



## OPEN ACCESS

## EDITED BY

Diana Boraschi,  
Shenzhen Institute of Advanced  
Technology (SIAT) (CAS), China

## REVIEWED BY

Sukhwinder Singh Sohal,  
University of Tasmania, Australia  
Shixue Gou,  
Guangzhou Institutes of Biomedicine  
and Health (CAS), China

## \*CORRESPONDENCE

Dan Liu  
liudan10965@wchscu.cn  
Lanlan Zhang  
llzhang@scu.edu.cn  
Xin Zhang  
xinzhang201618@gmail.com

## SPECIALTY SECTION

This article was submitted to  
Cytokines and Soluble  
Mediators in Immunity,  
a section of the journal  
Frontiers in Immunology

RECEIVED 23 June 2022

ACCEPTED 24 August 2022

PUBLISHED 29 September 2022

## CITATION

Zhang L, Tang C, Zhang M, Tong X,  
Xie Y, Yan R, Wang X, Zhang X, Liu D  
and Li S (2022) Single cell  
meta-analysis of EndMT  
and EMT state in COVID-19.  
*Front. Immunol.* 13:976512.  
doi: 10.3389/fimmu.2022.976512

## COPYRIGHT

© 2022 Zhang, Tang, Zhang, Tong, Xie,  
Yan, Wang, Zhang, Liu and Li. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use,  
distribution or reproduction is  
permitted which does not comply with  
these terms.

# Single cell meta-analysis of EndMT and EMT state in COVID-19

Lanlan Zhang<sup>1\*</sup>, Chuang Tang<sup>2</sup>, Min Zhang<sup>3</sup>, Xia Tong<sup>2,4</sup>,  
Yingying Xie<sup>5</sup>, Ruitong Yan<sup>6</sup>, Xiangjun Wang<sup>7</sup>, Xin Zhang<sup>2,4\*</sup>,  
Dan Liu<sup>1\*</sup> and Shasha Li<sup>8,9</sup>

<sup>1</sup>Division of Pulmonary Diseases, State Key Laboratory of Biotherapy, And Department of Respiratory and Critical Care Medicine, West China Hospital of Sichuan University, Chengdu, China, <sup>2</sup>Department of Gastroenterology, West China (Airport) Hospital, Sichuan University, Chengdu, China, <sup>3</sup>Oncology Business Department, Novogene Co., Ltd, Beijing, China, <sup>4</sup>Department of Gastroenterology, West China Hospital of Sichuan University, Chengdu, China, <sup>5</sup>Department of Nephrology, Seventh Affiliated Hospital Sun Yat-sen University, Shenzhen, China, <sup>6</sup>Sichuan University, Chengdu, China, <sup>7</sup>First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, <sup>8</sup>Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, United States, <sup>9</sup>Harvard Medical School, Boston, MA, United States

COVID-19 prognoses suggests that a proportion of patients develop fibrosis, but there is no evidence to indicate whether patients have progression of mesenchymal transition (MT) in the lungs. The role of MT during the COVID-19 pandemic remains poorly understood. Using single-cell RNA sequencing, we profiled the transcriptomes of cells from the lungs of healthy individuals (n = 45), COVID-19 patients (n = 58), and idiopathic pulmonary fibrosis (IPF) patients (n = 64) human lungs to map the entire MT change. This analysis enabled us to map all high-resolution matrix-producing cells and identify distinct subpopulations of endothelial cells (ECs) and epithelial cells as the primary cellular sources of MT clusters during COVID-19. For the first time, we have identified early and late subgroups of endothelial mesenchymal transition (EndMT) and epithelial-mesenchymal transition (EMT) using analysis of public databases for single-cell sequencing. We assessed epithelial subgroups by age, smoking status, and gender, and the data suggest that the proportional changes in EMT in COVID-19 are statistically significant. Further enumeration of early and late EMT suggests a correlation between invasive genes and COVID-19. Finally, EndMT is upregulated in COVID-19 patients and enriched for more inflammatory cytokines. Further, by classifying EndMT as early or late stages, we found that early EndMT was positively correlated with entry factors but this was not true for late EndMT. Exploring the MT state of may help to mitigate the fibrosis impact of SARS-CoV-2 infection.

## KEYWORDS

COVID-19, endothelial-mesenchymal transition (EndMT), epithelial-mesenchymal transition (EMT), single cell RNA sequencing, endothelial cells (ECs)

**Abbreviations:** COVID-19, Coronavirus disease 2019; MT, Mesenchymal transition; EndMT, Endothelial-mesenchymal transition; EMT, Epithelial-mesenchymal transition; IPF, Idiopathic pulmonary fibrosis.

## Introduction

Alveolar epithelial type II cells (AEC2s) produce surfactants and serve as progenitors (1). Type II alveolar epithelium is reduced due to massive necrosis with SARS-COV-2 infection in patients (2). This leads to decreased barrier permeability and accumulation of proteinaceous edema fluid in the alveolar cavity, resulting in hypoxemia and even acute respiratory distress syndrome (ARDS), with a prognosis including fibrosis formation (3). Epithelial-mesenchymal transition (EMT) is the process by which epithelial cells differentiate into mesenchymal (fibroblast-like) cells expressing mesenchymal biomarkers such as  $\alpha$ -Smooth Muscle Actin ( $\alpha$ -SMA), and N-cadherin (4, 5). Therefore, we speculated that EMT might underlie the mechanism of lung fibrosis observed with ARDS in COVID-19 patients.

Early-stage and late-stage mesenchymal transition (MT) processes are considered the most critical pathogenic mechanisms of fibrosis in diseases like idiopathic pulmonary fibrosis (IPF), pulmonary hypertension, coronary artery disease, and other vascular-related disorders (6). Once endothelial cells (ECs) are damaged and the collagen and matrix are exposed, in addition to activation of the coagulation system, immune cells are recruited, leading to the release of large amounts of inflammatory factors are then released (7). Endothelial-mesenchymal transition (EndMT) occurs when ECs respond to chronic inflammation, transforming themselves into a more aggressive mesenchymal state (8, 9). In the early stages, irreversible vascular damage or EndMT, mainly presents more robust endothelial markers (e.g., vascular endothelial cadherins (VE-cadherins), *CD31* and *Tie1/2*) and expresses weaker mesenchymal markers (e.g., *N-cadherin*, *fibroblast specific protein-1* or *S100A4*, *fibronectin*, *vimentin*, *SM22- $\alpha$* , *calponin*, and  *$\alpha$ -smooth muscle actin*) (10, 11). Otherwise, late-stage EndMT presents weaker endothelial markers but with more robust mesenchymal markers.

EMT also has a similar process with the presence of EndMT as well as MT markers in the early stages (12, 13). However, in the late stage, epithelial markers are lost, and fibroblast markers are predominant (12, 13). Therefore, we hypothesized that EMT and EndMT may underlie factors implicated in the prolonged ventilator dependence and high mortality rates observed in hospitalized COVID-19 patients.

Herein, we draft EndMT and EMT by integrated analysis of 167 single-cell and single-nucleus RNA-sequencing (scRNA-seq and snRNA-seq, respectively) samples. We performed the first single-cell RNA sequencing meta-analysis associating COVID-19 with underlying contributor genes for MT. We have also identified specific gene programs enriched in MT-associated genes with fibrosis and highlight other entry factors that are significantly expressed in the lungs, which may play a role in SARS-CoV-2 infection.

## Materials and methods

### Data collection

We obtained scRNA/snRNA sequencing data from Gene Expression Omnibus database (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The majority of the data used in this manuscript is publicly available from previously published studies: COVID-19:GEO accession GSE171524, GSE171668, GSE149878, GSE161382 and GSE163919 (14–17). IPF: GEO accession GSE136831 and GSE135893 (18, 19). The Bulk RNA data is also available in the GEO database (accession number GSE47460) (20).

### scRNA/snRNA-seq data processing

We analyzed the scRNA and snRNA sequencing data using R (version 4.1.3) and Seurat v 4.0.2 (21). To eliminate low quality/dead cells or empty droplets, any single cells containing less than 200 genes or greater than 5,000 genes, or cells with more than 10% mitochondrial genes were removed. We next removed doublet contamination. As a result. We detected 49,795 genes in a total of 612,151 cells. Next, we performed `LogNormalize()` to normalize the gene expression data, used the `FindVariableFeatures()` function to identify genes whose expression was highly variable between cells, and used `ScaleData()` to scale the data. The `"RunPCA"` function was used for principal component analysis (PCA). Using the default parameters, Harmony package (22) was utilized to combine data and eliminate batch effect, Using the default parameters, Harmony performs de-batching by single patients. Subsequently, the top 20 statistically significant principal components were used in the `"FindNeighbors"` function. Cells were clustered (Cluster resolution = 0.6) by using the `"FindClusters"` function and visualized with the UMAP method.

