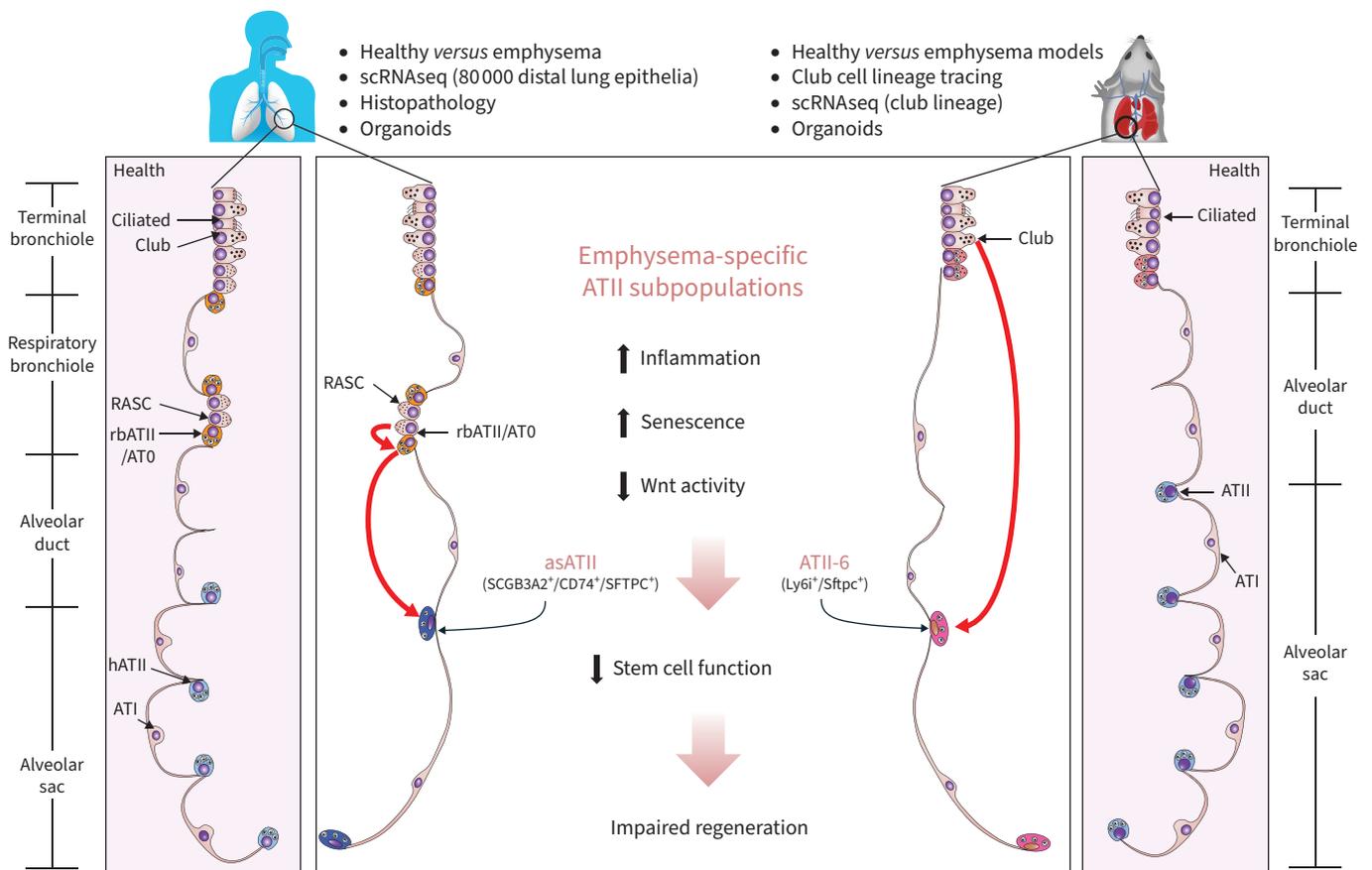


Airway-derived emphysema-specific alveolar type II cells exhibit impaired regenerative potential in COPD

Yan Hu , Qianjiang Hu , Meshal Ansari, Kent Riemondy, Ricardo Pineda , John Sembrat, Adriana S. Leme, Kenny Ngo, Olivia Morgenthaler, Kellie Ha, Bifeng Gao, William J. Janssen, Maria C. Basil , Corrine R. Kliment, Edward Morrissey, Mareike Lehmann , Christopher M. Evans, Herbert B. Schiller and Melanie Königshoff



GRAPHICAL ABSTRACT Main findings of the study. In human emphysema and mouse models, club cells give rise to an emphysema-specific alveolar type II (ATII) cell subpopulation in enlarged alveolar sacs, characterised by impaired regenerative function. scRNAseq: single-cell RNA sequencing; RASC: respiratory airway secretory cell.



Airway-derived emphysema-specific alveolar type II cells exhibit impaired regenerative potential in COPD

Yan Hu ¹, Qianjiang Hu ², Meshal Ansari ³, Kent Riemondy ⁴, Ricardo Pineda ², John Sembrat ², Adriana S. Leme ², Kenny Ngo ¹, Olivia Morgenthaler ¹, Kellie Ha ¹, Bifeng Gao ¹, William J. Janssen ⁵, Maria C. Basil ^{6,7,8}, Corrine R. Kliment ², Edward Morrissey ^{6,7,9}, Mareike Lehmann ^{3,10}, Christopher M. Evans ^{1,13}, Herbert B. Schiller ^{11,13} and Melanie Königshoff ^{2,12,13}

¹Division of Pulmonary Sciences and Critical Care Medicine, University of Colorado School of Medicine, Aurora, CO, USA. ²Center for Lung Aging and Regeneration (CLAR), Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA. ³Comprehensive Pneumology Center (CPC)/Institute of Lung Health and Immunity (LHI), Helmholtz Munich, Member of the German Center for Lung Research (DZL), Munich, Germany. ⁴RNA Bioscience Initiative, University of Colorado School of Medicine, Aurora, CO, USA. ⁵Department of Medicine, National Jewish Health, Denver, CO, USA. ⁶Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁷Penn-CHOP Lung Biology Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁸Penn Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA, USA. ⁹Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ¹⁰Institute for Lung Research, Philipps-University Marburg, Member of the German Center for Lung Research (DZL), Marburg, Germany. ¹¹Research Unit Precision Regenerative Medicine (PRM), Helmholtz Munich, Comprehensive Pneumology Center (CPC), Member of the German Center for Lung Research (DZL), Munich, Germany. ¹²Geriatric Research Education and Clinical Center (GRECC) at the VA Pittsburgh Healthcare System, Pittsburgh, PA, USA. ¹³Co-senior authors.

Corresponding author: Melanie Königshoff (koenigshoffm@upmc.edu)



Shareable abstract (@ERSpublications)

This study describes an ATII subpopulation derived from airway club cells with unique localisation, transcriptome and impaired progenitor cell function in emphysema, which contribute to impaired tissue regeneration <https://bit.ly/4da4fT8>

Cite this article as: Hu Y, Hu Q, Ansari M, *et al.* Airway-derived emphysema-specific alveolar type II cells exhibit impaired regenerative potential in COPD. *Eur Respir J* 2024; 64: 2302071 [DOI: 10.1183/13993003.02071-2023].

Copyright ©The authors 2024.

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

This article has an editorial commentary:
<https://doi.org/10.1183/13993003.01741-2024>

Received: 28 Nov 2023
Accepted: 25 July 2024

Abstract

Emphysema, the progressive destruction of gas exchange surfaces in the lungs, is a hallmark of COPD that is presently incurable. This therapeutic gap is largely due to a poor understanding of potential drivers of impaired tissue regeneration, such as abnormal lung epithelial progenitor cells, including alveolar type II (ATII) and airway club cells. We discovered an emphysema-specific subpopulation of ATII cells located in enlarged distal alveolar sacs, termed asATII cells. Single-cell RNA sequencing and *in situ* localisation revealed that asATII cells co-express the alveolar marker surfactant protein C and the club cell marker secretoglobin-3A2 (SCGB3A2). A similar ATII subpopulation derived from club cells was also identified in mouse COPD models using lineage labelling. Human and mouse ATII subpopulations formed 80–90% fewer alveolar organoids than healthy controls, indicating reduced progenitor function. Targeting asATII cells or their progenitor club cells could reveal novel COPD treatment strategies.

Introduction

COPD is a leading cause of morbidity and mortality worldwide [1]. One major phenotype of COPD is emphysema, which is characterised by airspace enlargement and loss of delicate gas exchange surfaces due to the progressive destruction of parenchymal lung tissue. Common causes of emphysema include chronic exposure to tobacco smoke, noxious gas, and particles, which trigger oxidative stress, inflammation, proteolytic tissue breakdown and loss of alveolar epithelial cells. Among these, tissue restoration may be a more effective strategy than targeting injurious and inflammatory stimuli, since most patients are diagnosed at advanced stages of disease when tissue damage is already significant. However, reversal of alveolar destruction in emphysema is presently unachievable with existing therapies. This lack of effective treatments is primarily due to our limited understanding of the cellular and molecular mechanisms underlying the failure of alveolar repair.



Airway club cells [2–9] and alveolar type II (ATII) cells [10–12] are known lung epithelial progenitor cells with the potential to induce alveolar repair, especially after acute injuries. However, their implication in human COPD and emphysema remains under-investigated. Recent single-cell RNA sequencing (scRNAseq) studies [13–17] have initiated characterisation of club/secretory cells (termed terminal airway-enriched secretory cells (TASCs) [17]/respiratory airway secretory cells (RASCs) [16]) and ATII cells (termed AT0 cell [15]) in the terminal/respiratory bronchioles of healthy and diseased human lungs, including COPD lungs [13, 14, 16, 17]. However, these studies lack sufficient numbers of ATII cells derived from severely emphysematous parenchyma in COPD, due to selection of airway dominant tissue [16, 17] or lack of epithelial cell enrichment [13, 14]. As such, the potential functional and molecular heterogeneities of ATII cells in the emphysematous alveolar sacs, which are four generations downstream from the respiratory bronchioles [18], remain unclear.

Here we report a scRNAseq dataset of enriched epithelial cells from emphysematous parenchymal lung tissue of COPD patients (n=6) and from non-lung disease controls (n=4). Our high-resolution analysis includes 78 699 ATII cells, a significant increase from the <5000 ATII cells reported in prior studies involving 62 patients [13, 14, 16, 17]. This expanded dataset revealed novel insights into emphysematous tissue-specific ATII cell subsets expressing the club cell marker gene secretoglobin-3A2 (SCGB3A2). We further investigated and characterised a subset of cells that localise to enlarged alveolar sacs in human COPD patients with severe emphysema. Additionally, we identified a similar subpopulation of ATII cells originating from club cells using mouse models of COPD. Mechanistically, our study also demonstrates reduced progenitor cell capacity of these SCGB3A2⁺ ATII cells from both human COPD patients and the mouse emphysema model.

Methods

Human lung tissue collection

Emphysematous parenchymal lung tissues were surgically dissected from explanted lung lobes of de-identified patients with COPD in Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage IV (very severe: forced expiratory volume in 1 s <30% predicted [19]) (n=6), who received lung transplantation. Protocol was approved by UHealth University of Colorado Hospital (COMIRB 11–1664). Healthy lung tissues of de-identified donors without lung diseases (n=4) were obtained from the human Lung Tissue Research Consortium at National Jewish Health (NJHIRB HS-1797). Details are provided in supplementary table S1.

