Title

Postnatal FGF-signalling establishes gradients of cell-identities along the proximal-distal axis of the lung airways

Authors

Alexandros Sountoulidis[1,2]*, Alexandra Firsova[1,2#], Andreas Liontos[1,2#], Jonas Theelke[1,2#], Janine Koepke[3], Pamela Millar-Büchner[3], Louise Mannerås-Holm[4], Åsa Björklund[2], Athanasios Fysikopoulos[3], Konstantin Gaengel[5], Fredrik Bäckhed[4], Christer Betsholtz[5], Werner Seeger[3,6], Saverio Bellusci[3] and Christos Samakovlis[1,2,3,6]*

[1]Stockholm University, Molecular Biosciences, The Wenner-Gren Institute (MBW)

[2]Science for Life Laboratory, Stockholm, Sweden

[3]Justus-Liebig University of Giessen, Medical Clinic II, Department of Internal Medicine, Excellence Cluster Cardio-Pulmonary System (ECCPS), Giessen, Germany.

[4]The Wallenberg Laboratory, Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[5]Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Sweden

[6]Max Planck Institute for Cardiopulmonary Research, Bad Nauheim, Germany

* Corresponding authors # equal contribution (alphabetic order)

Abstract

The spatial organization and developmental control of cell maturation in the branched airway epithelium are only partially understood. Here, we labelled the major adult secretory epithelial cell-types and analysed them at single-cell resolution. We found opposing and partially overlapping graded gene-expression programs along the proximal-distal axis of the airways, on top of a general program that relates to detoxification. The first program is proximally elevated and relates to innate immunity and the second is enriched distally and relates to lipid metabolism, ion transport and antigen presentation. Cells at intermediate locations express lower levels of both programs and show increased levels of differentiation potency in vitro, suggesting a graded distribution of functionalities along the airway network. Lineage-tracing combined with single-cell RNA sequencing of airway secretory cells at four developmental stages showed that these gradients are established sequentially in common postnatal progenitors. Fgfr2b expression is distally enriched and its inactivation in postnatal airways leads to a reduction in the levels of distal gene-expression programs and an expansion of proximal genes into distally located airway cells. This suggests a central role of FGF-signalling in tissue-scale airway patterning.

Introduction

The airway epithelium functions as a pathogen barrier and as a seamless conduit of air in the alveoli, where gas-exchange takes place[1]. Various epithelial cell-types are distributed along the airways and accomplish distinct functions, creating the mucociliary escalator[2]. Secretory cells contribute to the maintenance of respiratory homeostasis by producing antimicrobial proteins like secretoglobins[3, 4], lactotransferin and lysozyme and mediating immune responses upon exposure to viruses, bacteria and allergens by secretion of cytokines and chemokines[5, 6]. They are also responsible for the the detoxification of inhaled xenobiotics[7]. Ciliated cells are also numerous, considered post-mitotic and terminally differentiated and are responsible for propelling secretions and xenobiotics out of the airways[8-10]. In homeostasis, secretory cells self-renew and produce ciliated cells[11, 12] and after injury they show reparatory potential even for the alveolar epithelium[2, 13].

Subsets of airway secretory cells, such as the variant (v-) club cells[14, 15] and the *Upk3a*^{pos} (u-) club cells[16] near to neuroendocrine cells (NE), the bronchioalveolar stem cells (BASCs) in Terminal Bronchioles (TBs)[15, 17, 18], a β4^{pos} CD200^{pos} Scgb1a1^{pos} cell fraction in distal airways[13], in addition to cell in an activated transitional state (ADI[19], DAPT[20] and PATS[21]) all contribute to tissue repair upon airway and alveolar injury. The existence of such a large variety of airway secretory cell-states also indicates that they are highly heterogeneous and/or capable of adopting distinct stem-cell characteristics. Several of the described cell types emerge postnatally but their spatial coordinates in the tissue and their relationships are not well understood.

The airway network is anatomically divided in the extra-lobar (trachea and bronchi) and intra-lobar compartments[22] with distinct cell-compositions. Basal cells, for example, are predominantly localized in the extra-lobar compartment of the mouse airways[23] and gradually decrease along the proximal-distal (PD) axis of the human intra-lobar airways[24]. Advances in single-cell RNA Sequencing (scRNA-Seq) resulted in the detailed characterization of the epithelial cell-heterogeneity in mouse and human extra-lobar airways, capturing the gene-expression profiles of many cell-types with single cell resolution[25-27]. It was only recently shown that specific secretoglobins, mucins and surfactant proteins are differentially expressed within distinct positional zones of human intra-lobar small airways[27], even though pioneering experiments have reported two decades before[3].

We focused on intra-lobar, airway secretory epithelium in homeostasis and during development and captured the gene-expression profiles and topology of resident cells. Our work shows a general airway secretory gene-expression program in addition to at least two complementary and partially overlapping ones. These programs are genetically determined, are established postnatally in an initial embryonic airway progenitor, relate to cell-topology along the lung PD-axis and define distinct functional characteristics of the cells. We show that

FGF-signalling is a major determinant of the distal airway gene-expression program, controlling genes involved surfactant biosynthesis, mitochondrial function, ribosomal biogenesis and extracellular matrix formation. Our study provides a detailed characterization of the lung epithelial heterogeneity in time and space, suggests topology-related gene-expression patterns and functions and provides mechanistic information about the dynamic nature of their regulation.

Results

Analysis of lung epithelial cell heterogeneity in adults

To analyse the main secretory cell-types in the mouse lung, we used a double transgenic reporter mouse line that expressing green fluorescence protein (GFP) in alveolar secretory epithelium (Sftpc-GFP)[28] and tamoxifen-inducible red fluorescence protein (RFP)[29] in the airway secretory cells[12] (Scgb1a1creER-Ai14). We also took advantage of Pdpn positivity of proximal airway secretory cells[30] to label and isolate them by antibody staining (Fig. 1A).

We induced recombination in adult mice and three days later, we FACS-sorted the labelled cells for droplet-based scRNA-Seq (Fig. 1B, Suppl. Fig. 1A). We generated and analysed 12030 high quality cDNA libraries (Suppl. Fig. 1B-C). The (Uniform Manifold Approximation and Projection) UMAP-plot [31] and marker genes of the cell clusters were in good agreement with the FACS-sorting criteria of the cells. We annotated clusters according to the positivity of known lung epithelial cell markers (Fig. 1C-D, Suppl. Fig. 1D-F, Suppl. Table 1). A small cluster (cl-4) contained Foxj1^{pos} ciliated cells deriving from the *Scgb1a1*creER-Ai14^{pos}/Pdpn^{neg} cell sorting fraction (99.8%) suggesting that the moderate *Scgb1a1* expression in ciliated cells was sufficient to induce recombination and labelling in the 72-hours interval after the Tamoxifen pulse. As expected, cluster-4 cells uniquely expressed a large number of related genes to cilium organization (GO:0044782 FDR:2.57*10^{-63}).

The *Scgb1a1*^{pos} airway secretory cells separated in three clusters. Cluster-5 contains mainly Scgb1a1-Ai14^{pos} Pdpn^{pos} sorted cells that also express high levels of the proximal airway cell markers *Scgb3a1* and *Scgb3a2*[3] (Suppl. Fig. 1F). We annotated these cells as S1. Cluster 1 cells cluster (cl) -1 was almost exclusively (99.2%) composed of Scgb1a1creER-Ai14^{pos} Pdpn^{neg} cells (Suppl. Fig. 1E), which were also positive for *Kdr*[32] and other previously reported epithelial cell markers. We annotated these cells as S2 (Suppl. Fig. 1F, Suppl. Table 1). Cluster-3 contained 92.4% of Scgb1a1creER-Ai14^{pos} Sftpc-GFP^{pos} double-positive (DP) cells. The rest of these cells were evenly distributed among the two alveolar secretory (AEC2) clusters -0 and -2 (Suppl. Fig. 1D-E). This suggested that *Scgb1a1* is also expressed by some AEC2s. We did not detect any unique markers for DP-cells, suggesting that they are in an intermediate cell-state between airway and alveolar secretory epithelium.

The clusters of alveolar secretory cells AEC2 (cl-0, AEC2a) and (cl-2, AEC2b) differed in *Lyz1* expression in the AEC2b cluster (Suppl. Fig. 1G), as also reported previously[33]. Both alveolar clusters (AEC2) expressed high levels of genes, involved in lipid biosynthesis (GO:0008610 FDR:$6.83*10^{-10}$) and lipid transport (GO:0006869 FDR:$1.9*10^{-5}$), in addition to genes implicated in the regulation of leukocyte activation (GO:0002694 FDR:$1.15*10^{-5}$), such as various MHC class-II genes involved in antigen presentation (Suppl. Table 2).

Overall, this analysis defined two airway secretory cell-types, a group of Scgb1a1^{pos} Sftpc^{pos} cells, two secretory alveolar cell-identities that differ in the expression of *Lyz1* and a cluster of Foxj1^{pos} ciliated cells.

Gene-expression patterns suggest distinct secretory cell functions

The continuous arrangement of the lung secretory cells from the S1 to the AEC2b cluster in the UMAP-embedding suggests intermediate expression levels of gene programs shared among the secretory epithelial cell-types. To exploit intermediate cell states, we used diffusion maps[34] and trajectory analysis (Fig. 2A), similar to the pseudotemporal cell-ordering along a developmental trajectory. We ordered equal numbers of randomly selected cells from each cluster and identified 1563 differentially expressed genes (DEGs) along the trajectory. These genes can be grouped into 10 stable modules (Fig. 2B, Suppl. Table 3). The aggregated gene expression scores confirmed groups of co-expressed genes, which are either gradually reduced from S1- to S2-cells (modules-5 and-2) or other gene programs with the opposite pattern (modules-4 and -1). Module-3 genes showed equal activation in both S1 and S2, but also in ciliated cells, representing a general airway gene-expression program. Modules -7 and -8 were enriched in S2-cells, but only module-7 was expressed in ciliated cells. Module-9 genes were enriched in all but the S1-cells and modules-10 and -6 in all but S2 cells (Suppl. Fig. 1I). This analysis reveals, shared, distinct and graded gene expression programs in the adult airways.

