

RESEARCH ARTICLE

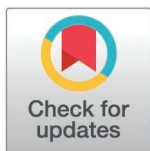
Single-cell transcriptomics of bronchoalveolar lavage reveals divergent macrophage subpopulations and trajectories in interstitial lung disease

Lai-Ying Zhang^{1,2,*}, Peter C. Allen^{3,4}, Viviana P. Lutzky^{1,2}, Simon H. Apte^{1,2}, Penny L. Groves¹, Maxine E. Tan¹, Tharushi De Silva⁵, Eleanor R. Spenceley^{4,6,7}, Rachael A. McCloy³, Quan H. Nguyen⁸, Joseph E. Powell^{3,7}, Daniel C. Chambers^{1,2}, Brendan J. O'Sullivan¹

1 Queensland Lung Transplant Service, The Prince Charles Hospital, Brisbane, Australia, **2** Faculty of Medicine, University of Queensland, Brisbane, Australia, **3** Translational Genomics Program, Garvan Institute of Medical Research, Sydney, Australia, **4** Faculty of Medicine and Health, University of New South Wales, Sydney, Australia, **5** Centre of Immunology and Infection Control, Queensland University of Technology, Brisbane, Australia, **6** Garvan-Weizmann Centre for Cellular Genomics, Garvan Institute of Medical Research, Sydney, Australia, **7** UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, Australia, **8** Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

☞ These authors contributed equally to this work.

* lai-ying.zhang@health.qld.gov.au



OPEN ACCESS

Citation: Zhang L-Y, Allen PC, Lutzky VP, Apte SH, Groves PL, Tan ME, et al. (2026) Single-cell transcriptomics of bronchoalveolar lavage reveals divergent macrophage subpopulations and trajectories in interstitial lung disease. PLoS One 21(4): e0347852. <https://doi.org/10.1371/journal.pone.0347852>

Editor: Gernot Zissel, Medical Center - University of Freiburg, GERMANY

Received: July 30, 2025

Accepted: April 8, 2026

Published: April 29, 2026

Copyright: © 2026 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The data is held in a public repository - (<https://doi.org/10.5281/zenodo.18626573>).

Funding: This work was supported by a Perpetual Grant obtained via The Prince Charles

Abstract

Rationale

Interstitial lung diseases (ILDs) encompass a diverse range of fibrotic conditions and contribute to significant respiratory morbidity and mortality. Assessment of the cellular composition of bronchoalveolar lavage (BAL) fluid is an important diagnostic test in people presenting with ILD, but BAL cellularity remains relatively uncharacterized at single-cell resolution.

Objective

To characterize immune cell populations in BAL across different ILDs and investigate the impact of shortened peripheral blood leukocyte telomere length on BAL immune profiles.

Methods

Single-cell RNA sequencing and downstream analysis were performed on BAL samples from 24 male patients with various ILDs, including idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis, sarcoidosis, and silicosis. Both normal-telomere and short-telomere patients were included. Additionally, we integrated our findings with IPF genome-wide association study (GWAS) data.

Hospital Foundation (IPAP2018/1315 to DC and JP).

Competing interests: The authors have declared that no competing interests exist.

Results

We identified sixteen distinct cell populations in BAL with notable differences across ILD subtypes. Analysis revealed six monocyte-like macrophage (MLM) subclusters following divergent trajectories: inflammatory *CXCL10*^{hi} MLMs predominated in hypersensitivity pneumonitis, while pro-fibrotic *SPP1*^{hi} MLMs were significantly expanded in IPF. Short-telomere patients showed a trend toward increased proportion of pro-fibrotic *SPP1*^{hi} MLMs, with enhanced expression of fibrotic genes compared to patients with normal telomere length. Integration with genomic data confirmed that *SPP1*^{hi} and *CCL2*^{hi} MLM subclusters harbour cells with the highest IPF disease relevance scores.

Conclusion

BAL-derived transcriptomics reveals distinct myeloid subpopulations across ILD subtypes, with specific populations associated with disease pathogenesis. These findings provide insight into ILD pathogenesis, motivate the development of more sophisticated diagnostic tests using BAL sampling, and highlight specific myeloid subpopulations as potential therapeutic targets.

Introduction

Interstitial lung diseases (ILD) encompass a less common but broad subset of chronic lung conditions, including idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis (HP), sarcoidosis, and silicosis, in which abnormal accumulation of stromal cells and extracellular matrix thickens and distorts the pulmonary interstitium. Over time, this causes impaired lung compliance, reduced pulmonary gas exchange, and chronic respiratory failure. While pathogenic molecular signatures in ILD have been shown to extend across multiple biological compartments in various transcriptomic studies, such as in lung parenchyma [1–3] and peripheral blood [4], its landscape in bronchoalveolar lavage (BAL) fluid remains relatively unprofiled at single-cell resolution.

While alveolar macrophages are the predominant cell type in healthy BAL, in pulmonary fibrosis, the cell type composition and gene expression of BAL varies depending on the underlying disease. Single-cell RNA sequencing (scRNAseq) of macrophages isolated from the BAL of fibrotic HP patients, for example, has demonstrated the upregulation of pro-inflammatory pathways compared to IPF [5]. Altered gene expression profiles have been identified in T-helper cells and alveolar macrophages in the BAL of patients with pulmonary sarcoidosis [6]. A recent study integrated the scRNAseq gene expression of peripheral blood monocytes with pulmonary macrophages in the BAL of IPF patients. It inferred a trajectory from classical blood monocytes toward *SPP1*+ve pro-fibrotic BAL macrophages [7], implying a blood-lung recruitment model in pulmonary fibrosis. These studies all however have smaller sample sizes and focus on a single ILD; a broader scRNAseq characterization of

multiple ILDs simultaneously is currently absent from the literature. Understanding the distinct cellular and molecular mechanisms that underpin different ILDs has important clinical implications, as these differences may inform diagnostic approaches, guide selection of disease-specific therapies such as antifibrotics versus immunosuppression, and provide insights into prognosis.

In this study, we utilize scRNAseq to characterize immune cell populations within the BAL of patients with a variety of ILDs, including IPF, HP, sarcoidosis, and silicosis, with a focus on myeloid cells. Given the emerging association between shortened telomere length and inferior outcomes in ILD, such as more rapid disease progression and lower transplant-free survival [8], we also describe alterations in BAL macrophages within a subgroup of patients with ILD and shortened age-adjusted peripheral blood leukocyte telomere length. Finally, we confirm the importance of these macrophages in fibrosis by integrating BAL scRNAseq with published IPF genome-wide association study (GWAS) data. This study characterizes the molecular mechanisms specific to various ILDs and defines cellular phenotypes at single-cell resolution, providing new insights into pathogenesis and potential therapeutic targets.

Methods

Patient recruitment and sample collection

ILD patients were recruited between 2018 and 2020 from a single tertiary ILD centre in Queensland, Australia, with all ILD diagnoses established through gold-standard, expert multidisciplinary discussion in accordance with international guidelines. BAL samples were collected at time of ILD diagnosis from either the right middle lobe or the lingula in accordance with the American Thoracic Society clinical practice guideline, and fresh-frozen in vapor phase for downstream genotyping and sequencing. Demographic and clinical information was obtained retrospectively on the 25th January 2024 from patient records, clinical letters, respiratory function testing, and radiology. Ethics approval was obtained through the Metro North Hospital and Health Services Human Research Ethics Committee under HREC/2023/MNHB/10028. All patients provided written consent for participation in this study, and all patient data was de-identified during and after data collection.

Patient relative telomere length phenotyping

Relative telomere length was assessed by flow fluorescence in situ hybridisation (Flow-FISH), using the peptide nucleic acid (PNA) Kit/FITC for Flow Cytometry (DAKO, Agilent Technologies, USA), according to the manufacturer's instructions. The kit uses fluorescein-conjugated PNA probes to detect telomeres in nucleated haematopoietic cells with the fluorescence intensity of the cells directly correlating with telomere length. Briefly, PBMCs were washed in phosphate-buffered saline (PBS) and mixed 1:1 with control cell line (1301). The samples' DNA were denatured in a heating block adjusted to 82°C for 10 minutes in hybridization solution with or without fluorescently labelled, telomere specific probe and then incubated overnight at room temperature. The following day, samples were washed, and DNA stained before running on a flow cytometer (LSRFortessa, BD Biosciences). The relative telomere length value was calculated as the ratio between the telomere signal of each sample and the control cell line. To determine the 10th centile cut-off, a standard curve was constructed by measuring the relative telomere length in 150 healthy volunteers (75 males and 75 females, aged 18–77 years old).

BAL preparation for single-cell RNA sequencing

Cryopreserved BAL cells were thawed and viable cells sorted on a BD Influx cell sorter (Becton Dickinson) using propidium iodide into Dulbecco's PBS + 0.04% bovine serum albumin and retained on ice. Sorted cells were counted and assessed for viability with Trypan Blue using a Countess automated counter (Invitrogen) and then resuspended at a concentration of 800–1,200 cells/μl.

Genotyping

To demultiplex the pooled single-cell RNA data, DNA was extracted using Qiagen QIAamp DNA mini kit and genotyping performed using the UKB Axiom array at Ramaciotti Centre for Genomics. Imputation was performed using the Michigan Imputation Server, with Minimac4 and the Haplotype Reference Consortium (HRC) panel.