### Cell type identification and data visualization

Cell type identification was mostly accomplished with the use of "FindAllMarkers," classical cell marker genes, R-packages (cluster, singscore), and cell type annotation using a mixture of known cell marker genes. *EPCAM*<sup>+</sup> epithelial cells, *CLDN5*<sup>+</sup> endothelial cells, *PTPRC*<sup>+</sup> immune cells, and *COL1A2*<sup>+</sup> (mesenchymal cells). Cell clustering was performed using the Seurat "FindNeighbors" and "FindClusters" functions, and the final Seurat objects were created after UMAP downsampling. The heatmap was created using the R package ggplot2 in Seurat, and umap maps, violin maps, bar maps, heat maps, and point maps were created.

## Enrichment analysis

Enriched gene ontology biological (<http://geneontology.org/>) processes with a false discovery rate less than 5 percent were identified using the Gene Ontology Resource. KEGG (<http://www.kegg.jp>) enrichment analyses were carried out with the Fisher's exact test, and FDR correction for multiple testing was also performed. Gene Set Enrichment Analysis (GSEA) were performed with clusterProfile. The three methods determine whether a set of genes has statistically significant differences between two biological states (23).

## Trajectory analysis

The R package monocle3 v1.0.0 was used to construct single cell trajectory analysis. Cells marked as EndMT/EMT and their subsets information were input and constructed into a monocle object (24). The left side of the heat map, created with the R package monocle2, shows the signaling pathways enriched to the gene set. Each row of the heat map (ranging from red to blue) represents a gene; each column represents a proposed time point; and the color represents the average expression value of the gene at the current time point (from high to low expression, respectively) (25).

## Bulk RNA data

Univariate analysis was performed by using the Pearson's correlation coefficient of gene expression data and diffusing capacity of the lung for carbon monoxide (DLCO). The data is available in the Gene Expression Omnibus database (accession number GSE47460).

## Statistical analysis of functional data

The R language or the GraphPad Prism software, version 9 (San Diego, California USA, [www.graphpad.com](http://www.graphpad.com)) were used to perform all computations and analyses. T-tests were used for comparisons between groups, and one-way ANOVA was used for comparisons between multiple groups. A statistically significant difference was defined as a p value <0.05. The Wilcoxon rank sum statistical test was used to examine differentially expressed genes in each cluster. The meta package was used to perform the meta-analysis in R.

## Results

### Single-cell transcriptome sequencing reveals the presence of fibrosis-associated cell subpopulations

To investigate the contribution of EndMT in SARS-CoV-2 infection in the lungs, we analyzed existing scRNA/snRNA-seq

datasets to assess which clusters express EndMT markers. In a previously unpublished dataset consisting of COVID-19 and IPF lung tissue ( $n=167$ ), we recovered at least 14 distinct major cell types, including macrophages, monocytes, neutrophils, T cells, B cells, DCs, granulocytes, endothelial cells, fibroblasts, SMC/pericytes, goblet cells, alveolar type 1 epithelial cells (AT1), alveolar type 2 epithelial cells(AT2), ciliated cell populations, and proportional analysis of subpopulations, of whole lung tissue (Figures 1A, B). After analyzing their differentially expressed genes (DEGs), compared to controls, we found that the DEGs in COVID-19 patients were mainly concentrated in macrophages and monocytes, and the DEGs in IPF patients were mainly concentrated in the epithelium (Figure 1C). Further, we analyzed the ratio of cell subpopulations and found that the ratio of ECs and epithelial was upregulated in COVID-19 patients, while higher than in pulmonary fibrosis (Figure 1D). Using correlation analysis, we found that lymphatic ECs, SMC/pericytes, myofibroblasts, rest ECs, late EMT, proliferative fibroblast, rest fibroblasts, late EndMT, *DKK2+* ECs, and activated ECs cells had a strong correlation. This phenomenon prompted us to question why fibroblasts and correlate strongly with ECs (Figure 1E). Therefore, we analyzed fibroblasts, and based on a literature search, we analyzed fibroblast subpopulations such as rest fibroblasts, myofibroblasts, and proliferative fibroblasts (Figures 1F, H). A comparison of COVID-19 and IPF samples revealed that the, myofibroblast proportion was significantly higher in patients with IPF compared to COVID-19 patients (Figure 1G). Interestingly, we also found that myofibroblast expression was much higher in male COVID-19 patients than male patients, but the opposite was true for rest fibroblasts (Figure 1I). We also analyzed the expression of entry factors in whole lung cells, and consistent with previous literature reports, *ACE2* was mainly concentrated in respiratory epithelium cells such as AT1 and AT2. However, *BSG*, *CTSL*, and *FURIN* were also abundantly expressed in macrophages and fibroblasts (Figure 1J). *TMPRSS2* and *ACE2* expression were upregulated in COVID-19 patients but *BSG*, *CTSL*, *FURIN* expression was higher in IPF patients, suggesting their role in increased myofibroblast expression in COVID-19 (Figure 1K). We also analyzed correlation analysis of entry factors and lung function test (%predicted DLCO) and found that lung function was negatively correlated with the normalized bulk RNA-seq gene expression of the entry factors (*CTSL*, *BSG*, *ACE2*) (Figure 1L).

### Single-cell transcriptome sequencing Atlas of EMT stage

EMT is a common type of transition, and we divided the epithelium into goblet, ciliated, EMT, AT1, and AT2 cell subgroups, in which the EMT proportion was significantly upregulated in COVID-19 and IPF patients, but there were no significant