The Gene-ontology (GO) analysis of the modules, identified selectively enriched biological processes in each of the cell-types (Fig. 2C-D, Suppl. Table 3). For example, module-5 includes genes related to innate immunity regulation (GO:0002682 *Reg3g*[35], *Ltf, Bpifb1*[36] *P2rx4*[37], *Il13ra1*[38] and *Mfge8*[39]) and its expression is restricted to S1-cells, only. Module-2 genes are also highly expressed in S1-cells but are gradually decreased in cells of the S2-and DP-clusters. These genes encode various metabolic enzymes (e.g., *Gsta1*, *Acsl1* and *Gstm5*), cytokines (*Cxcl1*, *Cxcl2*, *Cxcl5* and *Cxcl17*) and interferon induced antiviral proteins (*Ifitm1*, *Ifitm3* and *Ifit1*), suggesting differential functions of secretory cell-types in responses to chemicals (GO:0070887) and pathogens (GO:0034097). This is also supported by previous functional analysis of few module- 2 genes. The ones encoding the transcription factors (TFs)

Six1[40] and *Spedf*[41] have been implicated in airway inflammation and *Irf7* is upregulated in airway secretory cells upon RSV infection[42].

Module-3 genes are uniformly expressed in S1- and S2- cell clusters and primarily encode metabolic enzymes (e.g., *Cyp2f2*, *Aldh1a1* and *Gsta3*), supporting the notion that a general function of airway secretory epithelium is to respond to xenobiotics (GO:0009410) and activate glutathione metabolism (GO:0006749) to detoxify inhaled air.

The cells in the S2-cluster selectively express module-7 genes, which relate to the general term cellular development (GO:0048869). This module contains developmental genes, such as the *Shh*[43, 44] and the negative regulators of airway inflammation *Kdr*, *Sema3e*, *Sema3a* and *Nr1d1*[32, 45-47] (Suppl. Table 3).

Cells in the DP- and in both AEC2 clusters show a selective and gradual upregulation of module-1 genes, which were decreased in the S2 and S1 clusters. Module-1 contains genes encoding kinases like the *Fgfr2* and *Lrrk2*, required for maintenance and function of the alveolar type 2 cells[48-51]. Additionally, *Atp8a1*, *Abca3* and the Rab-family genes *Rab27a* and *Rab34* relate to positive regulation of phospholipid transport (GO:0051050) and vesicle-trafficking.

Finally, the cells in the two alveolar epithelial clusters upregulate module-4 genes, which encompasses previously reported regulators of alveolar cell differentiation and maintenance, like *Etv5*[52] and *Nkx2-1*[53, 54], together with genes related to lipid biosynthesis (GO:0008610, *Scd1*, *Elovl1*) and transport across the plasma membrane (GO:0098739). Notably, the enriched expression of *H2-Eb1, H2-DMb1, H2-Dma, H2-Aa* and *H2-Ab1* relating to the MHC class-II complex assembly (GO:0002399), suggesting a selective role of alveolar secretory cells in antigen-presentation to immune cells.

This analysis identifies several gene programs that are shared but also differentially expressed among the secretory cell clusters. The intensity of these patterns along a continuum of cell-states suggests that prominent biological processes of the lung epithelium relating to immune responses, detoxification, lipid biosynthesis and ion transport, are segregated in the airway and alveolar compartments but are also expressed in a graded fashion along the secretory trajectory.

Spatial analysis of the secretory-cell trajectories in the airways

To further investigate the graded gene expression patterns along the secretory cell trajectory between S1 and AEC2b cell-states we focused on their spatial analysis in the airways. We selected a panel of 18 DEGs (Fig. 2E) and detected their transcripts in situ by SCRINSHOT[55]. We quantified the signals in 5906 manually segmented airway epithelial cell-ROIs in distinct airway positions, based on the stereotyped airway branching pattern of the left lung lobe[22] (Fig. 2F, Suppl. Fig. 2A). Hierarchical clustering of the 3096 secretory cell-ROIs showed high

expression of the module-5 genes, *Scgb3a1* and *Muc5b* in the proximal domains (P and I1), whereas module-1 genes, *Sftpb* and *Atp8a1* were more abundantly expressed in the distal domains I3 and D. Interestingly, cells in the intermediately located I2 domain co-expressed lower levels of both module-1 and module-5 markers (Fig. 2G), resembling the opposing graded expression of module-1 and module-5 genes along the secretory cell-cluster trajectory (Fig. 3A). This suggests that higher differentiated secretory cell states S1 and AEC2 expressing high levels of unique markers are located at the proximal and distal sites of each airway branch and cells in intermediate positions of each branch are in an intermediate cell-state with moderate expression levels of both S1 and AEC2 gene modules. To further test this hypothesis, we analysed protein expression levels by immunofluorescence co-stainings for the robust S1 and S2 markers Scgb3a1, Muc5b, Hp and Atp8a1 relatively to E-cadherin, which is homogeneously expressed (Suppl. Fig. 2B-C). Also, this analysis revealed protein expression gradients, where S1 markers were highest in the P domain and gradually reduced in the I (1-3) and D domains. The relative levels of the S2 markers Hp and ATP8a1 formed an opposite gradient, highest in the D-domain and gradually reducing towards the P-domain.

Many of the module-5 genes relate to innate immune responses and they are predominantly expressed in proximal airway S1-cells. We therefore examined if their expression patterns are set by the lung microbiome. Comparison of the relative protein expression levels of the four robust markers in adult lungs of specific-pathogen free and germ-free mice showed similar selective distribution levels. This indicates that the identified gene-expression programs are genetically specified (Fig. 3A-D).

The graded expression values in the spatial (Fig. 2G) and in the scRNA-Seq diffusion-map analyses (Fig. 2E) indicates that the trajectory of secretory cells recapitulates the PD-pattern of gene expression in the airway secretory epithelium (Fig 3E). The expression values of the 1563 genes in the 10 modules may provide a conceptual ruler for the airway secretory epithelial differentiation state. The unaltered expression peak of innate immunity proteins in proximal airway cells and their decline in distal cells of germ free-mice suggest that airway cell patterning may be influenced but is not dependent on the microbiome.

Spatial analyses of the Scgb1a1^{pos} Sftpc^{pos} cells

We further examined the topology-related gene-expression markers in rare secretory cell-identities like the double positive (DP) Scgb1a1^{pos} Sftpc^{pos} cells using a panel of 16 DEGs along the secretory trajectory, together with the neuroendocrine (NE) cell markers, *Ascl1* and *Calca*. This confirmed the previously reported localization of the DP-cells in BADJs and close to NEBs[56] (Fig. 3F, Suppl. Fig. 2D) showing that they are mainly localized at the airway terminal bronchioles (TB) (42.3%), in the alveolar compartment (12.5%) close to BADJs. A small fraction of DP-cells (7.1%) was close to NE-cells, within a 20 µm radius surrounding the

neuroepithelial body (NEB) borders. We found that the cell-ROIs in the TB-part of BADJs express higher levels of the S2- (module-3) than AEC2-enriched markers (module-1 and -4), whereas DP-cell-ROIs in the alveolar-part of BADJs and alveoli showed the opposite profile. The SCRINSHOT results at protein level with immunofluorescence (Suppl. Fig. 2E).

We conclude that DP-cells are detected in three distinct locations along the epithelial PD-axis, expressing high levels of S2 or AEC2 markers, depending on their position. The DP-cells in the vicinity of NEBs may correspond to v-club cells[14, 15], and were found more distantly from NEBs than the *Upk3a*^{pos} u-club cells[16] (Suppl. Fig. 2F). We conclude that even rare secretory cell types express graded levels of S2 and AT2 cell markers depending on their position in the airway tree. Their dispersal and distinct identities may reflect unique signals from the NEB-microenvironment, such as the Notch-signalling through Dll-ligands[57].

Differentiation potential along airway epithelial trajectory

The graded activation levels of distinct gene-expression modules are expected to create a continuum of cell-states, along the airway length, with possible functional differences in cell properties like proliferation and differentiation potential in the proximal intermediate and distal regions. As a first test of this hypothesis, we isolated cells from double reporter mice (as in Suppl. Fig.1A) and cultured them in Matrigel together with cells from a mouse lung fibroblast cell-line (MLg-2908), as described before[58, 59]. The proximal-domain cells correspond to the sorted Scgb1a1creER-Ai14^{pos} Pdpn^{pos} fraction, the Scgb1a1creER-Ai14^{pos} GFP^{pos} (DP) to the airway distal end and the Scgb1a1creER-Ai14^{pos} Pdpn^{neg} cells derive from intermediate and distal airways. We initially assessed cell clonogenic potential and found that the two Scgb1a1creER-Ai14^{pos} GFP^{neg} cell fractions were more potent than DP-cells under our culture conditions (Fig. 4 A).

Our cultured cells produced three colony-types as previously described[58, 60]. The large cystic colonies (type-A: bronchiolar) expressed the airway secretory markers Scgb3a1, Muc5b, Hp and Scgb1a1 in addition to the ciliated marker acetylated tubulin[61] and the basal cell markers Krt5 and Pdpn[62]. The dense colonies (type-C: alveolar) were positive for the alveolar markers Sftpc and Ager and the type-B (bronchioalveolar) colonies showed mixed morphology and expression of both bronchiolar and alveolar markers. The Pdpn^{neg} cells produced predominantly bronchioalveolar and alveolar colonies, in contrast to the mainly bronchiolar of Pdpn^{pos} and the exclusively alveolar of GFP^{pos} cells (Fig. 4B-E).

Considering the clonogenic nature of the colonies[58], the increased potency of Pdpn^{neg} cells to produce bronchioalveolar colonies that contain positive cells for all the analysed airway secretory, ciliated, basal and alveolar markers (Fig. 4A) indicates a heterogeneous cell population in intermediate and distal airways with higher differentiation potential, than those of proximal airways and DP-cells, at least under the specific in vitro co-culture conditions. Our

results also indicate that the airway secretory epithelial cells have different properties that relate to their tissue topology and that they retain and transmit them to their progeny in vitro, similarly to human basal cells[27].