Human lung tissue processing and scRNAseq

Two to three pieces (2 cm×2 cm) of parenchymal lung tissues were sampled from regions with visible air trapping under pleura in each COPD lungs or comparable regions in donor lungs. Tissues without pleura were minced with sterile scissors and digested by collagenase and dispase for 1 h at 37°C (collagenase 0.1 Wünsch U·mL⁻¹; dispase 0.8 Wünsch U·mL⁻¹; Roche 11097113001). The homogenised tissues after digestion were filtered through 100 µm and 20 µm pore size sterile nylon mesh (Sefar) sequentially in a vacuum system and washed with DMEM/F12 media (Gibco, 11330–032) with PenStrep (100 U·mL⁻¹, Gibco 15140122) and DNaseI (0.1 mg·mL⁻¹, Roche 11284932001). The single-cell suspension was centrifuged at 300×g for 10 min at 15°C, followed by immune cell depletion with magnetic-activated cell sorting (MACS) beads conjugated with anti-human CD45 antibody (Miltenyi 130-045-801) and MACS LS column (Miltenyi 130-042-401). Then cells were stained for PerCP/Cyanine5.5 anti-human epithelial cell adhesion molecule (EpCAM) antibody (BioLegend 369804) and DAPI (4',6-diamidino-2-phenylindole, dihydrochloride) (Invitrogen D1306). Fluorescence-activated cell sorting (FACS) sorted viable epithelial cells (DAPI⁻/EpCAM⁺) were used for single cell capture by 10X Genomics Chromium Controller. scRNAseq libraries were generated using 10X Genomics Chromium single cell 3' Library & Gel Bead Kit v2. Sequencing was performed on illumine NovaSeq6000. Details of sequencing are provided in supplementary table S1.

Mouse COPD models

The *Scgb1a1-CreERTM*; *Rosa-tdTomato* mice were generated by crossing *Scgb1a1-CreERTM* (Jax 016225) [2] and Ai9 (Jax 007909) [20] mice. Both male and female mice aged 8–10 weeks were subject to intraperitoneal injections of tamoxifen (200 mg·kg⁻¹ bodyweight; Sigma, T5648) in corn oil (Sigma, C8267) every 2 days for three doses. 1 week after the last dose of tamoxifen, mice were treated with a single dose of porcine pancreatic elastase (PPE) (40 U·kg⁻¹ bodyweight in 80 µL volume; Sigma, 45124) or saline *via* oropharyngeal injection [21]. All animal PPE treatment experiments were performed according to the institutional and regulatory guidelines of University of Colorado institutional animal care and use committee.

All animal smoke exposure experiments were performed in accordance with the institutional animal care and use committee of the University of Pittsburgh. Animals were housed according to standard housing criteria. C57BL/6 mice obtained from Jackson Laboratories (female mice aged 10–12 weeks, $n=3-4$ per group) were subjected to the smoke of four unfiltered cigarettes per day (lot #3R4F; University of Kentucky, Lexington, KY, USA), 5 days a week for a duration of 6 months, using a smoking apparatus that delivers targeted cigarette smoke to single mice isolated in individual chambers [22, 23]. The controls in each group were exposed to room air alone. These mice were caged separately and housed in the same facility as their smoke-exposed counterparts. At the completion of each experiment, mice were euthanised by carbon dioxide inhalation, the chest was opened, and the trachea was cannulated. The lungs were perfused with 10 mL of ice-cold PBS. The lungs were then inflated with 10% buffered formalin at a constant pressure of 25 cmH₂O for 10 min. The lungs were then ligated, excised and fixed in formalin for 24 h before washing in PBS, storing in 70% ethanol and embedding in paraffin. Serial midsagittal sections were obtained for histological analysis.

Mouse lung epithelial cell isolation

At day 21 (and day 28 for detection of ATII-6 cells in alveolar sacs of emphysematous region by co-immunofluorescence (co-IF) for lymphocyte antigen 6 complex I (Ly6i)) after PPE or saline treatment, mouse lung lobes were inflated and submerged with dispase (Corning, 354235) at room temperature for 45 min. The digested lung lobes were further homogenised in DMEM/F12 with PenStrep (100 U·mL⁻¹, Gibco 15140122) and DNaseI (0.1 mg·mL⁻¹, Roche 11284932001). The single cell suspension was sequentially filtered through 70 µm and 40 µm cell strainers on top of 50 mL tubes, followed by centrifuge at 300×*g* for 10 min at 15°C. Epithelial cells were enriched with anti-mouse CD326 (Miltenyi 130-105-958) and MACS LS column (Miltenyi 130-042-401). Cells were stained with APC anti-mouse Epcam (BioLegend 118214) and DAPI for flow cytometric analysis and FACS sorting. For scRNAseq, FACS sorted viable tdTomato⁺ (endogenous)/Epcam⁺ cells from PPE ($n=3$) and saline ($n=2$) treated *Scgb1a1-CreERTM;Rosa-tdTomato* mice were used to generate scRNAseq libraries as described earlier.

Human single-cell RNA data analysis

FASTQ files of human single-cell RNAseq were processed and aligned to the hg38 human reference genome (GRCh38.97) using the Cell Ranger computational pipeline from 10X Genomics (v3.1.0, STAR v2.5.3a). Gene expression data of each sample were integrated and analysed using the Scanpy (v1.7.1). Cells with high counts were filtered out using sample specific thresholds (>12 000 counts, >18 000 counts or >25 000 counts). Then, cells with <350 genes and cells with >20% mitochondria reads were filtered out. Genes detected in fewer than three cells were further filtered out.

We performed doublet detection with Scrublet (v0.2.3) [24] on the individual sample level. Next, ambient RNA correction was performed for each sample separately with SoupX (v1.6.1, background contamination fraction was manually set to 0.3) [25] followed by scanpy's size factor based normalisation (v1.24.1) [26] using cell groupings obtained through Louvain cluster (resolution=2.0) as input for both. The normalised count matrix was log-transformed with scanpy.pp.log1p().

To account for individual variances, we first calculated the most variable genes for each patient *via* the Scanpy function scanpy.pp.highly_variable_genes(flavour="cell_ranger") and then selected those genes that ranked among the top 2000 most variable genes in at least five samples. After removal of cell cycle genes, we obtained 599 highly variable genes that were used as input for the principal component analysis. To further mitigate effect of individual patient variance, the k-nearest neighbours graph was constructed in a batch-balanced manner using BBKNN (v1.4.0) [27] with the parameters $n_pcs=50$, $batch_key="patient_ID"$ and $neighbours_within_batch=15$. The two-dimensional embedding was then generated with scanpy.tl.umap() and louvain clustering was performed at resolution 1 and 2. The annotation of the resulting clusters was based on canonical marker genes.

For trajectory analysis, we integrated our COPD ATII and club cells with data published by BASIL *et al.* [16] *via* Seurat's function IntegrateLayers (method=CCAIntegration) [28]. The neighbourhood graph (dim=30) and uniform manifold approximation and projection (UMAP) was constructed based on the shared embedding following standard procedure. Slingshot was used to infer the principal curve between the lineages.

To assess disease specificity of asATII cells, we used scArches (v0.5.10) [29] to map our human ATII subsets (query) to the integrated Human Lung Cell Atlas (HLCA) [30]. To assess the potential overlap of cell type population from our query and the HLCA reference, Leiden clustering at resolution 1 was performed on the integrated latent space.

Murine lung cell scRNAseq analysis

FASTQ files were processed and aligned to the mm10 reference genome using the cell ranger count pipeline (10X Genomics). Gene expression data were analysed using the Seurat R package. Cells from the first experiment (one saline-treated mouse and one PPE-treated mouse) were filtered to only include those with >500 UMI (unique molecular identifier) counts, <60 000 UMI counts, >1500 genes detected and <12.5% mitochondrial counts. Cells from the second experiment (one saline-treated mouse and two PPE-treated mice) were filtered to include those with >500 UMI counts, <20 000 UMI counts, >200 genes detected, <5000 genes detected and <3 mean absolute deviations and <20% mitochondrial counts. UMI counts were normalised by the total counts for the cell, multiplied by a scale factor (10 000), and log-transformed (NormalizeData). Normalised counts were scaled and centred (ScaleData) using the top 3000 variable features (FindVariableFeatures). The scaled data were used for PCA (RunPCA), and the first 20 principal components were used for clustering (FindNeighbors, FindClusters) and to calculate uniform manifold approximation and projection (RunUMAP). The two experiments were integrated using Harmony, with the θ parameter set to 1 [31]. Cell types were annotated using clustifyr with a murine lung regeneration single-cell dataset as the reference [8, 32]. Differentially expressed genes were identified for each cluster using the presto R package (<https://github.com/immunogenomics/presto>).

Data and code availability

The human lung scRNAseq data have been deposited in Gene Expression Omnibus with accession number GSE222374. Data processing and analysis code is provided at a GitHub repository (https://github.com/KonigshoffLab/HU_COPD_2023).

Murine scRNAseq data have been deposited at the Gene Expression Omnibus (GSE218813). Murine scRNAseq analysis code is provided at a Github repository (<https://github.com/rnabioco/Koenigshoff-murine-copd>).

Histology and immunofluorescence staining

Human parenchymal lung tissues were obtained as described earlier. Four to six 1.5 cm×1.5 cm pieces from each lung were fixed in 10% formalin overnight, followed by dehydration in 70% ethanol for ≥24 h before paraffin embedding and sectioning at thickness of 5 μ m. Mouse lungs were inflated and fixed with 4% paraformaldehyde (PFA) overnight at 4°C, followed by dehydration in 70% ethanol. Paraffin sections of mouse lungs at thickness of 5 μ m were generated as described earlier. For immunofluorescence staining, paraffin sections were deparaffinised and rehydrated as described previously [21]. Antigen retrieval was performed using 10 mM sodium citrate (pH 6.0) boiled in a pressure cooker. Sections were incubated with 1X Carbo-Free Blocking Solution (Vector Laboratories SP-5040-125) diluted in PBS with 0.1% Triton-X100 (PBST) for 30 min at room temperature, followed by incubation with primary antibodies at 4°C overnight. After washing three times with PBS, sections were incubated with secondary antibodies in darkness for 2 h at room temperature. Then slides were washed and mounted with anti-fade mounting media with DAPI (Invitrogen P36935). Immunofluorescence was visualised using an Olympus BX63 microscope and Olympus cellSens software. QuPath software [33] was used to measure the numbers of cells labelled by staining.

Antibodies

Primary antibodies used for immunofluorescence include rabbit anti-SPC (Millipore, AB3786), rabbit anti-SPC (Abcam, ab90716), mouse IgM anti-HTII280 (Terrace Biotech, TB-27AHT2-280), CD74 (Novus, NBP1-33109), intercellular adhesion molecule 1 (ICAM1) (Novus, BBA3), TM4SF1 (R&D Systems, MAB8164), SCGB3A2 (R&D Systems, AF3545), acetylated α -tubulin (Abcam, ab24610), tdTomato (Biorbyt, orb182397) and chicken anti-Ly6i antibody. The Ly6i antibody was customised polyclonal antibody generated and validated by Rockland targeting peptide sequence Ac-CPDEIEKKFILDPN TKM-amide. Secondary antibodies used for immunofluorescence include anti-rabbit-Alexa Fluor488 (Invitrogen, A-21206), anti-rabbit-Alexa Fluor647 (Invitrogen, A-31573), anti-rabbit-Alexa Fluor546 (Invitrogen, A-10040), anti-mouse-Alexa Fluor546 (Invitrogen, A-10036), anti-mouse-Alexa Fluor488 (Invitrogen, A-21202), anti-chicken-Alexa Fluor488 (Invitrogen, SA5-10070) and anti-goat-Alexa Fluor 647 (Invitrogen, A32849).

