Title page

Cultured alveolar basal cells display similar characteristics and functional properties to airway basal cells

Sabrina Blumer[1*], Petra Khan PhD[1*], Julien Roux PhD[1,2], Nataliia Artysh[3,4], Linda Plappert[3], Antje Prasse[1,3,4], Katrin E. Hostettler MD PhD[1#].

[1] Department of Biomedicine and Clinics of Respiratory Medicine, University Hospital Basel, University of Basel, Basel 4031, Switzerland;
[2] Swiss Institute of Bioinformatics, Basel, Switzerland;
[3] Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover 30625, Germany;
[4] Hannover Medical School, Department of Pulmonology & Infectious diseases, Hannover, Germany;

*Both authors equally contributed to this manuscript

[#]**Corresponding author:** Katrin E. Hostettler, Prof., Department of Biomedicine, University Hospital Basel, Hebelstrasse 20, 4031 Basel, Switzerland; E-mail: Katrin.Hostettler@usb.ch

ABSTRACT

Rationale: Basal cells (BC) appear ectopically within the lung parenchyma of interstitial lung disease (ILD) patients, potentially through migration of airway BC or though trans-differentiation of resident alveolar epithelial type 2 (AT2) cells. The exact origin, role, and function of these ectopic alveolar BC remains poorly understood. By comparing ectopic alveolar to "classical" airway BC, we aimed to get a better understanding of the origin and characteristics of alveolar BC in ILD. **Methods:** Alveolar and airway BC were isolated from transbronchial- and airway mucosal biopsies (MB) of the same ILD patients. BC were cultured on plastic under submerged conditions, on an air liquid interface (ALI), or in a 3D organoid model. Additionally, BC were intratracheally injected into bleomycin-challenged mice. Samples were analyzed by single cell RNA sequencing (scRNA-seq), TaqMan RT-PCR, immunocytochemistry/immunofluorescence (ICC/IF), or immunohistochemistry (IHC)/IF. **Results:** scRNA-seq analysis revealed several differentially expressed genes in alveolar compared to airway BC, which, however, did not include specific AT2 cell marker genes. Expression of BC markers and the capacity for wound repair, proliferation, organoid formation and differentiation were not significantly different between cultured BC originating from the different locations. Furthermore, there was no difference in engraftment and differentiation potential after instillation into bleomycin-injured mice between the two human BC types. **Conclusion:** Close similarities between alveolar and airway BC suggest that the specific pro-fibrotic microenvironment, rather than a cell intrinsic functional failure, determines alveolar BC characteristics and functions in ILD.

Keywords: Basal cells, interstitial lung diseases, IPF, scRNA-seq, organoids, air liquid interface

Introduction

In interstitial lung diseases (ILD) such as idiopathic pulmonary fibrosis (IPF) ectopic BC accumulate within the alveolar space, where they replace resident alveolar epithelial cells and contribute to the formation of pathological honeycomb cysts (HC). During normal lung homeostasis, the alveoli are lined with alveolar epithelial type

(AT)1 and 2 cells, whereas basal cells (BC) are restricted to proximal and distal airways. AT1 cells cover approximately 90% of the alveolar surface and facilitate efficient gas exchange. AT2 cells are critical for alveolar homeostasis through their production of pulmonary surfactant and their ability to self-renew and differentiate into AT1 cells [1]. BC maintain airway homeostasis by their ability to self-renew and to differentiate to other airway epithelial cells such as acetylated tubulin (AcTub)+, and FOXJ1+ ciliated, or SCGB1A1+, MUC5AC+, or MUC5B+ secretory epithelial cells [2].

Accumulation of ectopic BC in the lung parenchyma after severe injury was observed in mouse and human. In mouse, where BC are normally restricted to the trachea, ectopic KRT5+ BC originate from a heterogeneous pool of pre-existing injury-activated airway p63+/KRT5- progenitor cells [3, 4]. In contrast to mouse, BC are present throughout the airways all the way to the terminal bronchioles in human [5]. It is therefore likely that ectopic alveolar BC derive from airway BC that migrated to the alveoli post-injury. Alternatively, trans-differentiation of human AT2 cells to BC was hypothesised to contribute to the pool of ectopic alveolar BC in the fibrotic lung, which is supported by a study showing AT2 to BC trans-differentiation *in vitro* [6].
Earlier studies suggested a potential beneficial role of ectopic BC, as basal-like cells were shown to regenerate the injured alveolar epithelium by their ability to transdifferentiate to AT2 cells in animal models [7, 8]. However, other studies showed no contribution of ectopic BC to normal AT2 cells in mice [9, 10], and associated ectopic BC with increased mortality and pathological HC formation in IPF [11, 12], suggesting a potential detrimental role.

By comparing alveolar BC to their “normal” airway counterpart after expanding these two respective BC types *in vitro*, we aimed to get a better understanding of the origin and disease-specific characteristics of these cells, which may help to define specific pharmacological interventions to prevent accumulation of ectopic BC in the lung parenchyma of ILD patients.

Materials and Methods

Basal cell culture

Alveolar BC were cultured from transbronchial biopsies (TBB) and airway BC from airway mucosal biopsies (MB) as previously described [13-15]. Briefly, TBB or airway MB derived from the same ILD patient were placed into cell culture-treated plastic dishes containing an epithelial cell-specific growth medium (Cnt-PR-A). After 5-7 days, tissue was removed and the outgrown cells were trypsinized and expanded. All experiments in this study were performed with BC between passage 1-4 and were derived from tissue of several different ILD patients as indicated by the n-number. The local ethical committee of the University Hospital, Basel, Switzerland (EKBB05/06) approved the use of human tissue sections and the culture of human primary lung cells. ILD patient characteristics are outlined in supplement table 1. Cell culture material used in this study can be found in supplement table 2.

TaqMan RT-PCR, immunocytochemistry/immunofluorescence and immunohistochemistry/immunofluorescence

TaqMan RT-PCR, immunocytochemistry (ICC)/immunofluorescence (IF), immunohistochemistry (IHC)/IF or hematoxylin and eosin (H&E) stainings were performed as previously described [15-18]. Results of TaqMan PCR are expressed as -ΔCt values (-(target-GAPDH)). When the target genes were expressed at levels too low to be detected after 40 cycles, the undetermined -ΔCt values were arbitrarily set to -20. ICC/IF, IHC/IF or H&E images were acquired on a Nikon Ti2-E widefield microscope and stored on the client-server platform Omero (Omero.insight 5.5.17 and Omero.web 5.9.1). All primers and antibodies used in this study are listed in table S2.

Single-cell RNA-sequencing

Airway and alveolar BC were grown from airway MB and from TBB derived from three different ILD patients (Patient-ID 01, 02, and 03 in supplement table 1). scRNA-seq was performed on cultured cells at passage 1 as previously described, using the 10X Genomics technology [17]. A total of six single cell captures were loaded each onto a channel of a Chromium controller and cDNA and library preparation were performed with a Single Cell 3' v3 Reagent Kit (10X Genomics) according to the manufacturer's instructions. Sequencing was performed on one lane of a flow-cell of the Illumina Novaseq 6000 platform (with 101nt-long R2 reads) at the Genomics Facility Basel of the ETH Zurich, Basel.

The dataset was analyzed by the Bioinformatics Core Facility, Department of Biomedicine, University of Basel. STARsolo (version 2.7.9a) [19] was used to perform sample and cell demultiplexing, and alignment of reads to the human genome (hg38) and UMI counting on gene models from Ensembl 102 (with options "--outFilterType=BySJout --outFilterMultimapNmax=10 --outSAMmultNmax=1 --outFilterScoreMin=30 --soloCBmatchWLtype=1MM_multi_Nbase_pseudocounts --soloUMIfiltering=MultiGeneUMI_CR --soloUMIdedup=1MM_CR --soloType=CB_UMI_Simple –soloStrand=Reverse).

Processing of the UMI counts matrix was performed using the Bioconductor packages DropletUtils (version 1.16.0; function *emptyDrops* using 5000 iterations, the option test.ambient=TRUE, a lower threshold of 100 UMIs and an FDR threshold of 0.1%) [20, 21], scran (version 1.24.1) [22, 23] and scater (version 1.24.0) [24], following mostly the steps illustrated in the OSCA book (http://bioconductor.org/books/3.15/OSCA/) [23, 25]. Cells with less than 5% or more than 20% of UMI counts attributed to mitochondrial genes [26], with less than 10% or more than 30% of UMI counts attributed to ribosomal protein genes, with less than 4000 UMI counts, or with less than 2500 detected genes were excluded. The presence of doublet cells was investigated with the scDblFinder package (version 1.10.0) [27], and suspicious cells were filtered out (doublet score>0.9). Systematic differences between samples were removed using the fastMNN function (d = 50, k = 50) of the batchelor package (version 1.12.30) [28]. tSNE embeddings used for visualization of cells were calculated using the 500 most variable genes and the PCA coordinates or the batch-corrected matrix of low-dimensional coordinates for cells (using a perplexity of 30). Shared nearest-neighbor graph clustering was performed using the Louvain algorithm for finding community structure (k=20). The package SingleR (version 1.10.0) was used for cell-type annotation of the cells [29] using different public lung scRNA-seq datasets as references [6, 30-34] and with the pseudo-bulk aggregation option enabled (aggr.ref=TRUE").

Differential expression between alveolar and airway BC, controlling for donor-specific differences, was performed using a pseudo-bulk approach, summing the UMI counts of cells from each cluster (excluding cluster 5) in each sample. The aggregated samples were then treated as bulk RNA-seq samples [35]. The package edgeR

(version 3.38.4) [36] was used to perform TMM normalization [37] and to test for differential expression with the quasi-likelihood framework (*glmQLFit*). Genes with a false discovery rate (FDR) lower than 1% were considered differentially expressed. Gene set enrichment analysis was performed with the function camera [38] on gene sets from the Molecular Signature Database (MSigDB; version 7.5.1), notably the Hallmark collection and Reactome subset from the C2 collection [39, 40]. We retained only gene sets containing more than 5 genes, and gene sets with a FDR lower than 10% were considered as significant.

Air liquid interface culture

Alveolar and airway BC were cultured on an air liquid interface (ALI) platform as previously described [15]. After 23 days, the cells on the cell culture insert membranes were fixed in 4 % paraformaldehyde (PFA) and directly used for ICC/IF or they were embedded in Tissue-Tek® O.C.T. compound and cut into 4-7 μm frozen sections and then used for ICC/IF, or they were lysed for RNA expression analysis [15].

3D organoid culture

Cultured alveolar or airway BC were trypsinized and resuspended in 40 µl Matrigel (500 cells/µl) as previously described [15]. Differentiation media were slightly modified from a previously described protocol [41] and were composed as follows: Basic differentiation media (bDM): IMDM-Ham`s F12 (3:1 vol/vol) with B 27 (with retinoic acid) supplement (1 %), N 2 supplement (0.5 %), BSA (0.05 %), antibiotic-antimycotic 1x, glutamax 1x, ascorbic acid (50 μg/ml) and monothioglycerol ($4.5x10^{-4}$ M). Complete differentiation medium (cDM): bDM supplemented with FGF-2 (250 ng/ml), FGF-10 (100 ng/ml), EGF (5 ng/ml), dexamethasone (50 nM), 8-Bromo-cAMP (0.1 mM) IBMX (0.1 mM), and A83-01 (TGF-ß inhibitor, 1 μM). A 1:1 mixture of epithelial cell growth medium (Cnt-PR-A) and cDM with ROCK inhibitor (Y-27632 dihydrochloride, 10 μM) was added to the wells for the first 7 days. The medium was then changed to cDM for another 14 days. cDM was exchanged after 7 days. After 20 days, organoids were imaged at 10x with a z-stack (every 50 µm). The diameter and number of organoids per mm^3 (>50 μm) was determined by using Fiji software version 2.9.0 [42]. On day 21, the Matrigel was digested with Dispase II (2mg/ml) and organoids were fixed in 4 % PFA, embedded in pre-heated histogel, polymerized and stored in 50 % ethanol. Organoids were then dehydrated using an ethanol series

(50%, 70%, 95%, 100%) followed by Xylene incubation and then embedded in paraffin. Paraffin blocks were cut on a microtome (4 µm) and analysed by ICC/IF.