Single-cell RNA sequencing

Single-cell RNA sequencing was performed using the following methods. Single-cell suspensions for samples pertaining to *BAL_scRNA_sample#* were loaded onto 10X Genomics Single Cell 3' Chips along with the RT mastermix as per the manufacturer's protocol for the Chromium Single Cell 3' Library (v2, PN-120233; 10X Genomics), to generate single-cell gel beads in emulsion. RT was performed using a C1000 Touch Thermal Cycler with a Deep Well Reaction Module (Bio-Rad) as follows: 55°C for 2 h; 85°C for 5 min; hold 4°C. cDNA was recovered and purified with DynaBeads MyOne Silane Beads (catalog no. 37002D; Thermo Fisher Scientific) and SPRIselect beads (catalog no. B23318; Beckman Coulter). Purified cDNA was amplified as follows: 98°C for 3 min; 12 times (98°C for 15 s, 67°C for 20 s, 72°C for 60 s); 72°C for 60 s; hold 4°C. Amplified cDNA was purified using SPRIselect beads and sheared to ~200 bp with a Covaris S2 instrument using the manufacturer's recommended parameters. Sequencing libraries were generated with unique sample indices for each sample. Libraries for the respective samples were multiplexed respectively and sequenced on an Illumina NextSeq 500 (NextSeq control software v2.0.2/ Real Time Analysis v2.4.11) using a 150-cycle NextSeq 500/550 High Output Reagent Kit v2 (FC-404–2002; Illumina).

For samples pertaining to RZ7##, single-cell RNA sequencing was performed in similarly as above but according to the 10x Genomics 3' scRNAseq protocol, v3.1 and sequenced on an Illumina NovaSeq 6000 using a 200-cycle S4 Reagent Kit (Catalog no. 20028313; Illumina). Sequence reads were aligned using 10X Genomics's Cell Ranger software to genome reference GRCh38, Gencode release 44, Ensemble 110. CellBender [9] was also used to adjust for residual RNA in each library. Doublet detection was performed using Demuxafy [10] (v3.0.0) containing the doublet detection packages Demuxalot [11], Vireo [12], scDbfFinder [13], scds [14], and Solo [15]. Droplets identified as containing two or more cells were labelled doublets if multiple softwares labelled it so, and were excluded from downstream analyses. For demultiplexing pooled scRNA libraries, we additionally used Demuxafy containing the demultiplex packages Demuxalot and Vireo. Cells were assigned to their respective samples based on genotype information. Only cells with consensus sample assignments across the software tools were retained for further analysis.

Scanpy (v1.10.4) was used to preprocess the scRNA libraries. For quality control, we filtered out cells with fewer than 200 detected genes. Mitochondrial, ribosomal, and hemoglobin gene content was calculated as a percentage for each cell using Scanpy's *calculate_qc_metrics* function. Cells were filtered out if they fell outside 5 mean absolute deviations (MADs) for log total counts, log number of genes by counts or percentage of counts in the top 20 genes. Additionally, cells with more than 3 MADs for percent mitochondrial counts were removed.

We used the single-cell variational inference (scVI [16]) package to correct for batch effects. Raw counts for each cell were corrected for categorical variables such as pool and 10x chemistry (v2/v3) in addition to continuous variables (percentage of mitochondrial or ribosomal counts and total counts). Default parameters were used with early stopping enabled to prevent overfitting. The resulting latent representation was used to generate the Uniform Manifold Approximation and Projection (UMAP) embeddings for visualization.

Cell-type annotation was performed in a two-step process. In the first step, automated cell annotation methods Celltypist [17] and SCimilarity [18] classified cells using reference-based models for lung [1,19–21] and BAL [22]. Secondly, manual annotation was performed using canonical cell type markers as defined in literature [19,23,24] to ensure proper annotation. A list of these markers can be found in S1 Table.

Processed data can be found on <https://doi.org/10.5281/zenodo.18626573>.

Data analysis

Downstream data analyses were performed using Seurat [25] (v5). Cell type proportion analysis was performed using propeller [26], with differences in proportion assessed for statistical significance ($p < 0.05$) using Seurat's Wilcoxon rank-sum test. Differential gene expression analysis for each cluster was performed using a combination of the *FindAllMarkers* function in Seurat and edgeR [27] (version 4.0.16).

To further investigate macrophage subclusters, alveolar, monocyte-like, and interstitial-like macrophages were subset and re-clustered for further characterization using Seurat's *FindAllMarkers*. To identify significant pathways activated in the monocyte-like macrophage subsets, Fast Gene Set Enrichment Analysis [28] (FGSEA, version 1.32.4) was used with pathways listed from the MSigDB hallmark gene set collection [29,30]. Trajectory analysis was performed using slingshot [31] (version 2.14.0) to trace the lineages of MLM subsets. Cellular senescence scoring was calculated using a gene set enriched in senescent cells (SenMayo [32]) as input for the *AddModuleScore* function in Seurat.

Integrating public GWAS data with scRNA-seq

To identify disease-relevant cells, we applied the single-cell disease relevance score model (scDRS [33], v1.0.3) to our scRNA data. GWAS summary statistics from Allen_IPF [34], UKBB_LymphocyteCount, and UKBB_MonocyteCount were used to find gene sets consisting of top genes with the most significant gene-level p -values, which were calculated by combining p -values of SNPs nearby to the genes using MAGMA [35] (v1.10). The disease relevance score for each cell was calculated using the command-line interface version of scDRS with default parameters. Covariates for adjusting the scores included in the model for each cell were the number of genes and a binary IPF status.

Results were visualized using Python (v3.13.1) with the matplotlib library [36] (v3.10.0). For plotting disease relevance scores, we utilized UMAP embeddings calculated in the preprocessing step. These embeddings were used to visualize both the full dataset and the subset of monocyte-like macrophages, highlighting cells with high disease relevance scores.

All scripts used to perform preprocessing, integration, annotation, and generate figures can be found on our online GitHub repository (<https://github.com/powellgenomicslab/ILD-BAL-scRNA-Manuscript-2025/>).

Results

Patient demographics

BAL samples were obtained from 24 patients with interstitial lung disease (Table 1; further detail in S2 Table). Ten patients were diagnosed with idiopathic pulmonary fibrosis (IPF), four with sarcoidosis, three with hypersensitivity pneumonitis (HP), three with smoking-related ILD (SR-ILD), two with silicosis, one with non-specific interstitial pneumonia (NSIP), and one with pleuroparenchymal fibroelastosis (PPFE). Patients were all male, predominantly ex-smokers, and predominantly had mild restriction on spirometry with mild-to-moderate impairment of diffusing capacity, with the exception of the single patient with NSIP who had more severe disease. Non-IPF patients were younger than patients diagnosed with IPF or PPFE, consistent with expected demographics [37]. All patients were treatment-naïve at time of BAL (i.e., no immunosuppressants or anti-fibrotic therapies). Eighteen patients had peripheral blood stored at the time of BAL collection, and a relative telomere length was performed for these patients, with a short peripheral blood leukocyte telomere length defined as a length $<10^{\text{th}}$ centile adjusted for age.

Single-cell RNA sequencing of BAL fluid reveals myeloid, lymphoid, and epithelial cell populations present in interstitial lung disease

ScRNAseq of BAL from 24 patients with ILD yielded 92,164 individual cells after quality control and filtering. We identified sixteen distinct cell types including macrophages, lymphocytes, dendritic cells (DCs), Mast cells, and a small cluster of ciliated bronchial epithelial cells (Fig 1a, S1 Fig). Macrophages were further subtyped into alveolar macrophages (AMs),

Table 1. Demographics of patients presenting with interstitial lung disease included in this study.

Diagnosis (n)	Age (yrs, SD)	Ever-Smoker (%)	Short-Telomere (%)	FVC at Diagnosis (%pred, SD)	DLCO at Diagnosis (%pred, SD)
IPF					
IPF (n = 10)	70.5 (8.0)	90	20	82.8 (11.4)	58.5 (14.3)
Non-IPF					
HP (n=3)	61.0 (13.5)	0	67	75.7 (2.5)	53.0 (6.0)
Sarcoidosis (n=4)	44.0 (11.1)	75	0	97.0 (10.3)	94.5 (12.8)
SR-ILD (n=3)	52.0 (10.0)	100	0	83.3 (6.4)	64.0 (7.8)
Silicosis (n=2)	23.5 (0.7)	100	100	97.0 (4.2)	77.0 (0)
NSIP (n=1)	74.0 (-)	0	100	62.0 (-)	34.0 (-)
PPFE (n=1)	70.0 (-)	0	0	73.0 (-)	54.0 (-)

IPF: idiopathic pulmonary fibrosis. HP: hypersensitivity pneumonitis. PPFE: pleuroparenchymal fibroelastosis. NSIP: non-specific interstitial pneumonia. SR-ILD: smoking-related interstitial lung disease. FVC: forced vital capacity. DLCO: diffusion capacity of the lung for carbon monoxide. %pred: percentage of predicted for age/height.