Airway epithelial cells mature postnatally

To identify lineage relationships within the airway secretory cell-states and analyse when the observed gene-expression patterns become established, we induced the labelling of embryonic Scgb1a1^{pos} cells with a farnesylated GFP variant (Scgb1a1creER-fGFP)[12] and lineage-traced them. We induced recombination at the onset of Scgb1a1 expression[12], on embryonic day (E) 16 and FACS-sorted labelled cells and their progeny in different perinatal [E19.5 and postnatal day (P) 2] and postnatal stages (P21 and P60) for full-length scRNA-Seq[63] (Fig. 5A).

We analysed 354 high quality cells (Suppl. Fig. 3A) using diffusion maps and trajectory analysis and found that four distinct trajectories stem from the immature cells of cl-0, which contains E19.5 and P2 cells (Fig. 5B, Suppl. Fig. 3B). The trajectories end in mature P21 and P60 cells that correspond to the DP (cl-3), S2 (cl-1), S1 (cl-2) and ciliated cells (cl-4), as seen also in the adult-cell dataset.

Differential expression analyses between the clusters confirmed the perinatal expression of *Upk3a* and *Krt15*[35]. We also found upregulation of *Cccnd2*, which becomes induced after perinatal hyperoxia injury[64] and adult lung pneumonectomy[65], in addition to the WNT-signalling mediator *Fzd1*, the autophagy regulator *Itm2a*[66] and the AEC marker *Ager*[67] (Fig. 5C). The enriched gene sets in the mature cell-clusters are largely overlapping with those of the adult dataset, as expected (Suppl. Fig. 3C, Suppl. Table 5).

Next, we used GO-analysis to identify enriched biological processes in the clusters (Suppl. Fig. 3D and Suppl. Table 6) and scored the cells along the trajectories based on the expression of the corresponding genes. We found that the immature cells express high levels of ribosomal genes, indicating high translation (GO:0006412) and ribosome biogenesis, in the perinatal airway epithelium, which gradually decrease as the cells mature (Fig. 5D, Suppl. Fig. 3D, Suppl. Table 5-6).

The mature airway secretory gene modules, encoding detoxification, oxidative stress responses, xenobiotic and lipid metabolism were gradually established along the S1 and S2 trajectories in the domains that exclusively contain P21 and P60 cells, but not in DP-cells (Fig. 5D). This is evident by the expression levels of genes encoding metabolic enzymes, such as the *Aldh1a1*, *Fmo2* and *Gsta3* (Suppl. Fig. 3E). Interestingly, representative genes of the immune system process (GO:0002682) like*Scgb3a1*, *Tff2* and *Muc5b* reached high expression only at the end of the S1-trajectory (Suppl. Fig. 3F). This suggests either slower establishment of that gene-expression program or that mature cells from intermediate and

distal domains are erroneously placed along that trajectory, interfering with the results. To test this, we also analysed the cells according to their actual developmental stage and found that the P60 cells have generally higher expression levels e than the ones isolated at P21 (Suppl. Fig. 3G). The cells along the DP-trajectory gradually increase their ability for lymphocyte mediated immunity (GO:0002449), upregulating the *Cd74*, *Ctsc*, *Hc*, *Emp2*, *H2-Aa* and *H2-Ab1*. The middle part of that trajectory contains perinatal cells that likely contribute to the extracellular matrix (ECM) organization (GO:0030198), expressing high levels of genes like the *Col4a2*, *Spock2* and *Matn4* (Suppl. Fig. 3H).

In summary, we showed that heterogenous populations of adult airway secretory cells derive from a secretory Scgb1a1^{pos} progenitor. Differentiating cells mature postnatally, acquiring their functional characteristics during the first three weeks after birth. The perinatal airway epithelium shows high ribosomal biogenesis, which gradually decreases over time. The gene programs of innate immunity (S1-trajectory) are established later than those involved in xenobiotic metabolism and reduction of reactive lipid aldehydes (S1- and S2- trajectories), suggesting that the gene-expression programs are activated sequentially. In the developing distal lung, the differentiating DP-cells transiently contribute to ECM composition and gradually acquire the expression of antigen presentation genes, which are also expressed by the adult AEC2s (Fig. 5E).

Fgfr2 inactivation affects distal airway cell maturation and differentiation

Fibroblast growth factor (FGF) signalling is crucial for epithelial branching[68] and is later required for AEC2 differentiation and maintenance[48-50, 69-71]. Our gene expression analysis showed that *Fgfr2* is also expressed in the adult epithelial cells belonging to gene module-1, which shows high levels in AEC2 and DP-cells and gradually decreases in S2 and S1 airway cells (Suppl. Table 3, Suppl. Fig. 4A). We also detected *Fgfr2* expression in the perinatal airway secretory cells (Suppl. Fig. 4B) and confirmed its localized expression pattern at protein level by immunofluorescence (Suppl. Fig. 4C) in P2 lung sections. Co-stainings with an Fgfr2β(IIIb)-Fc chimeric protein to detect the spatial distribution of Fgfr2 ligands and used Fgfr2 antibodies showed a punctate staining for the ligand, which was higher around the TBs and distal epithelial cells compared to regions at proximal airways. This suggested stronger pathway activation in the distal airway regions[71] (Suppl. Fig. 4C).

To examine if Fgfr2-mediated signalling has any role in the postnatal establishment of the identified gene-expression gradients along the airway epithelium, we deactivated the receptor in Scgb1a1^{pos} cells, just after birth. We induced tamoxifen-mediated Fgfr2-inactivation[72] in the airway secretory epithelium[12] (Scgb1a1creER-Fgfr2KO) and utilized Rosa26-locus tdTomato expression[29] to detect the recombination and presumably mutant cells. After three Tamoxifen injections (P1-P3), we analysed the lungs at P7 by scRNA-Seq and histology (Fig. 6A). We

clustered and annotated 9911 high-quality scRNA-Seq libraries from Epcam^{pos} cells from wildtype (lirbrary-1), Epcam^{pos} mutant (library-2) lungs and Epcam^{pos} RFP^{pos} cells from mutant lungs (library-3). The UMAP-plot was consistent with the FACs-sorting cell fractionation and showed that *Fgfr2* inactivation affected only the positioning of S2-cells (cl-2) in the UMAP-embedding (Suppl. Fig. 4D-G). The alveolar (cl-0, -1, -5), basal (cl-7), NE (cl-8), ciliated (cl-6) and S1 (cl-4) clusters were composed of both wildtype and mutant intermingled cells, indicating that there is no significant effect of *Fgfr2* inactivation in cells that express undetectable levels of the receptor. We detected *Fgfr2*^{pos} cells within the *RFP*^{pos} libraries (Suppl. Fig. 4E) suggested that there were escaper cells, which recombined the Rosa26R-Ai14 (RFP) but failed to deactivate both Fgfr2 alleles. We validated that at protein level (Suppl. Fig. 4H) and filtered out the *Fgfr2*^{pos} S1- and S2-cells of library-3 for down-stream analyses to reduce noise in the scRNA-Seq analysis.

Re-clustering of equal airway secretory numbers of cells form wildtype (lib-1) and mutant (lib-3) libraries revealed eight clusters corresponding to S1-, S2- and DP-cells according to the DEGs (Fig. 6B, Suppl. Fig. 5A-C). Cluster correlation according to the aggregated scores of their statistically-significant DEGs (Suppl. Table 7B) showed that the clusters are grouped according to their genotype (Suppl. Fig. 6D), indicating significant gene-expression differences in S2 and DP clusters because of Fgfr2-inactivation. Importantly, there was no clear pairwise correlation between the wildtype and mutant clusters. To avoid comparing irrelevant cell-identities, we compared the wildtype and mutant airway secretory cells, regardless clustering and identified 240 statistically significant DEGs, which we categorized based on previous knowledge and GO-analysis (Suppl. Table 7C-G, Suppl. Table 8, Fig. 6C). We also related the expression levels of these genes with their levels in the lineage-tracing (Suppl. Fig. 5E) and in the Fgfr2-inactivation (Suppl. Fig. 5F) experiments. We included all cell-types regardless genotype to identify possible temporal defects in cell maturation and in cell-type differentiation, respectively.

We found a prominent reduction of AEC2 genes, like the *Sftpc*, *Napsa* and *Cd74* in addition to increase of the AEC1 TF *Hopx* [73] (Suppl. Fig. 5A). Mutant cells also upregulated a large number of nuclear encoded mitochondrial genes of the complexes I, III, IV and V in addition to sodium channel genes (*Commd3*, *Scnn1a*, *Scnn1b*[74]) and others that relate to autophagy and vesicle trafficking, like the *Creg1*[75], *Vamp8*, *Vamp5*[76], *Rab25* and *Rabac*, which are all normally enriched in P42 AEC1s[77] (Suppl. Fig. 5E and Suppl. Fig. 6A-C). The senescence related gene *Cdkn1a* (p21)[78, 79] and survival proteins *Bax* and *Bag*1[80] are also enriched in AEC1s and up-regulated in mutant airways secretory cells. Interestingly, *Hopx*, *Vamp5, Creg1* and *Scnn1b* are also detected in adult wild-type S2 cells, indicating their propensity to further activation of the AEC1 genes (Suppl. Fig. 6D). This suggests that Fgfr2 up-regulates AEC2-related genes that are responsible for surfactant biosynthesis and lamellar body formation

(Fig. 6D) and down-regulates numerous AEC1 genes, that relate to mitochondrial function, ion homeostasis, vesicle trafficking and cell-survival.

Fgfr2-inactivation in the airways also caused increased expression of the ECM protein encoding genes *Eln*, *Mgp* and *Mfap4*, which are transiently expressed along the DP-cell trajectory (Suppl. Fig. 5E) and in developing alveolar epithelium (Suppl. Fig. 6C) but become down-regulated in later stages. Also, mutant cells failed to downregulate the expression of numerous ribosomal-subunit genes which are highly expressed in all immature lung epithelial cells and gradually decrease as maturation proceeds (Suppl. Fig. 5E, 6C). These findings suggest that Fgfr2 is continuously required for the normal maturation process of distal airway secretory epithelium controlling the temporal ribosomal biogenesis and ECM composition.