Proliferation assay

Alveolar or airway BC were seeded into a 24-well plate. At about 70% confluence, cells representing timepoint 0 were stained with live cell stain NucBlue, imaged at 4x using a Nikon Ti2-E widefield microscope and counted using a general analysis NIS software version 4.2.1. After an additional 48 hours, the procedure was repeated on a different well with cells representing time point 48 hours.

Epithelial wound repair assay

Confluent alveolar or airway BC were mechanically wounded with a pipette tip and wound closure was monitored with a Nikon Ti2 2.3.PO widefield time lapse microscope and calculated with a Fiji plugin wound healing size tool [15, 43].

Intratracheal administration of human BC into bleomycin-challenged mice

The Hannover Medical School, Germany performed all mouse procedures in accordance with the German law for animal protection and the European Directive 2010/63/EU and an approval by the Lower Saxony State Office for Consumer Protection and Food Safety in Oldenburg/Germany (LAVES); AZ: 33.12-42502-04-15/1896 and AZ: 33.19-42502-04-15/2017), AZ:33.12-42502-04-17/2612). Human alveolar and airway BC were cultured as described earlier and transduced with a lentiviral vector (kindly provided by Dr. Axel Schambach, Hannover Medical School) for firefly-luciferase and enhanced green fluorescence protein (eGFP) [12]. A total of nine NRG mice received bleomycin at a dose of 1.2 mg/kg intratracheally at day 0. Three days later, three mice were intratracheally injected with vector-transduced human alveolar BC (0.3 X 10^5 cells per mouse), and three mice with vector-transduced human airway BC (0.3 X 10^5 cells per mouse). To image firefly-luciferase activity, mice were injected subcutaneously with XenoLight D-Luciferin-K+ Salt Bioluminescent Substrate (150 mg/kg) on day 4, 11 and 21, and bioluminescence was measured by an IVIS Lumina II and data analyzed by using LivingImage 4.5 [12]. At day 21, mice were sacrificed and mouse lungs were formalin-fixed and paraffin embedded for H&E or IHC/IF analysis. Fibrosis scoring was performed on lung tissue of mice treated with bleomycin only (n=3), bleomycin + human alveolar (n=3), or + human airway (n=3) BC

as previously described by Ashcroft [44]. Briefly, 20 non-overlapping 10x fields per sample were scored as followed: 0 = no fibrosis, 1 = mild thickening of alveolar septa, 2 = <10%-, 3 = 10–20%, 4 = 20–40%-, 5 = 40–60%-, 6 = 60–80%, 7 = >80% fibrotic area or 8 = complete fibrosis. The scores of all fields were then averaged for each sample.

Statistics

Paired or unpaired t-tests were performed using GraphPad Prism software version 9.1.1. p values < 0.05 were considered as statistically significant. Data were presented as mean ± SEM.

Results

Characteristics of cultured alveolar and airway basal cells

In airway MB, KRT5+/p63+ BC are lining the basement membrane of the airway epithelium, whereas in TBB, KRT5+/p63+ BC were found scattered within the lung parenchyma (Figure 1 A). Alveolar BC were cultured from TBB and airway BC from airway MB in cell culture dishes containing an epithelial cell growth medium (Cnt-PR-A) (Figure 1 B). Alveolar and airway BC were indistinguishable by their morphology (Figure 1 C) and expressed similar levels of the basal cell markers KRT5, KRT14, KRT17 or TP63 on the RNA (Figure 1 D) and protein level (Figure 1 E). Expression of mesenchymal cell marker CDH2 and that of secretory epithelial cell marker SCGB1A1 were measured to exclude the possibility of spontaneous epithelial to mesenchymal transition (EMT) or differentiation to secretory epithelial cells of BC under the here-described culture conditions. Both markers were absent or expressed at very low levels on the RNA level (Figure 1 D) and not detected on the protein level in cultured alveolar or airway BC (Figure 1 E).

Cultured alveolar and airway basal cells display transcriptomic differences

Cultured alveolar or airway BC derived from TBB or airway MB of the same ILD patients (n=3; Donors 1-3) were subjected to scRNA-seq using the 10X Genomics technology. After quality control and filtering of cells, 34,435 cells were retained (4,500 to 6,975 cells per sample). The major source of variation in the gene expression data

from single cells was the individual origin of the sample (Donors 1-3), followed by the biopsy sampling location (TBB or airway MB) (Figure 2 A; Supplement figure 1). After removal of these sources of variation through correction of sample-specific effects by matching mutual nearest neighbors, the cells were grouped into 7 different clusters encompassing between 7.6% and 20% of the cells (Figure 2 B). We annotated these clusters by comparing our cells to previously published scRNA-seq datasets (Figure 2 C [45]; Supplement Figure 2 [6, 30-33, 46-49]): the majority of cells in clusters 2-6 (84 to 97%) showed closest transcriptomic similarities to the BC subset 1 (canonical BC) described in [45]. Most cells (86%) from cluster 7 closely matched to BC subset 2 (cycling BC) [45]. Last, a large fraction (39%) of cells from cluster 1 matched to BC subset 3 (secretory-primed) [45]. This annotation of the clusters was supported by the expression of the best markers across clusters (Figure 2 D, Supplement Figure 3). More specifically, cluster 7 cells displayed high expression of proliferation genes such as *PTTG1*, *BIRC5*, *HMGB2*, *MIK67*, *CENPF* and *TOP2A* (Figure 2 D). Secretory-primed BC within cluster 1 were enriched with transcripts of members of the serpin family, *KRT16*, or *CSTA* (Figure 2 D). The markers of the other clusters seemed to illustrate a differentiation gradient from cluster 6 to the secretory-primed cluster 1, analogous to what we previously reported [15]. In line with that, BC within cluster 6 were enriched with transcripts highly expressed in undifferentiated BC [50] such as *AREG*, *CAV1*, *THBS1*, or *ITGB4*, whereas transcripts indicating BC to secretory epithelial cell differentiation (*CSTA*, *KRT16*, *SERPINB2*) [51] were expressed at low levels in cluster 6 BC, higher levels in cluster 3 and 4 BC, and highest in cluster 1 BC (Figure 2 D). Interestingly, despite of expressing lower levels of *AREG*, *CAV1*, *THBS1* and *ITGB4* than cluster 6 BC, cluster 2 BC showed much lower levels of *CSTA*, *KRT16*, *SERPINB2* transcripts than cluster 3 and 4 BC. Instead, cluster 2 BC were enriched with transcripts including several metallothioneins (eg. *MT1X*, *MT2A*, *MT1E*), *FOS*, *KRT15*, or *AGR2* (Figure 2 D). Cluster 5 showed notably higher expression of *FN1* and *VIM* (Figure 2 D), but it was majoritarily composed of cells from donor 1 (74% of cells, Supplement figure 1A). Of note, there was overall no detected or a very low expression of specific AT2 cell markers (e.g., *SFTPA-D*, *ABCA3*) across clusters in our dataset (not shown).

Comparing the relative abundance of alveolar compared to airway BC across clusters (excluding cluster 5 including cells mostly specific to donor 1), we observed a

significantly higher abundance of alveolar BC in cluster 2 (FDR=8.6%), along with a reduced abundance of alveolar BC in cluster 4 (FDR=8.6%) and cluster 1 (FDR=21%). Clusters 3, 6, and 7 did not display significant trends (Figure 2 E).

We next aimed at comparing gene expression between alveolar and airway BC. Testing for differences in expression state of cells independently within each cluster allowed us to account for the above-described differences in abundance, while the paired design allowed us to account for donor-specific effects. At an FDR of 1%, there were in total 40 genes significantly up or down-regulated in alveolar compared to airway BC in at least one of the clusters, shown in Figure 3 A. Cluster 2 and 6 (Figure 3 C) yielded the highest amount of differentially expressed genes (Figure 3 B and C; differential expression results in the remaining clusters are shown in Supplement Figure 4). Genes up-regulated in response to hypoxia (*PLIN2*, *S100A4*), or KRAS activation/mutation (*LMO3*, *IL33*), as well as genes involved in glycolysis (*EGLN3*), or ECM organization (*SPOCK3*, *MXRA7*) were amongst the up-regulated genes, whereas genes that are up-regulated in response to interferon gamma (*IFI27*, *HLA-DRB1*, *BST2*) were down-regulated in alveolar BC. Interestingly, the basal cell marker genes *KRT5*, *KRT17*, *KRT14* and *TP63* were consistently down-regulated (Supplement Figure 3), and genes for *FN1*, *Col1A1* or *VIM* were up-regulated in alveolar BC, although not reaching significance.

Functional enrichment analysis using the MSigDB Hallmark pathways (clusters 2 and 6 in Figure 3 D, E; clusters 1, 3, 4 and 7 in Supplement Figure 5) showed that pathways related to hypoxia, glycolysis, angiogenesis, KRAS signaling, and epithelial to mesenchymal transition (EMT) were up-regulated in alveolar BC. Furthermore, Reactome pathway analysis revealed an up-regulation of ECM-related pathways such as “ECM organization”, “collagen formation”, “collagen biosynthesis and modifying enzymes”, or “activation of matrix metalloproteinases” in alveolar BC (Supplement Figure 6). Interestingly, in line with the earlier described trend of a down-regulation of *KRT5*, *KRT14*, and *KRT17* transcripts in alveolar BC, the pathway “Keratinization” was significantly down-regulated in cluster 1, 6, and 7 alveolar BC (Supplement Figure 6).

Differential expression results were validated in alveolar and airway BC from donor 1-3 and three additional donors by TaqMan RT-PCR, confirming the up-regulation of *LMO3*, *S100A4*, *SPOCK3*, *HNMT* and down-regulation of *IFI27*, *BST2* in alveolar versus airway basal cells (Figure 3 F).

Cultured alveolar and airway BC display similar proliferation and wound healing capacities

Cultured alveolar and airway BC showed robust proliferation, which was reflected by a significant increase in cell numbers between 0-48 h (Figure 4 A). There was no significant difference in the proliferation rate comparing alveolar to airway basal cells (Figure 4 A). Furthermore, alveolar and airway BC closed a manually set scratch after 6 hours with no significant difference in wound healing capacity between alveolar and airway BC (Figure 4 B, C).

Cultured alveolar and airway basal cells display similar differentiation capacities when cultured on an ALI

To compare the differentiation potential of alveolar and airway BC, the cells were cultured on an ALI platform for 23 days (Figure 5 A). Alveolar and airway BC differentiated to secretory and ciliated epithelial cells after 23 days in ALI culture, which was reflected by a significant reduction of *KRT5* RNA and an increase in the RNA expression of the secretory epithelial cell markers *SCGB1A1*, *MUC5AC*, and *MUC5B* and the ciliated epithelial cell marker *FOXJ1* between day 0 and day 23 (Figure 5 B). RNA expression levels of these markers at day 23 were not significantly different between alveolar and airway BC (Figure 5 B).