<https://doi.org/10.1371/journal.pone.0347852.t001>

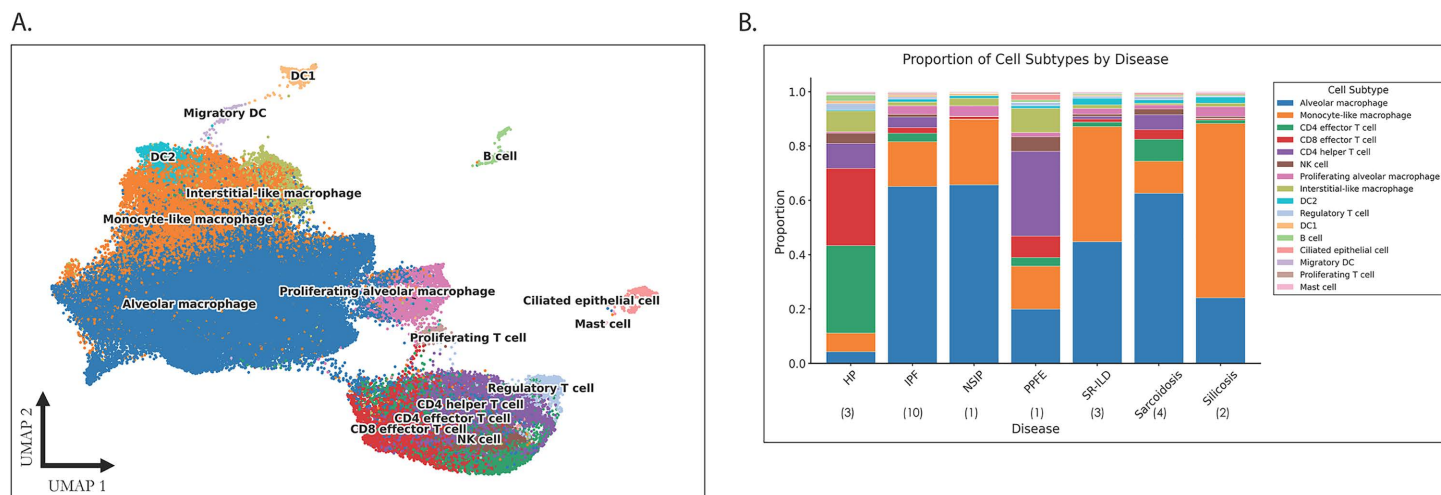


Fig 1. Cell subtypes identified in bronchoalveolar lavage fluid of patients with interstitial lung disease. A) Uniform manifold approximation and projection (UMAP) of all cells isolated from bronchoalveolar lavage. **B)** Cell type proportion of each cell subtype stratified by interstitial lung disease subtype. Number of patients pooled within each ILD subtype denoted in brackets. DC: Dendritic cell. ILM: Interstitial-like macrophage. NK: natural killer.

<https://doi.org/10.1371/journal.pone.0347852.g001>

defined by the expression of classic AM markers such as *MARCO*, *PPARG*, and *FABP4*; monocyte-like macrophages (MLMs), expressing both macrophage markers and markers of classical or non-classical monocytes (*VCAN*, *FCN1*, *TYMP*); interstitial-like macrophages (ILMs), expressing markers associated with tissue interstitial macrophages such as *LGMN* and *MARCKS*; and proliferating macrophages. Although previous studies have described monocyte-derived macrophages within BAL [38], analogous to the monocyte-like macrophages we identified, we opted to term them as being “monocyte-like” as it was not possible in our analysis to definitively infer that this population had derived from or differentiated from blood monocytes. Similarly, we describe “interstitial-like” macrophages as this population resembled previously described tissue interstitial perivascular macrophages [19], but were found within the airspace rather than within lung tissue. To compare our macrophage labels, we applied transfer learning to a recent IPF BAL dataset [7]. Our labels accurately mapped to their macrophage cluster, forming distinct groups (S2 Fig) that support the validity of our cell-type labels.

Lymphocytes were subtyped into either B cells or T cells, with the latter further subtyped into CD4 and CD8 effector T cells, CD4 helper T cells, regulatory T cells (Tregs), proliferating T cells, or natural killer (NK) cells based on characteristic gene expression profiles. Three subtypes of dendritic cell were identified, including a DC1 population clustered close to macrophages, a DC2 population expressing *CD1E* and *CLEC10A*, and a migratory DC population. Only airway epithelial cells were identified; we identified no alveolar epithelial cells within our analysis.

Immune cell profiles differ between various ILD subtypes, suggesting divergent mechanisms toward pulmonary fibrosis

Comparison of BAL cell type proportions revealed notable differences between ILD subtypes (Fig 1b, S3 Fig). In IPF patients, alveolar and monocyte-like macrophage populations predominated. In contrast, all three HP patients included in this study had ground-glass nodularity on cross-sectional imaging consistent with an inflammatory HP phenotype [39]; as expected in this context [40], a high proportion of CD4 and CD8 T lymphocytes were identified in BAL via scRNAseq. BAL from sarcoidosis and PPFE patients also revealed an expanded lymphocyte population with an increased CD4:CD8 ratio consistent with published literature [41,42], with a notable prominence of the CD4 helper T cell subtype in PPFE. In contrast, silicosis and SR-ILD BAL demonstrated a significantly increased proportion of monocyte-like macrophages compared to other ILDs, suggesting a role for MLMs in their pathogenesis.

Transcriptomic analysis of BAL macrophage subpopulations, especially monocyte-like macrophages, reveals distinct phenotypes correlating with ILD subtype

IPF and non-IPF macrophages demonstrated marked transcriptional differences that implied a divergence in macrophage phenotype and function between the two conditions (Fig 2a, S3 Table). We therefore sought to characterize this further by subclustering the monocyte-like macrophage, alveolar macrophage, and interstitial-like macrophage populations (Fig 2b, 2c).

We identified six different monocyte-like macrophage subclusters (Fig 3a, S4A Fig, S4 Table). Importantly, within these subclusters we were able to identify a group of airspace MLMs that upregulated known pro-fibrotic genes such as *SPP1*, *MMP7*, *MMP9*, and *CHI3L1* (denoted as *SPP1^{hi}* MLMs), and which have previously been described within fibrotic lung tissue [1–3]. Pathway analysis revealed that this subcluster was uniquely enriched in pathways associated with epithelial-to-mesenchymal transition (EMT) and angiogenesis, which have previously been implicated in fibrogenesis [1,43]. In contrast, two other MLM subclusters (*IL1B^{hi}* MLMs, *CXCL10^{hi}* MLMs) were found to have gene upregulation patterns associated with pro-inflammatory pathways. *IL1B^{hi}* MLMs demonstrated upregulated *AREG*, *NR4A3*, *SIK1* and other genes associated with canonical TNF α /NF κ B signalling, hypoxia, and apoptosis, while *CXCL10^{hi}* MLMs were enriched in IFN α , IFN γ , and the IL6/JAK/STAT3 signalling pathways.

Trajectory analysis was performed to gain insight into possible pathways cells from a specified cluster take when progressing from one cell state to either a terminal state or a more pathogenic state. Of note, a trajectory analysis of MLM subclusters found that pro-fibrotic *SPP1^{hi}* MLMs and inflammatory *CXCL10^{hi}* MLMs existed as terminal endpoints of a divergent trajectory (Fig 3b). Cell type proportion analysis found that these two terminal endpoints were differentially expanded in various ILDs; not only was the pro-fibrotic *SPP1^{hi}* MLM cluster significantly expanded in the BAL of patients with IPF in comparison to other ILD subtypes (Fig 3c), it also demonstrated significantly upregulated fibrotic gene expression (Fig 3d), while *CXCL10^{hi}* MLMs were expanded in the BAL of inflammatory HP patients. When scored for cellular senescence using the SenMayo gene set [32], cells with the highest senescence scores were found toward the terminal endpoints of both MLM trajectories, confirming trajectory directionality. In contrast, silicosis BAL demonstrated a significantly expanded population of early-trajectory MLM0 macrophages.

Subclustering of alveolar macrophages revealed nine subpopulations (S4b Fig and S5A, S5b Fig), including a *CCL24^{hi}* AM cluster which upregulated collagen genes (*COL22A1*, *COL23A1*, *COL24A1*), found to be significantly expanded in