RNA in situ hybridisation

RNA *in situ* hybridisation was performed on 5- μ m paraffin sections of human COPD and healthy lungs using RNAScope Multiplex Fluorescent Assay v.2 kit (ACD) following the manufacturer's instructions.

Tissue sections were deparaffinised in xylene (2×5 min), followed by 100% ethanol (2×2 min). Then, sections were treated with RNAScope hydrogen peroxide for 10 min at room temperature, before boiling in

target retrieval reagent for 15 min and protease plus treatment at 40°C in HybEZ Oven for 15 min. After pre-treatment, the tissue sections were incubated with probes for hsSCGB3A2 (ACD 549951), hsSCGB3A1 (ACD 593251) and hsSCGB1A1 (ACD 469971) at 40°C for 2 h, followed by signal amplification and incubation with HRP-C1-3, HRP blockers, and Opal Dyes (1:1000 dilution; Akoya, Opal 520 FP1487001KT, Opal 620 FP1495001KT, Opal 690 FP1497001KT) based on user manuals. For co-immunofluorescence staining for pro-surfactant protein C (SPC), tissue sections were processed for immunofluorescence from incubation with blocking solution as described earlier. ~60 images at 20× randomly distributed on stained tissue were obtained using an Olympus BX63 microscope and Olympus cellSens software. QuPath software [33] was used to measure the numbers of cells labelled by immunofluorescent staining.

Flow cytometry and FACS

For mouse lung cell flow cytometry analysis and sorting, cells were stained with allophycocyanin (APC) anti-mouse CD326 (BioLegend, 118214), APC-Cy7 anti-mouse Sca-1 antibodies (BioLegend, 108126), chicken anti-Ly6i antibody (Rockland, custom polyclonal antibody targeting peptide sequence Ac-CPDEIEKKFILDNPNTKM-amide) at room temperature followed by PBS wash and incubation with donkey anti-chicken IgY antibody conjugated with fluorescein isothiocyanate (FITC) (Sigma-Aldrich, AP194F) at room temperature for 20 min in darkness. Cells were washed and re-suspended in PBS with 0.1% fetal bovine serum (FBS). To detect dead cells, DAPI (Invitrogen, D1306) was added at a concentration of 2 µg·mL⁻¹ before analysis or sorting.

For human ATII cell flow cytometry analysis and sorting, lung epithelia cells were first stained with mouse IgM anti-HTII280 antibody (Terrace Biotech, TB-27AHT2-280) at 1:50 dilution at room temperature for 20 min. Cells were washed with PBS and incubated with goat anti-mouse IgM-FITC secondary antibody for 20 min at room temperature. Then cells were washed with PBS twice followed by incubation with PerCP/Cyanine5.5 anti-human CD326 antibody (Biolegend, 324214), APC anti-human CD74 antibody (Biolegend, 326812), PE anti-human CD54 antibody (Biolegend, 353106) and mouse IgG1 anti-TM4SF1 (R&D Systems, MAB8164) conjugated with APC-Cy7 using Lightning-Link Conjugation Kit (Abcam, ab102859) at room temperature for 20 min in darkness. Cells were washed and resuspended in PBS with 0.1% FBS with DAPI before analysis or sorting.

Flow cytometry analysis was performed using BD LSRFortessa™ Cell Analyzer (BD Biosciences). Cell sorting was performed using BD FACSAria™ Fusion cell sorter (BD Biosciences).

Organoid culture

For mouse ATII cell organoid cultures, sorted Ly6i⁺ and Ly6i⁻ ATII cell-derived club cells (DAPI⁻/Epcam⁺/tdTomato⁺/Sca1⁻) were isolated as described earlier. For organoid culture, 5000 sorted ATII cells were mixed with mitomycin C treated MRC5 fibroblast cells at a 1:1 ratio in 50 µL growth factor reduced Matrigel (diluted with DMEM/F12 to a concentration of 5 mg·mL⁻¹; Corning 354263) and plated in transwell inserts (Falcon 353095) for 24-well plates. The plates were incubated at 37°C for 10 min to solidify the Matrigel. Cultures were maintained in an air/liquid interface condition by supplying organoid culture media outside of the inserts, described previously [21, 34, 35]. The organoid culture media contained DMEM/F12 (Gibco, 11330-032) containing 5% FBS, penicillin/streptomycin (100 U·mL⁻¹), 1% GlutaMax (Life Technologies, 35050-061), 1X amphotericin B (Gibco, 15290018), 1X insulin-transferrin-selenium (Gibco #51300-044), recombinant mouse epidermal growth factor (EGF) (25 ng·mL⁻¹, Sigma, SRP3196), cholera toxin (0.1 µg·mL⁻¹, Sigma C8052) and bovine pituitary extract (30 µg·mL⁻¹, Sigma, P1476). Y-27632 (10 µM, Tocris, 1254) was added for the first 48 h of culture to prevent anoikis. Medium was renewed every 2 days for up to 4 weeks before fixation. Organoids with Feret diameter >25 µm were counted.

For human lung organoid culture, 10 000 sorted epithelial cells were mixed with 10 000 mitomycin C pre-treated with 10 000 MRC5 fibroblast cells in 100 µL Matrigel (diluted with DMEM/F12 to a concentration of 5 mg·mL⁻¹; Corning 354263) in transwell inserts, as described earlier. Human organoid culture media contains recombinant human EGF (25 ng·mL⁻¹, Peprotech) replacing recombinant mouse EGF and all the other components in the mouse organoid media. Organoids were maintained in culture for 4 weeks before fixation. Organoids with Feret diameter >30 µm were counted.

$$\text{Organoid forming efficiency (\%)} = \frac{\text{Number of organoids per insert} \times 100}{10000 \text{ (human) or } 5000 \text{ (mouse)}}$$

Immunofluorescence staining of organoids

For whole-mount immunofluorescence staining, mouse lung organoids in Matrigel were fixed by adding 4% PFA on top of Matrigel and outside inserts for 10 min at room temperature, followed by PBS wash and permeabilisation using acetone and methanol (1:1) at -20°C for 15 min. After PBS wash, 1X Carbo-Free Blocking Solution diluted in PBST was added to the top of Matrigel in inserts to block for 4 h at room temperature. Then organoids were incubated with primary antibodies diluted in PBST at 4°C overnight, followed by PBS washing three times (10 min each time). Finally, organoids were incubated with mixture of secondary antibodies and DAPI diluted in PBST for 2 h at room temperature or overnight at 4°C . Organoids were imaged within the culture plate using a Cytation1 Cell Imaging Multi-Mode Reader (BioTek), and images were obtained and analysed with Gen5 Image+ software (BioTek).

Results

Molecular heterogeneity of ATII cells in emphysema

To identify the molecular abnormalities of distal lung epithelial progenitor cells in emphysematous alveoli, we generated scRNAseq libraries of EPCAM⁺ live epithelial cells from emphysematous parenchymal tissue of six end-stage (GOLD IV) COPD patients and from comparable distal lung parenchymal regions of four age-matched non-lung disease never-smoker donors [36] (figure 1a and supplementary table S1). After quality control filtering, we analysed 46 471 EPCAM⁺ cells from COPD lungs and 35 914 EPCAM⁺ cells from controls, with an average of 1811 genes detected per cell (figure 1b and supplementary table S1).

Annotations based on known epithelial cell marker genes [37–39] identified eight major cell types. These included alveolar type II (ATII, SFTPC⁺) cells, alveolar type I (ATI, AGER⁺) cells, ciliated cells (FOXJ1⁺), club cells (MUC5B-/SCGB1A1⁺), mucous cells (MUC5B⁺/SCGB1A1⁺), basal cells (KRT5⁺/KRT17⁺), as well as basaloid cells (KRT5⁻/KRT17⁺), which were recently described in healthy and diseased lungs [36, 40] (figure 1c and d and supplementary figure S1). Due to their central roles as distal lung progenitor cells, we focused further on ATII cells.

Our dataset consists of a large number of ATII cells (n=10 donors, k=78 699 cells, 95.5% of total cells), enabling a comprehensive analysis of human lung ATII cell heterogeneity. Analysis of differentially expressed genes (DEGs) between emphysematous and healthy ATII cells identified 152 upregulated genes (\log_2 fold change >1) and 461 downregulated genes (\log_2 fold change <-1) (figure 1e, supplementary table S2). Gene Ontology enrichment analysis [41] using these DEGs identified inflammatory responses, tissue morphogenesis, epithelial–mesenchymal transition and stress response pathways that were increased in emphysematous ATII cells (supplementary figure S2a, supplementary table S3). By contrast, downregulated pathways included cellular metabolism, transmembrane transport, DNA repair and mitochondrial function (supplementary figure S2b, supplementary table S3).

Subclustering of ATII cells from healthy and COPD lungs (figure 1f,g) revealed expression of secretoglobin genes dominantly in COPD ATII cells, including SCGB1A1, SCGB3A1 and SCGB3A2 (figure 1h, supplementary figure S1g and h), which are canonical markers of airway club cells. Notably, SCGB3A2 was recently identified in club/secretory cells of respiratory bronchioles (RASCs) and in the ATII cells of alveoli in the respiratory bronchiolar region (AT0) in healthy and COPD lungs [15, 16] (figure 1f and g), suggesting that these COPD-specific ATII cell subsets might represent respiratory bronchiole-associated cells that could be involved in halted alveolar repair.

Identification of secretoglobin-expressing ATII subpopulations in human COPD lungs

To annotate the ATII subclusters in our scRNAseq data and to localise the ATII subpopulations expressing SCGB genes *in situ*, we performed multiplexed fluorescence *in situ* hybridisation for three SCGB mRNAs and co-immunofluorescence staining for the common ATII cell marker SPC in human COPD (n=9) and control lung tissue (n=5).