At the protein level, alveolar and airway BC did not express AcTub (ciliated epithelial cell marker), MUC5AC, or SCGB1A1 protein at day 0 (Figure 4 C), while after 23 days, AcTub, MUC5AC, MUC5B or SCGB1A1 protein expression was strongly induced and detected in cells on PFA-fixed (Figure 5 C) and/or cryo-sectioned (Figure 5 D) ALI membranes. Again, no significant difference in marker protein expression between alveolar and airway BC was apparent (Figure 5 C, D).

Alveolar and airway basal cell-derived organoids display similar characteristics

To assess characteristics of airway and alveolar BC-derived organoids, we cultured the cells in Matrigel for 21 days (Figure 6 A). Alveolar and airway BC started to form

organoids (>50µm diameter) after 7 days, which continued to grow in size and number over a period of 21 days (Figure 6 B). After 20 days, alveolar BC formed between 45-138 organoids/mm^3, and airway BC between 63-162 organoids/mm^3, with no significant differences comparing alveolar to airway BC (Figure 6 C). The size of alveolar and airway BC-derived organoids were not significantly different and ranked between 66-122 µm in alveolar BC and between 73-105 µm in airway BC-derived organoids (Figure 6 C). Cells within alveolar and airway BC-derived organoids after 21 days expressed similar levels of *TP63*, *KRT5*, *KRT17*, *SCGB1A1*, *MUC5AC*, and *MUC5B* RNA, whereas *FOXJ1* RNA was slightly, but significantly higher in alveolar BC-derived organoids when compared to those derived from airway BC (Figure 6 D). Similarly, cells within alveolar and airway BC-derived organoids expressed similar levels of KRT5, p63, KRT17, SCGB1A1, MUC5AC, MUC5B and AcTub protein (Figure 6 E).

Alveolar and airway BC show similar behavior after intratracheally injection into bleomycin-challenged mice

Mice were intratracheally injected with a low dose of bleomycin. After three days, human alveolar or airway BC with or without a luciferase and GFP encoding vectors were intratracheally administered and mice sacrificed after 21 days (Figure 7 A). Alveolar and airway BC engrafted and proliferated, which was reflected by increasing expression of luciferase in mice injected with transfected cells (Figure 7 B). As shown by H&E stainings, bleomycin induced mild fibrotic changes in the mouse lungs, which were slightly increased when alveolar and airway BC were injected after bleomycin-treatment (Figure 7 C), but according to fibrosis scores the increase was not significant (Figure 7 D). Mouse tissue sections were stained with the human-specific protein HNA, which again confirmed the engraftment of human cells and helped to differentiate human from mouse cells (Figure 7 E). Performing IHC/IF stainings with antibodies targeting basal- (KRT5, KRT17), secretory- (MUC5AC, MUC5B, SCGB1A1), or ciliated (AcTub) epithelial cell markers in mouse lung tissue sections, we observed that human alveolar and airway BC either maintained their basal cell identity or differentiated to MUC5AC+, MUC5B+, or SCGB1A1+ secretory epithelial cells (Figure 7 E). At day 21, we did not observe differences in fibrosis scores or differentiation capacity within mouse lung tissue, which were injected with either human alveolar or airway BC (Figure 7 D, E).

Discussion

In ILD such as IPF, a maladaptive injury repair mechanism occurs in which alveolar epithelial cells are depleted and ectopic BC accumulate within the alveolar space [11, 16, 30-32, 46, 52-55]. To get a better understanding of the origin and characteristics of the disease-related alveolar BC, we made use of a previously established *in vitro* BC culture method allowing us to isolate and expand the cells in sufficient numbers [15]. We studied the morphology, BC marker expression, transcriptomic landscape, and proliferation and differentiation capacities of cultures alveolar BC in comparison to classic airway BC derived from the same ILD patients. Our data show that alveolar BC display similar characteristics and functional properties like their airway counterparts: Alveolar and airway BC displayed similar morphology, BC marker expression, and proliferation and differentiation capacity *in vitro*. Furthermore, when instilled into bleomycin-challenged mice, human alveolar and airway BC showed similar capacity to engraft and to differentiate *in vivo*. scRNA-seq analysis revealed transcriptomic differences between cultured alveolar and airway BC, with an up-regulation of genes associated with fibrosis-related pathways such as hypoxia or EMT in alveolar BC, suggesting that the cells are imprinted by their fibrotic *in vivo* microenvironment.

Human AT2 cells were shown to trans-differentiate to BC through intermediate cells, co-expressing varying levels of AT2 cell- and BC marker genes *in vitro* [6]. In order to determine, if our cultured alveolar BC could originate from AT2 cells, we compared BC and AT2 cell marker gene expression in alveolar versus airway BC. IF stainings and TaqMan RT-PCR, using RNA isolated from whole cell lysates, showed no significant differences in BC marker expression in cultured alveolar versus airway BC. scRNA-seq analysis revealed an overall trend, although not significant, towards reduced BC marker gene expression in alveolar BC, which was, however, not accompanied with expression of specific AT2 cell marker genes. This observation tends to favour the hypothesis that the source of our cultured alveolar BC would be migrating airway BC. However, it is also possible that alveolar BC could have fully lost AT2 cell traits after trans-differentiation in the lung, or later after expanding them in culture. Furthermore, it is also possible that our cell culture method introduce a selection bias by favouring

the outgrowth of cells closely related to classic BC, but not that of intermediate cell types.

BC in the healthy airways are a heterogeneous cell population, including canonical, activated, proliferating and secretory-primed BC subsets [45, 51]. Our scRNA-seq data demonstrated that the majority of our cultured cells display close transcriptomic similarity to canonical, proliferating, and secretory-primed BC subsets in both, alveolar and airway-derived BC populations. Secretory-primed BC in our BC population likely arise through spontaneous BC differentiation in culture, following a previously described differentiation gradient [17], originating from a cluster of proliferating alveolar BC to secretory-primed BC. Interestingly, we observed differences in relative abundance between alveolar and airway-derived cells in some of the subpopulations of canonical and secretory-primed cell clusters. Notably, the secretory-primed BC cluster contained less alveolar than airway BC, while a subpopulation of canonical BS showing enrichment for several metallothioneins (MTs), FOS, AGR2 or KRT15, contained more. This may indicate an altered differentiation path in alveolar BC, resulting in less secretory-primed BC. This would be in line with previous studies, showing disturbed cell differentiation with an accumulation of disease-specific intermediate cell types such as the aberrant basaloid cells in the peripheral fibrotic lung [30, 31]. Interestingly, alveolar BC showed a trend towards reduced BC and increased mesenchymal marker expression similar to aberrant basaloid cells [30, 31]. Furthermore, alveolar BC tended to differentiate more towards cells that are enriched with the cancer-associated MTs [56], FOS [57], AGR2 [58] or KRT15 [59], which may indicate a contribution of ectopic BC to developing lung cancinoma [60], a risk that is increased in IPF patients [61]. However, these conclusions are still speculative and would need closer assessment in future studies.

Despite overall similarities between alveolar and airway BC, they considerably differed in their transcriptomic landscape. Cultured alveolar BC showed an up-regulation of genes associated with glycolysis, KRAS signalling, and hypoxia. Cellular glycolysis is activated under hypoxic conditions in order to maintain sufficient ATP production [62]. Furthermore, hypoxia activates KRAS signaling to increase survival of cancer cells [63]. As mentioned earlier, the amount of alveolar epithelial type 1 and 2 (AT1 and 2) cells in the IPF lung parenchyma is reduced and it is therefore not surprising that

hypoxia is increased in fibrotic areas of the IPF lung [3, 64, 65]. Ectopic BC specifically accumulate in hypoxic regions of the IPF lung parenchyma [3] and are thus likely imprinted by their hypoxic *in vivo* microenvironment. Interestingly, these imprinted changes persist after *in vitro* culture in our study system. Whether these transcriptomic differences result in altered cellular functions or may serve as specific drug targets remains to be determined in future studies.

BC seem to preferentially differentiate to MUC5AC and/or MUC5B expressing secretory epithelial cells, depending on their location within the healthy airways: The majority of secretory epithelial cells within the proximal airways co-express MUC5AC, MUC5B and SCGB1A1, whereas distal secretory epithelial cells mainly co-express MUC5B and SCGB1A1, but not MUC5AC [66]. Secretory cells in the terminal bronchioles express SCGB1A1, but not MUC5AC or MUC5B [66]. Alveoli within the healthy lung parenchyma do not contain BC or any other airway epithelial cells. In the IPF lung, BC line pathological honeycomb (HC) cysts within the alveolar region, where they appear to preferentially differentiate to MUC5B, but not MUC5AC-expressing cells [67]. In this study, we show that alveolar and airway BC differentiated to MUC5B and MUC5AC-expressing secretory epithelial cells when cultured under the same *in vitro* conditions, indicating that the cellular microenvironment determines BC fate decision, and that this was not influenced by the origin of BC (alveolar or airway). Alterations of the BC microenvironment may therefore serve as a novel pharmacological intervention to direct or inhibit BC differentiation.

Human fibrotic BC instilled into bleomycin-challenged mice have previously been shown to aggravate bleomycin-induced fibrosis [12] and to form HC-like structures within the mouse parenchyma [12, 15]. In this study, we instilled alveolar or airway BC derived from the same ILD patient and compared their behaviour in the mouse lung *in vivo*. There was a trend towards increased fibrosis after instillation of BC in the bleomycin-challenged mouse lung. Furthermore, both human alveolar and airway BC differentiated to secretory epithelial cells, forming HC-like structures within the mouse lung, without significant difference in the mediated effects of BC from the two origins.

Among weaknesses of this study, a small sample size for some of our experiments limits the power and robustness of the results. Furthermore, *in vitro* culture of BC may

alter their phenotype and may therefore not fully recapitulate BC characteristics and functions *in vivo*. However, our data provide novel insights into the origin and disease-specific characteristics of alveolar BC: BC isolated from the fibrotic alveolar space or from the airways display similar characteristics, suggesting that the here cultured alveolar BC are derived from migrating airway BC. The transcriptomic landscape and differentiation capacity of BC seems largely dependent from their specific microenvironment rather than their origin, indicating that altering the *in vivo* BC microenvironment may serve as a potential therapeutic approach to limit fibrogenesis in ILD.

Acknowledgments

The authors acknowledge the microscopy core facility at the Department of Biomedicine in Basel for their support and assistance in ICC/IF, IHC/IF image acquisition and the sciCORE (http://scicore.unibas.ch/) scientific computing center at the University of Basel for performing calculations.

Author Contributions

SB: Cell culture of human primary lung cells, ALI and organoid processing, ICC/IF, IHC/IF, image acquisition and image quantification in cultured cells, human IPF- and mouse tissue slides, immunoblotting, PCR, proliferation assay, scratch assay, data analysis and interpretation; **PK:** Conception and design of the study, cell culture of human primary lung cells, PCR, data analysis and interpretation, writing the manuscript; **JR:** Data analysis and interpretation of scRNA-seq data, revision and editing of the manuscript for important intellectual content; **NA:** bleomycin-mouse model, data analysis and interpretation, revision and editing of the manuscript for important intellectual content; **LP:** bleomycin-mouse model, data analysis and interpretation, revision and editing of the manuscript for important intellectual content; **AP:** bleomycin-mouse model, data analysis and interpretation, revision and editing of the manuscript for important intellectual content; **KEH:** Conception and design of the study, data analysis and interpretation, revision and editing of the manuscript for important intellectual content.