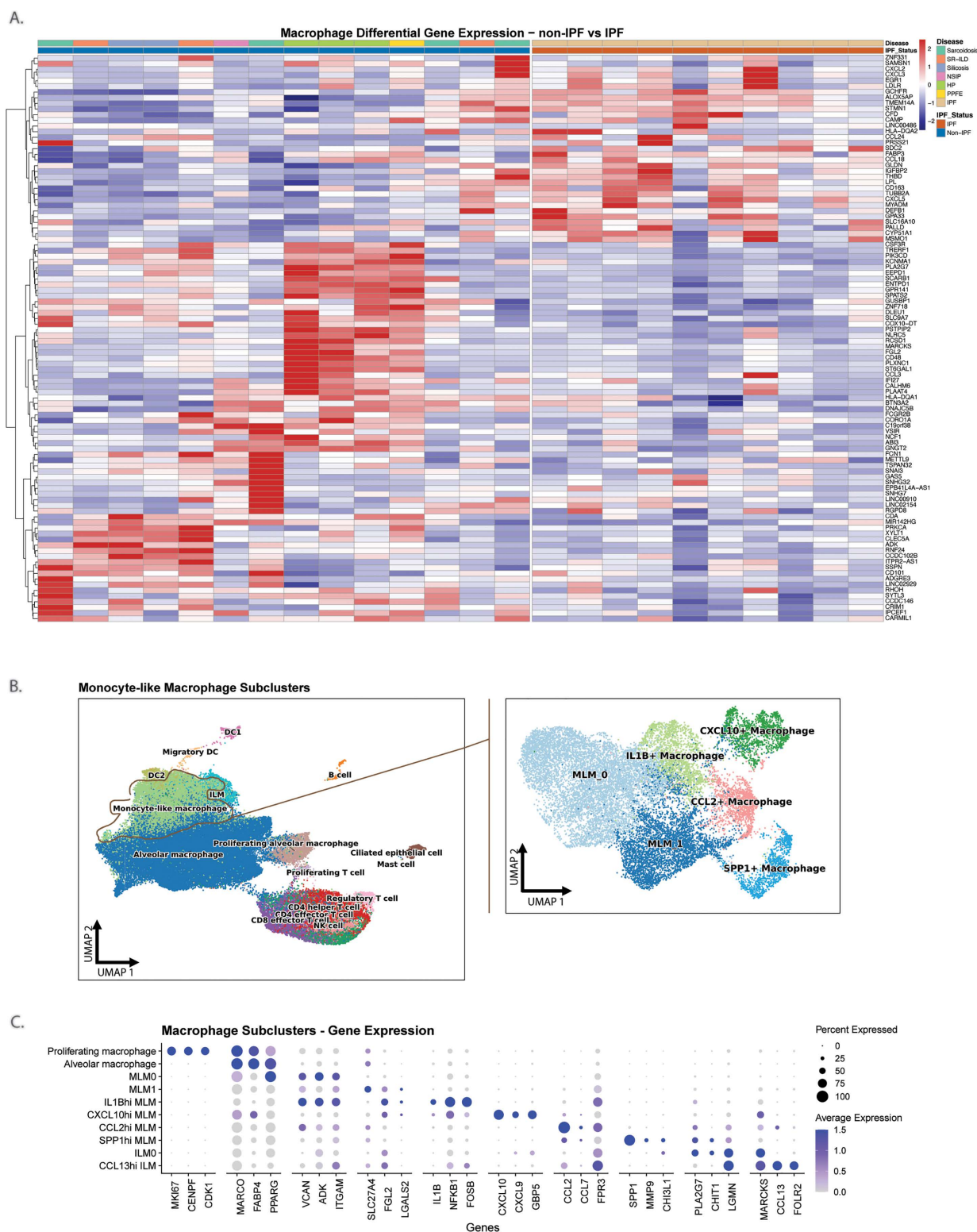


Fig 2. Characterization of macrophage populations in non-IPF and IPF bronchoalveolar lavage. A) Heatmap of top differentially expressed genes within BAL macrophages in non-IPF versus IPF patients. B) UMAP of monocyte-like macrophage subclusters. C) Dotplot of top genes expressed by each macrophage subcluster. MLM: monocyte-like macrophage. ILM: interstitial-like macrophage.

<https://doi.org/10.1371/journal.pone.0347852.g002>

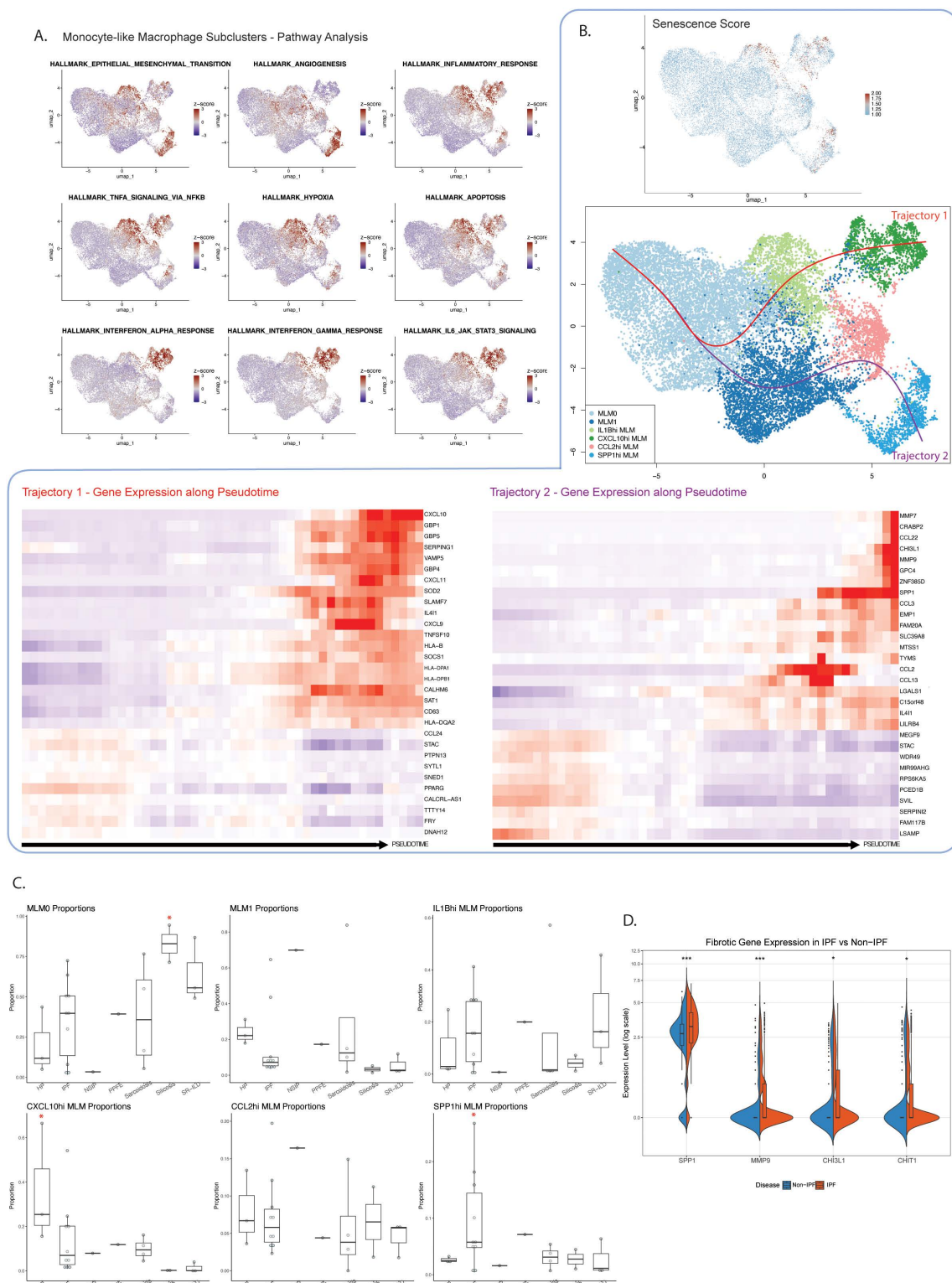


Fig 3. In-depth characterization of BAL monocyte-like macrophages in various ILD subtypes. A) MSigDB Hallmark pathway analysis of MLM subclusters, with each highlighted pathway significantly upregulated (adjusted $p < 0.05$) in each subcluster portrayed, as determined by Fast Gene Set Enrichment Analysis. **B)** Trajectory analysis of MLM subclusters demonstrating two differential trajectories, with senescence scoring via SenMayo gene

set analysis confirming trajectory directionality. Genes upregulated along each trajectory with progressing pseudotime portrayed via heatmaps. **C)** Cell type proportion analysis of each MLM subcluster (out of all MLMs) in each ILD subtype, with significance ($p < 0.05$) determined via Wilcoxon rank-sum testing. Red *: significantly increased proportion. **D)** Differential expression of selected fibrotic genes in the *SPP1*^{hi} MLM subcluster of non-IPF and IPF monocyte-like macrophages, with significance determined via Wilcoxon rank-sum testing. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

<https://doi.org/10.1371/journal.pone.0347852.g003>

silicosis BAL (Cluster 3 in **S5C Fig**; $p = 0.023$), and a *PPBP*^{hi} AM cluster enriched for various senescence-related genes including *CDK1* and *CCNA2* that was enriched in IPF (Cluster 5 in **S5C Fig**; $p = 0.0004$). We also identified two interstitial-like macrophage subclusters (**S4C Fig** and **S6 Fig**) that differentially enriched in inflammatory and adipogenesis pathways. Together, this data highlights the spectrum of phenotypes and functions embodied by various macrophage subsets in fibrotic lung disease.

Monocyte-like macrophages in short-telomere pulmonary fibrosis BAL may have increased fibrotic potential

Since leukocyte telomere length strongly associates with disease trajectory in ILDs, we next sought to characterize differences in BAL myeloid subpopulations between patients with short and normal age-adjusted peripheral blood leukocyte telomere length. When stratified for telomere length, patients with short-telomere-associated ILD had a reduced proportion of inflammatory MLMs (*IL1B*^{hi} MLMs, *CXCL10*^{hi} MLMs) in comparison with normal-telomere patients (**Fig 4a**). This effect was more prominent when only IPF patients were analyzed, with short-telomere IPF patients having an increased *SPP1*^{hi} MLM population compared to normal-telomere IPF (**Fig 4b**). Furthermore, differential gene expression analysis of the fibrotic *SPP1*^{hi} MLM population comparing short-telomere and normal-telomere IPF patients revealed enhanced expression of *SPP1* and *CCL2* in short-telomere patients, with a relative downregulation of more inflammatory genes such as *MALAT1* and *HLA-DPA1* (**Fig 4c**).

Integrating GWAS data with scRNA reveals increased IPF-associated signal in *SPP1*^{hi} and *CCL2*^{hi} monocyte-like macrophages

Multiple GWAS to date have identified IPF-associated variants in genes involved in mucin and surfactant production, telomere maintenance, and desmosome function [34,44,45]. Recently, a study linked variants found in these GWAS to gene