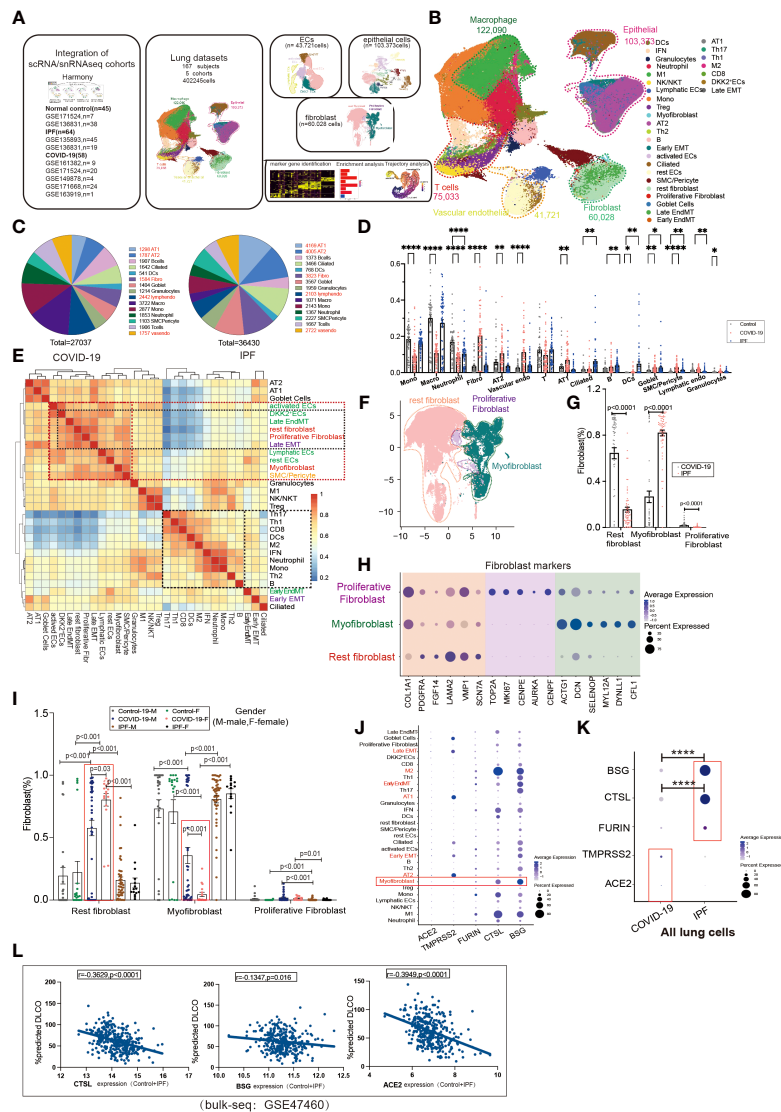


FIGURE 1

Subpopulation correlation analysis suggested a strong correlation between MT subpopulation and fibroblast subpopulation. (A) Scheme of this study. (B) Demonstration of HC, COVID-19, and IPF subgroups using UMAP, including Endothelial (vascular endothelial: rest ECs, EndMT, DKK2+ ECs, activated ECs; Lymphatic ECs); Epithelial (AT2, AT1, Goblet Cell, EMT, Ciliated); Fibroblast (Proliferative Fibroblast, Myofibroblast, Rest fibroblast, Late EndMT, Late EMT); SMC/Pericyte; T cells (Th1, Th17, Th2, Treg, NK/NKT, CD8, IFN); Macrophage (M2, M1); Granulocytes; DCs; Mono; Neutrophil. And the proportion of cell subpopulations in HC, COVID-19, IPF. Marker genes: Endothelial cells:PECAM1, VWF, CLDN5; Macrophages: MARCO, MSR1, MRC1; T cells:CD3E, CD3D, GZMH; Granulocytes:MS4A2, CPA3, TPSA1; B cells:MS4A1, BANK1, CD79A; Monocytes:CD14, FCN1; Neutrophil:S100A8, S100A9; Epithelial cells:EPCAM; Fibroblast:COL1A1, PDGFRA, ELN; SMC:SCGB1A1, RPL26). (C) Pie charts show the number of differentially expressed genes per cell type in COVID-19/IPF compared to controls. The DEGs of COVID-19 were mainly concentrated in macrophages and monocytes, and the DEGs of IPF were mainly concentrated in epithelium. (D) The proportion of cell subpopulations in HC, COVID-19, IPF. (E) The heat-map shows the correlation analysis for each cell subpopulation. The correlation between late-stage EMT, late-stage EndMT and fibroblast subpopulation was higher. (F) UMAP suggests a subpopulation distribution of fibroblasts, including Myofibroblasts, Proliferative fibroblasts and Rest fibroblast. (G) Ratio of fibroblast subpopulations in patients with COVID-19, IPF, showing myofibroblast ratio in IPF is higher than that in COVID-19, but Rest fibroblast ratio in IPF is lower than that in COVID-19. (H) Markers of fibroblast subpopulations, including Myofibroblast (ACTG1, DCN, SELENOP, MYL12A, DYNLL1, CFL1), Proliferative fibroblast (TOP2A, MKI67, CENPE, AUPKA, CENPF) and Rest fibroblast (COL1A1, PDGFRA, FGF14, LAMA2, VMP1, SCN7A). (I) The proportion of fibroblast subpopulations in COVID-19, IPF by female and male, showing Myofibroblast ratio in females is lower than that in males with COVID-19. (J) The DOT plot shows the expression of entry factors in subgroups of control, COVID-19, IPF. Subgroups including Rest fibroblast, Myofibroblast, Proliferative Fibroblast, Late EndMT, Late EMT, SMC/Pericyte; rest ECs, DKK2+ ECs, activated ECs, Lymphatic ECs, EndMT-early; AT1, AT2, Ciliated, Early EMT, Goblet Cell; Granulocytes; B, Th2, Th1, Th17, Treg, CD8, Mono, M1, M2, Neutrophil, DCs, NK/NKT. (K) The DOT plot shows the expression of entry factors in COVID-19, IPF. (L) Correlation analysis of entry factors (CTSL, BSG, ACE2) and lung function test (%predicted DLCO) showed that lung function was negatively correlated with the normalized bulk RNA-seq gene expression of the entry factors. \*P < 0.05, \*\*P < 0.01, \*\*\*\* P < 0.0001

differences between COVID-19 and IPF (Figures 2A, B). After excluding datasets with significant heterogeneity, we performed a meta-analysis using 3 GEO datasets with a pooled SMD of 0.99 and 95% CI (0.49; 1.49) (Figure 2C). Surprisingly, the EMT rate was much more significant in the middle-aged group than in the aged group. The middle-aged group was mainly enriched for inflammation-related signaling pathways, and then the aging adult group was enriched for more fibrosis-related signaling pathways (Figures 2D, E). Whether classified by smoking or not, we observed that EMT was higher in both COVID-19 patients than in controls, and then IPF was higher in the non-smoking group compared to samples. In COVID-19, smoking is enriched for more interferon-related signaling pathways compared to samples from non-smoking individuals (Figures 2F, G). We found that the proportion of EMT was higher in patients who smoked, while in COVID-19 patients, the proportion of EMT was more significant in non-smoking males than in females (Figures 2H, I). This phenomenon indicates that males contribute more to EMT progression. EMT is positively correlated with myofibroblasts in COVID-19 patients, and IPF patients also have the same expression trend. Males upregulated inflammation-related signaling pathways, however females were enriched for downregulated inflammatory signaling pathways (Figure 2J).

Trajectory analysis mentions early and late mesenchymal transitions, and its marker also suggests its transition status (Figure 3A). *EPCAM* expression was significantly decreased from early to late stages. *TAGLN* expression was increased in the early stages but decreased in the late stages. *COL3A1* expression was somewhat lower during the early stages but increased significantly during the late stages (Figures 3C, D, G). From early to late stages, Monocel2 displayed the primary enhanced signaling pathways concurrently (Figure 3B). *TAGLN* is a typical EMT marker gene that is elevated in COVID-19 patients. A meta-analysis of a dataset of five GEO resulted in a pooled SMD of 0.73, 95 CI% (0.30; 1.15) (Figures 3D–F). *COL3A1* is a well-characterized marker of late MT that is increased in COVID-19 patients. The dataset of five GEOs was analyzed using meta-analysis, resulting in a pooled SMD of 1.77, 95 CI% (0.82; 2.72) (Figures 3G–I). Further analysis of the correlation between invasion genes showed that the epithelial entry factor genes (*BSG*, *FURIN*, *CTSL*) were positively correlated with the ratio of early EMT in COVID-19, *TMPRSS2* was negatively correlated with the ratio of early EMT in COVID-19, and *FURIN* was positively correlated with the ratio of early EMT in IPF (Figure 3J). *TMPRSS2* was positively correlated with the ratio of late EMT in COVID-19 and *BSG* was negatively correlated with the ratio of late EMT in COVID-19 (Figure 3K).

## Single-cell transcriptome sequencing Atlas of EndMT stage

A further exploration of the potential cause of elevated fibroblast expression in COVID-19, revealed subpopulations of

endothelium cells based on the high correlation between fibroblasts and endothelial cells (Figure 1D), and we found a subpopulation of endothelial cells with EndMT as in IPF (Figure 4A), which highly expresses the marker gene for subpopulations (Figure 4B). We divided the subgroups of endothelial cells into lymphatic ECs, EndMT, *DKK2+* ECs, activated ECs, and rest ECs. Comparing the endothelium subgroup ratios in ECs, we found that the ratio of EndMT was increased in COVID-19 and higher in IPF than in the controls (Figure 4C). We showed common inflammation-associated cytokines through different subpopulations, implying an abundant expression of EndMT (Figure 4D). To exclude the effect of smoking, we selected nonsmoking patients for comparison, and the results suggested that the proportion of EndMT was also higher in males than in females in COVID-19 patients (Figure 4E). We compared the proportion of entry factor genes in Lymphatic ECs, EndMT, *DKK2+* ECs, activated ECs, and rest ECs in COVID-19 and IPF (Figures 4F, G). The findings suggest that EndMT plays a unique function in COVID-19, and meta-analysis of data from several GEO databases indicates that the EndMT proportion in COVID-19 vs. control pooled SMD was 0.64, 95% CI [0.16; 1.12] (Figure 4H). After dividing the data into two groups, EndMT<sup>low</sup> and EndMT<sup>high</sup>, the latter group was found to have a higher proportion of myofibroblasts (Figure 4I). A hypothesis was then raised about whether EndMT led to the increase in fibroblasts. Enrichment analysis of cytokine expression in the EndMT subpopulation revealed high expression of cytokines associated with fibrosis in COVID-19 and IPF (Figure 4J). The EndMT ratio was positively correlated with myofibroblast cells and negatively correlated with rest fibroblast (Figure 4K).