Funding

This study was supported by a project grant (310030_192536) by the Swiss National Research Foundation.

Conflict of interest

The authors have declared that no conflict of interest exists.

Data availability

The processed scRNA-seq dataset (the accession is GSE270495) can be viewed on the gene expression omnibus database (GEO).

References

1. Barkauskas, C.E., et al., *Type 2 alveolar cells are stem cells in adult lung.* J Clin Invest, 2013. **123**(7): p. 3025-36.
2. Rock, J.R., et al., *Basal cells as stem cells of the mouse trachea and human airway epithelium.* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12771-5.
3. Xi, Y., et al., *Local lung hypoxia determines epithelial fate decisions during alveolar regeneration.* Nat Cell Biol, 2017. **19**(8): p. 904-914.
4. Yang, Y., et al., *Spatial-Temporal Lineage Restrictions of Embryonic p63(+) Progenitors Establish Distinct Stem Cell Pools in Adult Airways.* Dev Cell, 2018. **44**(6): p. 752-761.e4.
5. Basil, M.C. and E.E. Morrisey, *Lung regeneration: a tale of mice and men.* Semin Cell Dev Biol, 2020. **100**: p. 88-100.
6. Kathiriya, J.J., et al., *Human alveolar type 2 epithelium transdifferentiates into metaplastic KRT5(+) basal cells.* Nat Cell Biol, 2021.
7. Kumar, P.A., et al., *Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection.* Cell, 2011. **147**(3): p. 525-538.
8. Yuan, T., et al., *FGF10-FGFR2B Signaling Generates Basal Cells and Drives Alveolar Epithelial Regeneration by Bronchial Epithelial Stem Cells after Lung Injury.* Stem Cell Reports, 2019. **12**(5): p. 1041-1055.
9. Kanegai, C.M., et al., *Persistent Pathology in Influenza-Infected Mouse Lungs.* Am J Respir Cell Mol Biol, 2016. **55**(4): p. 613-615.
10. Ray, S., et al., *Rare SOX2(+) Airway Progenitor Cells Generate KRT5(+) Cells that Repopulate Damaged Alveolar Parenchyma following Influenza Virus Infection.* Stem Cell Reports, 2016. **7**(5): p. 817-825.
11. Prasse, A., et al., *BAL Cell Gene Expression Is Indicative of Outcome and Airway Basal Cell Involvement in Idiopathic Pulmonary Fibrosis.* Am J Respir Crit Care Med, 2019. **199**(5): p. 622-630.
12. Jaeger, B., et al., *Airway basal cells show a dedifferentiated KRT17highPhenotype and promote fibrosis in idiopathic pulmonary fibrosis.* Nature Communications, 2022. **13**(1): p. 5637.
13. Hostettler, K.E., et al., *Multipotent mesenchymal stem cells in lung fibrosis.* PLoS One, 2017. **12**(8): p. e0181946.
14. Khan, P., et al., *Basal-Like Cell-Conditioned Medium Exerts Anti-Fibrotic Effects In Vitro and In Vivo.* Frontiers in Bioengineering and Biotechnology, 2022. **10**.
15. Blumer, S., et al., *The use of cultured human alveolar basal cells to mimic honeycomb formation in idiopathic pulmonary fibrosis.* Respiratory Research, 2024. **25**(1): p. 26.

16. Khan, P., et al., *Basal-Like Cell-Conditioned Medium Exerts Anti-Fibrotic Effects In Vitro and In Vivo.* Front Bioeng Biotechnol, 2022. **10**: p. 844119.
17. Khan, P., et al., *Alveolar Basal Cells Differentiate towards Secretory Epithelial- and Aberrant Basaloid-like Cells In Vitro.* Cells, 2022. **11**(11).
18. Seidel, P., et al., *Dimethylfumarate inhibits NF-{kappa}B function at multiple levels to limit airway smooth muscle cell cytokine secretion.* Am J Physiol Lung Cell Mol Physiol, 2009. **297**(2): p. L326-39.
19. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
20. Griffiths, J.A., et al., *Detection and removal of barcode swapping in single-cell RNA-seq data.* Nat Commun, 2018. **9**(1): p. 2667.
21. Lun, A.T.L., et al., *EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data.* Genome Biol, 2019. **20**(1): p. 63.
22. Vallejos, C.A., et al., *Normalizing single-cell RNA sequencing data: challenges and opportunities.* Nat Methods, 2017. **14**(6): p. 565-571.
23. Lun, A.T., D.J. McCarthy, and J.C. Marioni, *A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor.* F1000Res, 2016. **5**: p. 2122.
24. McCarthy, D.J., et al., *Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R.* Bioinformatics, 2017. **33**(8): p. 1179-1186.
25. Amezquita, R.A., et al., *Orchestrating single-cell analysis with Bioconductor.* Nat Methods, 2020. **17**(2): p. 137-145.
26. Ilicic, T., et al., *Classification of low quality cells from single-cell RNA-seq data.* Genome Biol, 2016. **17**: p. 29.
27. Germain, P.L., et al., *Doublet identification in single-cell sequencing data using scDblFinder.* F1000Res, 2021. **10**: p. 979.
28. Haghverdi, L., et al., *Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors.* Nat Biotechnol, 2018. **36**(5): p. 421-427.
29. Aran, D., et al., *Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage.* Nat Immunol, 2019. **20**(2): p. 163-172.
30. Adams, T.S., et al., *Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis.* Science Advances, 2020. **6**(28): p. eaba1983.
31. Habermann, A.C., et al., *Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis.* Science Advances, 2020. **6**(28): p. eaba1972.
32. Reyfman, P.A., 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): p. 1517-1536.
33. Travaglini, K.J., et al., *A molecular cell atlas of the human lung from single-cell RNA sequencing.* Nature, 2020. **587**(7835): p. 619-625.
34. Sikkema, L., et al., *An integrated cell atlas of the lung in health and disease.* Nat Med, 2023. **29**(6): p. 1563-1577.
35. Lun, A.T.L. and J.C. Marioni, *Overcoming confounding plate effects in differential expression analyses of single-cell RNA-seq data.* Biostatistics, 2017. **18**(3): p. 451-464.

36. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.
37. Robinson, M.D. and A. Oshlack, *A scaling normalization method for differential expression analysis of RNA-seq data.* Genome Biol, 2010. **11**(3): p. R25.
38. Wu, D. and G.K. Smyth, *Camera: a competitive gene set test accounting for inter-gene correlation.* Nucleic Acids Res, 2012. **40**(17): p. e133.
39. Liberzon, A., et al., *The Molecular Signatures Database (MSigDB) hallmark gene set collection.* Cell Syst, 2015. **1**(6): p. 417-425.
40. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-15550.
41. Jacob, A., et al., *Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells.* Nature Protocols, 2019. **14**(12): p. 3303-3332.
42. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
43. Suarez-Arnedo, A., et al., *An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays.* PLoS One, 2020. **15**(7): p. e0232565.
44. Ashcroft, T., J.M. Simpson, and V. Timbrell, *Simple method of estimating severity of pulmonary fibrosis on a numerical scale.* J Clin Pathol, 1988. **41**(4): p. 467-70.
45. Carraro, G., et al., *Transcriptional analysis of cystic fibrosis airways at single-cell resolution reveals altered epithelial cell states and composition.* Nature Medicine, 2021. **27**(5): p. 806-814.
46. Strunz, M., et al., *Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis.* Nat Commun, 2020. **11**(1): p. 3559.
47. Kadur Lakshminarasimha Murthy, P., et al., *Human distal lung maps and lineage hierarchies reveal a bipotent progenitor.* Nature, 2022. **604**(7904): p. 111-119.
48. Sikkema, L., et al., *An integrated cell atlas of the lung in health and disease.* Nature Medicine, 2023. **29**(6): p. 1563-1577.
49. Deprez, M., et al., *A Single-Cell Atlas of the Human Healthy Airways.* Am J Respir Crit Care Med, 2020. **202**(12): p. 1636-1645.
50. Hackett, N.R., et al., *The human airway epithelial basal cell transcriptome.* PLoS One, 2011. **6**(5): p. e18378.
51. Carraro, G., et al., *Single-Cell Reconstruction of Human Basal Cell Diversity in Normal and Idiopathic Pulmonary Fibrosis Lungs.* Am J Respir Crit Care Med, 2020. **202**(11): p. 1540-1550.
52. Xu, Y., et al., *Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis.* JCI Insight, 2017. **1**(20).
53. Seibold, M.A., et al., *The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium.* PLoS One, 2013. **8**(3): p. e58658.
54. Smirnova, N.F., et al., *Detection and quantification of epithelial progenitor cell populations in human healthy and IPF lungs.* Respiratory Research, 2016. **17**(1): p. 83.
55. Kathiriya, J.J., et al., *Human alveolar Type 2 epithelium transdifferentiates into metaplastic KRT5+ basal cells during alveolar repair.* bioRxiv, 2020: p. 2020.06.06.136713.
56. Si, M. and J. Lang, *The roles of metallothioneins in carcinogenesis.* J Hematol Oncol, 2018. **11**(1): p. 107.

57. Kuonen, F., et al., *c-FOS drives reversible basal to squamous cell carcinoma transition.* Cell Rep, 2021. **37**(1): p. 109774.
58. Moidu, N.A., et al., *Secretion of pro-oncogenic AGR2 protein in cancer.* Heliyon, 2020. **6**(9): p. e05000.
59. Chen, W. and C. Miao, *KRT15 promotes colorectal cancer cell migration and invasion through β-catenin/MMP-7 signaling pathway.* Med Oncol, 2022. **39**(5): p. 68.
60. Königshoff, M., *Lung Cancer in Pulmonary Fibrosis: Tales of Epithelial Cell Plasticity.* Respiration, 2011. **81**(5): p. 353-358.
61. Stella, G.M., et al., *The oncogenic landscape of the idiopathic pulmonary fibrosis: a narrative review.* Transl Lung Cancer Res, 2022. **11**(3): p. 472-496.
62. Kierans, S.J. and C.T. Taylor, *Regulation of glycolysis by the hypoxia-inducible factor (HIF): implications for cellular physiology.* The Journal of Physiology, 2021. **599**(1): p. 23-37.
63. Zeng, M., et al., *Hypoxia activates the K-ras proto-oncogene to stimulate angiogenesis and inhibit apoptosis in colon cancer cells.* PLoS One, 2010. **5**(6): p. e10966.
64. Tzouvelekis, A., et al., *Comparative expression profiling in pulmonary fibrosis suggests a role of hypoxia-inducible factor-1alpha in disease pathogenesis.* Am J Respir Crit Care Med, 2007. **176**(11): p. 1108-19.
65. Bodempudi, V., et al., *miR-210 promotes IPF fibroblast proliferation in response to hypoxia.* Am J Physiol Lung Cell Mol Physiol, 2014. **307**(4): p. L283-94.
66. Okuda, K., et al., *Localization of Secretory Mucins MUC5AC and MUC5B in Normal/Healthy Human Airways.* Am J Respir Crit Care Med, 2019. **199**(6): p. 715-727.
67. Plantier, L., et al., *Ectopic respiratory epithelial cell differentiation in bronchiolised distal airspaces in idiopathic pulmonary fibrosis.* Thorax, 2011. **66**(8): p. 651-7.