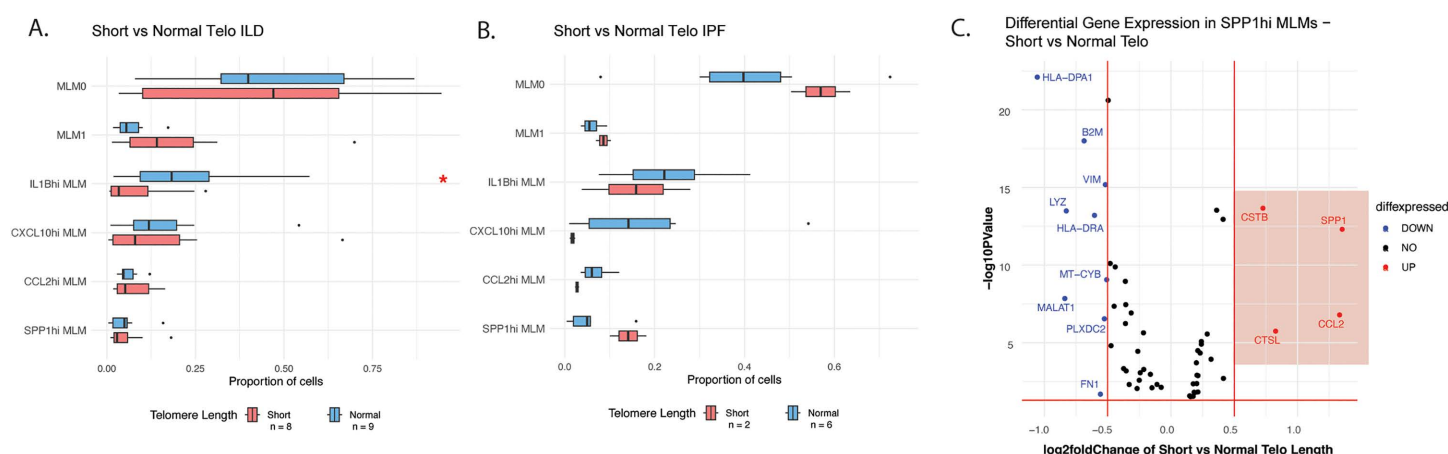


Fig 4. BAL macrophages in short-telomere pulmonary fibrosis are characterized by increased fibrotic potential. Cell type proportion analysis of monocyte-like macrophage subclusters by telomere length, in **A)** all ILD subtypes (via Wilcoxon rank-sum testing) and **B)** IPF only. *: $p < 0.05$. **C)** Volcano plot of differential gene expression of the *SPP1*^{hi} MLM cluster in short versus normal telomere length IPF, with red genes comparatively upregulated in short-telomere BAL (FDR < 0.05 and log2foldchange > 0.5), and blue genes comparatively downregulated.

<https://doi.org/10.1371/journal.pone.0347852.g004>

expression in specific lung cell types through expression quantitative trait loci (eQTL) analysis [46]; however, the sparsity of the data and relatively small sample size makes it challenging to identify true associations between specific cell types and IPF. We applied the single-cell disease relevance score (scDRS) model [33] to our scRNAseq data to identify IPF-relevant cell types in BAL samples to further address this. By intersecting the Allen *et al.* (2022) GWAS with our scRNAseq data, we identified a subset of proliferating alveolar macrophages and monocyte-like macrophages with high disease relevancy scores (Fig 5); further subsetting of the monocyte-like macrophages revealed that pro-fibrotic *SPP1*^{hi} and *CCL2*^{hi} MLM subclusters harbored cells with the highest scores. These findings further suggest that these specific macrophage subsets may be crucial to IPF pathogenesis.

Discussion

Bronchoalveolar lavage is a simple, fast, and safe procedure currently recommended as an important diagnostic tool to differentiate various ILDs based on cell cytology and counts [47]. Unlike surgical or bronchoscopic lung biopsy, BAL is minimally invasive and results in considerably less morbidity, without the risk of triggering a life-threatening ILD exacerbation. The potential utility of BAL-derived cells as a platform for elucidating fibrotic pathophysiological mechanisms, providing insights into disease progression, and identifying biomarkers with prognostic relevance, warrants further investigation. In this study we have successfully sampled and profiled key immune populations within the alveolar space using scRNAseq of BAL cells across a diverse range of ILDs, in patients with both short and normal peripheral blood telomere length. Our findings reveal distinct cellular compositions across different fibrotic conditions, highlighting the heterogeneity of airspace macrophages and their potential contribution to disease pathophysiology.

Previous transcriptomic studies of fibrotic lung tissue have identified a population of pro-fibrotic, monocyte-derived macrophages believed to promote fibrogenesis through interactions with other culprit cells within the fibrotic niche such as aberrant basaloid cells, fibroblasts, and abnormal secretory epithelial cells [48]. In our data, we identify an analogous macrophage subpopulation in BAL. These *SPP1*^{hi} macrophages upregulate similar pro-fibrotic genes to their tissue counterpart, are preferentially enriched in IPF, and particularly in those with telomere shortening. The correlation of our single-cell data with GWAS data confirms a significant association between this *SPP1*^{hi} BAL macrophage subcluster and IPF. Through trajectory analysis, we demonstrate that these macrophages exist in a terminally differentiated state originating from a cluster of macrophages that co-express monocyte markers (MLM0); interestingly, we find that these same MLM0 macrophages can also proceed down a different trajectory toward a more inflammatory terminus (*CXCL10*^{hi} MLMs), especially prominent in HP. In contrast, in silicosis and SR-ILD, both of which show a significant increase in the overall proportion of BAL MLMs, this increase seems to be due to higher proportions of early-trajectory MLM0 macrophages rather than distal-trajectory subclusters.

These findings reveal that lung macrophage populations exhibit remarkable heterogeneity, adopting distinct phenotypes and gene expression profiles – potentially in response to microenvironmental cues or intercellular signaling – suggesting a specialized division of labor among these immune cells. This granular characterization of macrophage subpopulations can provide important insights into immune cell function and behavior that may fundamentally enhance our understanding of ILD disease pathophysiology, or aid in the identification of drug targets. For example, macrophages and monocytes with similar gene expression profiles to the *CXCL10*^{hi} MLMs we describe in inflammatory HP BAL have been previously observed in HP lung tissue [49] and peripheral blood mononuclear cells [7], and are believed to be responsible for CD8 effector T cell chemotaxis into the lung to perpetuate alveolitis [50]. In IPF, conversely, multiple clusters of both alveolar and monocyte-like macrophages in our data are preferentially enriched for fibrotic and senescent, rather than inflammatory, pathways; this effect appears to be most prominent in short-telomere IPF patients, in which there is almost a complete absence of inflammatory *CXCL10*^{hi} MLMs. Clinically, this may partially explain why immunosuppression has been repeatedly demonstrated to be detrimental in IPF and, especially, in short-telomere IPF [51,52], and why short-telomere IPF patients tend to manifest a more aggressive fibrotic phenotype [53].

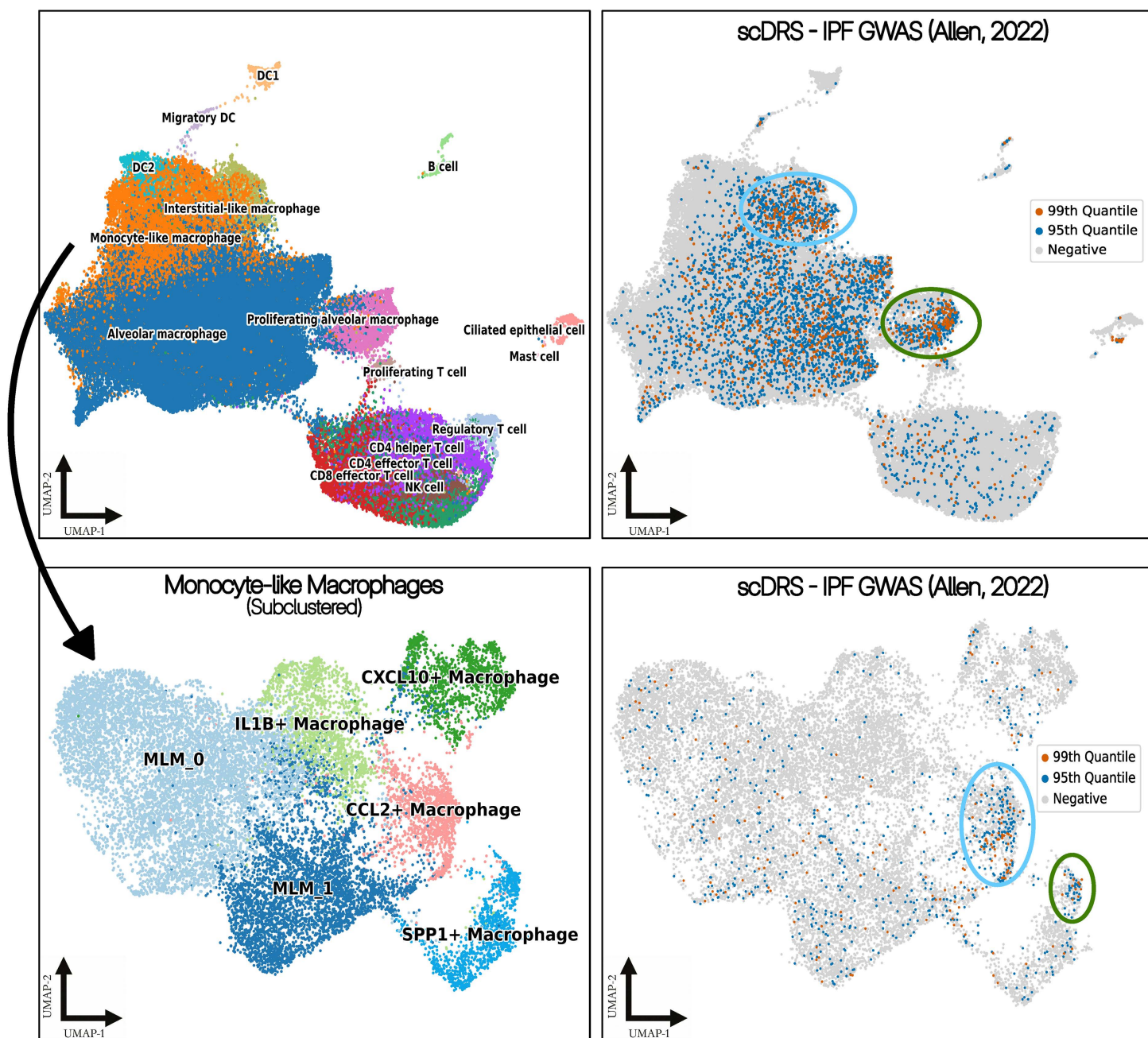


Fig 5. BAL macrophage subpopulations associated with IPF through integration with the Allen *et al.* 2022 genome wide association study (GWAS). Single-cell disease relevance score (scDRS) of each cell to IPF, across all BAL cell types (upper) and in monocyte-like macrophages only (lower), with cells at the 95th percentile in blue and those at the 99th percentile in red. Circles highlight subpopulations with the highest concentration of 95th and 99th percentile scores.