We then tried to determine whether EndMT has a series of genes that can affect COVID-19 and IPF MT by trajectory analysis and classified EndMT into early, intermediate, and late stages (Figure 5A). The enrichment of genes in early and late stages demonstrates that the interferon-gamma signaling pathway is upregulated in early stages. Simultaneously, late stages exhibit increased expression of the ECM-receptor interaction signaling pathway (Figure 5B). The expression of marker genes associated with EndMT also suggested early, intermediate, and late stages (Figure 5C). *ACTA2* expression was low in the intermediate stage but rapidly increased in the late stage (Figure 5D). Calculating the fraction of *ACTA2* present throughout the early stages of the disease reveals that the expression in COVID-19 is more significant than in controls (Figure 5E). By screening the marker genes of early stage EndMT, the marker gene at the early stage was finally identified as *ACTA2* with an SMD of 0.87, 95% CI (0.23, 1.51) for COVID-19 vs. control in 5 GEO datasets following meta-analysis (Figure 5F). Subsequently, *MMP2* and *COL3A1* were used to identify the late-stage marker gene (Figure 5G, J). *COL3A1* expression was increased in COVID-19 compared to controls (Figure 5K), and although all GSE statistics of *MMP2*-

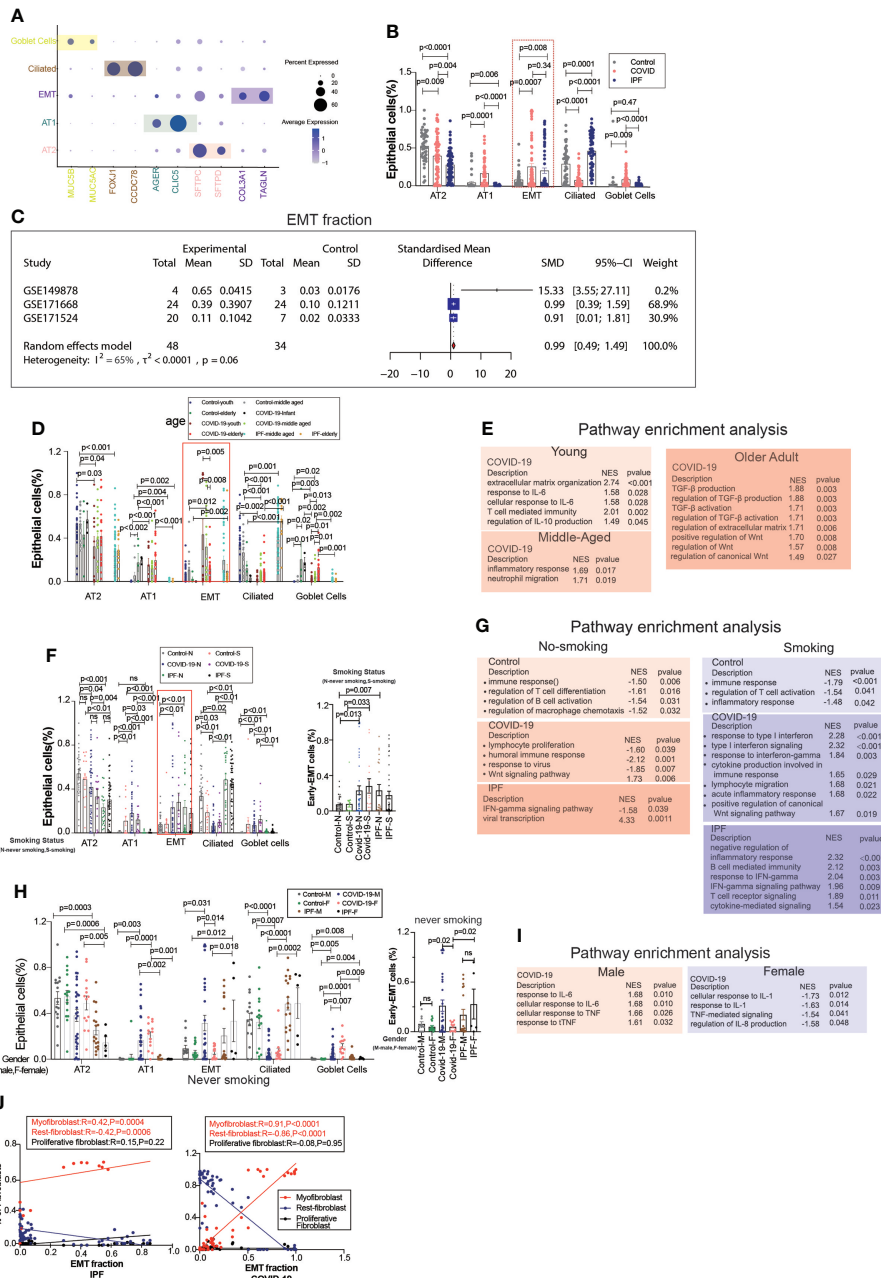


FIGURE 2

EMT by age, smoking status, and gender. (A) Dot chart shows the marker of epithelial subpopulations, including AT1(AGER, CLIC5), AT2(SFTPC, SFTPD), Ciliated (FOXJ1, CCDC78), Goblet cells (MUC5B, MUC5AC), EMT(TAGLN, COL3A1). (B) Comparison of the ratios of HC, COVID-19, IPF subpopulations of epithelium. EMT proportion was significantly upregulated in COVID-19 and IPF. (C) Forest plot of studies with EMT ratio on the COVID-19 and HC, after excluding a study with only one case and a high heterogeneity study. The analysis included data from 3 studies with a total of 48 COVID-19 and 34 controls. p value for heterogeneity was 0.06,  $I^2$  was 65%. (SD, Standard deviation). (D) Comparison of the ratios of HC, COVID-19, IPF subpopulations of epithelium by age. (E) GO (Gene ontology) enrichment analysis was performed on the genes highly expressed in EMT in normal control, COVID-19 and IPF groups in patients of different ages, respectively, and the graphs show the signaling pathways enriched to EMT cell populations in different groups. (F) Comparison of the ratios of HC, COVID-19, IPF subpopulations of epithelium by smoking. (G) GO enrichment analysis was performed to enrich for genes that were highly expressed in EMT of smoking and non-smoking patients in normal control, COVID-19 and IPF groups, respectively, and the graphs show the signaling pathways enriched in EMT cell populations of different groups. (H) Comparison of the ratios of HC, COVID-19, IPF subpopulations of epithelium by sex. (I) GO enrichment analysis of the genes highly expressed in EMT in normal control, COVID-19 and IPF groups by sex, respectively. (J) Correlation analysis of EMT with myofibroblast, rest fibroblast and proliferative fibroblast in COVID-19 and IPF patients (Pearson test), showing EMT is positively correlated with myofibroblasts.