Figure legends

Figure 1: Cultured alveolar and airway BC do not differ in their morphology or BC marker expression. KRT5+ and TP63+ BC in formalin fixed paraffin embedded (FFPE) transbronchial biopsies (TBB) or airway mucosal biopsies (MB) of the same ILD patient (n=5) (A). To culture alveolar or airway BC, tissue specimens were placed in cell culture dishes containing Cnt-PR-A as illustrated in B. Representative phase contrast images showing the morphology of confluent alveolar (n=5) or airway BC (n=5) (C). RNA expression (mean ± SEM) of KRT5, KRT14, KRT17, TP63, CDH2 or SCGB1A1 in alveolar (n=6) or airway BC (n=6) (D). Representative IF stainings for KRT5/p63, KRT17/KRT14, CDH2, SCGB1A1 KRT5/TP63 (E) in cultured alveolar (n=3) and airway BC (n=3) (E).

Figure 2: Alveolar and airway BC differ in their transcriptome. T-SNE plot (uncorrected data) showing cultured alveolar and airway BC derived from donor 1-3 (A). T-SNE (after MNN correction) showing the separation of cells into 7 clusters (B).

T-SNE showing the best-matching reference label in the transcriptome comparison to a publicly available scRNA-seq dataset [41] (C). Heatmap of the relative expression across clusters, donors, and sample location of the strongest cluster-specific genes. The log- normalized expression levels were averaged across cells from the same cluster, donor, and sample location, corrected for donor-specific effects to mirror the MNN correction performed before clustering and the inclusion of donor as a covariate in the differential expression analysis, and averaged values were then centered and scaled per gene. Marker genes are highlighted in bold (D). Proportion of alveolar and airway BC across donors in clusters 1, 2, 3, 4, 6, and 7 (E).

Figure 3: Differential genes and pathways in alveolar versus airway BC. Heatmap showing the differential expressed genes in alveolar versus airway BC across all tested clusters. The plotted relative expression values were calculated similarly to Figure 2D. Annotation colors for rows indicate in which specific cluster each gene was found significantly differentially expressed at an FDR threshold of 1% (A). Volcano plots showing the differential expression analysis results between alveolar and airway BC in cluster 2 (B) and cluster 6 (C). Genes colored (beige=decreased, and blue=increased in alveolar BC relative to airway BC) are significant at an FDR threshold of 1% (B, C). Functional enrichment analysis amongst differentially expressed genes in cluster 2 (D) and cluster 6 (E) using MSigDB Hallmark gene sets. Six differentially expressed genes (*LMO3*, *S100A4*, *SPOCK3*, *HNMT*, *IFI27*, *BST2*) were selected and analyzed by Taqman RT-qPCR (F). Data points in red represent results from cells derived from donor 1-3. Data points in black were derived from cells of three additional donors. Data are presented as mean ± SEM (paired t-test, ** indicates $p < 0.01$, * indicates $p < 0.05$).

Figure 4: Alveolar and airway basal cells display similar proliferation and wound healing capacities. Automated cell counts of cultured alveolar and airway BC at 0 and 48 hours (n=4) (A). Representative phase contrast images showing alveolar and airway BC wound closure between 0-24 h (B). Bar chart displaying BC wound closure (%) at the indicated time points. Dots (alveolar BC) or squares (airway BC) within the

bars represent datapoints generated for alveolar or airway BC from each patient (n=4). Data are expressed as mean ± SEM (C).

Figure 5: Alveolar and airway basal cells have similar differentiation capacities. Illustration of alveolar and airway BC ALI culture (A). RNA expression levels (-ΔCt) of KRT5, SCGB1A1, MUC5AC, MUC5B, or FOXJ1 in ALI-cultured alveolar (n=4) and airway (n=4) BC at day 0 and 23. Data are expressed as mean ± SEM (paired t-test, **** indicates $p < 0.0001$, *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$) (B). Representative ICC/IF images of alveolar (n=3) and airway (n=3) BC incubated with antibodies detecting acetylated-tubulin (AcTub), MUC5AC or SCGB1A1 at day 0 and 23 (C). Representative images of cryosections incubated with antibodies detecting KRT5, p63, SCGB1A1, KRT17, MUC5AC, MUC5B, or AcTub of ALI-cultured alveolar (n=3) or airway (n=3) BC at day 23 (D).

Figure 6: Alveolar and airway basal cells form organoids containing basal, secretory- and ciliated epithelial cells. Alveolar and airway BC were resuspended in Matrigel and cultured for 21 days as illustrated (A). Representative phase contrast images of alveolar (n=7) or airway (n=7) BC organoid formation at day 2, 7, 14, and 21 (B). Number of alveolar and airway BC-derived organoids/mm^3 and the distribution of their diameters (>50μm) after 20 days (n=7; mean ± SEM; unpaired t-test, ns=not significant) (C). RNA expression levels (-ΔCt) of TP63, KRT5, KRT17, SCGB1A1, MUC5AC, MUC5B, or FOXJ1 in organoids after 21 days (n=4; mean ± SEM; unpaired t-test, * values ≤ 0.05, ns=not significant) (D). Representative H&E and ICC/IF images of alveolar (n=3) or airway (n=3) BC-derived organoids incubated with antibodies detecting KRT17, KRT5, p63, SCGB1A1, AcTub, MUC5AC, or MUC5B (E).

Figure 7: Human alveolar or airway basal cells in bleomycin-challenged mouse lungs. NRG mice were treated with bleomycin (n=3) +/- human alveolar (n=3) or airway (n=3) BC carrying a luciferase encoding vector as illustrated (A). Luciferase expression in mice treated with bleomycin +/- human alveolar or airway BC carrying a luciferase encoding vector and the corresponding bioluminescence measurements (B) at day 4, 11, and 21. Representative images of H&E stainings in mice treated with bleomycin +/- human alveolar or airway BC (C). Fibrosis score for mice treated with

bleomycin +/- human alveolar or airway BC (D). Representative IHC/IF images of mouse lung tissue (mice treated with bleomycin + human alveolar or airway BC) incubated with antibodies detecting human-specific HNA with or without KRT5, KRT17, SCGB1A1, MUC5AC, MUC5B, or AcTub. Nuclei were counter-stained with DAPI. White squares indicate regions imaged at higher magnification (E).

Supplement figure 1: tSNE showing the distribution of alveolar and airway BC from donor 1 (airway BC-green; alveolar BC- grey), donor 2 (airway BC- dark green; alveolar BC- light blue) and donor 3 (airway BC- lilac; alveolar BC-purple) (A). Percentage of scRNA-seq reads originating from mitochondrial genes (B). Number of detected genes across cells (number of genes with at least one UMI count, on a log-scale) (C).

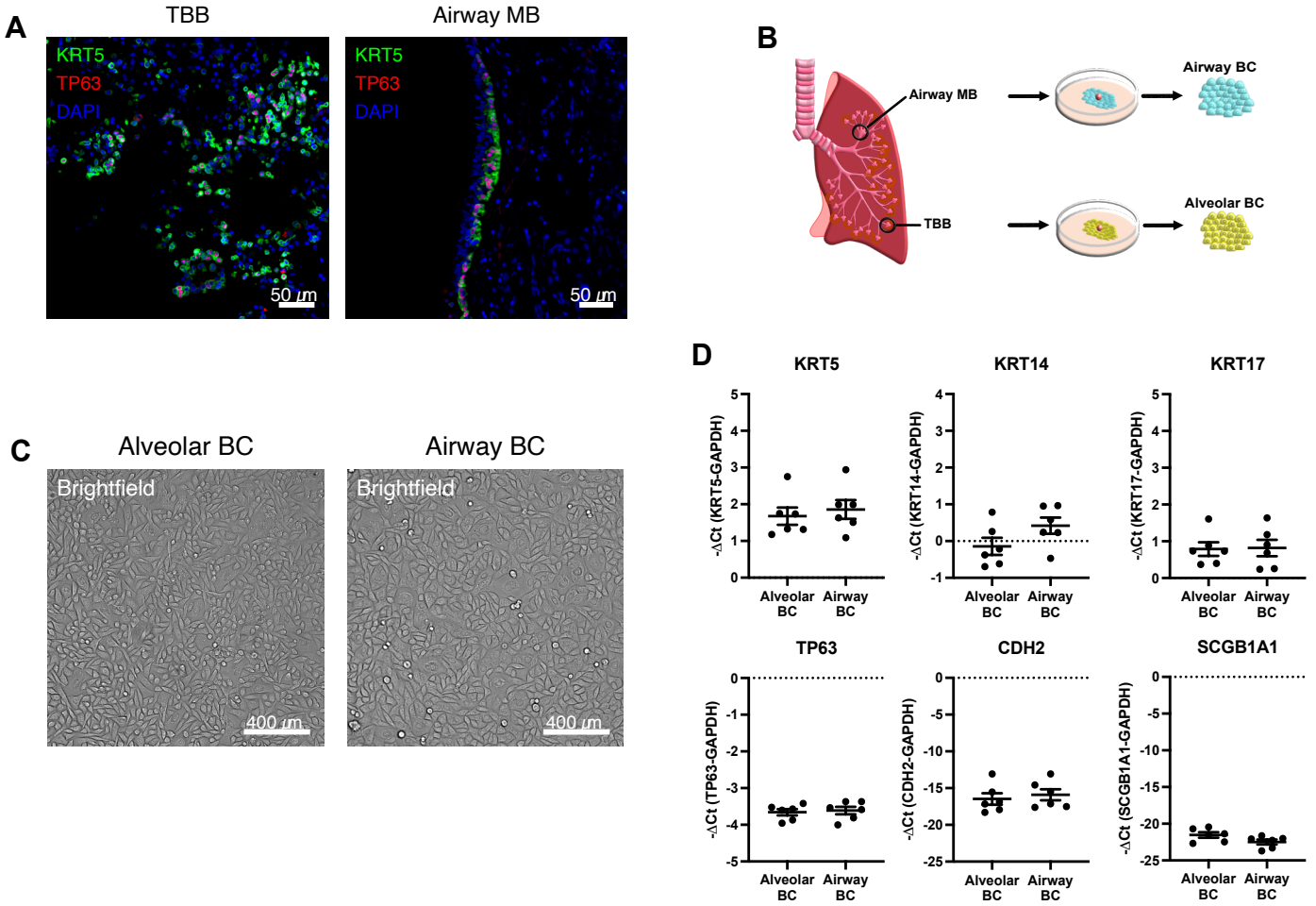
Supplement figure 2: tSNE showing the best matching cell type for each cell from the *SingleR* analysis using as reference the scRNA-seq datasets from Adams et al. [30] (A), Habermann et al. [31] (B), Reyfmann et al. [32] (C), Travaglini et al. [33] (D), Kathiriya et al. [6] (E), Sikkema et al. [48] (F), Murthy et al. [47] (G), Strunz et al. [46] (H) or Deprez et al. [49] (I). Cell types represented by no more than 5 cells in our dataset were pooled into the category “Others”.