<https://doi.org/10.1371/journal.pone.0347852.g005>

Additionally, we find that the *SPP1*^{hi} MLM cluster in short-telomere IPF BAL upregulates not only *SPP1*, suggesting increased fibrotic potential, but also upregulates *CCL2*, a chemokine known to increase monocyte and macrophage recruitment [54], in comparison to the analogous cluster in normal-telomere IPF; potentially this might imply a mechanism by

which fibrotic macrophages in short-telomere IPF further self-perpetuate fibrosis through a positive feedback loop, recruiting additional monocyte-derived macrophages to the lung to be subsequently directed down a fibrotic trajectory. Elevated levels of *CCL2* have previously been found in the BAL and/or serum of patients with IPF [55], fibrotic HP [49], post-Covid-19 pulmonary fibrosis [56], and connective tissue disease-associated ILD [57], further implying a role in fibrotic progression. While we did not observe a specific variability in the cell type proportion of the *CCL2*^{hi} MLM subcluster between ILDs, this subcluster is present along the fibrotic MLM trajectory. Our analysis of GWAS data found a correlation between IPF and both the *CCL2*^{hi} and *SPP1*^{hi} MLM subclusters, suggesting that both may play an important role in developing and promoting fibrosis.

In silicosis, alveolar macrophages are involved in the phagocytosis of silica particles, which is believed to ultimately trigger apoptosis due to lysosomal disruption [58]; our BAL cell type proportion data correspondingly demonstrates a reduced proportion of AMs and a significantly increased recruitment of likely monocyte-derived macrophages to the lung consistent with published literature in mouse models [59], that are early trajectory and not heavily polarized toward an inflammatory or fibrotic phenotype. As both silicosis patients in our study had early uncomplicated silicosis, the question naturally arises as to whether the early-trajectory MLM0/MLM1 cells seen in their BAL might eventually polarize later in the disease course such as in the development of progressive massive fibrosis, and whether these shifts down particular undesirable trajectories might be capable of modulation or reversal. A recently published study [60] has indeed suggested that airspace macrophages retrieved from patients with silicosis possess significant plasticity and are capable of polarizing toward and away from a pro-fibrotic phenotype when cultured with various drugs *in vitro*, confirming the potential therapeutic relevance of our findings.

Our study presents the first in-depth characterization of BAL macrophage subpopulations across several different ILDs. We provide a detailed description of multiple alveolar, monocyte-like, and interstitial-like macrophage subclusters. We also explore the divergence of monocyte-like macrophages down distinct trajectories into either inflammatory or fibrotic phenotypes depending on the underlying ILD. We are the first to compare, on a single-cell level, BAL from short- and normal-telomere patients with ILD, and demonstrate that the airspace macrophages of short-telomere IPF patients appear to have a greater fibrotic potential.

These findings have several potential implications for ILD management. Our demonstration that BAL-derived scRNA-seq can capture disease-relevant macrophage populations analogous to those identified in lung tissue suggests several potential clinical applications. The disease-specific macrophage signatures we identify, particularly the differential enrichment of *SPP1*^{hi} versus *CXCL10*^{hi} MLM subclusters across ILDs, could potentially refine diagnostic approaches in cases where conventional BAL cytology and radiology remain indeterminate. Beyond initial diagnosis, BAL cellular profiling could serve as a platform to inform clinical decision-making in situations where treatment options are unclear or when patients demonstrate unexpected disease progression despite therapy, potentially providing molecular insights into treatment failure or guiding selection between competing therapeutic strategies. The correlation of specific macrophage subclusters with IPF in GWAS data further suggests these populations may have biomarker potential. Additionally, our trajectory analysis identifies a branch point at the MLM0 cluster from which monocyte-like macrophages diverge toward either inflammatory or fibrotic phenotypes; this suggests the existence of a plastic state during which macrophage polarization might potentially be modulated or redirected through therapeutic intervention.

Our study does however have important limitations. Our study cohort is comprised entirely of male patients, which reflects the real-world demographics of patients who underwent BAL during our study period but nevertheless represents an important limitation that restricts the generalizability of our findings. Given the known sex-related differences in ILD prevalence and outcomes [61,62], with male gender especially associated with greater disease progression and worse survival [63], future studies with a more balanced gender representation are needed. Additionally, our study is limited by a small sample size for rarer ILDs, such as NSIP and PPFE (n=1), which limits robust characterization and comparison for these conditions. We were also only able to analyze two IPF patients with short peripheral blood telomere length which limited our statistical power; a future study with a larger, more balanced cohort of patients would be important to determine whether the findings of

our study are generalizable. Furthermore, since our BAL samples were obtained at initial diagnosis, we cannot characterize changes in cell type proportions or phenotypes following disease progression or in response to treatments such as anti-fibrotics or immunosuppression. It is also important to note that the 10x Genomics technology we utilized is unable to capture neutrophils, which unfortunately limits our understanding of how this cell type contributes to fibrosis. Previous studies have described elevated BAL neutrophil proportions in various ILDs such as IPF [64], SR-ILD [65], and scleroderma-associated ILD [66], with BAL neutrophilia identified as a potential biomarker for reduced survival in progressive ILDs [67]. Indeed, we identified a *PPBP*^{hi} alveolar macrophage subcluster enriched in IPF within our study (S5 Fig), further implicating BAL neutrophils in fibrosis given *PPBP* is a neutrophil chemoattractant [68]. Therefore, our current study likely does not comprehensively describe the pathophysiology of these conditions and future studies should be employed to more definitively investigate the transcriptome of neutrophils in ILD BAL. Finally, while it was not our study's objective, we could not directly correlate gene expression or cell type proportions to clinical metrics such as lung function decline over time, which would be important to explore in future studies. It is important to note that our study is exploratory and cross-sectional in nature; we have not directly correlated cellular phenotypes with clinical outcomes such as lung function decline, treatment response, or survival. Whether the macrophage subpopulations and trajectory patterns we describe can reliably predict disease progression, guide treatment selection, or serve as pharmacodynamic biomarkers will require validation in larger prospective cohorts with longitudinal clinical follow-up. Furthermore, the extent to which macrophage phenotypes are modifiable through therapeutic intervention, and whether such modification translates to clinical benefit, remains to be determined.

Conclusion

Our single-cell RNA sequencing analysis of bronchoalveolar lavage fluid cells reveals distinct cellular compositions across various ILDs, demonstrating that BAL can provide valuable molecular insights through minimally invasive sampling. We were able to characterize diverse macrophage subpopulations and describe airspace monocyte-like macrophages diverging toward inflammatory phenotypes in hypersensitivity pneumonitis, versus fibrotic phenotypes in IPF. Short-telomere IPF patients displayed expanded *SPP1*^{hi} macrophage subpopulations with enhanced fibrotic gene expression, potentially explaining their more aggressive disease course. Integration with GWAS data further highlighted the relevance of specific fibrotic macrophage subclusters in IPF pathogenesis. These findings establish BAL-derived transcriptomics as a powerful approach for understanding fibrotic lung disorders, and suggest specific macrophage subpopulations as potential targets for therapeutic intervention and biomarker development.

Supporting information

S1 Table. Cell type markers used in manual annotation of cell subtypes in BAL.

(DOCX)

S2 Table. Detailed demographics of patients presenting with interstitial lung disease included in this study. IPF: idiopathic pulmonary fibrosis. HP: hypersensitivity pneumonitis. PPFE: pleuroparenchymal fibroelastosis. NSIP: non-specific interstitial pneumonia. SR-ILD: smoking-related interstitial lung disease. FVC: forced vital capacity. DLCO: diffusion capacity of the lung for carbon monoxide. GORD: gastro-oesophageal reflux disease. %pred: percentage of predicted for age/height.

(DOCX)

S3 Table. Differential gene expression between IPF and non-IPF macrophages. Top 20 genes with greatest absolute log2FoldChange between IPF and non-IPF macrophages, with associated adjusted *p*-values included. Positive log2FoldChange indicates relative upregulation in IPF macrophages compared to non-IPF, negative log2FoldChange indicates relative downregulation.

(DOCX)

S4 Table. Top 20 marker genes of each monocyte-like macrophage cluster.

(DOCX)

S1 Fig. Heatmap of top marker gene expression in each cell subtype, in A) myeloid, and B) lymphoid compartments.

(TIF)

S2 Fig. Label transfer of bronchoalveolar lavage macrophage annotations to an independent dataset of IPF BAL (Zhao *et al*, 2024) demonstrates accurate mapping.

(TIF)

S3 Fig. Characterization of T and NK cells in fibrotic BAL. Uniform manifold approximation and projection (UMAP) of T and NK cells isolated from ILD BAL, and cell type proportion analysis of each subcluster (out of all T/NK cells) in each ILD subtype. NK: natural killer.

(TIF)

S4 Fig. Correlation heatmap plots of gene expression between macrophage subsets in A) monocyte-like macrophages (MLM), B) alveolar macrophages (AM), and C) interstitial-like macrophages (ILM). Red indicates positive correlation, blue indicates negative correlation.

(TIF)

S5 Fig. Characterization of alveolar macrophage subclusters in fibrotic BAL. A) Uniform manifold approximation and projection (UMAP) of alveolar macrophage (AM) subclusters isolated from ILD BAL. B) Heatmap of top differentially expressed genes within each AM subcluster. C) Cell type proportion analysis of each alveolar macrophage subcluster (out of all AMs) in each ILD subtype, with significance determined via Wilcoxon rank-sum testing. Only subclusters with significantly different proportions between ILDs are shown. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

(TIF)

S6 Fig. Characterization of interstitial-like macrophage subclusters in fibrotic BAL. Uniform manifold approximation and projection (UMAP) of interstitial-like macrophage (ILM) subclusters isolated from ILD BAL, and cell type proportion analysis of each ILM subcluster (out of all ILMs) in each ILD subtype.