Surprisingly, cells co-expressing SCGB3A2 and SFTPC mRNA (3A2⁺/SPC⁺) were localised to enlarged alveolar sacs within emphysematous regions specifically in the COPD lungs (figure 2a), but they were rare or absent in alveolar sacs of healthy lungs (figure 2c, supplementary figure S3a). Based on their tissue localisation, we annotated these 3A2⁺/SFTPC⁺ ATII cells as alveolar sacs (as)ATII in our scRNAseq data (figure 2d, supplementary figure S4). Distinct from asATII, we found ATII cells co-expressing SCGB3A2 together with SCGB1A1 or SCGB3A1 in individual alveoli of respiratory bronchioles, similar to the localisation of AT0 cells identified recently in the healthy lung [15] (figure 2b and c, supplementary figure S3b and c). Here we annotate these multi-SCGB^{pos} ATII cells as respiratory bronchiole (rb)ATII cells (figure 2d, supplementary figure S4). Importantly, in COPD scRNAseq samples, asATII (30.4%) and rbATII (66.1%) cells were highly enriched populations (figure 2e). The ATII cells without SCGB gene

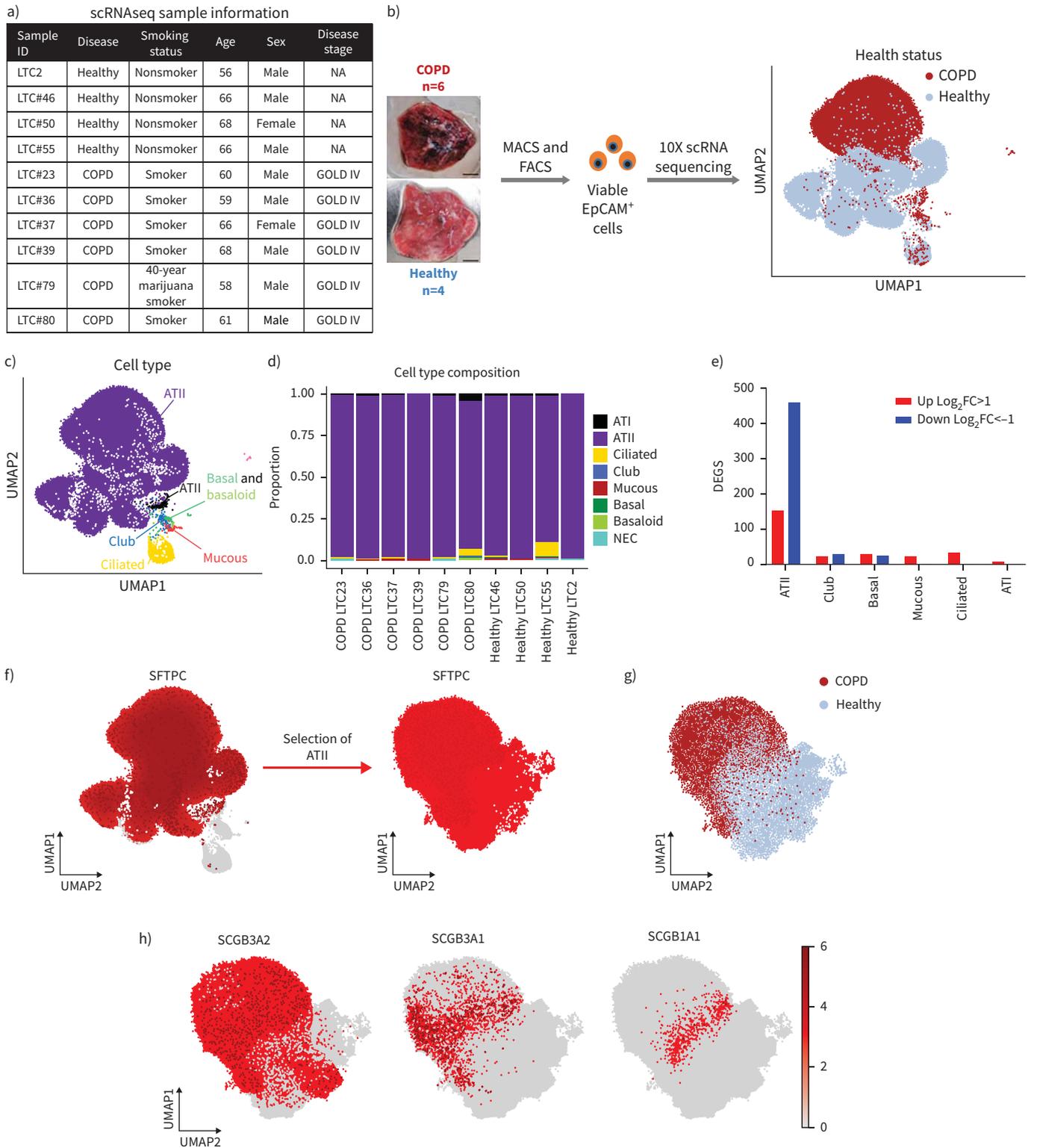


FIGURE 1 Alveolar epithelial cell heterogeneity in COPD. **a)** Sample information of single-cell RNA sequencing (scRNAseq) data. **b)** Schematic showing workflow of sampling parenchymal tissue from emphysematous (n=6) and healthy (n=4) lungs, viable epithelial cell isolation and scRNAseq using the 10X Genomics platform. Uniform manifold approximation and projection (UMAP) shows cells obtained from COPD patients and healthy donors. **c)** UMAP visualisation of annotated lung epithelial cell types. **d)** The cellular composition of epithelial cells from individual COPD patients and healthy donors. **e)** The numbers of genes upregulated (log₂ fold change (fc) >1, p<0.05) and downregulated (log₂fc <-1, p<0.05) in each cell type from COPD patients, compared to those of healthy donor. **f)** UMAP visualisation of alveolar type II (ATII) cells marked by expression of surfactant protein C (SFTPC). **g)** UMAP showing ATII cells from emphysema patients and healthy donors. **h)** Feature plots showing expression of

secretoglobins *SCGB3A2*, *SCGB3A1* and *SCGB1A1* in ATII cells. NA: not applicable; GOLD: Global Initiative for Chronic Obstructive Lung Disease; MACS: magnetic-activated cell sorting; FACS: fluorescence-activated cell sorting; EpCAM: epithelial cell adhesion molecule; NEC: nonepithelial cells; DEGs: differentially expressed genes.

expression are predominantly from the healthy control lungs (88%; figure 2e), which were termed hATII cells (figure 2d, supplementary figure S4). Notably, both asATII and rbATII cells exhibited pronounced ageing characteristics compared to hATII cells, as demonstrated by higher levels of senescence marker genes (figure 2f), elevated senescence scores, and reduced mitochondrial function scores. (supplementary figure S5).

To test whether asATII and rbATII cells are specific to COPD, we mapped [29] our human ATII subsets to the HLCA [30] data, which contains idiopathic pulmonary fibrosis (IPF) and interstitial lung disease (ILD) samples, in addition to healthy and COPD samples (supplementary figure S6a and b). After unsupervised clustering of ATII cells in the merged data (supplementary figure S6c), we identified both asATII and rbATII cells in a dominant ATII cluster (cluster 20) specific to COPD (supplementary figure S6d and f). However, neither asATII nor rbATII cells were found in IPF/ILD samples (supplementary figure S6e and g).

In addition, we tested expression of AT0 [15] and RASC-derived LAMP3⁺ ATII cells [16] marker genes in our ATII subpopulations. Neither asATII nor rbATII showed significant enrichment of those markers (supplementary figure 6h), suggesting that asATII and rbATII cells are distinct from previously described cell subpopulations.

Next, we aimed to further identify unique marker genes of asATII and rbATII cells (figure 2g) for functional characterisation. Compared to hATII and asATII cells, we found the highest expression of *ICAM1* and the Wnt active ATII progenitor marker *TM4SF1* [11] in rbATII cells. In asATII cells, we observed elevated levels of the Wnt inhibitor *WIF1*, the oxidative stress marker *SOD2*, and inflammatory genes, such as *CD74* and *CXCL8* [42–46]. These results indicated potentially distinct functions for asATII and rbATII cells.

Club cells give rise to emphysema-specific ATII subpopulations in human and mouse COPD lungs

An important gap in the field is understanding how bronchiolar airway club cells contribute to alveolar repair upon chronic injury in COPD. From our human single cell data, trajectory inference algorithms (Slingshot and PAGA velocity) support the hypothesis that club or rbATII cells may differentiate to asATII cells in COPD (figure 3a). To test whether RASCs contribute to asATII cells, we integrated [28] RASC and secretory/goblet cell data from healthy and COPD lungs [16] to our ATII subsets and club cells (figure 3b). Slingshot analysis indeed supports the notion that RASCs contribute to the asATII population.

To further confirm and study if the club cell lineage can differentiate to emphysema-specific ATII cells in COPD, we utilised the established COPD mouse model of porcine pancreatic elastase (PPE) instillation to induce acute emphysematous lung injury [47]. Briefly, *Scgb1a1-Cre^{ERT};Rosa-tdTomato* mice were first treated with tamoxifen to label club cells for lineage tracing *in vivo*. 2 weeks later, this was followed by a single challenge with PPE (40 U·kg⁻¹, oropharyngeal instillation) (figure 3c). On day 21 post-challenge, mice receiving PPE had enlarged alveolar sacs and altered pressure–volume relationships consistent with emphysema (supplementary figure S7a and b) [21].

We then analysed expansion and localisation of tdTomato-labelled (tdTomato^{pos}) epithelial cells obtained from whole lung digests using flow cytometry. PPE-treated mice showed a 50% increase in tdTomato^{pos} cells (38.0±2.2% after PPE *versus* 25.3±1.3% after saline) (supplementary figure S7c), suggesting expansion of club lineage cells. To test whether this was accompanied by increased differentiation of club cells to ATII cells, we quantified ATII cells along with co-expression of tdTomato using co-immunofluorescence staining. However, there was no significant change in the percentage of tdTomato^{pos} cell among total SPC⁺ ATII cells in the alveolar region of PPE-treated lungs, compared to saline controls (supplementary figure S7d and e). Taken together, these data demonstrate that club-to-ATII differentiation was not increased in emphysema induced by PPE challenge in mice.

To further identify cell types derived by club cells upon emphysematous lung injury, we performed scRNAseq on tdTomato^{pos} epithelial cells from PPE- (n=3) and saline- (n=2) treated mice. We found six main epithelial cell types, including ATII, ATI, club (*Scgb1a1⁺Muc5b⁻*), mucous (*Scgb1a1⁺Mu5b⁺*), club-to-ciliated, ciliated (*Foxj1⁺*) and proliferating cells (*Mki67⁺*). Among these, ATII cells comprised the biggest population (figure 3d, supplementary figure S8a).