Fgfr2-inactivation caused expression changes in many genes that are involved in lipid synthesis and lipid droplet formation (*Mboat1*[81], *Acsm1*[82], *Ces1d*[83]), lipid trafficking (*Fabp5*[84], *Lpcat1*[85], *Pi4k2b*[81]) and adipogenesis (*Cebpb*[86], *Noct*[87], *Klf3*[88], *Arxes2*[83]). These results indicate altered lipid and fatty acid biosynthesis and metabolism that could be also supported by the up-regulated enzymes *Taldo1* and *Tkt*[89] that participate in the pentose phosphate pathway for NADPH production.

The most significantly reduced TF in mutant S2-cells is *Myc*, which is highly expressed in normal immature airway secretory cells and becomes downregulated in mature S1 and S2 cells (Suppl. Fig. 5E). In the nasal epithelium *Myc* together with *Ccn1* and *Ccnd1* induce cell proliferation[90]. This suggests that mutant cells are defective in cell-cycle regulation (Fig. 6F, Suppl. Tables 7). Other significantly changed TFs are the down-regulated *Atf4* and *Ets1* and the up-regulated *Cebpb*. *Myc* induces *Atf4* expression in cancer cells and they co-operatively regulate the promoters of various target genes, including *Cebpb*[91]. In return, *Cebpb* directly binds on *Myc* promoter and inhibits its expression, suggesting a negative feedback loop[92]. The reduced expression of both *Atf4* and *Ets1* likely reduces *Slc7a11*, affecting the detoxification ability of the cells by compromising their ability to exchange intracellular glutamate with extracellular cystine for glutathione synthesis[93]. This is also supported by the reduced expression of *Gclc* a rate-limiting enzyme of the first step of glutathione biosynthesis[94]. The reduced glutathione levels might explain the up-regulation of other detoxification-related genes, like the *Sod1, Aldha1*, *Gsta3* and *Gstm1* as a compensatory mechanism to protect the cells from reactive oxygen species and lipid peroxidation.

Additionally, Fgfr2-inactivation formed an intermediate S1/S2 (cl-4), which expressed moderate levels of S1-markers, such as the *Scgb3a1*, *Bpifb1* and *Reg3g* (Fig. 6F). This suggested spatial expansion of the S1-related gene-expression program, that was confirmed by the increase of Scgb3a1 protein signal in distal airway epithelium (Fig. 6G).

In summary, perinatal Fgfr2-inactivation in airway secretory epithelium halts its maturation and differentiation, causing extensive gene-expression and patterning changes. The extension of

the S1 gene-expression program in distal airways, down-regulation of AEC2 genes in parallel to up-regulation of AEC1 ones suggests that FGFs are a major signal, setting the gene expression differentiation gradients Mutant cells failed to reduce perinatally expressed ribosomal-subunit and ECM-genes suggesting a temporal role of FGF signaling in the progression of differentiation (Fig. 7).

Discussion

Here, we focused on the intra-lobar, airway secretory cells of the mouse lung to provide a detailed description of cell-states and their gene-expression profiles. Using trajectory analyses with cells from both adult and developing lungs, we showed that there is no abrupt zonation pattern along the proximal-distal axis of the airway network. Our results indicate that the characteristics of the airway epithelial cells, as indicated by their gene-expression profile, are defined by a uniformly activated program of co-expressed genes that relates to cell-responses to xenobiotics and two complementary and partially overlapping graded programs, encoding innate immunity, cytokine production and response to cytokines in proximal airways and lipid synthesis and surfactant production in the distal ones (Figure 7A). This topology-related differential activation of the identified programs is likely genetically determined and becomes gradually established in the postnatal lung, in an asynchronous manner with the lipid and xenobiotic metabolic functions being established first. Graded expression patterns of mainly secreted proteins, have been recently shown in the human lung airway epithelium[27, 95, 96], even though their spatial validation is limited to the analysis of small biopsies from distinct positions, because of the large size of the organ that prohibits the in situ detection of graded gene-expression patterns.

Our analysis allowed the identification MAPK-signalling between the pathways that are gradually activated in secretory cells, along proximal to distal airways and alveoli (Fig. 2C). Its compromised activation by Fgfr2-inactivation in airway secretory epithelium (Fig. 7B) affected cell differentiation, showing spatial extension of the S1-related gene-expression program, in agreement with the previously reported Muc5b expression in adult Fgfr2-mutant distal airway secretory cells upon naphthalene-induced injury[97]. Fgfr2-inactivation caused down-regulation of involved genes in surfactant production and endosome formation, which normally are highly activated in AEC2s but are also expressed along the airways in moderate levels. In parallel, we detected activation of numerous AEC1-enriched genes, including several nuclear encoded mitochondrial genes in addition to others that are involved in autophagy and vesicle trafficking. This phenotype agrees with the perinatal Fgfr2-inactivation in AEC2s that reprograms them to AEC1s[49]. Thus, Fgfr2 is required for activation of the surfactant-related genes not only in alveolar but also airway epithelium and its deactivation is sufficient to induce various aspects of the AEC1 gene-expression program in both secretory cell-types.

Interestingly, the activation of the AEC1-enriched genes has been also reported after bleomycin-induced lung-injury in mouse airway secretory cells[19] and AEC2s[20, 21]. AEC2s of COPD-patient lungs also up-regulate AEC1-genes including several proteins of the mitochondrial complexes I, III, IV and V[98] and *Cebpb,* the most up-regulated TF upon Fgfr2-inactivation in our study, is enriched in distal airways of advanced-COPD patients[99]. These striking similarities between the Fgfr2-inactivation phenotype and the activated cell-states could suggest a causative role of compromised Fgfr2-mediated RTK-signalling in lung pathology not only in mouse models[71, 100] but also in human. The involvement of RTK-signalling in proper patterning of the airway epithelium has been also indicated by the role of Kdr in the patterning of the S1 gene-expression program in the proximal part of the airway network[32]. That also suggests that qualitative (distinct ligands and receptors) and quantitative (signal strength and duration) differences in that pathway-activation results to different cell properties.

The lineage tracing analysis of the airway secretory cells and the meta-analysis of the enriched for alveolar-epithelial cells dataset GSE149563[77] (Suppl. Fig. 5E-6C) indicate that there is a perinatal gene-expression increase of ribosomal proteins, which is normally down-regulated in the mature cells. The fact that Fgfr2-mutant cells failed to down-regulate these genes likely indicates impaired maturation. To our knowledge impaired ribosomal biogenesis has been reported only to various cancer types[101] but no other lung diseases. The correlation of the Fgfr2-inactivation dataset with the other two also showed that the developing distal-lung epithelial DP-cells, AEC2s and AEC1s normally express ECM protein genes, which are supressed in the mature lung. Importantly, Fgfr2-inactivation caused ectopic expression of these genes in S1 and S2 cells and up-regulation in DP-cells suggesting problems in both cell differentiation and maturation. The Fgfr2-inactivation in adult AEC2s causes increased collagen deposition in alveoli, accompanied by a significant reduction of the number of mutant cells[71] and a subsequent compensatory regeneration by the remaining AEC2s that escape Fgfr2-inactivation[49]. In our study we did not show signs of cell-death, such as exfoliated cells and/or with condensed and fragmented nuclei or apoptosis-induced genes in the scRNA-Seq analysis, suggesting that perinatal airway epithelial cells tolerate Fgfr2-inactivation similarly to the adult airway[97] and perinatal alveolar epithelium[49].

Here, we used mouse genetics, scRNA-Seq and spatially-resolved transcriptomics to systematically analyse and correlate gene-expression patterns with topology, using lung sections with large parts of the airway network, that allowed consistent identification and comparison of similar airway domains. Following that approach, allowed us to provide evidence that the lung secretory cell-types and states result from the differential activation of specific gene-expression programs that relate to cell-topology with both airway secretory and DP-cells following that rule. This concept of activation of gene-expression programs on

permissive cell-types (e.g., ability of Nkx2-1 to regulate both AEC2 and AEC1 gene promoters[54]) could explain the formation of previously reported lung cell-states like the SFTPB^{pos} SCGB3A2^{pos} and SFTPB^{pos} SCGB3A2^{pos} SCGB1A1^{pos} in distal human airways[27] and the proximal-airway and general ciliated cells[102]. It might also be medically relevant providing candidates that drive phenotypic changes in various cell-types, like the COPD AEC2s that activate aspects of AEC1 gene-expression program[98] and the ciliated cells in asthma that activate a mucus production program (mucus-ciliated)[103].

The proposed concept of gene-expression patterns is based on the identification of genes that are co-expressed in the same cell-state and change in the same way in a topology and maturation-state manner. This does not mean that the expression of these genes is governed by the same regulatory mechanisms. Computational tools that predict gene regulatory networks (GRN) mainly focus on positively co-regulated genes[104] and use previous knowledge for TF-motif binding sites, based on different and possibly irrelevant cell-types. This highlights the need for better approaches that will facilitate the efficient definition of gene co-regulation programs in a cell-type specific manner. Another challenging task in our study was the ordering of the cells along trajectories that are able to distinguish the different maturation states and tissue topologies of the analysed scRNA-Seq datasets that are both biology-relevant components. More advanced, future computational tools and datasets with tighter sampling and higher number of cells might provide a better overview of the cell maturation timeline integrating in a more efficient way the spatial-component. Another important challenge for the systematic correlation of our results with previous knowledge was the absence of a common definition of “proximal, intermediate and distal” airway domains even if the branching patterning in the mouse lung has been described almost 15 years ago[22]. Considering that there is a topology-related cell heterogeneity along the PD-axis of the airway network, it is of great importance to agree for a consensus definition for future studies.

