the staining for the spatial expression of the COL17A1 gene revealed the presence of both lung fibroblast and fibroblast 2. Furthermore, signaling genes for focal adhesion function were expressed along the neighborhood of fibroblasts, indicating that focal adhesion function was present in the snRNA-seq fibroblast 2 cell population. (Figure 5F). We further found that fibroblasts 2 cells were significantly enriched in focal adhesion, ECM-receptor interactions, and vascular smooth muscle contraction. Fibroblast 2 cell edge genes were enriched in fatty acid metabolism, aldosterone-regulated sodium reabsorption and PPAR signaling pathway (Figure 5L and 5M).



(A) Staining of tissue slices with hematoxylin and eosin (H&E). (B, C) Violin plots of the Fibroblasts and Epithelial cells subpopulation scores for individual places produced by snRNA-seq data for each group cluster. Clusters with the highest average score are shown by dotted boxes. (D, E, F, G) Spatial marker plots of fibroblasts subpopulation and epithelial subpopulation-high cluster, fibroblasts subpopulation, marker COL3A1 and LRRK2 expression in tissue sections. (I) Cluster spot annotated with the lung's neighborhood and isolated neighborhood spots indicated by the cluster. (J) Spot spatial edge ratio about the cluster: (K) Spot Spatial edge count about the cluster: (L) Heat map of the top 500 differentially expressed genes between cluster3 and no\_cluster3. (M) A dot plot of selected GO terms linked to differentially expressed genes in each cluster; as well as projections of non-fibroblast and epithelial subpopulation leading edge-related clusters. The total amount and proportions of spots at the group cluster leading edge are shown in bar graphs.

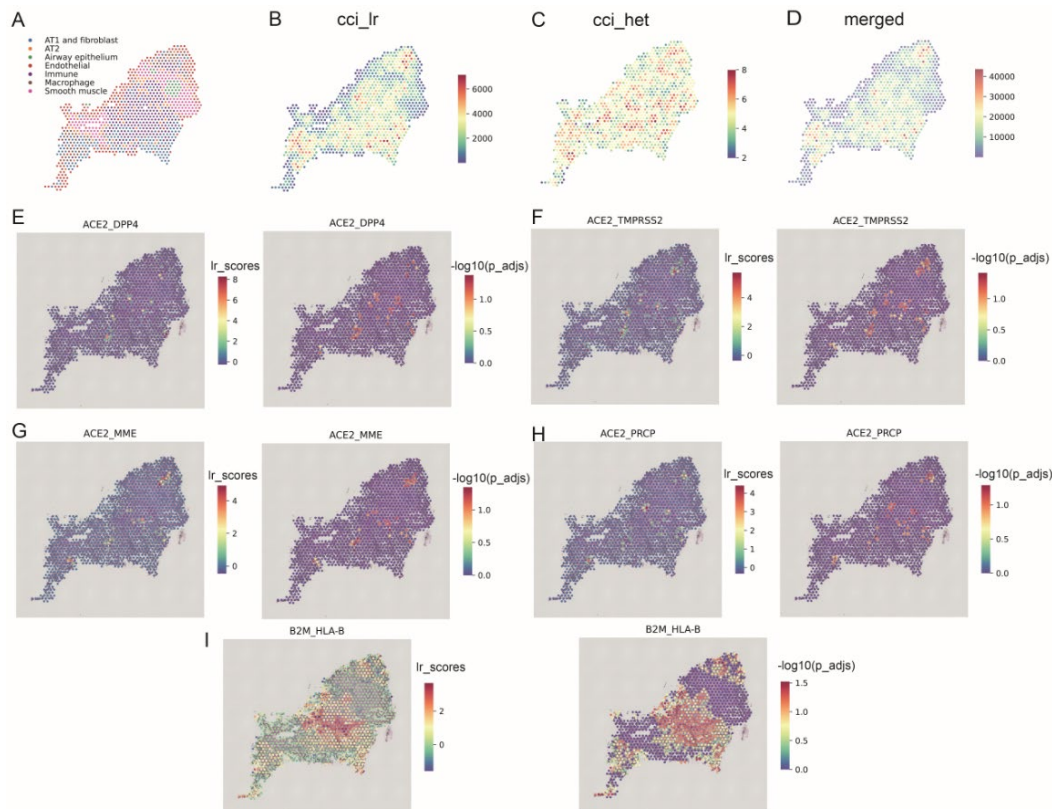
Figure 5: Spatial transcriptomics reveals significant heterogeneity.