(TIF)

Author contributions

Conceptualization: Simon H. Apte, Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers, Brendan J. O'Sullivan.

Data curation: Lai-Ying Zhang, Peter C. Allen, Brendan J. O'Sullivan.

Formal analysis: Lai-Ying Zhang, Peter C. Allen, Brendan J. O'Sullivan.

Funding acquisition: Joseph E. Powell, Daniel C. Chambers.

Investigation: Lai-Ying Zhang, Peter C. Allen, Viviana P. Lutzky, Simon H. Apte, Eleanor R. Spenceley, Rachael A. McCloy, Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers, Brendan J. O'Sullivan.

Methodology: Lai-Ying Zhang, Peter C. Allen, Viviana P. Lutzky, Penny L. Groves, Maxine E. Tan, Tharushi De Silva, Eleanor R. Spenceley, Rachael A. McCloy, Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers, Brendan J. O'Sullivan.

Project administration: Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers.

Resources: Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers.

Supervision: Simon H. Apte, Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers, Brendan J. O'Sullivan.

Validation: Viviana P. Lutzky, Penny L. Groves.

Visualization: Lai-Ying Zhang, Peter C. Allen, Brendan J. O'Sullivan.

Writing – original draft: Lai-Ying Zhang, Peter C. Allen.

Writing – review & editing: Lai-Ying Zhang, Peter C. Allen, Viviana P. Lutzky, Simon H. Apte, Penny L. Groves, Maxine E. Tan, Tharushi De Silva, Eleanor R. Spenceley, Quan H. Nguyen, Joseph E. Powell, Daniel C. Chambers, Brendan J. O'Sullivan.

References

- 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. <https://doi.org/10.1126/sciadv.aba1983> PMID: 32832599
- Reyffman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med*. 2019;199(12):1517–36. <https://doi.org/10.1164/rccm.201712-2410OC> PMID: 30554520
- Morse C, Tabib T, Sembrat J, Buschur KL, Bittar HT, Valenzi E, et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J*. 2019;54(2):1802441. <https://doi.org/10.1183/13993003.02441-2018> PMID: 31221805
- Unterman A, Zhao AY, Neumark N, Schupp JC, Ahangari F, Cosme C Jr, et al. Single-cell profiling reveals immune aberrations in progressive idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2024;210(4):484–96. <https://doi.org/10.1164/rccm.202306-0979OC> PMID: 38717443
- Unterman A, Bar-Lev TH, Stein Y. Single-cell RNA sequencing of bronchoalveolar lavage reveals pro-inflammatory macrophage activation in fibrotic hypersensitivity pneumonitis. *Eur Respir J*. 2023;62(suppl 67):OA895.
- Schupp J, Adams T, Jäger B, Deluiis G, Ahangari F, Kaminski N. Single cell RNA sequencing profiles distinct BAL cell subpopulations in sarcoidosis. *Eur Respir J*. 2019;54(suppl 63):OA5154.
- Zhao AY, Unterman A, Abu Hussein NS, Sharma P, Nikola F, Flint J, et al. Single-cell analysis reveals novel immune perturbations in fibrotic hypersensitivity pneumonitis. *Am J Respir Crit Care Med*. 2024;210(10):1252–66. <https://doi.org/10.1164/rccm.202401-0078OC> PMID: 38924775
- Courtwright AM, El-Chemaly S. Telomeres in interstitial lung disease: the short and the long of it. *Ann Am Thorac Soc*. 2019;16(2):175–81. <https://doi.org/10.1513/AnnalsATS.201808-508CME> PMID: 30540921
- Fleming SJ, Chaffin MD, Arduini A, Akkad A-D, Banks E, Marioni JC, et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods*. 2023;20(9):1323–35. <https://doi.org/10.1038/s41592-023-01943-7> PMID: 37550580
- Neavin D, Senabouth A, Arora H, Lee JTH, Ripoll-Cladellas A, sc-eQTLGen Consortium, et al. Demuxify: improvement in droplet assignment by integrating multiple single-cell demultiplexing and doublet detection methods. *Genome Biol*. 2024;25(1):94. <https://doi.org/10.1186/s13059-024-03224-8> PMID: 38622708
- Rogozhnikov A, Ramkumar P, Shah K, Bedi R, Kato S, Escola GS. Demuxalot: scaled up genetic demultiplexing for single-cell sequencing. 2021.
- Huang Y, McCarthy DJ, Stegle O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol*. 2019;20(1):273. <https://doi.org/10.1186/s13059-019-1865-2> PMID: 31836005
- Germain P-L, Lun A, Garcia Meixide C, Macnair W, Robinson MD. Doublet identification in single-cell sequencing data using scDblFinder. *F1000Res*. 2021;10:979. <https://doi.org/10.12688/f1000research.73600.2> PMID: 35814628
- Bais AS, Kostka D. scds: computational annotation of doublets in single-cell RNA sequencing data. *Bioinformatics*. 2020;36(4):1150–8. <https://doi.org/10.1093/bioinformatics/btz698> PMID: 31501871
- Bernstein NJ, Fong NL, Lam I, Roy MA, Hendrickson DG, Kelley DR. Solo: doublet identification in single-cell RNA-seq via semi-supervised deep learning. *Cell Syst*. 2020;11(1):95–101.e5. <https://doi.org/10.1016/j.cels.2020.05.010> PMID: 32592658
- Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods*. 2018;15(12):1053–8. <https://doi.org/10.1038/s41592-018-0229-2> PMID: 30504886
- Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science*. 2022;376(6594):eab15197. <https://doi.org/10.1126/science.ab15197> PMID: 35549406
- Heimberg G, Kuo T, DePianto DJ, Salem O, Heigl T, Diamant N, et al. A cell atlas foundation model for scalable search of similar human cells. *Nature*. 2025;638(8052):1085–94. <https://doi.org/10.1038/s41586-024-08411-y> PMID: 39566551
- Sikkema L, Ramírez-Suástegui C, Strobl DC, Gillett TE, Zappia L, Madissoon E, et al. An integrated cell atlas of the lung in health and disease. *Nat Med*. 2023;29(6):1563–77. <https://doi.org/10.1038/s41591-023-02327-2> PMID: 37291214
- 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. <https://doi.org/10.1126/sciadv.aba1972> PMID: 32832598

21. Madisson E, Oliver AJ, Kleshchevnikov V, Wilbrey-Clark A, Polanski K, Richoz N, et al. A spatially resolved atlas of the human lung characterizes a gland-associated immune niche. *Nat Genet.* 2023;55(1):66–77. <https://doi.org/10.1038/s41588-022-01243-4> PMID: [36543915](#)
22. Li X, Kolling FW, Aridgides D, Mellinger D, Ashare A, Jakubick CV. ScRNA-seq expression of IFI27 and APOC2 identifies four alveolar macrophage superclusters in healthy BALF. *Life Sci Alliance.* 2022;5(11):e202201458. <https://doi.org/10.26508/lsa.202201458> PMID: [35820705](#)
23. Franzén O, Gan LM, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database.* 2019;2019.
24. CZI Cell Science Program, Abdulla S, Aevermann B, Assis P, Badajoz S, Bell SM, et al. CZ CELLxGENE discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *Nucleic Acids Res.* 2025;53(D1):D886–900. <https://doi.org/10.1093/nar/gkae1142> PMID: [39607691](#)
25. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* 2024;42(2):293–304. <https://doi.org/10.1038/s41587-023-01767-y> PMID: [37231261](#)
26. Phipson B, Sim CB, Porrello ER, Hewitt AW, Powell J, Oshlack A. propeller: testing for differences in cell type proportions in single cell data. *Bioinformatics.* 2022;38(20):4720–6. <https://doi.org/10.1093/bioinformatics/btac582> PMID: [36005887](#)
27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616> PMID: [19910308](#)
28. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. 2016.
29. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417–25. <https://doi.org/10.1016/j.cels.2015.12.004> PMID: [26771021](#)
30. Mesirov P, Jill, Tamayo P. Molecular signatures database (MSigDB). 2023. <https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>
31. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics.* 2018;19(1):477. <https://doi.org/10.1186/s12864-018-4772-0> PMID: [29914354](#)
32. Saul D, Kosinsky RL, Atkinson EJ, Doolittle ML, Zhang X, LeBrasseur NK, et al. A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. *Nat Commun.* 2022;13(1):4827. <https://doi.org/10.1038/s41467-022-32552-1> PMID: [35974106](#)
33. Zhang MJ, Hou K, Dey KK, Sakaue S, Jagadeesh KA, Weinand K, et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet.* 2022;54(10):1572–80. <https://doi.org/10.1038/s41588-022-01167-z> PMID: [36050550](#)
34. Allen RJ, Stockwell A, Oldham JM, Guillen-Guio B, Schwartz DA, Maher TM, et al. Genome-wide association study across five cohorts identifies five novel loci associated with idiopathic pulmonary fibrosis. *Thorax.* 2022;77(8):829–33. <https://doi.org/10.1136/thoraxjnl-2021-218577> PMID: [35688625](#)
35. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* 2015;11(4):e1004219. <https://doi.org/10.1371/journal.pcbi.1004219> PMID: [25885710](#)
36. Hunter J. Matplotlib: A 2D graphics environment. *Comput Sci Eng.* 2007;9:90–5.
37. Patterson KC, Shah RJ, Porteous MK, Christie JD, D'Errico CA, Chadwick M. Interstitial lung disease in the elderly. *Chest.* 2017;151(4):838–44.
38. Zhao AY, Unterman A, Abu Hussein NS, Sharma P, Nikola F, Flint J, et al. Single-cell analysis reveals novel immune perturbations in fibrotic hypersensitivity pneumonitis. *Am J Respir Crit Care Med.* 2024;210(10):1252–66. <https://doi.org/10.1164/rccm.202401-0078OC> PMID: [38924775](#)
39. Hamblin M, Prosch H, Vařáková M. Diagnosis, course and management of hypersensitivity pneumonitis. *Eur Respir Rev.* 2022;31(163):210169. <https://doi.org/10.1183/16000617.0169-2021> PMID: [35140104](#)
40. Patolia S, Tamae Kakazu M, Chami HA, Chua A, Diaz-Mendoza J, Duggal A, et al. Bronchoalveolar lavage lymphocytes in the diagnosis of hypersensitivity pneumonitis among patients with interstitial lung disease. *Ann Am Thorac Soc.* 2020;17(11):1455–67. <https://doi.org/10.1513/AnnalsATS.202005-420OC> PMID: [32757946](#)
41. Shen Y, Pang C, Wu Y, Li D, Wan C, Liao Z, et al. Diagnostic performance of bronchoalveolar lavage fluid CD4/CD8 ratio for sarcoidosis: a meta-analysis. *EBioMedicine.* 2016;8:302–8. <https://doi.org/10.1016/j.ebiom.2016.04.024> PMID: [27428439](#)
42. Portillo K, Guasch I, Becker C, Andreo F, Fernández-Figueras MT, Ramirez Ruz J, et al. Pleuroparenchymal fibroelastosis: a new entity within the spectrum of rare idiopathic interstitial pneumonias. *Case Rep Pulmonol.* 2015;2015:810515. <https://doi.org/10.1155/2015/810515> PMID: [26380141](#)
43. Hanumegowda C, Farkas L, Kolb M. Angiogenesis in pulmonary fibrosis: too much or not enough?. *Chest.* 2012;142(1):200–7. <https://doi.org/10.1378/chest.11-1962> PMID: [22796840](#)
44. Partanen JJ, Häppölä P, Zhou W, Lehisto AA, Ainola M, Sutinen E, et al. Leveraging global multi-ancestry meta-analysis in the study of idiopathic pulmonary fibrosis genetics. *Cell Genom.* 2022;2(10):100181. <https://doi.org/10.1016/j.xgen.2022.100181> PMID: [36777997](#)
45. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet.* 2013;45(6):613–20. <https://doi.org/10.1038/ng.2609> PMID: [23583980](#)
46. Natri HM, Del Azodi CB, Peter L, Taylor CJ, Chugh S, Kendle R, et al. Cell-type-specific and disease-associated expression quantitative trait loci in the human lung. *Nat Genet.* 2024;56(4):595–604. <https://doi.org/10.1038/s41588-024-01702-0> PMID: [38548990](#)
47. Meyer KC, Raghu G, Baughman RP, Brown KK, Costabel U, du Bois RM, et al. An official American Thoracic Society clinical practice guideline: the clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease. *Am J Respir Crit Care Med.* 2012;185(9):1004–14. <https://doi.org/10.1164/rccm.201202-0320ST> PMID: [22550210](#)