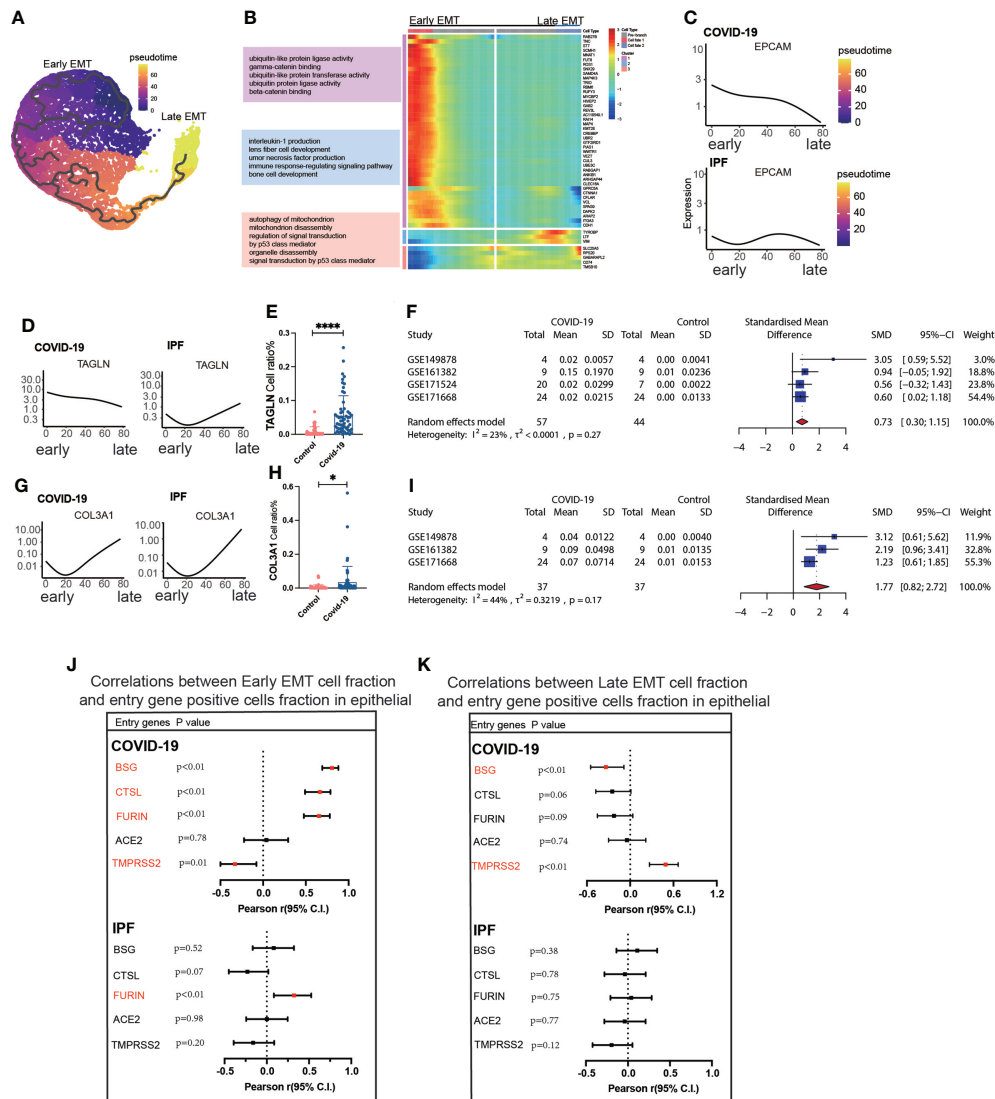


FIGURE 3

Characteristics of early late-stage EMT. (A) Pseudotime projection analysis showing early-stage EMT could evolve to late-stage EMT. (B) Single cell proposed time branch point analysis showing genes with progressively lower and higher expression from Early EMT to Late EMT. Each row of the heat map on the right represents a gene, each column is a proposed time point, and the color represents the average expression value of the gene at the current time point, with the color decreasing from red to blue. The left side shows the signaling pathways that the gene set is enriched. (C) Proposed time traces of individual genes, showing the change in gene expression from the beginning to the end of the proposed time course for epithelial cell marker genes (EPCAM) in COVID-19 and IPF patients, respectively. (D) Time-sensitive trajectories of individual genes, showing the change of Early EMT marker gene (TAGLN) expression from the beginning to the end of the proposed time course in COVID-19 and IPF patients, respectively. (E) TAGLN gene positive epithelial cell ratio is higher in COVID-19 than that in HC. (F) Forest plot of studies with TAGLN gene positive epithelial cell ratio on the COVID-19 and HC, after excluding a study with only one case. The analysis included data from 4 studies with a total of 57 COVID-19 and 44 controls. p value for heterogeneity was 0.27, I<sup>2</sup> was 23%. (G) Temporal trajectories of individual genes, showing changes in gene expression of Late EMT marker gene (COL3A1) in COVID-19 and IPF patients by the start to the end of the proposed time course, respectively. (H) COL3A1 gene positive epithelial cell ratio is higher in COVID-19 than that in HC. (I) Forest plot of studies with COL3A1 gene positive epithelial ratio on the COVID-19 and HC, after excluding a study with only one case and a high heterogeneity study. The analysis included data from 3 studies with a total of 37 COVID-19 and 37 controls. p value for heterogeneity was 0.17, I<sup>2</sup> was 44%. (J) Forest plot showing the correlation between Early EMT cell fraction and entry gene positive cells fraction in epithelial cells. And entry factors (BSG, FURIN, CTSL) were positively correlated with the expression of Early EMT in COVID-19 (p < 0.05, Pearson's r > 0). TMPRSS2 was negatively correlated with the expression of Early EMT in COVID-19 (p < 0.05, Pearson's r < 0). FURIN was positively correlated with the expression of Early EMT in IPF (p < 0.05, Pearson's r > 0). (K) Forest plot showing the correlation between Late EMT cell fraction and entry gene positive cells fraction in epithelial cells. And entry factor (TMPRSS2) was positively correlated with the expression of Late EMT in COVID-19 (p < 0.05, Pearson's r > 0). BSG was negatively correlated with the expression of Late EMT in COVID-19 (p < 0.05, Pearson's r < 0). \*P < 0.05, \*\*\*\* P < 0.0001.

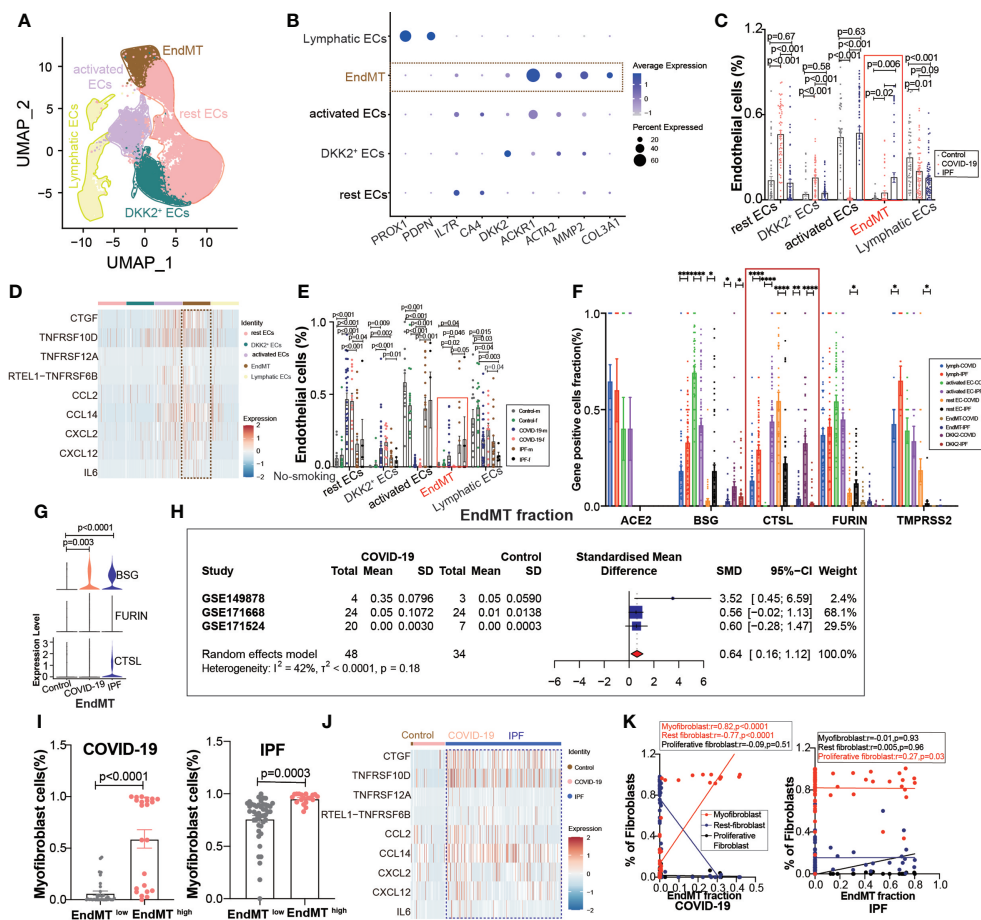


FIGURE 4

EndMT proportion is upregulated in COVID-19. (A) The expression of pulmonary ECs in patients with health control, COVID-19, IPF was demonstrated using UMAP plots with subpopulations marked by a color code. ECs subpopulations including rest ECs, EndMT, *DKK2*<sup>+</sup> ECs, activated ECs; Lymphatic ECs. (B) Dot chart shows the marker of EndMT, mainly focusing on *ACTA2*, *MMP2*, *COL3A1*; rest ECs (*IL7R*, *CA4*), *DKK2*<sup>+</sup> ECs (*DKK2*), activated ECs (*ACKR1*); Lymphatic ECs (*PROX1*, *PDPN*). (C) Proportion of ECs among normal control group, COVID-19 group and IPF group, the ratio of EndMT was increased in COVID-19 and IPF than controls. One-way ANOVA was used to compare multiple groups ( $p < 0.05$ ). (D) Heatmap shows a comparison of the expression of major cytokines in the cell subpopulations of ECs shows that EndMT is enriched in more cytokines compared to the other groups. (E) Comparison of the ratios of HC, COVID-19, IPF subpopulations of ECs by gender in nonsmokers. EndMT ratio in females is lower than that in males with COVID-19. (F) Comparison of entry factors (*ACE2*, *BSG*, *FURIN*, *CTSL*, *TMPRSS2*) positive cell fraction in ECs subtype of COVID-19 and IPF. (G) Comparison of EndMT proportion of entry factors (*BSG*, *FURIN*, *CTSL*) in HC, COVID-19, IPF. (H) Forest plot of studies with lung scRNA data on the COVID-19, after excluding a study with only one case and a high heterogeneity study. The analysis included data from 3 studies with a total of 48 COVID-19 and 34 controls.  $p$  value for heterogeneity was 0.18,  $I^2$  was 42%. (SD: Standard deviation). (I) Ratio of myofibroblasts in COVID-19 and IPF for EndMT<sup>low</sup> versus EndMT<sup>high</sup>, showing myofibroblast ratio in EndMT<sup>high</sup> is higher than that in EndMT<sup>low</sup>. (J) A comparison of the expression of major cytokines in the EndMT shows that COVID-19 and IPF are enriched in more profibrotic cytokines compared to HC group. (K) Correlation analysis of EndMT with myofibroblast, rest fibroblast and proliferative fibroblast in COVID-19 and IPF patients (Pearson test), showing EndMT is positively correlated with myofibroblasts in COVID-19 patients. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

positive cells were not significant (Figure 5H), *MMP2*-positive cells were statistically differentially upregulated in late stage of EndMT by meta-analysis of GSE data. Using meta-analysis, we demonstrated that the *MMP2* pooled SMD in 5 distinct GEO datasets was 0.54, 95% CI (0.13,0.96) (Figure 5I), and ultimately, meta-analysis revealed that the *COL3A1* pooled SMD of late stages of COVID-19 was 0.84, 95% CI (0.24,1.44) when

compared to controls (Figure 5L). Further analysis of the correlation with invasion genes showed that the entry factor genes (*BSG*, *FURIN*, and *CTSL*) were positively correlated with the ratio of early EndMT in COVID-19, and *CTSL* was positively correlated with the ratio of early EndMT in IPF (Figure 5M). *ACE2* was positively correlated with the ratio of late EndMT in IPF (Figure 5N).