### 3.6. Spatial cell communication of ACE2 in the lung

Subsequently, we conducted snRNA-seq and ST data integration annotations to describe the spatial location of cell types (Figure 6A). We expanded our study of human lung ST-seq data using a stLearn to investigate spatial cell communication between lung cells [15]. The lung comprises immune cells, alveolar epithelial type I and II cells (AT1 and AT2), airway epithelium, and fibroblasts, surrounded by an endothelial capsule bordered by macrophages (Figure 6A). The stLearn tool was used to examine

contact between podocytes and mesangial cells as ligand-receptor (L-R) gene co-expression inside ST-spots (Fig. 6B-D). First, we analyzed L-R pairings selected in the CellphoneDB database [16, 18]. Using stLearn [15], we mapped L-R gene spatial co-expression inside and between ST-spots of lung regions, discovering hundreds of possible connections (Figure 6B-D).

Next, we tested ACE2 gene published L-R pairings, including the ACE2\_DPP4, ACE2\_TMPRSS2, ACE2\_MME, and ACE2\_PRCP, which govern Fibroblast and Epithelial cell behavior (Figure 6E-H). The genes encoding these two proteins were co-expressed in several ST spots that covered a bigger area than COVID-19 clusters had previously revealed. In addition, B2M\_HLA-B had the greatest number of spots involved in overall distal lung slice space cell communication. Finally, we found that key gene locations for ACE2 interaction were associated with annotated fibroblasts.



(A) Spatial spots were assigned to different cell types via label transfer<sup>18</sup> from a published human lung dataset [13], which identified immune cells, endothelial cells, smooth muscle, macrophage, AT2 and airway epithelium within the tissue area; Spatial spots annotate each cluster used in the published articles. (B) Ligand-receptor co-expression in the neighboring spots. (C) Count cell type diversity for between-spot mode. (D) Calculation of the merged (B) and (C) spatial cell communication scores (between-spots). (E, F, G, H) COVID-19-specific receptor ligand spatial cell communication score and  $-\log(p\_adj)$ . (I) Lung receptor ligand spatial cell communication score and  $-\log(p\_adj)$ .

Figure 6: Lung L-R interactions are used to infer space cell communication relationships.

#### 4. Discussion

In this study, we used snRNA-seq to create a single-cell transcriptome atlas of cell subpopulations in healthy and COVID-19 lung states and integrated spatial profiling with ST and snRNA-seq to characterize the cell living environment. Our integrated approach to spatial profiling overcame the limitations of each method alone. Future efforts may reach the size of ST with the resolution of a single cell, enabling a comprehensive examination of structures. Our additional inference of the interaction between spatial cell communication and spatial neighborhood analysis provided a granular view of the deleterious effects of SARS-CoV-2 infection on the lung.

Toni et al. uncovered substantial remodeling in the lung epithelial, immune and stromal compartments, with evidence of multiple paths of failed tissue regeneration, including defective alveolar type 2 differentiation and expansion of fibroblasts and putative TP63<sup>+</sup> intrapulmonary basal-like progenitor



cells [19]. Our findings are consistent with these findings to a certain extent. We expanded on this idea by analyzing a distinct cell subset in the lungs of three COVID-19 and three non-COVID infected cases, with one subpopulation causing extensive lung-specific gene expression. The finding of Fibroblast1-enriched genes involved in the lung microenvironment and intercellular communication with epithelial and myeloid cells suggests that fibroblasts generate the fibroblast and endothelial states.

Our work provides novel insight into the host's responses to deadly SARS-CoV-2 infection, although the sample size was limited. In addition, we studied only part of the possible disease phenotype since our study was limited to lung tissue associated with COVID-19 death. Nonetheless, certain breakthroughs, such as the fast development of and differentiation of pulmonary fibroblasts and epithelial cells [20], may be important for patients who have survived severe COVID-19 and may help us understand the long-term problems observed in these patients.

Finally, we integrated a single-cell transcriptome and spatial transcriptome lung atlas from lung tissue samples and revealed COVID-19 receptor and ligand spatial interaction. However, some SARS-CoV-2 variants resulted in higher resistance to available vaccines or treatments [21-23]. We believe this cell communication approach is critical for studying the spatial location of disease-related receptor and ligand interactions and distribution on a physiological and vaccine molecular basis. This atlas is a valuable resource for studying host responses to SARS-CoV-2 and understanding possible long-term pulmonary sequelae from COVID-19, as well as laying the groundwork for developing future therapies.