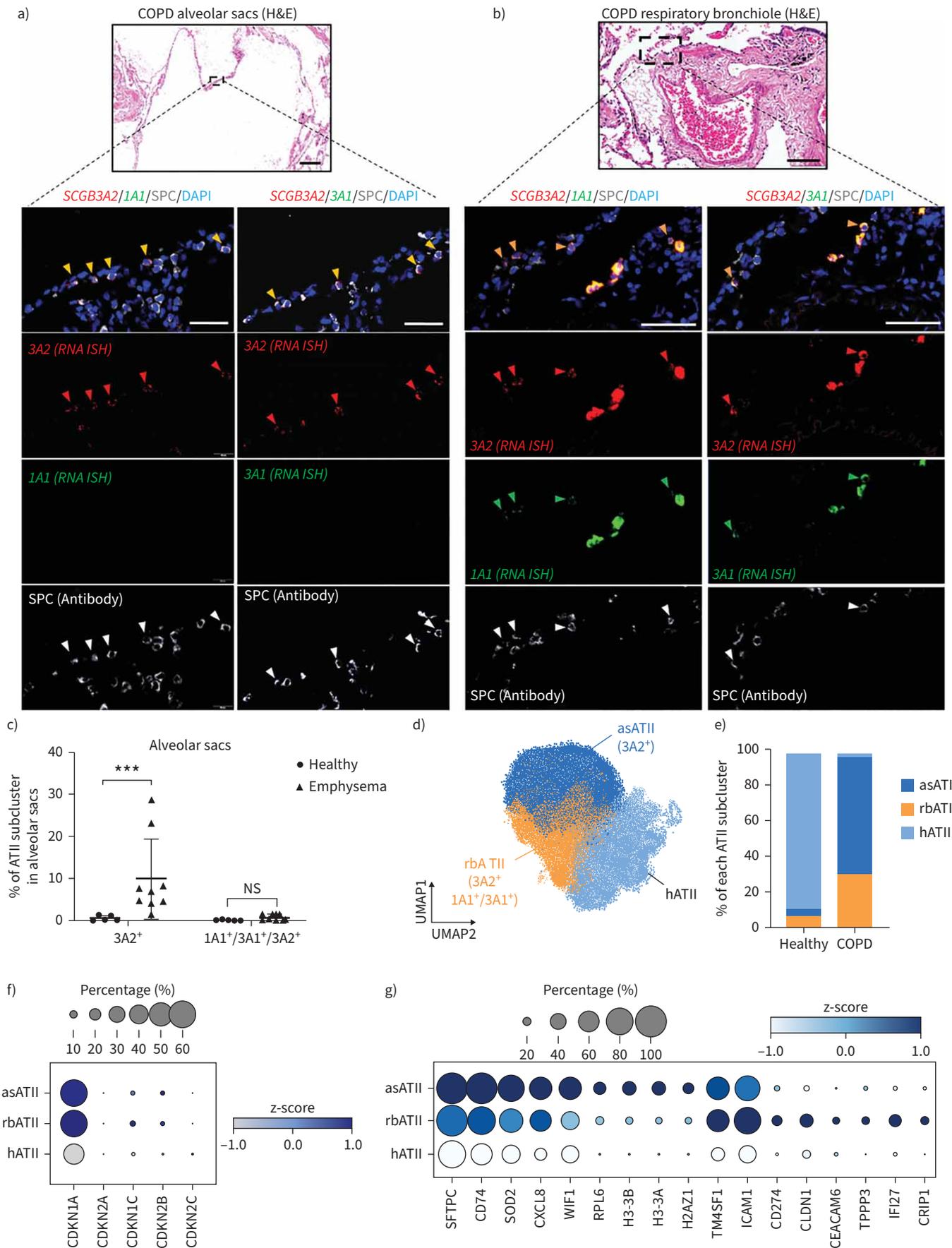


FIGURE 2 Identification of secretoglobin (SCGB^{pos}) alveolar type II (ATII) cell subpopulations in human COPD lungs. **a)** Haematoxylin and eosin (H&E) staining of emphysematous parenchymal lung tissue from COPD patient at low magnification. Scale bar=200 μ m. Within thin alveolar sacs (inset), fluorescent RNAScope for *SCGB3A2* (red), *IA1* (green), *3A1* (green) and immunofluorescence (IF) staining for surfactant protein C (SPC) (white) identifies *3A2*⁺ ATII cells (arrowheads). Scale bars=40 μ m. **b)** H&E staining of respiratory bronchiole region (inset) in COPD lung tissue. Scale bar=100 μ m. Multiplex fluorescence RNAScope for *SCGB3A2* (red), *IA1* (green), *3A1* (green), and co-IF for SPC (white) identifies ATII cells co-expressing multiple SCGB genes (arrowheads) in this region. Scale bars=40 μ m. **c)** Quantification of percentages of *3A2*⁺ and *IA1*⁺/*3A1*⁺/*3A2*⁺ ATII cells detected by RNAScope and IF in alveolar sacs of healthy (n=5) and emphysematous (n=9) lung tissue. **d)** Uniform manifold approximation and projection (UMAP) showing annotation of ATII subclusters based on expressions of SCGB genes and tissue localisations. ATII cells dominantly presenting in alveolar sacs of COPD samples and expressing *SCGB3A2* were identified as asATII. ATII cells expressing multiple SCGBs and localised in respiratory bronchioles were annotated as rbATII. The remaining of ATII cells, dominantly presented in healthy samples, were marked as hATII. **e)** Percentages of ATII subclusters in healthy and COPD samples. **f)** Dot plot of selected senescence marker genes in asATII, rbATII and hATII cells. **g)** Dot plot of selected marker genes identified in asATII and rbATII cells. ns: nonsignificant. ***: $p < 0.001$. ns: $p > 0.05$, Mann-Whitney test.

To investigate whether there are mouse ATII subpopulations similar to human asATII cells, we subclustered mouse club cell derived ATII cells into six groups (ATII-1 to -6; figure 3d). Among these, ATII-6 cells were significantly enriched after PPE challenge (13% of tdTomato^{pos} ATII in PPE, 2.2% of tdTomato^{pos} ATII in saline) (figure 3e). We next used matchScore analysis to quantify shared marker genes in mouse and human ATII subpopulations (supplementary table S5) [48]. Notably, mouse ATII-6 cells and human asATII cells had the highest matchScore (figure 3f). Moreover, we show that ATII-6 and asATII share many marker genes, including CD74 (figure 3g), which is not seen in the ATII clusters with a matchScore of 0, such as mouse ATII-5 and human hATII cells (figure 3h). Indeed, gene set enrichment analysis (GSEA) of marker genes of both mouse ATII-6 and human asATII cells further revealed a strong overlap in biological processes across species, including inflammation, cell adhesion and hypoxia responses (supplementary figure S8b and c and supplementary table S6). These overlaps were not found in the GSEA of ATII-2, which share a matchScore of 0.01 with human asATII (supplementary figure S8d and supplementary table S6). Collectively, our data reveal club cell derived ATII subpopulations that are associated with alveolar tissue destruction and emphysema in both species.

Mouse ATII-6 and human asATII cells display limited progenitor function

To test the overall progenitor function of club cell-derived tdTomato^{pos} cells, we performed lung organoid assays. We found that tdTomato^{pos} cells isolated from PPE-treated mouse lungs formed fewer alveolar organoids than those from saline controls (figure 4a and b). Having revealed impaired alveolar regeneration function in club cell derived subsets, we next sought to determine whether this could be explained by a specific emphysema-enriched ATII cell subpopulation identified in our scRNAseq analysis. We focused on the mouse ATII-6 cluster since it is enriched in PPE-treated samples and is similar to human asATII cells. Mouse ATII-6 cells selectively express lymphocyte antigen 6 complex i (*Ly6i*) [49] as a surface marker (figure 4c). We generated an antibody targeting the extracellular region of *Ly6i*, confirmed that *Ly6i* localised to tdTomato/SPC dual-positive alveolar epithelial in PPE-challenged mouse lungs and employed *Ly6i* for cell sorting (figure 4d and e).

Flow cytometry analysis of whole lung cells from PPE-treated mice demonstrated that 6.5% of epithelial cells were tdTomato⁺/*Ly6i*⁺ (supplementary figure S9). We used *Ly6i* in combination with *Sca1* to specifically demarcate ATII cells, which are known to be *Sca1*-negative (*Sca1*⁻) [3]. We quantified tdTomato⁺/*Sca1*⁻/*Ly6i*⁺ ATII-6 cell populations in the PPE challenge mouse lungs (figure 4e). ATII-6 cell numbers increased in proportion to total tdTomato⁺ ATII cells on days 21 (4.48 \pm 2.88%) and 28 (15.05 \pm 3.96%) after PPE, but they were rare in saline controls (0.6 \pm 0.53%) (figure 4f). By contrast, ATII cells expressing *Ly6i*, but lacking the club cell reporter tdTomato, were not detected (supplementary figure S9c, figure 4f). Thus, ATII-6 cells in emphysematous mouse lung tissues are derived from club cells and not the result of off-target tdTomato reporter expression.

Additionally, we tested whether *Ly6i*⁺ ATII-6 cells exist in the lungs of C57BL/6 mice challenged with cigarette smoke (CS), which develops mild emphysema with >6 months of treatment [50, 51]. Immunofluorescence co-labelling for *Ly6i* and SPC was performed on lung tissue slides from mice treated with CS or filtered air for 6 months, respectively (supplementary figure S10a). Quantification showed an increased percentage of *Ly6i*⁺/SPC⁺ cells in total SPC⁺ ATII cells in CS-treated lungs (6.2 \pm 3.2%) compared to controls (1.0 \pm 0.7%) (supplementary figure S10b).

Functionally, ATII-6 cells isolated from mouse PPE model demonstrated poor alveolar progenitor functions evident by the significantly lower alveolar organoid-forming capacity of *Ly6i*⁺ ATII-6 cells compared to

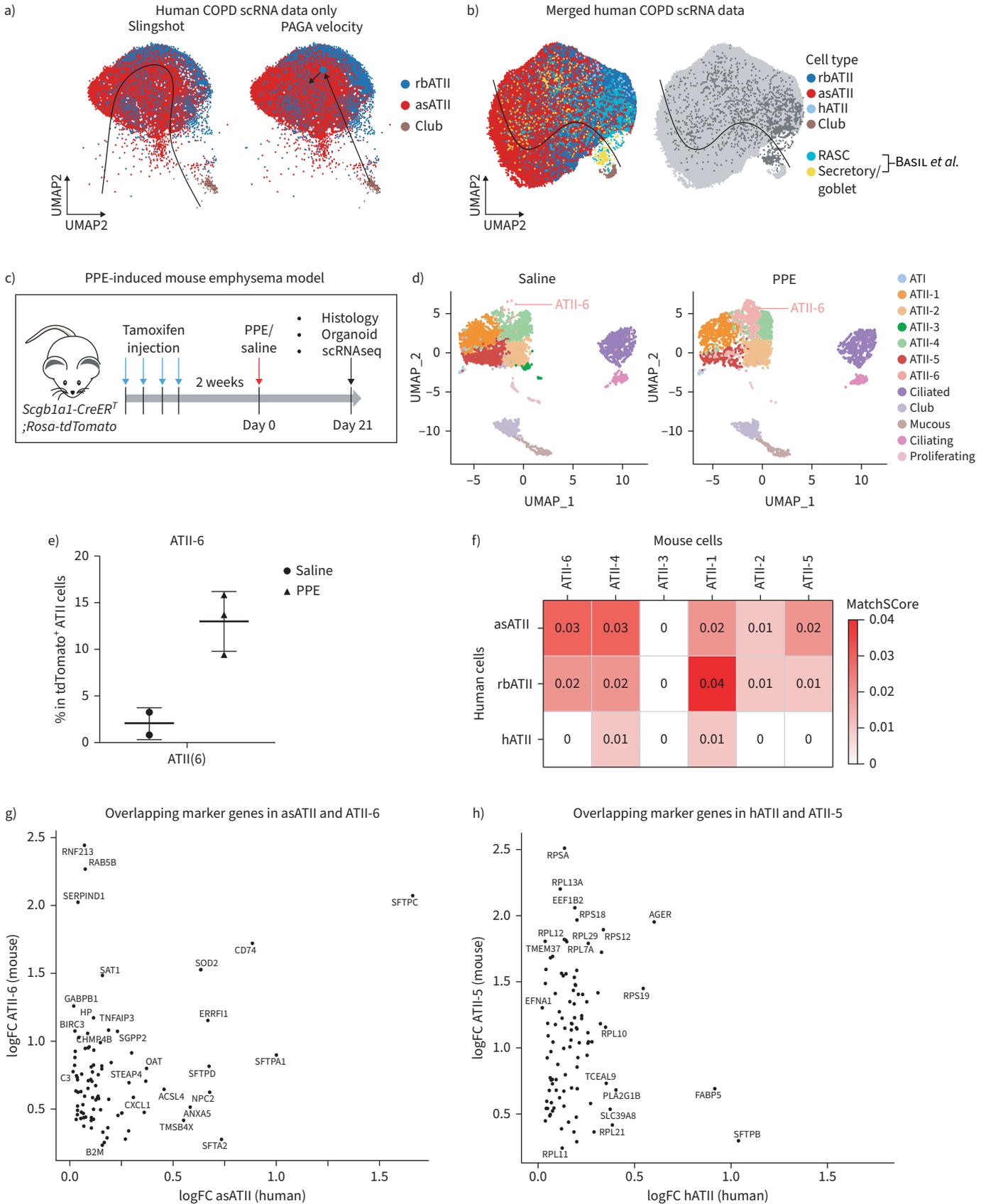


FIGURE 3 Club cells give rise to emphysema-specific alveolar type II (ATII) cell subpopulations in human and mouse COPD lungs. **a)** Principal curve representing lineage between non-mucin club cells and alveolar sac (as)ATII cells using Slingshot, and partition-based graph abstraction (PAGA)