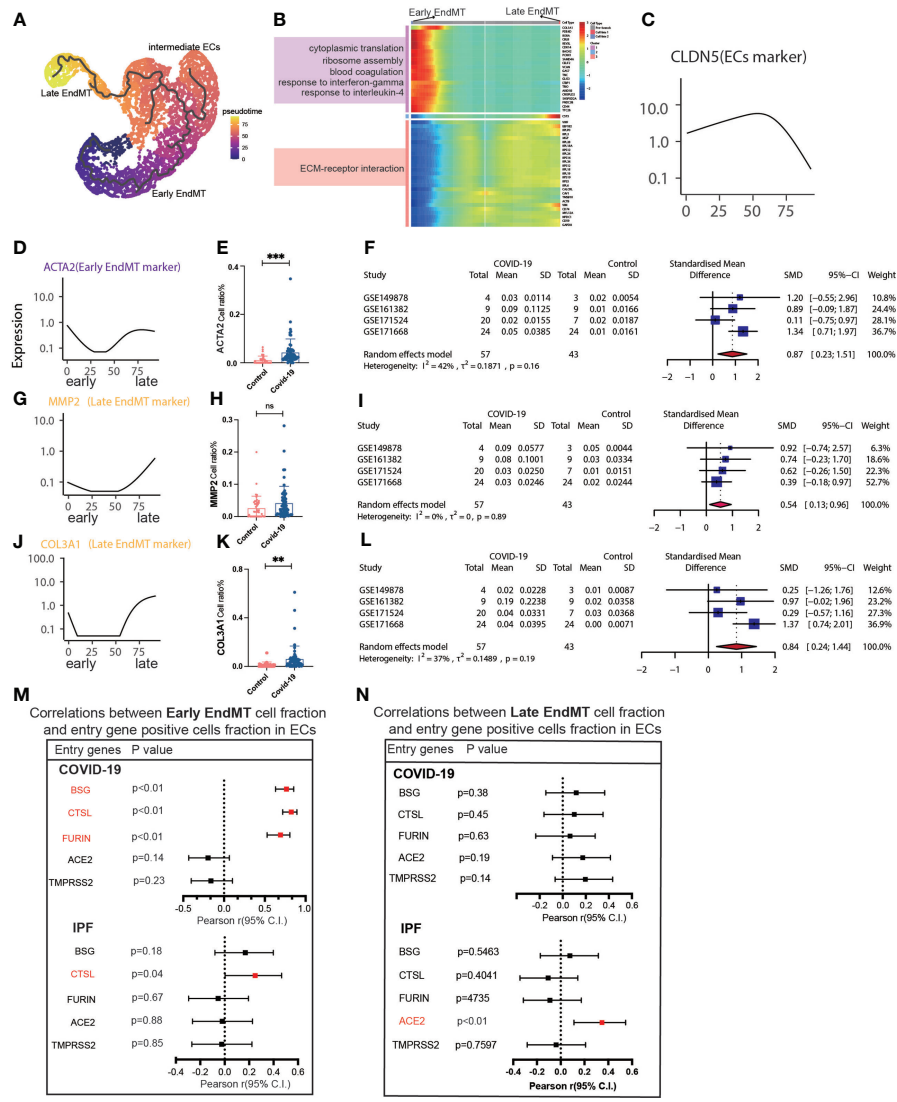


FIGURE 5

Characteristics of early and late stage EndMT. **(A)** Pseudotime projection analysis showing early-stage EndMT could evolve to late-stage EndMT. **(B)** Single cell proposed time branch point analysis showing genes with progressively lower and higher expression from Early EndMT to Late EndMT. Each row of the heat map on the right represents a gene, each column is a proposed time point, and the color represents the average expression value of the gene at the current time point, with the color decreasing from red to blue. The left side shows the signaling pathways that the gene set is enriched to. **(C)** The proposed time trajectory of a single gene, showing the change in gene expression of the ECs marker gene (CLDN5) from the beginning to the end of the proposed time course. **(D)** Trajectories of individual genes, showing changes in gene expression of Early EndMT marker genes (ACTA2) from the beginning to the end of the proposed time course. **(E)** ACTA2 gene positive ECs ratio is higher in COVID-19 than that in HC. **(F)** Forest plot of studies with ACTA2 gene positive ECs ratio on the COVID-19 and HC, after excluding a study with only one case. The analysis included data from 4 studies with a total of 57 COVID-19 and 43 controls. p value for heterogeneity was 0.16,  $I^2$  was 42%. **(G)** Temporal trajectories of individual genes showing changes in gene expression of Late EndMT marker gene (MMP2) from the beginning to the end of the proposed time course. **(H)** MMP2 gene positive ECs ratio is higher in COVID-19 than that in HC. **(I)** Forest plot of studies with MMP2 gene positive ECs ratio on the COVID-19 and HC, after excluding a study with only one case. The analysis included data from 4 studies with a total of 57 COVID-19 and 43 controls. p value for heterogeneity was 0.89,  $I^2$  was 0%. **(J)** Trajectories of individual genes showing changes in gene expression of the Late EndMT marker gene (COL3A1) from the beginning to the end of the proposed time course. **(K)** COL3A1 gene positive ECs ratio is higher in COVID-19 than that in HC. **(L)** Forest plot of studies with COL3A1 gene positive ECs ratio on the COVID-19 and HC, after excluding a study with only one case. The analysis included data from 4 studies with a total of 57 COVID-19 and 43 controls. p value for heterogeneity was 0.19,  $I^2$  was 37%. **(M)** Forest plot showing the correlation between Early EndMT cell fraction and entry gene positive cells fraction in ECs. Entry factors (BSG, FURIN, CTSL) was positively correlated with the expression of Early EndMT in COVID-19 ( $p < 0.05$ , Pearson's  $r > 0$ ). CTSL was positively correlated with the expression of Early EndMT in IPF ( $p < 0.05$ , Pearson's  $r > 0$ ). **(N)** Forest plot showing the correlation between Late EndMT cell fraction and entry gene positive cells fraction in ECs. Entry factors (ACE2) was positively correlated with the expression of Late EndMT in COVID-19 ( $p < 0.05$ , Pearson's  $r > 0$ ). \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; NS,  $P > 0.05$ .

## Discussion

To our knowledge, this first single-cell meta-analysis that describes MT-related clusters in COVID-19 patients. When we examined the MT transition state in COVID-19 using IPF with MT and a high number of fibroblasts as a control, we identified a substantial link between fibroblasts and MT across all subpopulations. By analyzing the correlation of each subgroup, we screened the relevant subgroups of MT (i.e., EndMT and EMT). COVID-19 patients exhibited an increased MT than healthy controls, but less than IPF patients. We further found that male patients had a higher proportion of cells in EMT and EndMT. Several studies have shown a higher prevalence of pulmonary fibrosis after COVID-19 in males. This is likely because males are more exposed to fibrotic triggers, such as occupational agents (26–28).

In EMT, young people are more enriched in inflammation such as *IL-6* signaling pathways, aging adults are enriched with more fibrosis-promoting signaling pathways such as TGF- $\beta$ . Increased levels of TGF- $\beta$ , an anti-inflammatory but profibrotic cytokine, might be the leading cause of EMT in aging adults (29). Smoking COVID-19 patients showed a higher percentage of EMT, but were more enriched in the interferon signaling pathway, compared to non-smoking COVID-19 patients. Despite early reports to the opposite, there is mounting evidence that individuals with severe COVID-19 have a strong type I interferon response, as opposed to the delayed, potentially suppressed response observed early in infection (30). Through various pathways, a potent type I interferon response might increase hyperinflammation in the progression to severe COVID-19 (31). Insights into the therapeutic use of type I interferon in patients with MT will come from an improved knowledge of the functions of type I interferon at various stages of the disease and in patients who are non-smokers vs. smokers.