48. Mayr CH, Santacruz D, Jarosch S, Bleck M, Dalton J, McNabola A, et al. Spatial transcriptomic characterization of pathologic niches in IPF. *Sci Adv.* 2024;10(32):eadl5473. <https://doi.org/10.1126/sciadv.adl5473> PMID: 39121212
49. Wang J, Zhang L, Luo L, He P, Xiong A, Jiang M, et al. Characterizing cellular heterogeneity in fibrotic hypersensitivity pneumonitis by single-cell transcriptional analysis. *Cell Death Discov.* 2022;8(1):38. <https://doi.org/10.1038/s41420-022-00831-x> PMID: 35091537
50. Agostini C, Calabrese F, Poletti V, Marcer G, Facco M, Miorin M, et al. CXCR3/CXCL10 interactions in the development of hypersensitivity pneumonitis. *Respir Res.* 2005;6(1):20. <https://doi.org/10.1186/1465-9921-6-20> PMID: 15725351
51. Newton CA, Zhang D, Oldham JM, Kozlitina J, Ma S-F, Martinez FJ, et al. Telomere length and use of immunosuppressive medications in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2019;200(3):336–47. <https://doi.org/10.1164/rccm.201809-1646OC> PMID: 30566847
52. Molina-Molina M. Telomere shortening is behind the harm of immunosuppressive therapy in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2019;200(3):274–5. <https://doi.org/10.1164/rccm.201812-2330ED> PMID: 30624965
53. Newton CA, Batra K, Torrealba J, Kozlitina J, Glazer CS, Aravena C, et al. Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive. *Eur Respir J.* 2016;48(6):1710–20. <https://doi.org/10.1183/13993003.00308-2016> PMID: 27540018
54. Deshmane SL, Kremlev S, Amini S, Sawaya BE. Monocyte chemoattractant protein-1 (MCP-1): an overview. *J Interferon Cytokine Res.* 2009;29(6):313–26. <https://doi.org/10.1089/jir.2008.0027> PMID: 19441883
55. Baran CP, Opalek JM, McMaken S, Newland CA, O'Brien JM Jr, Hunter MG, et al. Important roles for macrophage colony-stimulating factor, CC chemokine ligand 2, and mononuclear phagocytes in the pathogenesis of pulmonary fibrosis. *Am J Respir Crit Care Med.* 2007;176(1):78–89. <https://doi.org/10.1164/rccm.200609-1279OC> PMID: 17431224
56. Bailey JI, Puritz CH, Senkow KJ, Markov NS, Diaz E, Jonasson E, et al. Profibrotic monocyte-derived alveolar macrophages are expanded in patients with persistent respiratory symptoms and radiographic abnormalities after COVID-19. *Nat Immunol.* 2024;25(11):2097–109. <https://doi.org/10.1038/s41590-024-01975-x> PMID: 39367123
57. Matsuda S, Kotani T, Kuwabara H, Suzuka T, Kiboshi T, Fukui K, et al. CCL2 produced by CD68+/CD163+ macrophages as a promising clinical biomarker of microscopic polyangiitis-Interstitial lung disease. *Rheumatology (Oxford).* 2021;60(10):4643–53. <https://doi.org/10.1093/rheumatology/keab064> PMID: 33493350
58. Tan S, Chen S. Macrophage autophagy and silicosis: current perspective and latest insights. *Int J Mol Sci.* 2021;22(1):453.
59. Mukherjee A, Epperly MW, Fisher R, Hou W, Shields D, Wang H, et al. Silica Induced Lung Fibrosis Is Associated With Senescence, Fgr, and Recruitment of Bone Marrow Monocyte/Macrophages. *In Vivo.* 2021;35(6):3053–66. <https://doi.org/10.21873/in vivo.12601> PMID: 34697137
60. Groves PL, Hockey L, O'Sullivan BJ, Zhang L-Y, Xiong Z, Nguyen QH. Transcriptomic plasticity of human alveolar macrophages revealed by single-cell RNA sequencing following drug exposure: implications for therapeutic development. 2025.
61. Zaman T, Moua T, Vittinghoff E, Ryu JH, Collard HR, Lee JS. Differences in Clinical Characteristics and Outcomes Between Men and Women With Idiopathic Pulmonary Fibrosis: A Multicenter Retrospective Cohort Study. *Chest.* 2020;158(1):245–51. <https://doi.org/10.1016/j.chest.2020.02.009> PMID: 32084394
62. Xiang P, Wang L, Feng X, Guo Q, Xie G, Sheng L, et al. Sexual dimorphism of lung immune-regulatory units imprint biased pulmonary fibrosis. *Cell Mol Immunol.* 2025;22(7):743–59.
63. Tondo P, Scioscia G, De Pace CC, Murgolo F, Maci F, Stella GM, et al. Gender differences are a leading factor in 5-year survival of patients with idiopathic pulmonary fibrosis over antifibrotic therapy reduction. *Life (Basel).* 2025;15(1):106. <https://doi.org/10.3390/life15010106> PMID: 39860046
64. Ryu YJ, Chung MP, Han J, Kim TS, Lee KS, Chun E-M, et al. Bronchoalveolar lavage in fibrotic idiopathic interstitial pneumonias. *Respir Med.* 2007;101(3):655–60. <https://doi.org/10.1016/j.rmed.2006.06.003> PMID: 16919929
65. Caminati A, Cavazza A, Sverzellati N, Harari S. An integrated approach in the diagnosis of smoking-related interstitial lung diseases. *Eur Respir Rev.* 2012;21(125):207–17. <https://doi.org/10.1183/09059180.00003112> PMID: 22941885
66. Goh NSL, Veeraraghavan S, Desai SR, Cramer D, Hansell DM, Denton CP, et al. Bronchoalveolar lavage cellular profiles in patients with systemic sclerosis-associated interstitial lung disease are not predictive of disease progression. *Arthritis Rheum.* 2007;56(6):2005–12. <https://doi.org/10.1002/art.22696> PMID: 17530640
67. Watase M, Mochimaru T, Kawase H, Shinohara H, Sagawa S, Ikeda T, et al. Diagnostic and prognostic biomarkers for progressive fibrosing interstitial lung disease. *PLoS One.* 2023;18(3):e0283288. <https://doi.org/10.1371/journal.pone.0283288> PMID: 36930615
68. Smith NLD, Bromley MJ, Denning DW, Simpson A, Bowyer P. Elevated levels of the neutrophil chemoattractant pro-platelet basic protein in macrophages from individuals with chronic and allergic aspergillosis. *J Infect Dis.* 2015;211(4):651–60. <https://doi.org/10.1093/infdis/jiu490> PMID: 25193981