analysis of asATII, respiratory bronchiole (rb)ATII, and nonmucin club cells from COPD lungs. Arrows mark the predicted trajectory from nonmucin club cells to asATII cells. **b**) Data from **a**) are merged with respiratory airway secretory cells (RASCs) and secretory/goblet (Sec/Gob) cells published by BASIL *et al.* [16] to perform Slingshot analysis. **c**) Experimental schematics of lineage tracing of club cells in the porcine pancreatic elastase (PPE)-induced mouse emphysema model. **d**) Uniform manifold approximation and projection (UMAP) visualisation and annotation of sorted tdTomato⁺ epithelial cells from saline- (n=2) and PPE- (n=3) treated Scgb1a1-CreERT;Rosa-tdTomato mice, respectively. **e**) Percentage of ATII-6 subcluster cells in tdTomato⁺ ATII cells in **(d)**. **f**) Similarity measurement of human (h)ATII subpopulations and mouse ATII subclusters based on expression of common marker genes. Higher matchScore indicates higher similarity. **g**) Scatter plots showing marker genes driving similarity between human asATII cells and mouse ATII-6 in **(f)**. **h**) Scatter plot showing a lack of common marker gene in human hATII and mouse ATII-5 cells. scRNAseq: single-cell RNA sequencing.

the rest of tdTomato⁺ ATII cells from PPE-challenged mouse lungs. Notably, tdTomato⁺ ATII cells from healthy lungs, which lack ATII-6 cells, showed the highest organoid-forming efficiency (figure 4g and h). These results demonstrate a limited alveolar regeneration capacity of club cell derived ATII-6 cells in a mouse model of emphysema.

Finally, we validated the murine model results in human organoid assays by comparing alveolar organoid forming capacities of human asATII, rbATII cells and healthy ATII cells. We used CD74 as surface marker for asATII cells (figures 2f and 5a), and ICAM1 and TM4SF1 for rbATII cells (figures 2f and 5b). Flow cytometry analysis identified HTII280⁺/CD74⁺ asATII cells in emphysematous parenchymal tissue samples, but they were rare in healthy controls (figure 5c and d). Functionally, asATII cells had the lowest alveolar organoid-forming efficiency compared to rbATII cells from emphysematous tissue and healthy ATII cells from control tissues (figure 5e and f). Taken together, we identified asATII cells that are enriched in the distal emphysematous airspace in human COPD lungs and have significant defects in progenitor cell function, which could contribute to impaired tissue repair in COPD.

Discussion

Pathogenesis of emphysema involves chronic inflammation and oxidative stress in lung parenchymal tissue, as well as impaired endogenous tissue regeneration, probably through dysfunctional epithelial progenitor cells [52, 53]. However, precise molecular mechanisms controlling progenitor functions in disease remains elusive. Here we report a scRNAseq dataset with a large number of ATII cells from human lung tissue with severe emphysema and from healthy lungs, thereby allowing deep profiling of emphysematous ATII cells. We identified a unique ATII subpopulation present in emphysematous alveolar sacs and expression of *SCGB3A2*, named asATII. In healthy lungs, asATII cells were rare. Trajectory inference analysis suggested a potential contribution of club cells, including recently published RASCs, to this emphysema-specific ATII subpopulation. Importantly, through lineage tracing of club cells in a mouse emphysema model, we showed that club cells indeed gave rise to an ATII subcluster (ATII-6) with a similar transcriptomic profile and tissue localisation as human asATII cells. Notably, these cells were also found to be increased in CS-exposed mouse lungs, which represents a clinically relevant COPD model. Moreover, functional analyses using organoids demonstrated that these newly identified cell ATII subpopulations in human and mouse emphysema exhibit impaired alveolar regeneration capacity. Taken together, we have unravelled a novel mechanism of club cells differentiating to transcriptionally distinct ATII cell populations in emphysema. These may help to explain impaired tissue regeneration during COPD pathogenesis and to develop early therapeutics.

SCGB3A2 is a marker for respiratory bronchiolar club-like cells, including recently identified RASCs and TASCs [16, 17]. Additionally, *SCGB3A2* expression was identified in a subpopulation of ATII-like cells, termed AT0, within single alveoli attached to the respiratory bronchioles in healthy human, primate and ferret lungs [13]. On average, the respiratory bronchioles lead to up to 11 alveolar ducts, which in turn derive approximately six alveolar sacs [18]. It is crucial to highlight that *SCGB3A2* expression in alveolar sacs has not been reported previously. Thus, the asATII subpopulation reported in our study is distinct from the AT0 based on localisation and gene expression.

Our high-resolution analysis of emphysematous epithelial cells with scRNAseq and usage of sensitive RNA *in situ* hybridisation allowed us to identify the *SCGB3A2*⁺/*SFTPC*⁺ asATII cells within the enlarged alveolar sacs distal to respiratory bronchioles in severe emphysematous regions. Nonetheless, our DEG analysis highlights several key features of the asATII cells, including senescence, mitochondrial dysfunction, inflammation, oxidative stress [54], apoptosis and inhibition of Wnt signalling, which are known to be critical for stem cell activation, but are downregulated in COPD lungs [55, 56]. Stem cell

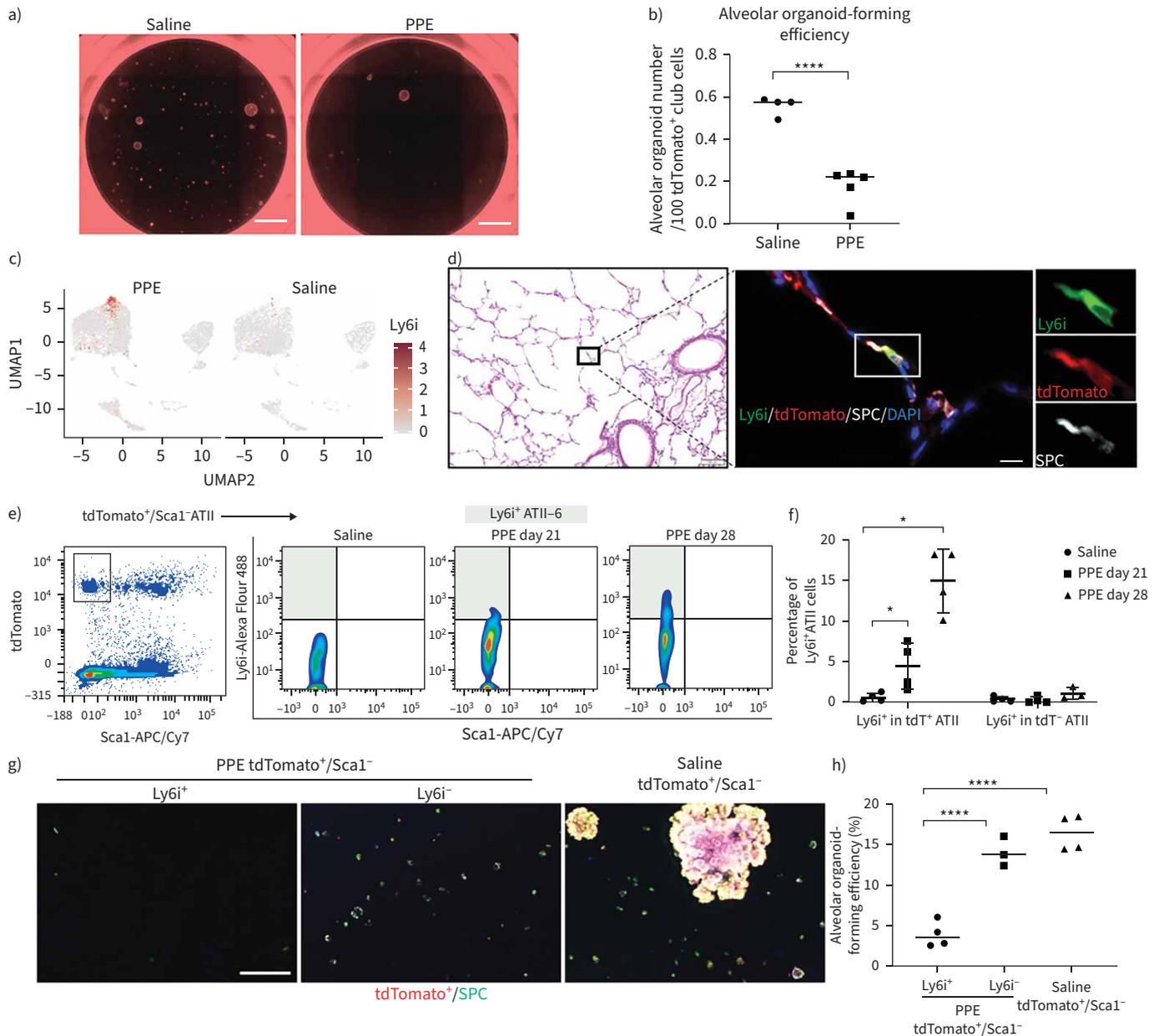


FIGURE 4 Identification of emphysema-specific alveolar type II (ATII) cells in a porcine pancreatic elastase (PPE)-induced mouse COPD model. **a)** Representative images of organoids from tdTomato⁺ (red) epithelial cells of saline- and PPE-treated Scg1a1-CreERT;Rosa-tdTomato mice on day 21 post-treatment. Scale bar=1000 μm. **b)** Quantification of alveolar organoids in Saline-treated mice (n=4), PPE-treated mice (n=5). Welch's t-test. ****: p<0.0001. **c)** Feature plot showing lymphocyte antigen 6 complex i (Ly6i) as a unique marker gene of mouse ATII-6 cells. **d)** Detection of ATII-6 cells in alveolar sacs of emphysematous region (black square in left picture with haematoxylin and eosin staining) by co-immunofluorescence for Ly6i (right, green), tdTomato (right, red), and surfactant protein C (SPC) (right, white). Scale bar=20 μm. **e)** Detection of tdTomato⁺/Sca1⁻/Ly6i⁺ ATII cells in viable epithelial cells (see gating strategy in supplementary figure S7a) from PPE-treated mouse lungs by flow cytometry. **f)** Percentages of Ly6i⁺ cells in the tdTomato⁺/Sca1⁻ and tdTomato⁻/Sca1⁻ ATII cells from mice treated with saline or PPE on days 21 and 28 after treatment. Mann-Whitney test. * p<0.05. **g)** Representative pictures of organoids from fluorescence-activated cell sorted ATII-6 (tdTomato⁺/Sca1⁻/Ly6i⁺, left) and the rest (tdTomato⁺/Sca1⁻/Ly6i⁻, middle) ATII cells from PPE-treated mice at day 21 post-treatment and ATII (tdTomato⁺/Sca1⁻) cells from saline-treated mice. Whole-mount immunofluorescence staining for pro-SPC (green) is used to identify alveolar organoids. Endogenous tdTomato is in red. 4',6-Diamidino-2-phenylindole (DAPI) in blue shows nuclei. Scale bar=1000 μm. **h)** Quantification of alveolar (SPC⁺) organoid-forming efficiency in Ly6i⁺, Ly6i⁻, and Saline tdTomato⁺/Sca1⁻ cells. ****: p<0.0001, Mann-Whitney test.

exhaustion is a hallmark of ageing and might be driven by cellular senescence [57], which is enriched in asATII cells. Further studies investigating the role of senescence and mitochondrial dysfunction in asATII cells are needed.