MT is involved in pulmonary fibrosis and vascular remodeling in the pathogenesis of IPF. Several triggers and pathways are associated with the EMT and EndMT (32). Archana et al. have shown that EndMT markers (N-cadherin, S100A4, and vimentin) are increased in the arterial layers (intima, media, and adventitia) of IPF patients (33). Similarly, during the EndMT process, active ECs express adhesion molecules, such as intercellular adhesion molecule-1 (*ICAM-1*) and vascular cell adhesion molecule-1 (*VCAM-1*), which enhance EndMT formation in our results. Moreover, *ICAM-1* is also an adhesion molecule for another virus, such as human rhinovirus (34), influenza virus (35), and HIV (36). It is reasonable to speculate that with the activation of *ICAM-1*, the risk of associated secondary infection may also increase.

The mesenchymal cells we identified are divided into early and late stage. Compared to IPF, these cells are related to the expression of entry factors such as *BSG*, *CTSL*, and *FURIN*. Sohal

et al. indicate that ILDs, especially IPF, have a higher risk of developing severe COVID-19 infection and post-COVID-19 interstitial pulmonary fibrosis. They found TGF- $\beta$ 1 and  $\alpha$ -smooth muscle actin (myofibroblast marker) are in similar areas as COVID-19 markers. They suggest that myofibroblasts and surrounding tissue secrete growth factors which could further affect COVID-19 adhesion proteins/cofactors and post-COVID-19 interstitial pulmonary fibrosis (37). Our study confirms that entry factors are negatively correlated with early EndMT cell fraction and are positively correlated with fibroblasts. This suggests that the conversion of activated ECs to EndMT is linked to mesenchymal conversion of ECs in the early phases. Therefore, this provides a strategy for the early prevention of pulmonary fibrosis in COVID-19 patients.

Improvements in patient stratification and therapeutic approaches for lung fibrosis will be contingent on the capacity to reprogram EMT (38). EMT has been hypothesized as a source of myofibroblast also in SSc (39) and EMT-related pathways (40). Early EMT leads to activation fibrosis as a result of the downregulation and/or destruction of junctional components like *TAGLN*. The late EMT mesenchymal marker *COL3A1* increases along with the slow loss of *EPCAM*, a marker of epithelial cells (41). Late EMT has been classically associated with fibrosis in organs such as the kidneys, liver, and lungs. It is interesting to note that, in consistent with our data, the EMT status of primary tumors was not consistently associated with poor patient survival (38). Our study suggests that late EMT and invasive genes are almost independent, and early EMT is associated with invasive genes. The process of EndMT can also represent a therapeutic target in the early stages of the disease (42). Early, immature stress fibers are present in the early stages of EndMT. Late stages of EndMT, however, may be identified by matured stress fibers formed by microfilaments (43). The inflammatory leukocytes engaged in the early stages of the illness trigger the EndMT, which has also been discovered in lung fibrosis (44). Therefore, late EndMT could contribute to aggravating the pro-fibrotic signaling in lung fibrosis. Late EndMT, then, could contribute to exacerbating the pro-fibrotic TGF- $\beta$  signaling in lung fibrosis. According to this principle, we chose *ACTA2* as the early marker and also *MMP2* and *COL3A1* as the late markers. *ACTA2* was only transiently displayed in the early EndMT, but *MMP2* and *COL3A1* were subsequently upregulated through further remodeling (45).

It is important to note that our study has some limitations: 1. First, we not count whole organ MT because this included all the up-regulation of the entry factor. This data will be included in our subsequent studies. Second, no in vitro or in vivo animal experiments were performed to validate the key genes. Third, we did not validate COVID-19 patient samples. Fourth, no SARS-CoV-2 positive cells were extracted for further analysis. Finally, the correlation between MT and entry factor was inadequate and required further exploration.

## Conclusion

In conclusion, for the first time, we have analyzed ECs as well as epithelial MT subpopulations in COVID-19 patients, establishing the correlation between their MT subpopulations with irreversible myofibroblasts. However, although entry factors are also highly expressed in IFP, it is not associated with the MT of IPF. Thus, we further explored the mesenchymal-associated cluster's genes and found that some genes are expressed both in COVID-19 patients and IPF patients. This suggests potent targets that could reverse MT in COVID-19 and IPF.

## Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## Author contributions

LZ, XZ, and DL conceived the topic for this study. LZ, XZ, MZ, and XT performed single-cell RNA sequencing. LZ, XZ, XW, SL, RY, MZ, YX, XT, CT, and DL contributed to the writing and revising of the manuscript. DL, and CT critically revised the manuscript. All authors listed have read and approved the manuscript before submission.

## Funding

Postdoctoral Funding of the Personnel Office-Special Fund for Postdoctoral New Crown Epidemic Prevention and Control (0040204153344), National Science Foundation for Young Scientists of China (81900065), Project supported by the Natural Science Foundation of Sichuan province

## References

- Jacob A, Vedaie M, Roberts DA, Thomas DC, Villacorta-Martin C, Alysandratos KD, et al. Derivation of self-renewing lung alveolar epithelial type ii cells from human pluripotent stem cells. *Nat Protoc* (2019) 14(12):3303–32. doi: 10.1038/s41596-019-0220-0
- Merad M, Martin JC. Pathological inflammation in patients with covid-19: A key role for monocytes and macrophages. *Nat Rev Immunol* (2020) 20(6):355–62. doi: 10.1038/s41577-020-0331-4
- Bost P, Giladi A, Liu Y, Bendjelal Y, Xu G, David E, et al. Host-viral infection maps reveal signatures of severe covid-19 patients. *Cell* (2020) 181(7):1475–88.e12. doi: 10.1016/j.cell.2020.05.006
- Kim KK, Kugler MC, Wolters PJ, Robillard L, Galvez MG, Brumwell AN, et al. Alveolar epithelial cell mesenchymal transition develops in vivo during

(2022NSFSC1394), the Youth Innovation Project of Sichuan Medical Association (Q21018), and Chengdu Medical Research Projects (2022516).

## Acknowledgments

The authors thank Dr. Jianming Zeng (University of Macau), and all the members of his bioinformatics team, biotrainee, for generously sharing their experience and codes.

## Conflict of interest

Author MZ is employed by the Oncology Business Department, Novogene Co., Ltd, Beijing, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.976512/full#supplementary-material>

pulmonary fibrosis and is regulated by the extracellular matrix. *Proc Natl Acad Sci United States America* (2006) 103(35):13180–5. doi: 10.1073/pnas.0605669103

5. Willis BC, Liebler JM, Luby-Phelps K, Nicholson AG, Crandall ED, Du Bois RM, et al. Induction of epithelial-mesenchymal transition in alveolar epithelial cells by transforming growth factor-B1: Potential role in idiopathic pulmonary fibrosis. *Am J Pathol* (2005) 166(5):1321–32. doi: 10.1016/S0002-9440(10)62351-6

6. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* (2014) 15(3):178–96. doi: 10.1038/nrm3758