Materials and Methods

Animal models and Tamoxifen administration

All mouse experiments were performed according to Swedish animal welfare legislation and German federal ethical guidelines. The Northern Stockholm Animal Ethics Committee approved the project (Ethical Permit numbers N254/2014 and 15196-2018). The Research Animal Ethics Committee in Gothenburg approved the analyses of germfree (GF) mice (Ethical Permit number 4805-23). The GF mice were maintained in flexible film isolators (Class Biologically Clean, Madison, WI, USA). GF status was monitored regularly by aerobic and anaerobic culturing as well as by PCR for bacterial 16S rRNA. All mice were group housed in a controlled environment (room temperature of 22 ± 2 °C, 12 h daylight cycle, lights off at 7 pm), with free access to autoclaved chow diet (#T.2019S; Envigo) and water. All mouse

Breedings and experiments performed in Germany were under the Ethical Permit with number GI 20/10, Nr. G 21/2017. For the lineage-tracing experiments, we used Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-fGFP$^{pos/neg}$ [12] mice. Noon of the day of the vaginal plug was considered as embryonic day (E) 0.5. Tamoxifen-induced recombination was done with one oral dose (gavage) of Tamoxifen solution in corn oil (30mg/kg body weight) on E16.5, as described previously[12]. For the analysis of adult-lung epithelial heterogeneity and organoid cultures, we used Scgb1a1-CreER^{het};Rosa26-Ai14^{het};Sftpc-fGFP^{het} adult mice[28, 29] and administered one Tamoxifen dose (100mg/kg body weight), 72 hours prior tissue collection. Experiments for Fgfr2-inactivation were performed using Scgb1a1-CreER$^{T2\ pos/neg}$;RosaAi14$^{pos/pos}$;Fgfr2b$^{fl/fl}$ and Scgb1a1-CreER$^{T2\ neg/neg}$;RosaAi14$^{pos/pos}$;Fgfr2b$^{fl/fl}$ mice. Tamoxifen-induced recombination was performed subcutaneously (87 mg/ kg body weight) on P1, P2 and P3, aiming for more efficient recombination than one pulse.

Tissue collection

Animals were euthanized by an intraperitoneal injection of anaesthesia overdose (ketamine 200mg/kg and xylazene 13.3mg/kg), followed by incision of the abdominal vein. For embryonic lungs, we did not perform heart perfusion. For postnatal lungs, the chest was opened and the left atrium was excised. Lungs were perfused through the right ventricle of the heart with ice-cold PBS 1X pH7.4, using a 26G needle and 5ml syringe until they became white. Lungs were inflated with a mixture of 4% PFA:OCT (2:1 v/v), using a 20G catheter (Braun, 4251130-01), until accessory lobe was expanded. Trachea was ligated (silk 5/0 Vömel, 14739) and tissues were later fixed. For histological analysis, tissues were collected on E16.5 (without Tamoxifen administration), on E19.5, on post-natal day 2 (P2), 5 (P5), 21 (P21) and 60 (P60). Embryonic and P2 tissues were fixed with freshly prepared 4% Paraformaldehyde (Merck, 104005) solution in PBS 1X pH7.4 (Ambion, AM9625) for 4 hours. Later stages were fixed for 8 hours. Thereafter, the tissues were placed in OCT: 30% sucrose in PBS (2:1 v/v) over-night (O/N) at 4°C with gentle shaking and frozen in OCT (Leica Surgipath, FSC22), using plastic molds (Leica Surgipath, 3803025), by placing them in isopentane and dry ice. Tissue-OCT blocks were kept at -80°C until sectioning.

Tissue dissociation and cell isolation

For full-length (Smart-Seq2) library preparation[63], the left lungs were used for enzymatic digestion and the right lungs were treated as described above for histological analysis. For cell culture and droplet based scRNA-Seq, both lungs were processed for digestion. Briefly, we cut the lungs in small pieces using a razor blade and digested them with elastase (Worthington, LS002292) and DNase I 0.5 mg/ml (Sigma-Aldrich, DN25) in HBSS (Gibco, 14175), at 37°C for 1 hour with rotation. Equal volume of HBSS++ [HBSS (Gibco, 14175),

supplemented with 2% 0.2µm filtered FCS (Gibco, 10500), 0.1 M HEPES (Sigma-Aldrich, H0887), antibiotics (Gibco, 15240096) and EGTA 2 mM] was added and the suspension was mixed gently. Then, the cells were centrifuged at 800g for 10 minutes at 4ºC. Supernatant was removed with a serological pipette and cells were resuspended in HBSS++. Viability was tested using trypan blue (Sigma-Aldrich, T8154) (1:1 dilution) and the presence of fluorescence positive cells was evaluated using a normal fluorescence microscope. Just before sorting, cells were passed through a 100 µm BD Filcon (BD Biosciences, 340610) in order to remove cell aggregates.

For E16.5 lungs, after digestion, centrifuge and resuspension in HBSS++, cells were passed through a 100 µm BD Filcon (BD Biosciences, 340610) and counted. We resuspended them in HBSS++ to obtain $20x10^6$ cells/ml cell concentration. 100 µl of cell suspension were stained with 0.5µl of anti-EpCam-PE antibody (Biolegend, 118205) and 0.5µl anti-CD45-APC antibody (Biolegend, 103112). Replicates of the above reactions were set in separate tubes; in order to prevent aggregate formation, which is typical when large number of epithelial cells are centrifuged.

For cell-sorting, we used a BD FACSARIA III cell-sorter, with 100 µm nozzle using single-cell sorting purity. For E16.5 lungs, $Epcam^{pos}$ $CD45^{neg}$ cells were sorted. Cells from all other stages were isolated according to GFP and/or Tomato (for Rosa-Ai14 mice) expression. Non-transgenic and single-transgene (either Scgb1a1-CreER;Rosa-Ai14 or Sftpc-GFP) positive animals were used for instrument calibration. For cell culture experiments, cells were sorted in HBSS++ medium and for droplet-based sequencing, we omitted EGTA and HEPES, according to 10x Genomics instructions.

Droplet based scRNA-Seq of adult lung cells

Droplet based scRNA-Seq was carried out with Chromium Next GEN Single Cell 3' Kit version 3 (10x Genomics), at the Eukaryotic Single Cell Genomics Facility at SciLifeLab, Sweden. The samples were processed with cellranger-4.0.0 pipeline (10x Genomics). The reads were mapped to a custom mouse (GRCm38) reference genome that contained GFP and Ai14 cassette (RFP and WPRE sequences) sequences. The reference genome was created with the 10x Genomics “cellranger mkref” and the mapping of the reads was done with the “cellranger count” function, using default settings.

Droplet based scRNA-Seq analysis of adult lung cells

For the analysis of the droplet based scRNA-Seq dataset, we initially applied filtering criteria to filter out low quality cells and contaminants (Sftpc-GFP^{pos} library: GFP-UMIs>4, number of detected genes > 2500 and <5500 and percent of mitochondrial genes >0 and <7.5; Scgb1a1-CreER:Rosa-$Ai14^{pos}$ $Pdpn^{neg}$ library: RFP-UMIs>4, number of detected genes > 2500 and

<5500 and percent of mitochondrial genes >0 and <7.5; Scgb1a1-CreER:Rosa-Ai14^{pos} Sftpc-GFP^{pos} library: RFP-UMIs>4 and GFP-UMIs>4, number of detected genes > 2500 and <5500 and percent of mitochondrial genes >0 and <7.5; Scgb1a1-CreER:Rosa-Ai14^{pos} Pdpn^{pos} library: RFP-UMIs>4, number of detected genes > 3000 and <5500 and percent of mitochondrial genes >0 and <5). Genes with less than 50 counts in all cells were removed and the counts were transformed using the SCTransform[105] function in Seurat[106], with 4000 variable genes and regressing out the number of counts and detected genes and the percent of mitochondrial counts. For dimension reduction and clustering, the first 50 principal components were used, setting the number of neighbours to 25 and resolution to 0.2. MAST[107] was used to identify DEGs after library normalization to 10.000 and $\log_2$-transformation.

For the trajectory-analysis, we used equal number of cells/cluster and ran diffusion maps with Destiny[34], implemented with scMEGA[108]. We used the 16 first principal components and k=25. For the visualization and down-stream analyses, we used the three first diffusion-map components. We calculated the principal curves ("getCurves" function), the pseudotime estimates ("slingPseudotime" function) and the lineage assignment weights ("slingCurveWeights" function) with Slingshot[109]. We identified differentially expressed genes with the "fitGAM" function of tradeSeq[110]. For multiple trajectories we used the "patternTest" and for one the "associationTest" functions. The genes were ordered based on the hierarchical clustering ward.D2 method, using "hclust" function in fastcluster package[111] and plotted using a custom script. The "clusterboot" function of fpc package[112] was used to calculate stability values of gene-modules. GO-analyses were done at http://geneontology.org/ selecting as organism the Mus musculus and using default settings. The Fisher's Exact test was used to calculate the False Discovery Rate (FDR). Aggregated gene-expression scores of genes in modules and biological-processes were calculated with the "AddModuleScore" function in Seurat[106]. For Balloon-plots and heatmaps we used the "DotPlot" and "DoHeatmap" functions in Seurat, in addition to the pheatmap-package[113].

Full length scRNA-Seq

Single-cell library preparation was done according to Smart-Seq2 protocol[63, 114] with some modifications. Cells were sorted in 96-well plates (Piko PCR Plates 24-well, Thermo Scientific, SPL0240 and Plate Frame for 24-well PikoPCR Plates, Thermo Scientific, SFR0241). Each well contained Triton-X100 (0.2%) (Sigma-Aldrich, T9284-100ML), ERCC RNA Spike-In Mix (1:400.000) (Life-Technologies, 4456740), Oligo-dT30 VN (1.25µM) AAGCAGTGGTATCAACGCAGAGTAC(30 x T)VN, dNTPs (2.5mM/each) (Thermo Scientific-Fermentas, R0192) and Rnase Inhibitor (1U/µl) (Clontech, 2313A) in 4 µl final volume. After sorting, strips were covered with Axygen PCR-tube caps (VWR, PCR-02-FCP-C), centrifuged and placed on dry ice until storage at -80ºC for further use. For cell culture, cells were sorted

into HBSS++ buffer and kept on ice until they were processed for culture. To optimize Smart-Seq2[63] for mouse primary lung cells that are small and contain small amount of RNA, we used 50% less Oligo-dT30 VN and the cDNA synthesis was divided into two steps, the first was without TSO LNA and the second contained 1µM TSO LNA and additional 40 U of SuperScript II RT (Thermo-Fisher Scientific, 18064071). Reaction lasted 30 minutes at 42ºC. Then enzyme was deactivated at 70ºC for 15 minutes. For Pre-Amplification PCR, we used the KAPA HiFi Hotstart ReadyMix (2x) (KAPA Biosystems, KK2602) and the ISPCR-primer AAGCAGTGGTATCAACGCAGAGT. PCR included 21 cycles and the total volume increased to 50µl in order to reduce the concentration of the unused Oligo-dT30 VN and TSO-LNA primers.