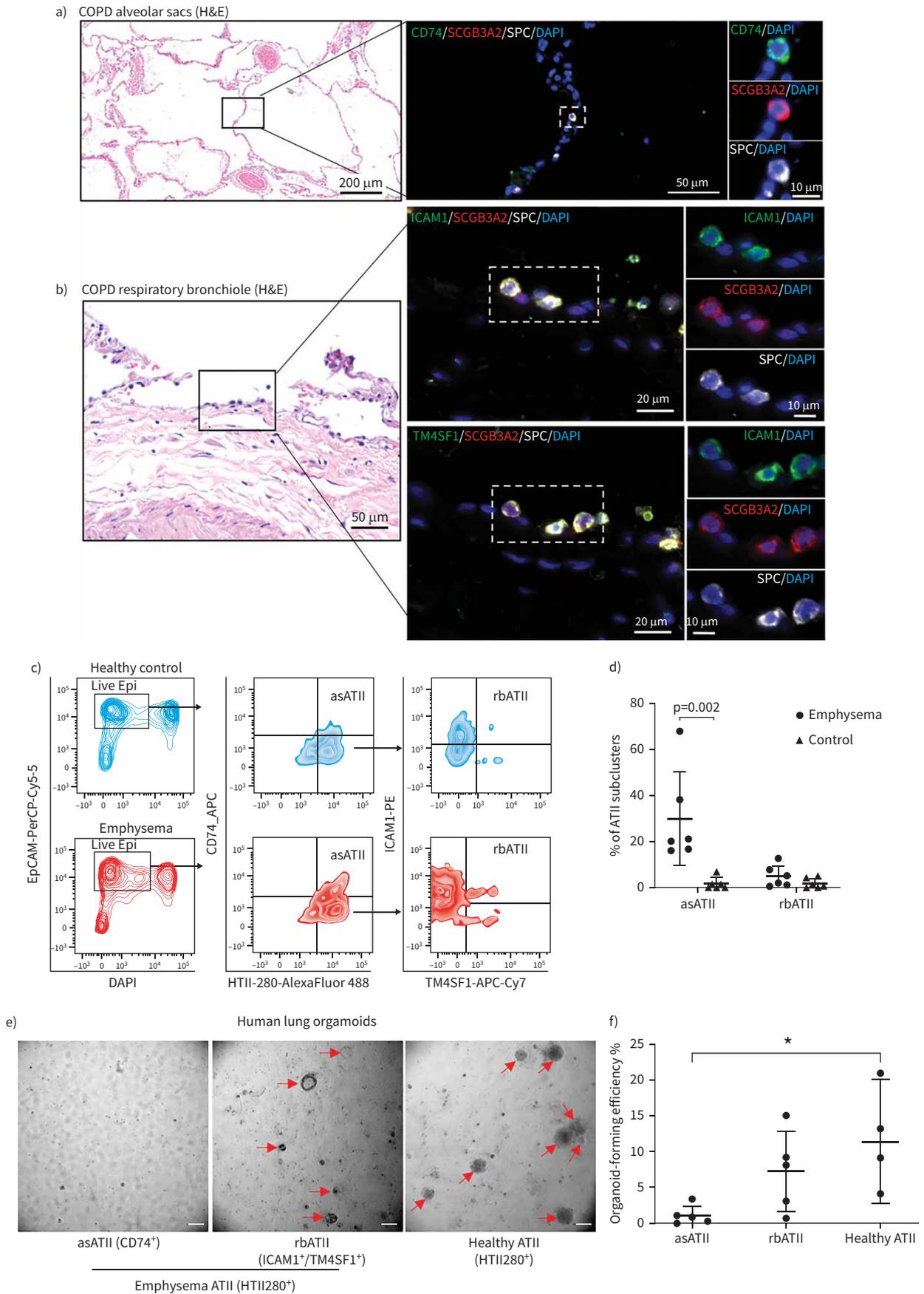


FIGURE 5 Decreased progenitor cell function of emphysema-specific alveolar type II (ATII) cells in human COPD lungs. **a)** Detection of asATII cells in alveolar sacs of emphysematous region (black square in left picture with haematoxylin and eosin (H&E) staining) by co-immunofluorescence (co-IF) for CD74 (green), secretoglobin-3A2 (SCGB3A2) (red), and surfactant protein C (SPC) (white). Scale bars=200 μ m, 50 μ m and 10 μ m. **b)** Detection of rbATII cells in respiratory bronchiole of emphysematous region (black square in left picture with H&E staining) by co-IF for intercellular adhesion molecule 1 (ICAM1) (green, top), TM4SF1 (green, bottom), SCGB3A2 (red), and SPC (white). Scale bars=50 μ m, 20 μ m and 10 μ m. **c)** Flow cytometry analysis of asATII and rbATII using CD74 (centre), TM4SF1 and ICAM1 (right) in healthy control sample (n=6, top) and emphysematous parenchymal tissue from COPD patients (n=6, bottom). **d)** Quantification of the percentages of asATII (CD74⁺) and rbATII (ICAM1⁺/TM4SF1⁺) cells in total ATII cells (HTII280⁺/epithelial cell adhesion molecule (EPCAM)⁺/4',6-diamidino-2-phenylindole (DAPI)⁻) in **a)**. Mann–Whitney test. **e)** Organoids formed by fluorescence-activated cell sorted asATII (HTII280⁺/CD74⁺) and rbATII (HTII280⁺/ICAM1⁺/TM4SF1) cells from human emphysematous parenchymal tissue (n=5) and by ATII cells (HTII280⁺) from healthy control (n=3). Scale bar=100 μ m. **f)** Quantification of organoid forming efficiency of **c)**. Values are presented as mean \pm SEM. Mann–Whitney test. *: p<0.05.

Transcriptionally, asATII cells show similar expression patterns to ATII_B cells identified by SAULER *et al.* [14] (supplementary figure S6h), including increased HHIP, NPC2, SERPINA1, SFTPA1 and SFTPD and reduced NUPR1. This similarity confirmed COPD specificity of asATII cells. However, their dataset contained non-emphysematous samples, mixed tissue regions and large populations of nonepithelial cells. Moreover, the localisation and regenerative function of ATII_B cells have not been reported. Here, we assessed the progenitor functions of the asATII cells through organoid assays, which revealed a significantly diminished organoid forming capacity of asATII cells. The small number of organoids derived by asATII has limited our characterisation of proliferation and differentiation potential of the asATII cells, specially the asATII-to-ATI differentiation, which needs to be further investigated in other organoid or tissue culture systems.

Additionally, we have identified a small fraction of rbATII cells expressing multiple SCGB genes and *SFTPC* in alveoli attached to the respiratory bronchioles in both COPD and healthy human lungs. Our RNA velocity analysis showed a trajectory from club cells to asATII cells, with the rbATII cells serving as a potential transitional population, suggesting a similar localisation and progenitor function to the recently reported AT0 cells in healthy lungs [15]. In COPD, the rbATII cells were detected in damaged respiratory bronchioles lining in perivascular regions (figure 2b), similar to the RASC-derived LAMP3⁺/3A2⁺ cells reported by BASIL *et al.* [16]. However, rbATII cells do not express the full scope of RASC derived ATII and AT0 marker genes [15, 16] (supplementary figure S6h). These distinct transcription patterns may be due to differences in tissue regions collected and disease subtype and stages in those studies and ours.

Furthermore, an intermediate population during club-to-ATII differentiation is missing in our human lung scRNAseq study, probably due to the late stage of emphysema in our samples. Therefore, a disease model is needed to trace the fate of club cells during emphysema-like injury. Here, we first employed a PPE-induced mouse emphysema model [47]. PPE induces severe and progressive destruction of alveoli in mouse lungs in 3 weeks. Using the PPE model, we identified a club cell derived ATII subpopulation in the mouse lung (ATII-6) that is similar to human asATII cells in terms of transcriptional profile, localisation in emphysematous alveolar sacs, and impaired regeneration capacity in organoid assays. Importantly, we also observed the ATII-6 subpopulation in a cigarette smoke induced COPD model [50, 51]. Collectively, these findings establish and validate the significant roles of ATII-6 cells in the pathogenesis of COPD-like lung injury in mice.

It is possible that the terminal bronchiolar club cells directly migrate to alveolar sacs and differentiate into ATII-6 cells in response to emphysema-like injury during early stages of disease. Alternatively, it is possible that the club cells may attempt to repair alveolar tissue by differentiating to normal ATII progenitors, which then shift to ATII-6 cells during late stage of emphysema. Further time-course studies are needed to test these hypotheses and to investigate the mechanisms controlling how club cells sense and respond to injuries.

Although we showed a trajectory from club-like cells in the respiratory bronchiole (RASC) to asATII cells in scRNAseq analysis, our mouse lineage-tracing model is limited in its ability to validate this due to the lack of respiratory bronchioles in mouse lungs. Larger animal (*e.g.* ferret) models of emphysema, along with disease modelling using patient-derived precision cut lung slices, could help evaluate the formation of similar abnormal ATII populations by RASCs, identify molecular mechanisms controlling differentiation, and investigate functional shifts of progenitor cells across stages of emphysema pathogenesis. Another limitation of our study is the restricted analysis of end-stage tissue. The localisation, number and secreted factors of asATII cells at early stages of COPD, including recently proposed pre-COPD stage [58], will be

helpful to facilitate early diagnosis of emphysema. Ultimately, determining how reprogramming of asATII cells and club cells promotes failed tissue repair in emphysema will shed light on novel therapeutic treatment strategies.

In summary, our study identified distinct ATII subpopulations in the alveolar sacs of emphysematous lung parenchyma, and we revealed a previously unknown contribution of club cells to this ATII subpopulation with impaired regeneration capacity during pathogenesis of emphysema. These studies pave the way for the development of early diagnostic tools and novel therapeutics for COPD.

Acknowledgements: We thank Anastasia van den Berg, Naoko Liu, James Needell, Ana Harder (Division of Pulmonary Sciences and Critical Care Medicine, University of Colorado Anschutz Medical School, Aurora, CO, USA) and Byron Chuan (Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA) for their technical support.

Data availability: The human lung single cell RNAseq data has been deposited in Gene Expression Omnibus with accession number GSE222374. Data processing and analysis code is provided at a GitHub repository (https://github.com/KoenigshoffLab/HU_COPD_2023). Murine single cell RNAseq data has been deposited at the Gene Expression Omnibus (GSE218813). Murine scRNAseq analysis code is provided at a Github repository (<https://github.com/rnabioco/Koenigshoff-murine-copd>).

Conflict of interest: C.M. Evans reports grants from Cystic Fibrosis Foundation, Department of Defense, and Enterprise Therapeutics, outside the submitted work, and royalties from Eleven P15 consulting. C.R. Kliment serves on an advisory board of Verona Pharmaceuticals. The remaining authors have no conflicts of interest.