7. Legein B, Temmerman L, Biessen EA, Lutgens E. Inflammation and immune system interactions in atherosclerosis. *Cell Mol Life Sci* (2013) 70(20):3847–69. doi: 10.1007/s00018-013-1289-1
8. Fang S, Guo H, Cheng Y, Zhou Z, Zhang W, Han B, et al. Circhectd1 promotes the silica-induced pulmonary endothelial–mesenchymal transition via Hectd1. *Cell Death Dis* (2018) 9(3):1–16. doi: 10.1038/s41419-018-0432-1
9. Hashimoto N, Phan SH, Imaizumi K, Matsuo M, Nakashima H, Kawabe T, et al. Endothelial–mesenchymal transition in bleomycin-induced pulmonary fibrosis. *Am J Respir Cell Mol Biol* (2010) 43(2):161–72. doi: 10.1165/rcmb.2009-0031OC
10. Dejana E, Hirschi KK, Simons M. The molecular basis of endothelial cell plasticity. *Nat Commun* (2017) 8(1):1–11. doi: 10.1038/ncomms14361
11. Eapen MS, Lu W, Gaikwad AV, Bhattarai P, Chia C, Hardikar A, et al. Endothelial to mesenchymal transition: A precursor to post-Covid-19 interstitial pulmonary fibrosis and vascular obliteration? *Eur Respir J* (2020) 56(4):2003167. doi: 10.1183/13993003.03167-2020
12. Kong D, Li Y, Wang Z, Sarkar FH. Cancer stem cells and epithelial-to-Mesenchymal transition (Emt)-phenotypic cells: Are they cousins or twins? *Cancers* (2011) 3(1):716–29. doi: 10.3390/cancers30100716
13. Singh M, Yelle N, Venugopal C, Singh SK. Emt: Mechanisms and therapeutic implications. *Pharmacol Ther* (2018) 182:80–94. doi: 10.1016/j.pharmthera.2017.08.009
14. Melms JC, Biermann J, Huang H, Wang Y, Nair A, Tagore S, et al. A molecular single-cell lung atlas of lethal covid-19. *Nature* (2021) 595(7865):114–9. doi: 10.1038/s41586-021-03569-1
15. Delorey TM, Ziegler CGK, Heimberg G, Normand R, Yang Y, Segerstolpe Å, et al. Covid-19 tissue atlases reveal sars-Cov-2 pathology and cellular targets. *Nature* (2021) 595(7865):107–13. doi: 10.1038/s41586-021-03570-8
16. Xu G, Liu Y, Liao M, Gou J, Wang X, Yuan J, et al. Persistent viral activity, cytokine storm, and lung fibrosis in a case of severe covid-19. *Clin Transl Med* (2020) 10(7):e224. doi: 10.1002/ctm2.224
17. Wang A, Chiou J, Poirion OB, Buchanan J, Valdez MJ, Verheyden JM, et al. Single-cell multiomic profiling of human lungs reveals cell-Type-Specific and age-dynamic control of sars-Cov2 host genes. *Elife* (2020) 9:e62522. doi: 10.7554/eLife.62522
18. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, et al. Single-cell rna-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* (2020) 6(28):eaba1983. doi: 10.1126/sciadv.aba1983
19. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, et al. Single-cell rna sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv* (2020) 6(28):eaba1972. doi: 10.1126/sciadv.aba1972
20. Kim S, Herazo-Maya JD, Kang DD, Juan-Guardela BM, Tedrow J, Martinez FJ, et al. Integrative phenotyping framework (Ipf): Integrative clustering of multiple omics data identifies novel lung disease subphenotypes. *BMC Genomics* (2015) 16:924. doi: 10.1186/s12864-015-2170-4
21. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. *Cell* (2019) 177(7):1888–1902.e21. doi: 10.1016/j.cell.2019.05.031
22. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* (2019) 16(12):1289–96. doi: 10.1038/s41592-019-0619-0
23. Yu G, Wang L-G, Han Y, He Q-Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
24. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* (2019) 566(7745):496–502. doi: 10.1038/s41586-019-0969-x
25. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* (2017) 14(10):979–82. doi: 10.1038/nmeth.4402
26. Aul DR, Gates DJ, Draper DA, Dunleavy DA, Ruickbie DS, Meredith DH, et al. Complications after discharge with covid-19 infection and risk factors associated with development of post-covid pulmonary fibrosis. *Respir Med* (2021) 188:106602. doi: 10.1016/j.rmed.2021.106602
27. Marvisi M, Ferrozzi F, Balzarini L, Mancini C, Ramponi S, Uccelli M. First report on clinical and radiological features of covid-19 pneumonitis in a Caucasian population: Factors predicting fibrotic evolution. *Int J Infect Dis* (2020) 99:485–8. doi: 10.1016/j.ijid.2020.08.054
28. McGroder CF, Zhang D, Choudhury MA, Salvatore MM, D'Souza BM, Hoffman EA, et al. Pulmonary fibrosis 4 months after covid-19 is associated with severity of illness and blood leucocyte telomere length. *Thorax* (2021) 76(12):1242–5. doi: 10.1136/thoraxjnl-2021-217031
29. Pawelec G. Aging as an inflammatory disease and possible reversal strategies. *J Allergy Clin Immunol* (2020) 145(5):1355–6. doi: 10.1016/j.jaci.2020.02.022
30. Lee JS, Shin EC. The type I interferon response in covid-19: Implications for treatment. *Nat Rev Immunol* (2020) 20(10):585–6. doi: 10.1038/s41577-020-00429-3
31. Kim YM, Shin EC. Type I and iii interferon responses in sars-Cov-2 infection. *Exp Mol Med* (2021) 53(5):750–60. doi: 10.1038/s12276-021-00592-0
32. Gaikwad AV, Eapen MS, McAlinden KD, Chia C, Larby J, Myers S, et al. Endothelial to mesenchymal transition (Endmt) and vascular remodeling in pulmonary hypertension and idiopathic pulmonary fibrosis. *Expert Rev Respir Med* (2020) 14(10):1027–43. doi: 10.1080/17476348.2020.1795832
33. Gaikwad AV, Lu W, Dey S, Bhattarai P, Chia C, Larby J, et al. Vascular remodelling in idiopathic pulmonary fibrosis patients and its detrimental effect on lung physiology: Potential role of endothelial-to-Mesenchymal transition. *ERJ Open Res* (2022) 8(1):00571–2021. doi: 10.1183/23120541.00571-2021
34. Atto B, Eapen MS, Sharma P, Frey U, Ammit AJ, Markos J, et al. New therapeutic targets for the prevention of infectious acute exacerbations of copd: Role of epithelial adhesion molecules and inflammatory pathways. *Clin Sci (Lond)* (2019) 133(14):1663–703. doi: 10.1042/CS20181009
35. Othumpangat S, Noti JD, McMillen CM, Beezhold DH. Icam-1 regulates the survival of influenza virus in lung epithelial cells during the early stages of infection. *Virology* (2016) 487:85–94. doi: 10.1016/j.virol.2015.10.005
36. Yu X, Shang H, Jiang Y. Icam-1 in hiv infection and underlying mechanisms. *Cytokine* (2020) 125:154830. doi: 10.1016/j.cyto.2019.154830
37. Lu W, Eapen MS, Singhera GK, Markos J, Haug G, Chia C, et al. Angiotensin-converting enzyme 2 (Ace2), transmembrane peptidase serine 2 (Tmprss2), and furin expression increases in the lungs of patients with idiopathic pulmonary fibrosis (Ipf) and lymphangioleiomyomatosis (Lam): Implications for sars-Cov-2 (Covid-19) infections. *J Clin Med* (2022) 11(3):777. doi: 10.3390/jcm11030777
38. Mandal S, Tejaswi T, Janivara R, Srikrishnan S, Thakur P, Sahoo S, et al. Transcriptomic-based quantification of the epithelial-Hybrid-Mesenchymal spectrum across biological contexts. *Biomolecules* (2021) 12(1). doi: 10.3390/biom12010029
39. Nikitorowicz-Buniak J, Denton CP, Abraham D, Stratton R. partially evoked epithelial-mesenchymal transition (Emt) is associated with increased tgfb $\beta$  signaling within lesional scleroderma skin. *PLoS One* (2015) 10(7):e0134092. doi: 10.1371/journal.pone.0134092
40. Lee DW, Lee WJ, Cho J, Yun CO, Roh H, Chang HP, et al. Inhibition of wnt signaling pathway suppresses radiation-induced dermal fibrosis. *Sci Rep* (2020) 10(1):13594. doi: 10.1038/s41598-020-70243-3
41. Zeisberg M, Neilson EG. Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest* (2009) 119(6):1429–37. doi: 10.1172/jci36183
42. Mendoza FA, Píera-Velázquez S, Farber JL, Feghali-Bostwick C, Jiménez SA. Endothelial cells expressing endothelial and mesenchymal cell gene products in lung tissue from patients with systemic sclerosis-associated interstitial lung disease. *Arthritis Rheumatol* (2016) 68(1):210–7. doi: 10.1002/art.39421
43. Ciszewski WM, Wawro ME, Sacewicz-Hofman I, Sobierajska K. Cytoskeleton reorganization in emdt: the role in cancer and fibrotic diseases. *Int J Mol Sci* (2021) 22(21). doi: 10.3390/ijms222111607
44. Thuan DTB, Zayed H, Eid AH, Abou-Saleh H, Nasrallah GK, Mangoni AA, et al. A potential link between oxidative stress and endothelial-to-Mesenchymal transition in systemic sclerosis. *Front Immunol* (2018) 9:1985. doi: 10.3389/fimmu.2018.01985
45. Zhang S, Li Y, Huang X, Liu K, Wang QD, Chen AF, et al. Seamless genetic recording of transiently activated mesenchymal gene expression in endothelial cells during cardiac fibrosis. *Circulation* (2021) 144(25):2004–20. doi: 10.1161/circulationaha.121.055417