Tagmentation and indexed library amplification were done with Nextera® XT DNA Library Preparation Kit (Illumina, FC-131-1096) and Nextera® XT Index Kit (96 indexes, 384 samples), (Illumina, FC-131-1002) according to the manufacturer protocol (with 2.5 x volume reduction in all reactions). For tagmentation, we used 50pg of the libraries, as it was indicated by the 500-9000 bp fraction of the library (Bioanalyzer).

Sequencing was done with Illumina 2500 HiSeq Rapid mode using paired-end (2x125 bp) and single-end (1x50 bp) reading. For downstream analyses, we used the one strand of paired-end libraries and trimmed the reads to 50 bp.

Single-cell RNA Sequencing bioinformatics analysis of Smart-seq2 dataset

We initially kept the libraries with >40% uniquely mapped reads to a reference genome that contained GFP and ERCC sequences and removed *Esr1*, as artifact because of sequence similarities with Scgb1a1-CreER^{T2} transgene and *Xis*t. Individual sequencing datasets were filtered regarding the number of detected genes (lower threshold: 2000 genes and upper threshold 10000 (P2272, P2661) and 6000 (P3504, P7657)). Then, we filtered out the libraries with more than 200 counts of Pecam1 as not epithelial contaminants. Finally, we removed libraries with more than 7.5% of mitochondrial gene counts, resulting to 354 libraries for downstream analysis.

We used SCT-transformation in Seurat with 3000 variable genes and regressed out the number of counts and detected genes and the percent of mitochondrial counts. The 20 first principal components were used for dimension reduction, setting the number of neighbours to 12 and resolution to 1. Diffusion maps were produced as in the adult dataset using the first 12 principal components and k=12. For the identification of DEGs, we used the MAST analysis in Seurat. For the trajectory analyses, we used Slingshot, setting as root the cluster-0 (embryonic) and end-point clusters the -2 (S1), -1 (S2), -4 (ciliated) and -3 (DP). The diffusion-map 3D-plots were created with scatter3D function of scatterplot3d[115]. GO-analyses and aggregated scores were produced as in the adult dataset.

Droplet based scRNA-Seq of *Fgfr2*-inactivated airway epithelial cells

For tissue dissociation and cell-isolation, we followed the same procedure as in the other FACs-sorting experiments. Single cell suspensions from three P7 Scgb1a1-CreER$^{T2\ neg/neg}$; Rosa26-Ai14$^{pos/pos}$; Fgfr2$^{fl/fl}$ mice were pooled and used as negative control sample. Three P7 and Scgb1a1-CreER$^{T2\ pos/neg}$; Rosa26-Ai14$^{pos/pos}$; Fgfr2$^{fl/fl}$ mice were combined and used as experimental groups. The same approach was used for two P21 negative control lungs and three experimental. Cells were counted with a Biorad cell counter, blocked with TruStain FcX™ PLUS (Biolegend, 156604), stained with a PE/Cyanine7 anti-mouse CD326 antibody and washed according manufacturer's protocol (Biolegend, 118216). The negative control samples were sorted based on Epcam positivity and from the experimental groups, we isolated Epcam^{pos}-Ai14^{neg} and Epcam^{pos}-Ai14^{pos} cells. The isolated cells were processed with the Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (10xGenomics), following manufacturer's instructions and targeting 7000 cells/well. The produced libraries were sequenced with a NovaSeq 6000 in two runs, one for each time point.

scRNA-Seq analysis of the *Fgfr2*-inactivated epithelial cells

We initially processed all cells and filtered out genes that are expressed in fewer than 5 cells and followed the same analysis approach as in the lineage-tracing dataset, using 4000 variable genes. We removed Krt13^{high} oesophageal/tracheal basal cells, Ptprc^{pos} immune and Col1a2^{pos} mesenchymal cells, as contaminants. Filtered cells were re-clustered after filtering out genes that are expressed in less than 20 cells and selecting the 5000 most variable. We used the 20 first principal components, k=15 and resolution=0.2. Then, we selected the airway secretory clusters for downstream analyses as described for the other datasets, using 600 variable genes, 15 principal components, resolution = 0.99 and k=8.

scRNA-Seq analysis of the GSE149563

For the analysis of the publicly available GSE149563 scRNA-Seq dataset, we analysed each timepoint individually with Seurat, using 4000 variable genes and 50 first principal components. We used DoubletFinder[116] package in R to identify and remove multiplets. The postnatal datasets were integrated and processed for clustering and differential expression analysis as in the other datasets. The epithelial clusters were further filtered to remove possible endothelial (Pecam1pos) and mesenchymal (Col1a2pos) cells and re-clustered selecting the 4000 most variable genes. We used the 50 first principal components, k=25 and resolution=0.6.

Organoid cultures

Lung digestion and cell sorting was performed as above, including the Dead cell stain (NucRed, Thermo) to sort out dead cells. Sorted epithelial cells (100-600 cells/well) were mixed with the Mlg2908 (ATCC, CCL-206) mouse lung fibroblasts (10^4 cells/well), as described before[117]. Colonies were then fixed in 4% PFA O/N and placed in 30% sucrose solution for 24 hours. Freeze-thawing and gentle pipetting was performed twice in order to remove Matrigel. Colonies were then incubated with 30% sucrose and 30% OCT overnight and embedded in OCT. Blocks were cut at 12-14 µm for immunofluorescence.

Immunofluorescence

The tissues were sectioned with a cryostat (Leica CM3050S). 10 µm thick sections were placed on poly-lysine slides (Thermo Scientific, J2800AMNZ), kept at room temperature (RT) for 3 hours with silica gel (Merck, 101969) to completely dry and then stored at -80ºC until use. All antibodies are described in Supplementary Table 16. All secondary antibodies were used at a dilution of 1:300-1:400.

For antigen retrieval (when necessary, see Supplementary Table 16), slides were placed in plastic jars, with the appropriate solution and warmed at 80ºC for 30min in a water-bath. Then, the jars were placed in ice for 30min to cool. Blocking was done with 5% donkey serum (Jackson Immuno-research, 017-000-121) for 1 hour at RT and the primary antibodies were incubated at 4ºC O/N. After washes the secondary antibodies were applied on the sections at RT for 1 hour in the dark. The nuclei were counterstained with a DAPI solution 0.5µg/µl (Biolegend, 422801) in PBS 1X Triton-X100 0.1% and for mounting we used the ProLong Gold Antifade Reagent (Thermo-Fischer Scientific, P36934).

Images were acquired with Zeiss LSM780, LSM800 confocal microscopes (Carl Zeiss Microscopy GmbH, Jena, Germany) and Zeiss Axio Observer Z.2 fluorescent microscope with Colibri2 or Colibri7. Image analysis was done using Fiji[118] and Zeiss Zen Blue 2.5.

Immunofluorescence-based quantification of S1 and S2 markers

For the quantification of S1 or S2 marker co-expression, we acquired 5 confocal microscopy images from P, I1-3 and D domains, from one P60 mouse lung section. Cell counting was performed using a custom pipeline at Cell Profiler 3.1. "*Global*" threshold strategy, "*otsu*" threshold method and three classes of thresholding were used. False positive and negative cells, because of problematic segmentation were manually curated.

SCRINSHOT spatial analyses

For spatial analysis of the identified cell-types, we applied SCRINSHOT[55], without modifications. The utilized padlock and detection probes are summarized in Supplementary Table 17. Images were captured with a Zeiss Axio Observer Z.2 fluorescent microscope with

Colibri2 or Colibri7, equipped with a Zeiss AxioCam 506 Mono digital camera and an automated stage.

We used DAPI to align the images of the same areas between the hybridizations. We created multi-channel *.czi files with the signal of each detected gene as a unique channel and exported them as images (16-bit *.tiff format), using Zen Blue 2.5 (Carl Zeiss Microscopy, GmbH). Images were tiled with Matlab with Image Analysis toolbox (The MathWorks, Inc.). Manual nuclei segmentation was done with Fiji ROI Manager[118] and signal-dot counting was performed with Cell-Profiler 3.15[119]. Annotation of signal-dots to the cells (2µm expanded nuclei) was done with Fiji. All scripts are found in https://github.com/AlexSount/SCRINSHOT and detailed instructions for the analysis have been previously described[55].

S1 and S2 cell spatial analyses

For the positional analysis of S1- and S2-cells, we targeted the module-5 secreted proteins *Scgb3a1, Reg3g*[3], *Bpifb1*[36], *Tff2*[120] and *Muc5b*[121], the receptor *Lgr6* that has previously reported to be expressed by distinct epithelial and mesenchymal lung cell-types[122] and the goblet-cell transcription factor *Pax9*[25], which are enriched in S1-cells (Suppl. Fig. 2B). For the distal lung compartment (DP, S2 and alveolar cells), we used the module-2 markers *Hp*[35] and *Scgb1a1* in addition to the module-3 surfactant proteins *Sftpc* and *Sftpb*[123], the enzyme *Atp8a1*[124], the IGF-signaling regulator *Igbp6*[125] and the advanced glycosylation end products receptor *Ager*[67, 126] (Suppl. Fig. 2C). We also targeted the ciliated-cell markers *Foxj1* and *Tuba1a* to recognize ciliated cells.

P, I1, I2, I3, D domains (how many for each) from three P60 mice were analysed for the selected panel of markers. Nuclei were manually segmented in the acquired images based on DAPI and manually curated based on E-cadherin antibody staining resulting to 6915 nuclei. Nuclear ROIs for each animal were expanded and filtered, keeping those with size between Mean cell-ROI size ± 2 standard deviations. Cell-ROIs with dots for only 1 analysed marker were removed. We further filtered the Cell-ROIs keeping those with total number of dots between the Mean number of dots ± 2 standard deviations. Cell-ROIs from all images were merged and $\log_2$-transformed [$\log_2$(dots + 1)]. After principal component analysis (PCA) the top up- or down-regulated genes of the first two principle components were used for clustering of the Cell-ROIs with clusterboot, using the ward.D2 method. The heatmap of the analysed cells was done with pheatmap-package in R. The balloon-plots of the expression levels (color intensity) and the percent of positive cells (size) were produced with ggpubr-package in R.

