Cell isolation

The human PASC explant tissues were processed as described [1]. Firstly, PASC lung tissue was finely minced, washed in Hank’s balanced salt solution (HBSS) at 4°C for 5 minutes with rocking, and then centrifuged for 5 minutes at 500g and 4°C. This process was repeated as necessary to remove as much blood as possible. Subsequently, the minced cleaned tissue was incubated in HBSS containing 1X Liberase (0.25U/ml) and DNaseI (0.5mg/ml), and agitated at 1000 rpm at 37°C for 30 minutes using a Thermomixer (Eppendorf) to dissociate the tissue chunks. The single-cell preparations obtained were then further dissociated by passing them through 16 G syringe needles and agitated gently at 1000 rpm at 37°C on the Thermomixer for another 15 minutes.

For enriching epithelial cells and immune cells, the dissociated single-cell preparations were stained with antibodies against EPCAM, CD235a, CD45, and CD31 (Biolegend) and sorted using fluorescence-activated cell sorting (FACS) with a BD Influx cell sorter (Becton Dickinson). The viability of the cells was determined by staining the cell preparations with DAPI (ThermoFisher Scientific) 15 minutes prior to cell sorting.

scRNA-Seq

CD45-CD31-EPCAM+ lung epithelial cells and CD45+CD31-EPCAM- immune cells were FACS sorted from the PASC explant samples and subjected to scRNA-Seq using the 10x Genomic Chromium system. Single cells were captured using a 10X Chromium device (10X Genomics), and the libraries were prepared according to the Chromium Next GEM Single Cell 3’ v3.1Reagent Kits User Guide (10X Genomics). The barcoded sequencing libraries were quantified using the KAPA Library Quantification Kit (KAPA Biosystems, Wilmington, MA). The sequencing libraries were then sequenced using Novaseq 6000 (Illumina) with 150 bp pair-end to obtain a sequencing depth of ~5x104 reads per cell.

Data analysis

The CellRanger v6.0.0 software was used with the default settings for demultiplexing, aligning reads with STAR software to a human GRCh38-2020A transcriptome reference genome (provided by 10X Genomics) https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest.The single-cell analysis R package Seurat v4.3 was used for further data analysis [2]. For quality control and filtering out low quality cells, only cells expressing more than 200 genes (defined as genes detected in at least 3 cells) and fewer than 10% mitochondrial genes were selected. To minimize doublet contamination, cells with a high number of detected genes were removed using a fit model generated from the suggested “multiplet rate”/ “number of cells recovered” ratio, as stated in the 10X Genomics user manual [3]. The UpdateGenesSeurat function from the Seurat.utils R package (DOI:10.5281/zenodo.7228243) was used to ensure consistency in the HGNC symbol for GSE135893 and GSE146981, which are aligned with an earlier version of the reference genome with the new HGNC symbol with the latest version reference genome.

We employed default normalization and data scaling procedures from the Seurat package, which involve log normalization and linear model for data scaling. For data integration, we utilized the batch correction package Harmony [4]. Principal Component Analysis (PCA) was performed using the 3000 most variable genes, and the first 30 independent components were used for downstream unbiased clustering with a resolution of 0.6. The Uniform Manifold Approximation and Projection (UMAP) method was utilized for unsupervised clustering visualization. The cell cluster identities were determined using known gene markers of major epithelial and immune cell types [5].

To identify differentially expressed genes between different clusters and groups, we utilized the Mode-based Analysis of Single-cell transcriptomics (MAST) method[6]. The differentially expressed genes of PASC vs control, or IPF vs control within respective cell types were analyzed using Ingenuity Pathway Analysis (Qiagen).

References:

1. Yao, C.; Guan, X.; Carraro, G.; Parimon, T.; Liu, X.; Huang, G.; Mulay, A.; Soukiasian, H.J.; David, G.; Weigt, S.S.; et al. Senescence of Alveolar Type 2 Cells Drives Progressive Pulmonary Fibrosis. Am J Respir Crit Care Med 2021, 203, 707-717, doi:10.1164/rccm.202004-1274OC.

2. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M., 3rd; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. Cell 2019, 177, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031.

3. Carraro, G.; Mulay, A.; Yao, C.; Mizuno, T.; Konda, B.; Petrov, M.; Lafkas, D.; Arron, J.R.; Hogaboam, C.M.; Chen, P.; et al. Single-Cell Reconstruction of Human Basal Cell Diversity in Normal and Idiopathic Pulmonary Fibrosis Lungs. Am J Respir Crit Care Med 2020, 202, 1540-1550, doi:10.1164/rccm.201904-0792OC.

4. Korsunsky, I.; Millard, N.; Fan, J.; Slowikowski, K.; Zhang, F.; Wei, K.; Baglaenko, Y.; Brenner, M.; Loh, P.R.; Raychaudhuri, S. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 2019, 16, 1289-1296, doi:10.1038/s41592-019-0619-0.

5. Habermann, A.C.; Gutierrez, A.J.; Bui, L.T.; Yahn, S.L.; Winters, N.I.; Calvi, C.L.; Peter, L.; Chung, M.I.; Taylor, C.J.; Jetter, C.; et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Sci Adv 2020, 6, eaba1972, doi:10.1126/sciadv.aba1972.

6. Finak, G.; McDavid, A.; Yajima, M.; Deng, J.; Gersuk, V.; Shalek, A.K.; Slichter, C.K.; Miller, H.W.; McElrath, M.J.; Prlic, M.; et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 2015, 16, 278, doi:10.1186/s13059-015-0844-5.