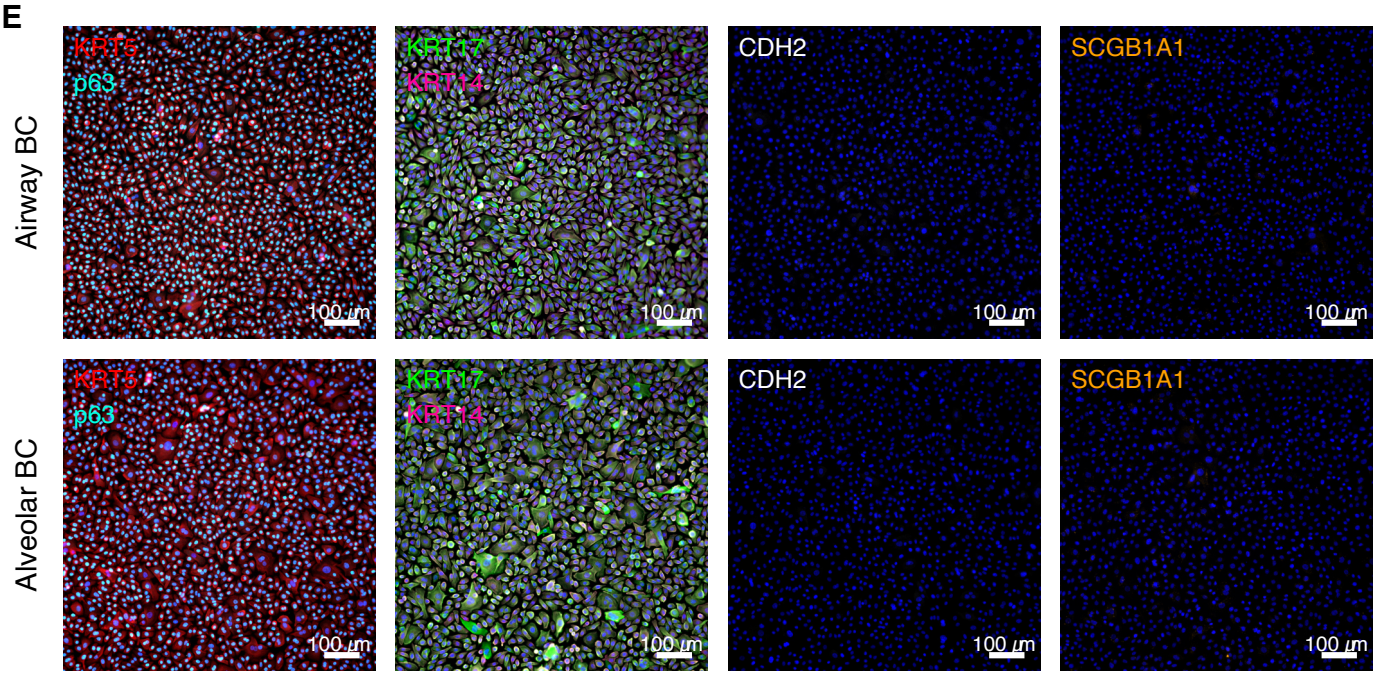
Supplement figure 3: Heatmap of the relative expression across clusters, donors, and sample location of selected marker genes.

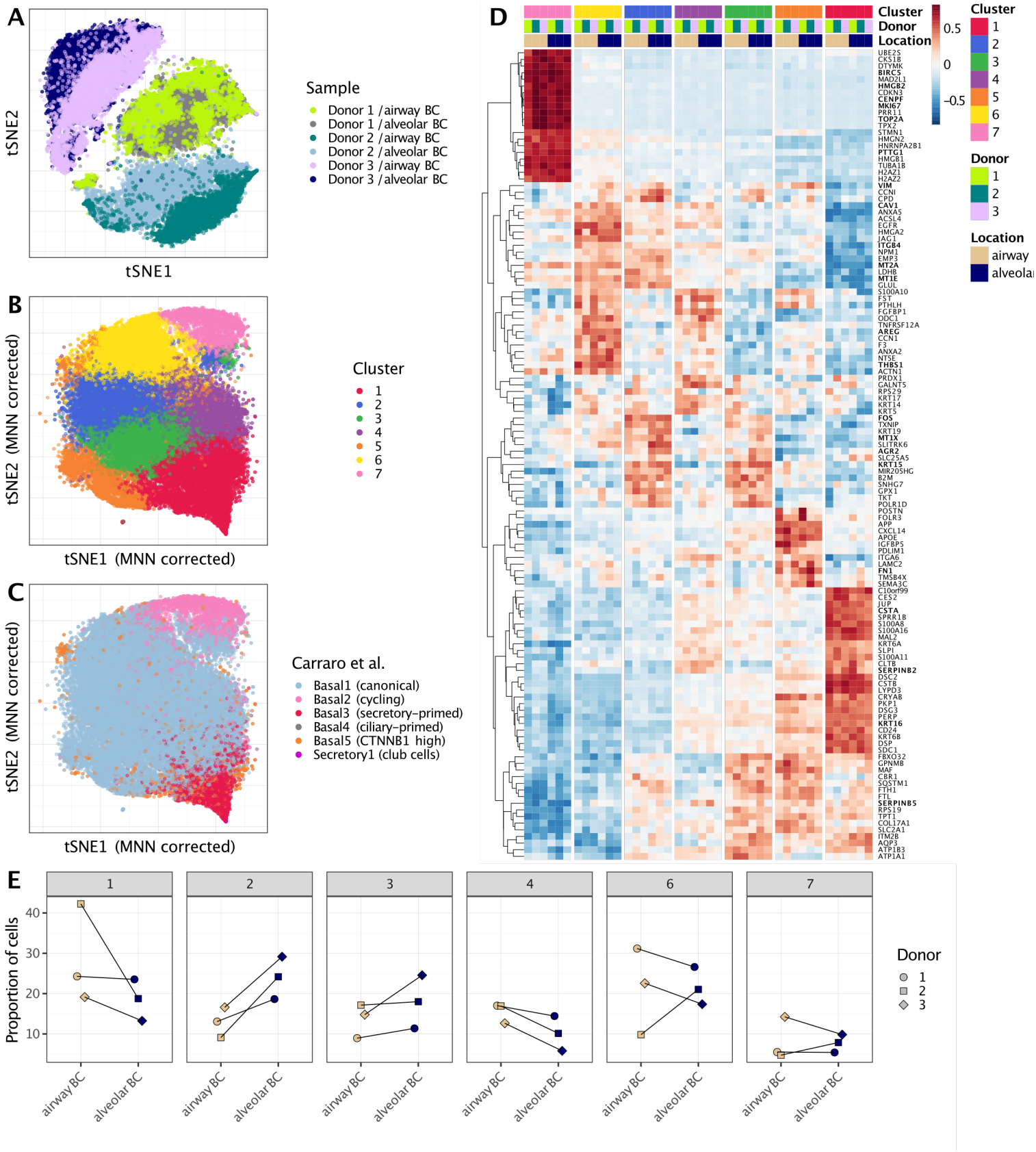
Supplement figure 4: Volcano plots showing the differential expression analysis results between alveolar and airway BC in cluster 1 (A), cluster 3 (B), cluster 4 (C), and cluster 7 (D). Genes colored (beige=decreased, and blue=increased in alveolar BC relative to airway BC) are significant at an FDR threshold of 1%.

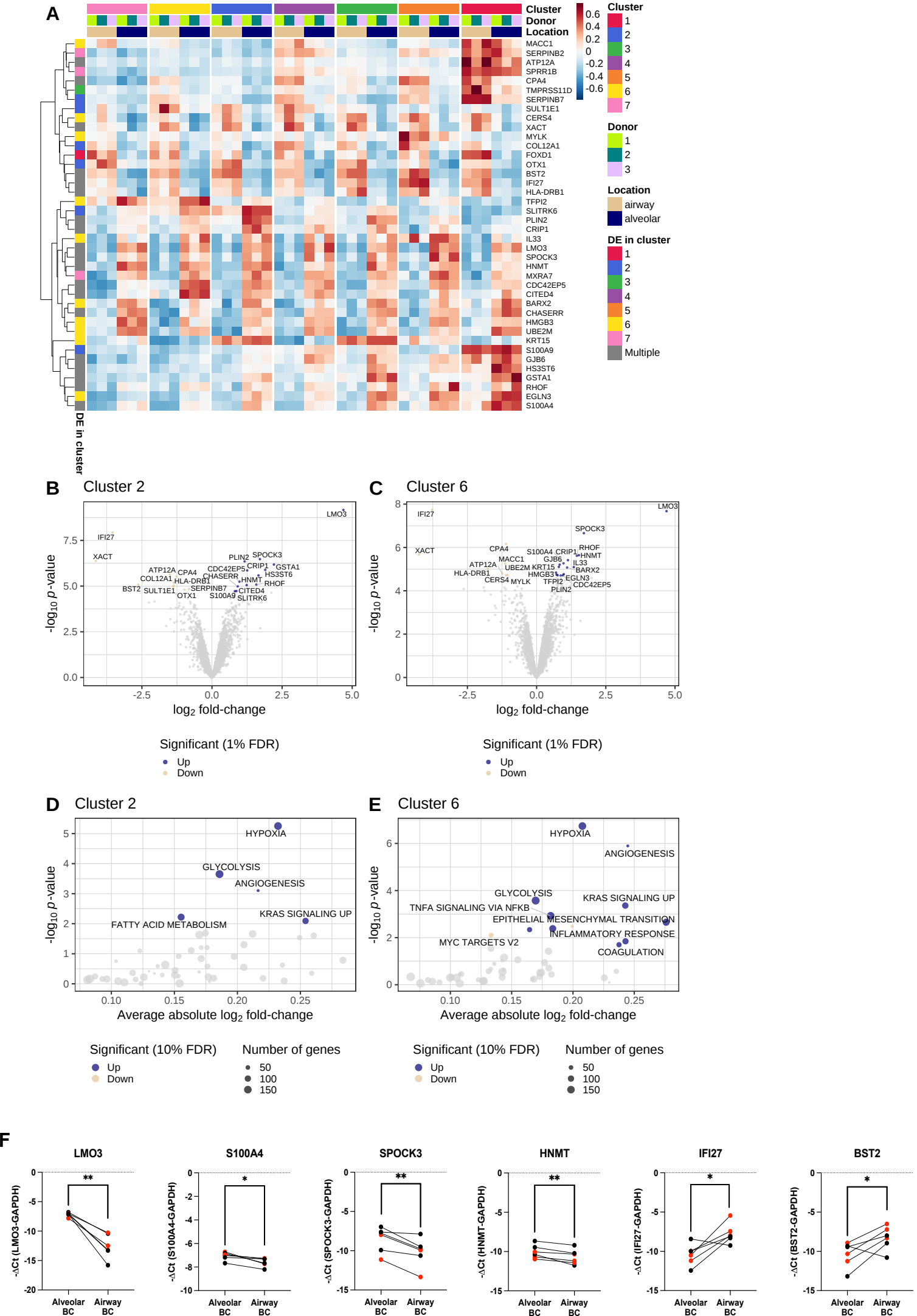
Supplement figure 5: Functional enrichment analysis amongst differentially expressed genes in cluster 1 (A), cluster 3 (B), cluster 4 (C), and cluster 7 (D) using MSigDB Hallmark gene sets.

Supplement figure 6: Functional enrichment analysis amongst differentially expressed genes in cluster 1 (A), cluster 2 (B), cluster 3 (C), cluster 4 (D), cluster 6 (E) and cluster 7 (F) using the MSigDB Reactome gene sets from the C2 collection.