DP-cell spatial analysis

For the selection of the gene-panel for the spatial analysis of DP-cells, we used differential expression analyses to compare them with S2-cells and AEC2a. The panel included *Scgb1a1*

and *Sftpc* (the positivity of both defines the DP-cells/BASCs[15, 17, 18]) and the cytochrome genes *Cyp2f2* and *Cyp4b1* that showed high expression in S1- and S2-cells, moderate in DP and low in AEC2 cells. Based on the average fold change and the percentage of positive cells, we additionally selected the secreted proteins *Lyz2*[127] and *Lgi3*[128], the extracellular matrix proteins *Egfl6*, *Npnt* and *Col4a2*[129-131], the enzyme *Napsa*[132], the surface molecules *Cd74*[56] and *Cldn18*[133], the transcription factors *Etv5*[134] and *Rbpjl*[135] and the negative regulator of Wnt-signalling *Nkd1*[136]. We also included *Ager,* as a distal epithelial marker[67] and the NE-cell markers, *Ascl1*[137, 138] and *Calca* (Cgrp)[139]. We analysed several lung areas from three different adult (P60) mice in two independent experiments and manually segmented 58072 nuclei. Cell-ROIs with size outside 2 standard deviations from the average size of all Cell-ROIs from each lung were excluded from the analysis. DP-cells were extracted by the selection of ROIs with more than 24 dots of *Scgb1a1* and *Sftpc*, that have also *Scgb1a1*-dots ≥ 10 and *Sftpc*-dots ≥ 10. The balloon-plots of the expression levels (colour intensity) and the percent of positive Cell-ROIs (size) were produced with ggpubr-package[140] in R.

Statistical analyses

Statistical analysis of the results was done with two-way multiple comparisons test in Graphpad Prism (GraphPad Software, Inc) or the MAST and ROC analyses in Seurat. In Graphpad Prism, multiple comparisons were performed using Tukey statistical hypothesis testing. Adjusted p-values in MAST were calculated based on Bonferroni correction.

Data Availability

scRNA-Seq data are available in GEO (lineage-tracing dataset of Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-fGFP$^{pos/neg}$ cells: GSE215957, adult dataset of Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-Ai14$^{pos/neg}$ cells: GSE216210 and Fgfr2-inactivation dataset of Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-Ai14$^{pos/pos}$; Fgfr2$^{fl/fl}$ cells: GSE216451). The scripts and the RAW-image data of the study can be found in the Zenodo.

Acknowledgements

We would like to thank Athanasios Fysikopoulos for his technical contribution in the Fgfr2-inactivation experiments. We thank SciLifeLab long-term bioinformatics support. We acknowledge resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973 (projects b2015134 and SNIC 2021/22-431). We acknowledge European Respiratory Society-EMBO for the European Respiratory Society Long-Term Research Fellowship to AS (Reference Number: LTRF 2014 – 3565).

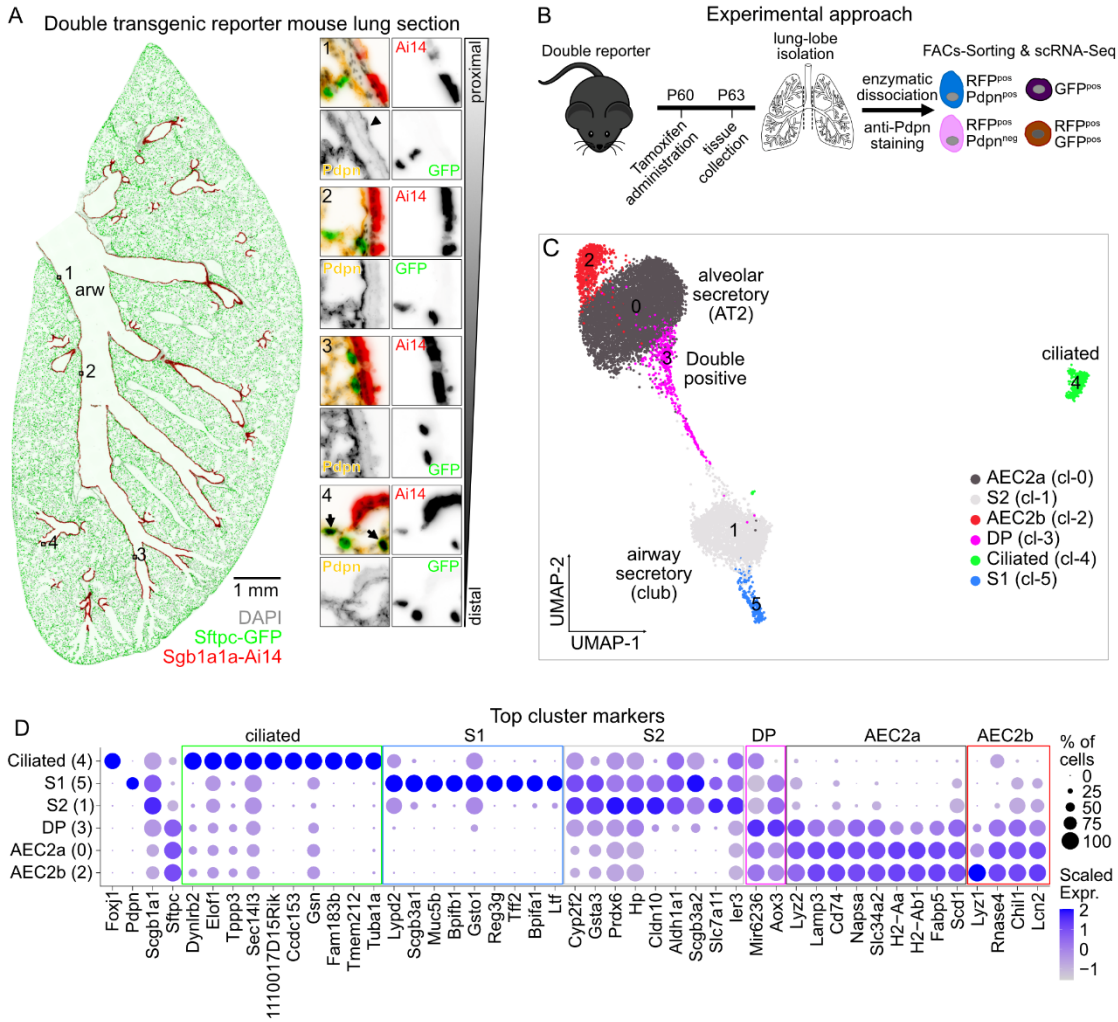

Figure 1. Characterization of lung secretory cell heterogeneity. (A) Colour-inverted fluorescence image of whole left-lung section from an Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-Ai14$^{pos/neg}$;Sftpc-GFP P63 reporter mouse, 72 hours after Tamoxifen administration. DAPI (nuclei): grey, RFP (Ai14): red, GFP (Sftpc): green. Scale-bar: 1mm. Inserts correspond to the numbered ROIs. **(B)** Experimental design for the isolation and single-cell RNA sequencing of the labelled cells from the reporter mouse. **(C)** UMAP-plot of 12030 cells, grouped into six clusters. Cluster (cl) 0: AEC2a (alveolar epithelial cell type-2a) (dark grey), cl-2: AEC2b (red), cl-5: airway secretory cell type-1 (S1) (blue), cl-1: S2 (light grey), cl-3: Scgb1a1^{pos} Sftpc^{pos} double positive (DP) (magenta) and cl-4: ciliated cells (green). **(D)** Balloon-plot of known epithelial markers. Ciliated: *Foxj1*[61], Club cells: *Scgb1a1*[3], Proximal-airway secretory[55] and AEC1[141],: *Pdpn*, AEC2 and DP-cells: *Sftpc*[15, 142], in addition to the expression of the top-10 differentially expressed genes of each cluster. The genes were filtered according to average log2 Fold-change (>0.5), adjusted p-value (<0.001) and percent of positive cells (>0.25) and the top-10 markers were plotted. Gene order follows the cluster order. Balloon size: percent of positive cells. Colour intensity: scaled expression (blue: high and grey: low).