### Acknowledgement

This work was financially supported by the Shangluo City Science and Technology Plan Project (grant numbers: 2023-N-0028). Xi'an Medical College University Research Fund Program (2024BS09); Shaanxi Provincial Sports Bureau Scientific Research Program (20240429).

### References

- [1] *Worldometer - real time world statistics*. <https://www.worldometersinfo/coronavirus>.
- [2] Jester, J.V., et al., *Radial keratotomy in non-human primate eyes*. *Am J Ophthalmol*, 1981. 92(2): p. 153-71.
- [3] Bellani, G., et al., *Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries*. *JAMA*, 2016. 315(8): p. 788-800.
- [4] Brault, C., et al., *COVID-19- versus non-COVID-19-related Acute Respiratory Distress Syndrome: Differences and Similarities*. *Am J Respir Crit Care Med*, 2020. 202(9): p. 1301-1304.
- [5] Kerchberger, V.E. and J.A. Bastarache, *Pulmonary Vasculopathy in COVID-19 ARDS: A Step Closer to the Full Picture*. *Am J Respir Crit Care Med*.2022 Oct 1; 206(7):809-810.
- [6] Antoni, G., et al., *In Vivo Visualization and Quantification of Neutrophil Elastase in Lungs of COVID-19 Patients - A First-In-Human Positron Emission Tomography Study with (11) C-GW457427*. *J Nucl Med*, 2023, 64(1):145-148.
- [7] Muus, C., et al., *Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics*. *Nat Med*, 2021. 27(3): p. 546-559.
- [8] Blanco-Melo, D., et al., *Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19*. *Cell*, 2020. 181(5): p. 1036-1045 e9.
- [9] Wilk, A.J., et al., *A single-cell atlas of the peripheral immune response in patients with severe COVID-19*. *Nat Med*, 2020. 26(7): p. 1070-1076.
- [10] Melms, J.C., et al., *A molecular single-cell lung atlas of lethal COVID-19*. *Nature*, 2021. 595(7865): p. 114-119.
- [11] Wu, T., et al., *Cluster Profiler 4.0: A universal enrichment tool for interpreting omics data*. *Innovation (N Y)*, 2021. 2(3): p. 100141.
- [12] Aibar, S., et al., *SCENIC: single-cell regulatory network inference and clustering*. *Nat Methods*, 2017. 14(11): p. 1083-1086.
- [13] Kadur Lakshminarasimha Murthy, P., et al., *Human distal lung maps and lineage hierarchies reveal a bipotent progenitor*. *Nature*, 2022. 604(7904): p. 111-119.
- [14] Hao, Y., et al., *Integrated analysis of multimodal single-cell data*. *Cell*, 2021. 184(13): p. 3573-3587 e29.
- [15] Duy Pham, X.T., Jun Xu, Laura F. Grice, Pui Yeng Lam, Arti Raghobar, Jana Vukovic, Marc J. Ruitenberg, Quan Nguyen, *stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues*. *bioRxiv*, 2020.

- [16] Efremova, M., et al., *CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes*. *Nat Protoc*, 2020. 15(4): p. 1484-1506.
- [17] Rendeiro, A.F., et al., *The spatial landscape of lung pathology during COVID-19 progression*. *Nature*, 2021. 593(7860): p. 564-569.
- [18] Browaeys, R., W. Saelens, and Y. Saeys, *NicheNet: modeling intercellular communication by linking ligands to target genes*. *Nat Methods*, 2020. 17(2): p. 159-162.
- [19] Delorey, T.M., et al., *COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets*. *Nature*, 2021. 595(7865): p. 107-113.
- [20] John, A.E., et al., *COVID-19 and pulmonary fibrosis: A potential role for lung epithelial cells and fibroblasts*. *Immunol Rev*, 2021. 302(1): p. 228-240.
- [21] Senevirathne, TH., *COVID-19: From emerging variants to vaccination*. *Cytokine Growth Factor Rev*. 2024 Apr; 76: 127-141
- [22] Huang, TC., *Structure-based approaches against COVID-19*. *J Chin Med Assoc*. 2024 Feb 1; 87(2):139-141.
- [23] Beladiya J., et al., *Safety and efficacy of COVID-19 vaccines: A systematic review and meta-analysis of controlled and randomized clinical trials*. *Rev Med Virol*. 2024 Jan; 34(1):e2507.