Support statement: Funding sources include: NIH NHLBI K99 HL161446 and NIH NHLBI F32 HL149290 to Y. Hu; NIH NHLBI R01 HL141380, NIH NIAID U54AG075931 and LF20146.1.14.009 to M. Königshoff; NIH NHLBI R01 HL080396, R01 HL130938, P01 HL162607 and VA Merit 101BX005343 to C.M. Evans; LongFonds Accelerate BREATH Program to M. Königshoff; CU Anschutz Pulmonary Division and Cancer Center Genomics Shared Resource Junior Investigator Pilot Award to Y. Hu; CU Anschutz Cancer Center Support Grant (P30CA046934); CU RNA Bioscience Initiative Single cell RNA sequencing Award to M. Königshoff and Y. Hu; German Center for Lung Research (DZL), the Helmholtz Association, the European Union's Horizon 2020 research and innovation programme, grant agreement 874656—project discovair to H.B. Schiller; the Chan Zuckerberg Initiative (CZF2019-002438, project Lung Atlas 1.0) to H.B. Schiller; Burroughs Wellcome Fund – Career Award for Medical Scientists (CAMS) to C.R. Kliment; German Center for Lung Research (DZL), Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 512453064 to M. Lehmann; Federal Institute for Risk assessment (Bundesinstitut für Risikoforschung, BfR) 60-0102-01.P588 to M. Lehmann; NIH HL152194, HL132999 and HL164929 to E. Morrisey; NIH K08 HL163398 and Burroughs Wellcome Fund Career Award for Medical Scientists to M.C. Basil. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Battaglia S, Basile M, Scichilone N, *et al.* Prevalence of co-morbidities and severity of COPD. *COPD* 2015; 12: 390–394.
- 2 Rawlins EL, Okubo T, Xue Y, *et al.* The role of Scgb1a1⁺ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. *Cell Stem Cell* 2009; 4: 525–534.
- 3 Kim CF, Jackson EL, Woolfenden AE, *et al.* Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* 2005; 121: 823–835.
- 4 Guha A, Deshpande A, Jain A, *et al.* Uroplakin 3a⁺ cells are a distinctive population of epithelial progenitors that contribute to airway maintenance and post-injury repair. *Cell Rep* 2017; 19: 246–254.
- 5 Liu Q, Liu K, Cui G, *et al.* Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. *Nat Genet* 2019; 51: 728–738.
- 6 Salwig I, Spitznagel B, Vazquez-Armendariz AI, *et al.* Bronchioalveolar stem cells are a main source for regeneration of distal lung epithelia *in vivo*. *EMBO J* 2019; 38: e102099.
- 7 Liu K, Meng X, Liu Z, *et al.* Tracing the origin of alveolar stem cells in lung repair and regeneration. *Cell* 2024; 187: 2428–2445.
- 8 Strunz M, Simon LM, Ansari M, *et al.* Alveolar regeneration through a Krt8⁺ transitional stem cell state that persists in human lung fibrosis. *Nat Commun* 2020; 11: 3559.
- 9 Kathiriya JJ, Brumwell AN, Jackson JR, *et al.* Distinct airway epithelial stem cells hide among club cells but mobilize to promote alveolar regeneration. *Cell Stem Cell* 2020; 26: 346–358.
- 10 Barkauskas CE, Cronce MJ, Rackley CR, *et al.* Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* 2013; 123: 3025–3036.

- 11 Zacharias WJ, Frank DB, Zepp JA, *et al.* Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* 2018; 555: 251–255.
- 12 Nabhan AN, Brownfield DG, Harbury PB, *et al.* Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* 2018; 359: 1118–1123.
- 13 Watanabe N, Fujita Y, Nakayama J, *et al.* Anomalous epithelial variations and ectopic inflammatory response in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2022; 67: 708–719.
- 14 Sauler M, McDonough JE, Adams TS, *et al.* Characterization of the COPD alveolar niche using single-cell RNA sequencing. *Nat Commun* 2022; 13: 494.
- 15 Kadur Lakshminarasimha Murthy P, Sontake V, Tata A, *et al.* Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. *Nature* 2022; 604: 111–119.
- 16 Basil MC, Cardenas-Diaz FL, Kathiriya JJ, *et al.* Human distal airways contain a multipotent secretory cell that can regenerate alveoli. *Nature* 2022; 604: 120–126.
- 17 Rustam S, Hu Y, Mahjour SB, *et al.* A unique cellular organization of human distal airways and its disarray in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2023; 207: 1171–1182.
- 18 Suarez CJ, Dintzis SM, Frevert CW. Respiratory. In: Treuting PM, Dintzis SM, eds. *Comparative Anatomy and Histology: a Mouse and Human Atlas*. London, Academic Press, 2012; pp. 121–134.
- 19 Agustí A, Celli BR, Criner GJ, *et al.* Global Initiative for Chronic Obstructive Lung Disease 2023 report: GOLD executive summary. *Am J Respir Crit Care Med* 2023; 207: 819–837.
- 20 Muzumdar MD, Tasic B, Miyamichi K, *et al.* A global double-fluorescent Cre reporter mouse. *Genesis* 2007; 45: 593–605.
- 21 Hu Y, Ng-Blichfeldt JP, Ota C, *et al.* Wnt/ β -catenin signaling is critical for regenerative potential of distal lung epithelial progenitor cells in homeostasis and emphysema. *Stem Cells* 2020; 38: 1467–1478.
- 22 Kliment CR, Nguyen JMK, Kaltreider MJ, *et al.* Adenine nucleotide translocase regulates airway epithelial metabolism, surface hydration and ciliary function. *J Cell Sci* 2021; 134: jcs257162.
- 23 Hautamaki RD, Kobayashi DK, Senior RM, *et al.* Requirement for macrophage elastase for cigarette smoke-induced emphysema in mice. *Science* 1997; 277: 2002–2004.
- 24 Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 2019; 8: 281–291.
- 25 Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* 2020; 9: gaa151.
- 26 Lun AT, Bach K, Marioni JC. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol* 2016; 17: 75.
- 27 Polański K, Young MD, Miao Z, *et al.* BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* 2020; 36: 964–965.
- 28 Hao Y, Stuart T, Kowalski MH, *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* 2024; 42: 293–304.
- 29 Lotfollahi M, Naghipourfar M, Luecken MD, *et al.* Mapping single-cell data to reference atlases by transfer learning. *Nat Biotechnol* 2022; 40: 121–130.
- 30 Sikkema L, Ramírez-Suástegui C, Strobl DC, *et al.* An integrated cell atlas of the lung in health and disease. *Nat Med* 2023; 29: 1563–1577.
- 31 Korsunsky I, Millard N, Fan J, *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019; 16: 1289–1296.
- 32 Fu R, Gillen AE, Sheridan RM, *et al.* clustifyr: an R package for automated single-cell RNA sequencing cluster classification. *F1000Res* 2020; 9: 223.
- 33 Bankhead P, Loughrey MB, Fernández JA, *et al.* QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; 7: 16878.
- 34 Ng-Blichfeldt JP, de Jong T, Kortekaas RK, *et al.* TGF- β activation impairs fibroblast ability to support adult lung epithelial progenitor cell organoid formation. *Am J Physiol Lung Cell Mol Physiol* 2019; 317: L14–L28.
- 35 Conlon TM, John-Schuster G, Heide D, *et al.* Inhibition of LT β R signalling activates WNT-induced regeneration in lung. *Nature* 2020; 588: 151–156.
- 36 Heinzlmann K, Hu Q, Hu Y, *et al.* Single-cell RNA sequencing identifies G-protein coupled receptor 87 as a basal cell marker expressed in distal honeycomb cysts in idiopathic pulmonary fibrosis. *Eur Respir J* 2022; 59: 2102373.
- 37 Habermann AC, Gutierrez AJ, Bui LT, *et al.* Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv* 2020; 6: eaba1972.
- 38 Adams TS, Schupp JC, Poli S, *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* 2020; 6: eaba1983.
- 39 Xu Y, Mizuno T, Sridharan A, *et al.* Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight* 2016; 1: e90558.
- 40 Deprez M, Zaragosi LE, Truchi M, *et al.* A single-cell atlas of the human healthy airways. *Am J Respir Crit Care Med* 2020; 202: 1636–1645.

- 41 Zhou Y, Zhou B, Pache L, *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; 10: 1523.
- 42 Leng L, Metz CN, Fang Y, *et al.* MIF signal transduction initiated by binding to CD74. *J Exp Med* 2003; 197: 1467–1476.
- 43 Shi X, Leng L, Wang T, *et al.* CD44 is the signaling component of the macrophage migration inhibitory factor-CD74 receptor complex. *Immunity* 2006; 25, 595–606.
- 44 Donnelly SC, Haslett C, Reid PT, *et al.* Regulatory role for macrophage migration inhibitory factor in acute respiratory distress syndrome. *Nat Med* 1997; 3: 320–323.
- 45 Calandra T, Roger T. Macrophage migration inhibitory factor: a regulator of innate immunity. *Nat Rev Immunol* 2003; 3: 791–800.
- 46 Sauler M, Bucala R, Lee PJ. Role of macrophage migration inhibitory factor in age-related lung disease. *Am J Physiol Lung Cell Mol Physiol* 2015; 309: L1–L10.
- 47 Kaplan PD, Kuhn C, Pierce JA. The induction of emphysema with elastase. I. The evolution of the lesion and the influence of serum. *J Lab Clin Med* 1973; 82: 349–356.
- 48 Mereu E, Lafzi A, Moutinho C, *et al.* Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat Biotechnol* 2020; 38: 747–755.
- 49 Pflugh DL, Maher SE, Bothwell AL. Ly-6I, a new member of the murine Ly-6 superfamily with a distinct pattern of expression. *J Immunol* 2000; 165: 313–321.
- 50 Polverino F, Doyle-Eisele M, McDonald J, *et al.* A novel nonhuman primate model of cigarette smoke-induced airway disease. *Am J Pathol* 2015; 185: 741–755.
- 51 Churg A, Cosio M, Wright JL. Mechanisms of cigarette smoke-induced COPD: insights from animal models. *Am J Physiol Lung Cell Mol Physiol* 2008; 294: L612–L631.
- 52 Kneidinger N, Yildirim A, Callegari J, *et al.* Activation of the WNT/ β -catenin pathway attenuates experimental emphysema. *Am J Respir Crit Care Med* 2011; 183: 723–733.
- 53 Lam AP, Gottardi CJ, Tudor R. Regenerative pathways and emphysema: a new paradigm? *Am J Respir Crit Care Med* 2011; 183: 688–690.
- 54 Barnes PJ. Oxidative stress in chronic obstructive pulmonary disease. *Antioxidants* 2022; 11: 965.
- 55 Stabler CT, Morrisey EE. Developmental pathways in lung regeneration. *Cell Tissue Res* 2017; 367: 677–685.
- 56 Baarsma HA, Königshoff M. ‘WNT-er is coming’: WNT signalling in chronic lung diseases. *Thorax* 2017; 72: 746–759.
- 57 Meiners S, Eickelberg O, Königshoff M. Hallmarks of the ageing lung. *Eur Respir J* 2015; 45: 807–827.
- 58 Han MK, Agusti A, Celli BR, *et al.* From GOLD 0 to pre-COPD. *Am J Respir Crit Care Med* 2021; 203: 414–423.