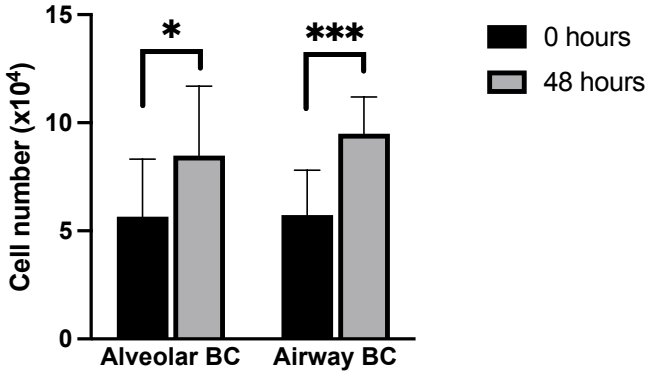

A
TBB
Airway MB
KRT5
TP63
DAPI
50 μm
B
Airway MB
TBB
Airway BC
Alveolar BC
C
Alveolar BC
Airway BC
Brightfield
400 μm
D
KRT5
KRT14
KRT17
TP63
CDH2
SCGB1A1
-ΔCt (KRT5-GAPDH)
-ΔCt (KRT14-GAPDH)
-ΔCt (KRT17-GAPDH)
-ΔCt (TP63-GAPDH)
-ΔCt (CDH2-GAPDH)
-ΔCt (SCGB1A1-GAPDH)
Alveolar BC
Airway BC


E
Airway BC
Alveolar BC
KRT5
p63
KRT17
KRT14
CDH2
SCGB1A1
100 μm

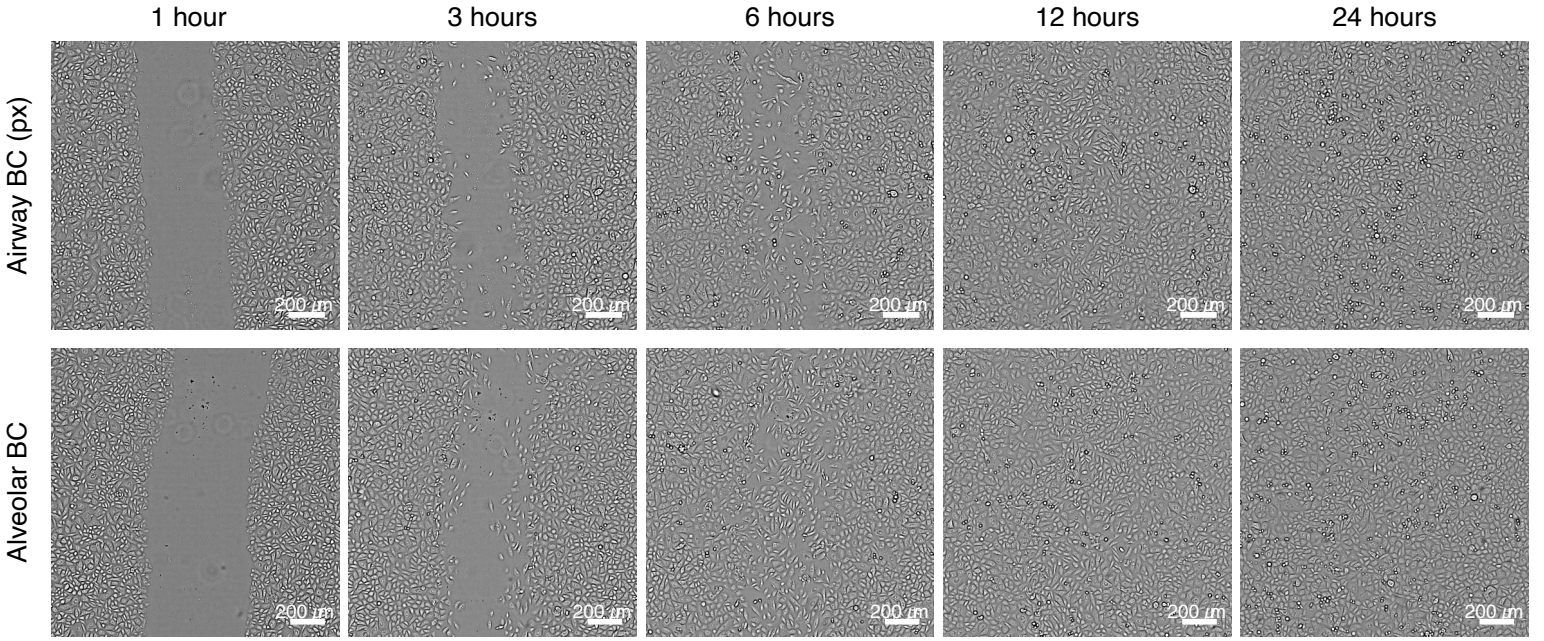

A
tSNE2
tSNE1
Sample
Donor 1 / airway BC
Donor 1 / alveolar BC
Donor 2 / airway BC
Donor 2 / alveolar BC
Donor 3 / airway BC
Donor 3 / alveolar BC
B
tSNE2 (MNN corrected)
tSNE1 (MNN corrected)
Cluster
1
2
3
4
5
6
7
C
tSNE2 (MNN corrected)
tSNE1 (MNN corrected)
Carraro et al.
Basal1 (canonical)
Basal2 (cycling)
Basal3 (secretory-primed)
Basal4 (ciliary-primed)
Basal5 (CTNNB1 high)
Secretory1 (club cells)
D
Cluster
Donor
Location
0.5
0
−0.5
Cluster
1
2
3
4
5
6
7
Donor
1
2
3
Location
airway
alveolar
E
Proportion of cells
40
30
20
10
1
2
3
4
6
7
airway BC
alveolar BC
Donor
1
2
3


A
Cluster
Donor
Location
DE in cluster
MACC1
SERPINB2
ATP12A
SPRR1B
CPA4
TMPRSS11D
SERPINB7
SULT1E1
CERS4
XACT
MYLK
COL12A1
FOXD1
OTX1
BST2
IFI27
HLA-DRB1
TFPI2
SLITRK6
PLIN2
CRIP1
IL33
LMO3
SPOCK3
HNMT
MXRA7
CDC42EP5
CITED4
BARX2
CHASERR
HMGB3
UBE2M
KRT15
S100A9
GJB6
HS3ST6
GSTA1
RHOF
EGLN3
S100A4
Cluster
1
2
3
4
5
6
7
Donor
1
2
3
Location
airway
alveolar
DE in cluster
1
2
3
4
5
6
7
Multiple
B
Cluster 2
-log10 p-value
log2 fold-change
Significant (1% FDR)
Up
Down
C
Cluster 6
-log10 p-value
log2 fold-change
Significant (1% FDR)
Up
Down
D
Cluster 2
HYPOXIA
GLYCOLYSIS
ANGIOGENESIS
KRAS SIGNALING UP
FATTY ACID METABOLISM
-log10 p-value
Average absolute log2 fold-change
Significant (10% FDR)
Up
Down
Number of genes
50
100
150
E
Cluster 6
HYPOXIA
ANGIOGENESIS
GLYCOLYSIS
KRAS SIGNALING UP
TNFA SIGNALING VIA NFKB
EPITHELIAL MESENCHYMAL TRANSITION
INFLAMMATORY RESPONSE
MYC TARGETS V2
COAGULATION
-log10 p-value
Average absolute log2 fold-change
Significant (10% FDR)
Up
Down
Number of genes
50
100
150
F
LMO3
S100A4
SPOCK3
HNMT
IFI27
BST2
-ΔCt (LMO3-GAPDH)
-ΔCt (S100A4-GAPDH)
-ΔCt (SPOCK3-GAPDH)
-ΔCt (HNMT-GAPDH)
-ΔCt (IFI27-GAPDH)
-ΔCt (BST2-GAPDH)
Alveolar BC
Airway BC

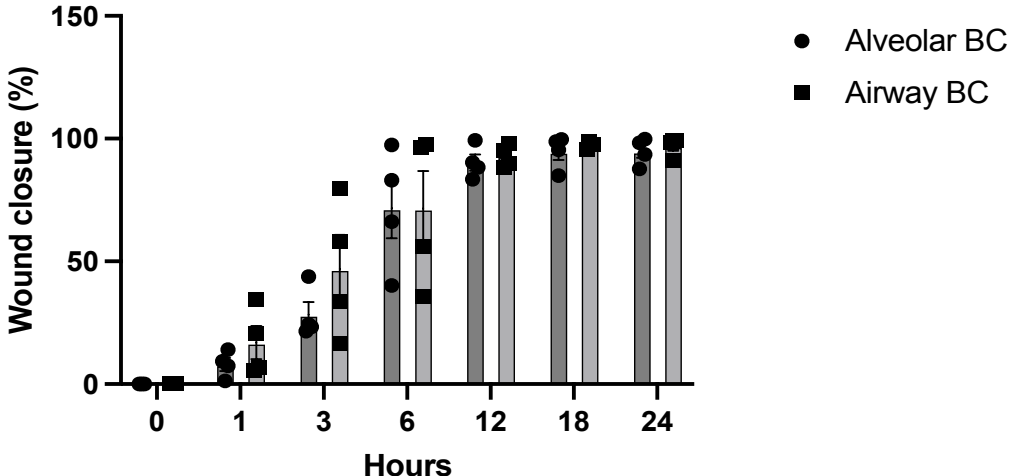
A

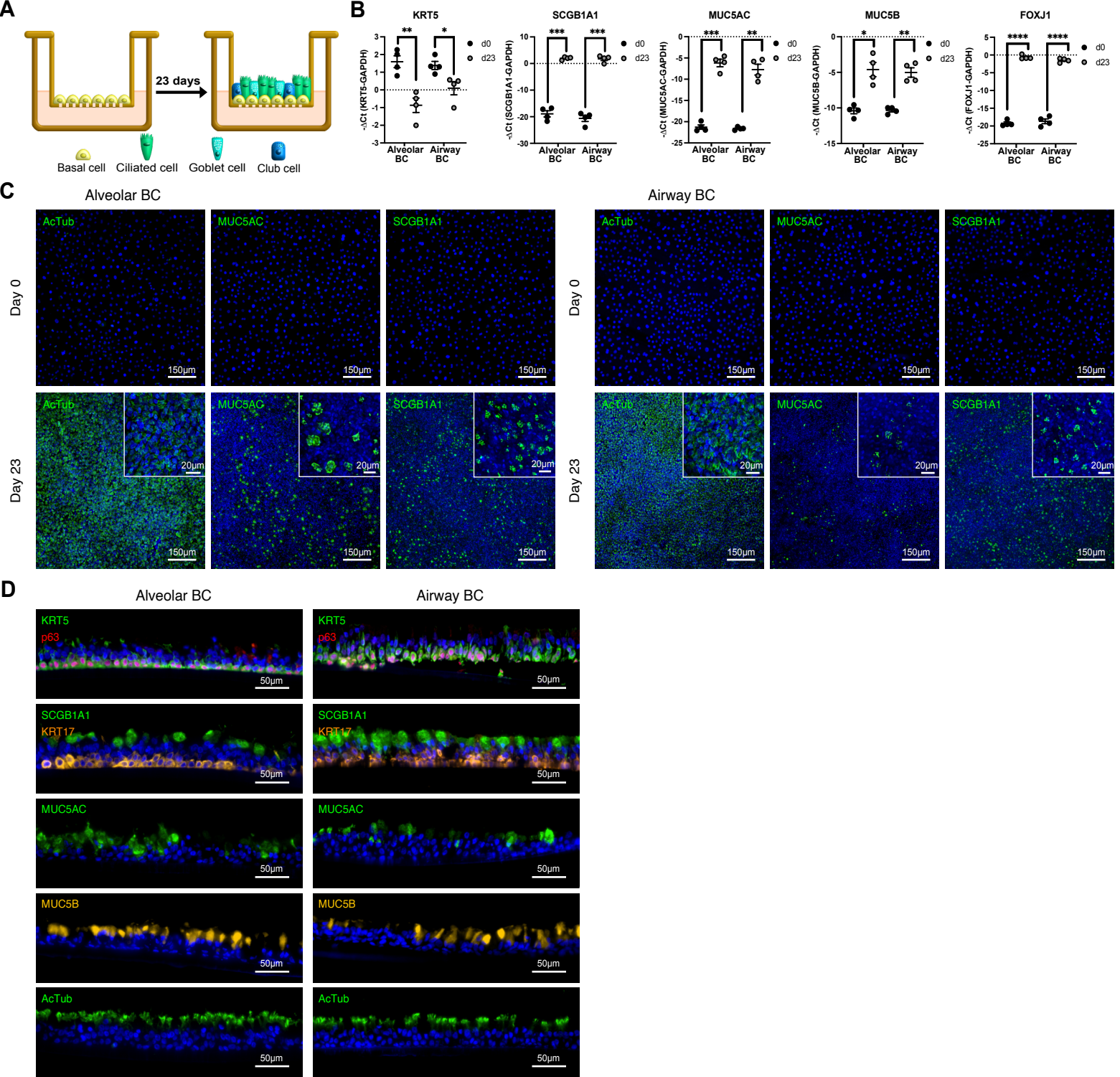

15
10
5
0
Cell number (x10⁴)
*
***
0 hours
48 hours
Alveolar BC
Airway BC

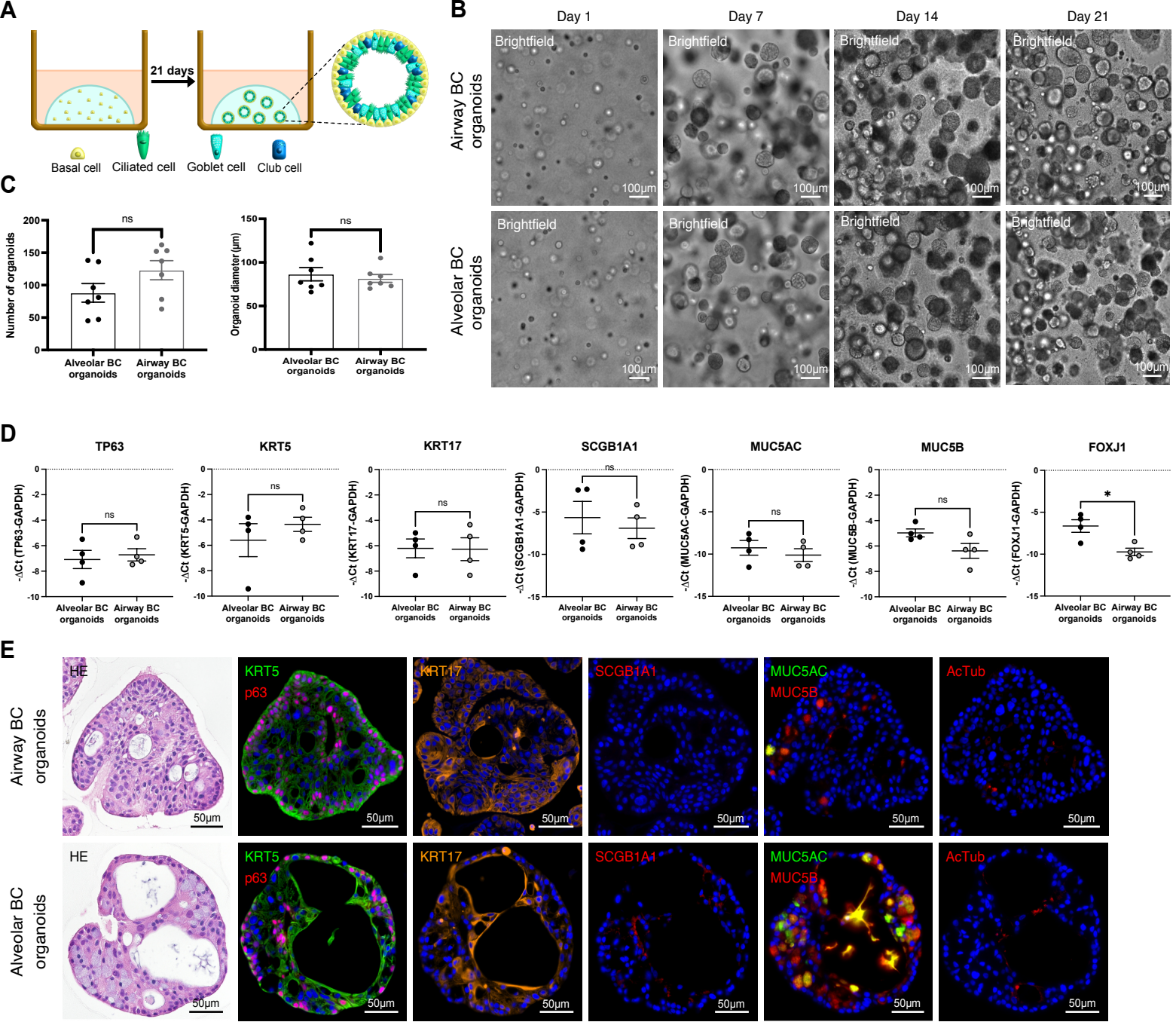
B

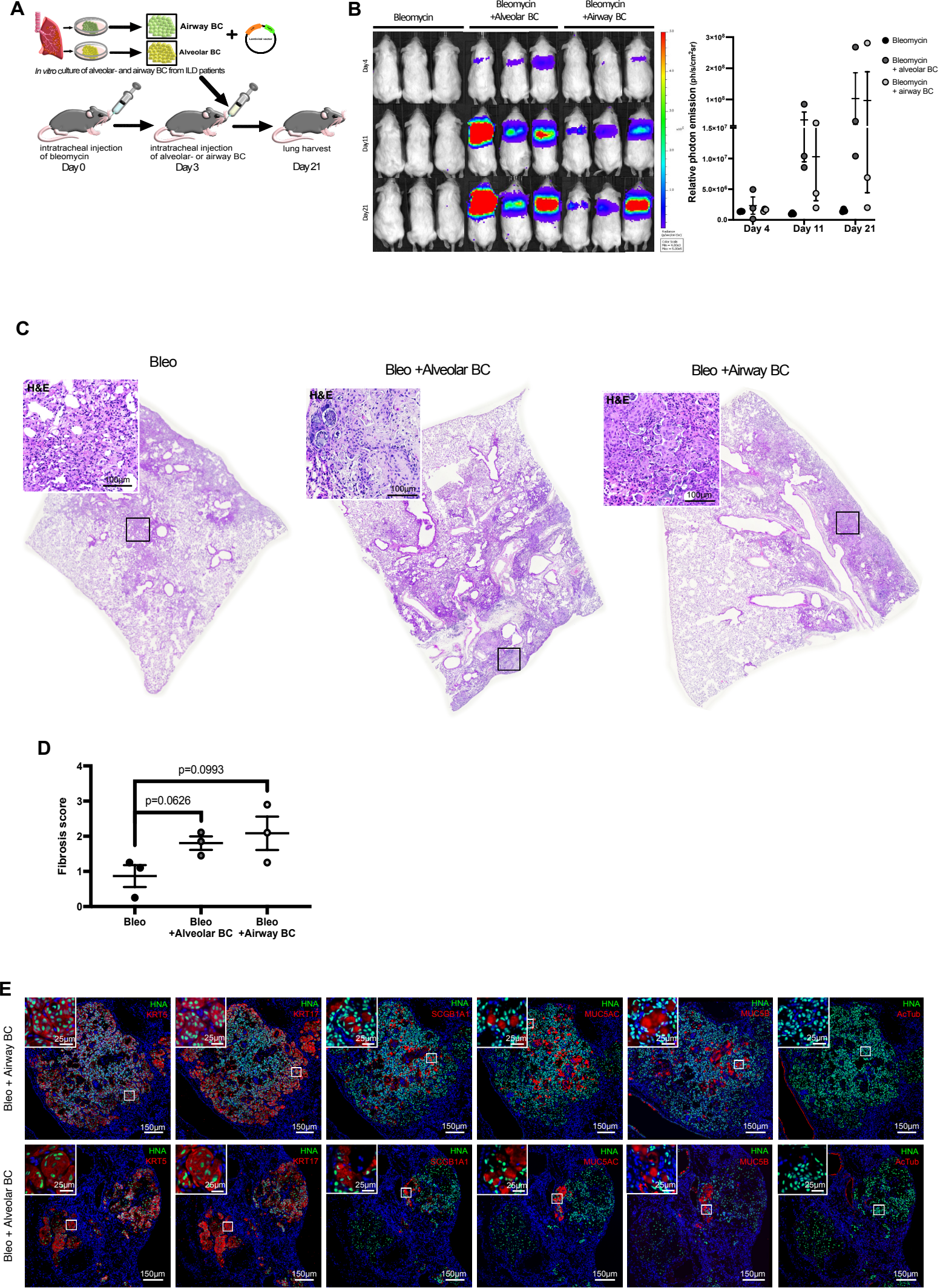

1 hour
3 hours
6 hours
12 hours
24 hours
Airway BC (px)
Alveolar BC
200 µm

C


150
100
50
0
Wound closure (%)
0
1
3
6
12
18
24
Hours
Alveolar BC
Airway BC


A
23 days
Basal cell
Ciliated cell
Goblet cell
Club cell
B
KRT5
SCGB1A1
MUC5AC
MUC5B
FOXJ1
-ΔCt (KRT5-GAPDH)
-ΔCt (SCGB1A1-GAPDH)
-ΔCt (MUC5AC-GAPDH)
-ΔCt (MUC5B-GAPDH)
-ΔCt (FOXJ1-GAPDH)
d0
d23
Alveolar BC
Airway BC
C
Alveolar BC
Airway BC
Day 0
Day 23
AcTub
MUC5AC
SCGB1A1
150μm
20μm
D
Alveolar BC
Airway BC
KRT5
p63
SCGB1A1
KRT17
MUC5AC
MUC5B
AcTub
50μm


A
21 days
Basal cell
Ciliated cell
Goblet cell
Club cell
B
Day 1
Day 7
Day 14
Day 21
Airway BC organoids
Alveolar BC organoids
Brightfield
100µm
C
Number of organoids
Organoid diameter (µm)
ns
Alveolar BC organoids
Airway BC organoids
D
TP63
KRT5
KRT17
SCGB1A1
MUC5AC
MUC5B
FOXJ1
-ΔCt (TP63-GAPDH)
-ΔCt (KRT5-GAPDH)
-ΔCt (KRT17-GAPDH)
-ΔCt (SCGB1A1-GAPDH)
-ΔCt (MUC5AC-GAPDH)
-ΔCt (MUC5B-GAPDH)
-ΔCt (FOXJ1-GAPDH)
*
E
HE
KRT5
p63
KRT17
SCGB1A1
MUC5AC
MUC5B
AcTub
50µm


A
Airway BC
Alveolar BC
Lentiviral vector
In vitro culture of alveolar- and airway BC from ILD patients
intratracheal injection of bleomycin
Day 0
intratracheal injection of alveolar- or airway BC
Day 3
lung harvest
Day 21
B
Bleomycin
Bleomycin +Alveolar BC
Bleomycin +Airway BC
Day 4
Day 11
Day 21
Relative photon emission (ph/s/cm²sr)
Bleomycin
Bleomycin + alveolar BC
Bleomycin + airway BC
C
Bleo
Bleo +Alveolar BC
Bleo +Airway BC
H&E
100µm
D
Fibrosis score
p=0.0993
p=0.0626
Bleo
Bleo +Alveolar BC
Bleo +Airway BC
E
Bleo + Airway BC
Bleo + Alveolar BC
HNA
KRT5
KRT17
SCGB1A1
MUC5AC
MUC5B
AcTub
25µm
150µm