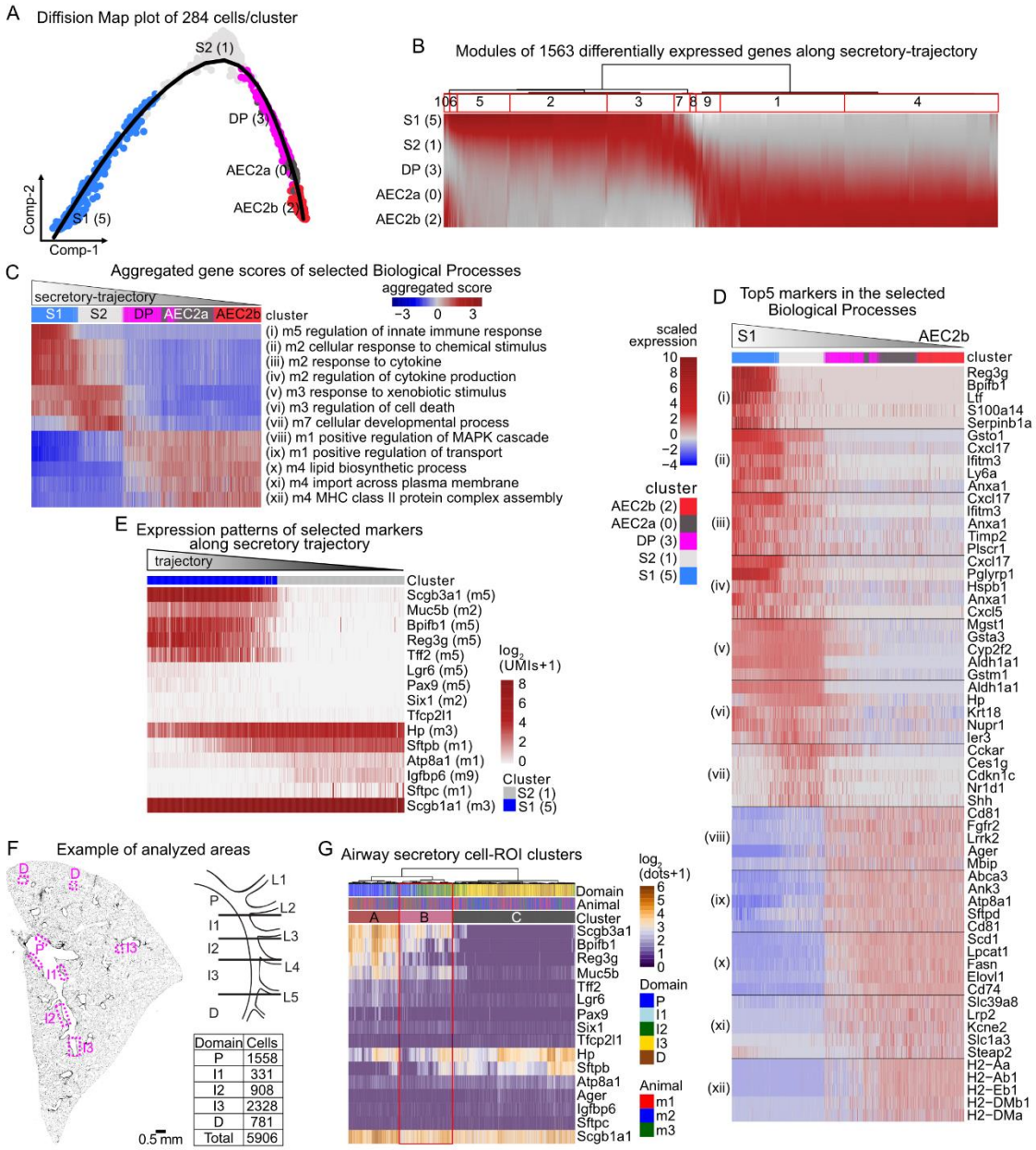

Figure 2. scRNA-Seq trajectory recapitulates the airway proximal-distal pattering. (A) Diffusion-map of secretory cell clusters. Colours and numbers as in Fig. 1 C. Line: estimated pseudotime-trajectory by Slingshot. **(B)** Heatmap of the 1563 differentially expressed genes (FDR<0.001 and meanLogFC>1) along pseudotime, based on tradeSeq. The dendrogram of hierarchical clustering (left) indicates 10 stable gene-modules. Bootstrapping values: module-1: 0.65, module-2: 0.62, module-3: 0.61, module-4: 0.74, module-5: 0.68, module-6: 0.75, module-7: 0.8, module-8: 0.7, module-9: 0.64, module-10: 0.9. Colour Intensity: scaled expression. Dark red: high, Gray: low. **(C)** Heatmap of the aggregated expression scores of the genes in the indicated biological processes (Suppl. Table3). The number after "m" indicates the module in "B" that contains the genes. Cells were ordered according to pseudotime (Fig. 2A). red: high, blue: low. "FDR": false discovery rate, "n": number of genes.

(D) Heatmap of the top-5 genes (according to “waldStat” score) of the indicated biological processes in “C”. Cells were ordered according to pseudotime. Colour: scaled expression (red: high, blue: low). **(E)** Heatmap of the S1 (cluster-5) and S2 (cluster-1) scRNA-Seq clusters, being ordered along pseudotime and showing the expression of the selected markers for spatial analysis. Expression levels: log2(normalized UMI-counts+1) (library size was normalized to 10.000). **(F)** Left: Representative adult mouse lung section counterstained with DAPI (grey) showing examples of imaged areas along the PD-axis (3 lungs have been analysed with similar results). Right-up: Cartoon of airway domain classification approach. Right-bottom: Synopsis of analysed cell-ROIs from three animals, for indicated domains. **(G)** Heatmap of 3096 analysed airway secretory cell-ROIs, showing the log2(SCRINSHOT dots +1) signal. Cell-ROI ordering is based on hierarchical clustering. Annotation bars show the (i) airway domains of the cell-ROIs, (ii) the analysed mouse and (iii) the indicated cluster. Cluster-A: P-domain cells 38.61 ± 6.89%, I1-domain cells 21.61 ± 2.46%. Cluster-B: I2-domain cells 37.79 ± 7.54%. Cluster-C: I3-domain cells 52.08 ± 6.57%, D-domain cells 69.49 ± 2.20%.

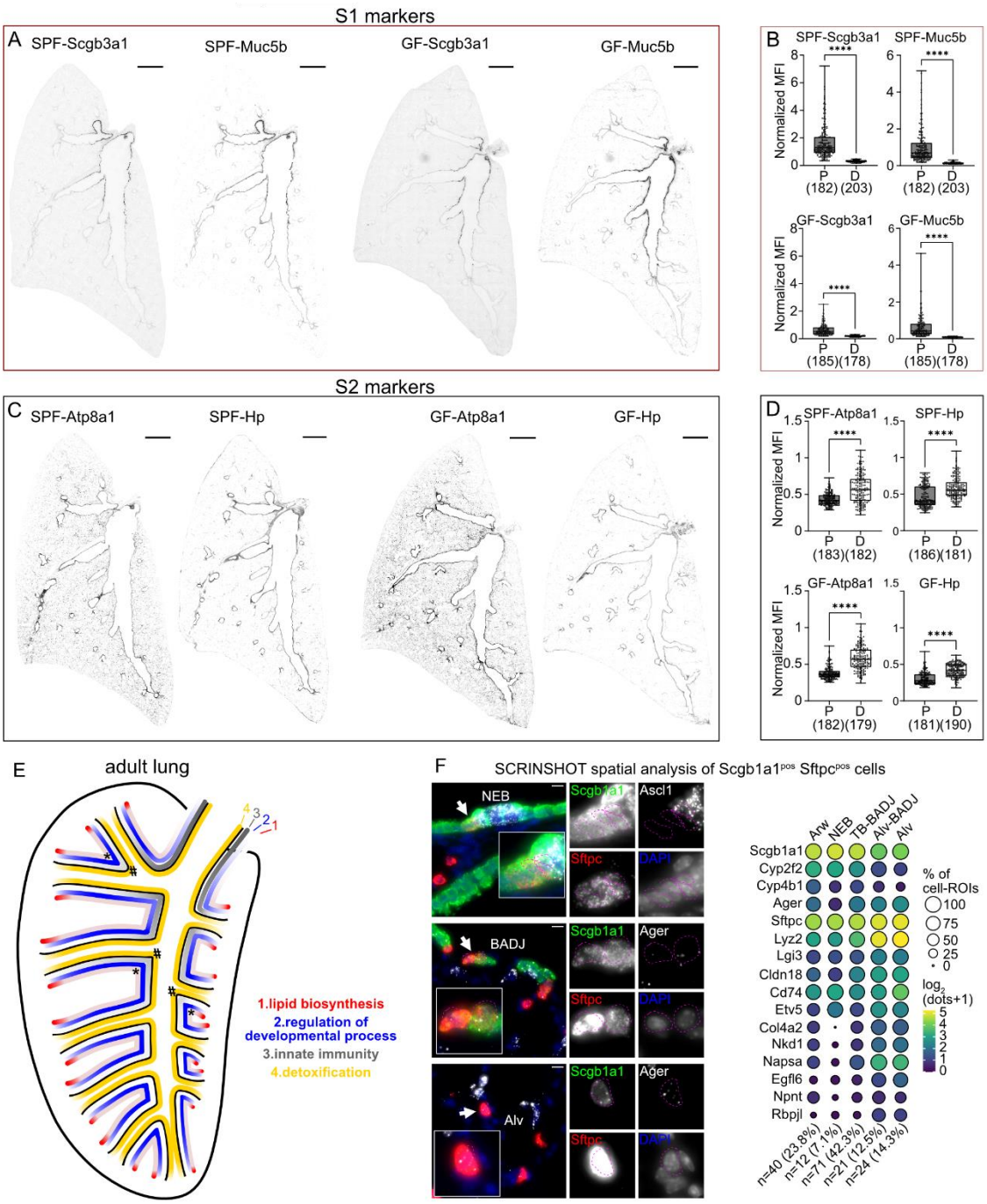

Figure 3. Gene-expression patterns along the airway PD-axis. (A) Immunofluorescence stainings of whole lung sections of specific pathogen free (SPF) and germ-free (GF) 2 months-old mice. Left: Scgb3a1, Muc5b and Tff2 S1-markers. **(B)** Quantification of immunofluorescence mean fluorescence intensity (MFI) of the indicated protein, normalized to the E-Cadherin signal. Numbers in parentheses: number of analysed proximal and distal cell-ROIs. Hp and Atp8a1 S2 markers. Statistics with Student's t-test: **** $p<0.0001$. Scale-bars: 1000 µm. **(C-D)** As in "A-B" for the Hp and Atp8a1 markers that are highly expressed in distal airways. **(E)** Graphical representation of the representative activated gene-expression programs (as in Fig. 2C) along the proximal-distal axis of the adult mouse lung

airways. Colour intensity: activation level. Dark: high, Fade: low. Asterisks (*) and hashes (#) represent exceptions of the lipid metabolism and detoxification expression programs activation that relate to neuroepithelial body topology. **(F)** Left: SCRINSHOT analysis images for Scgb1a1^{pos} Sftpc^{pos} cells (arrows) close to neuroendocrine (NE) cells (upper panel), terminal bronchioles (TB) (middle panel) and Alveoli (lower panel). Magenta dotted lines: outlines of 2µm-expanded Scgb1a1^{pos} Sftpc^{pos} nuclei. Arrows: Scgb1a1^{pos} Sftpc^{pos} cells. Sftpc: red, Scgb1a1: green, Ascl1: grey and DAPI: blue. Right: Balloon plot of the 16 analysed genes (module-3: *Scgb1a1, Cyp2f2* and *Cyp4*b1, module-1: *Ager, Sftpc* and *Lyz2* and module-4: *Lgi3, Cldn18, Cd74, Etv5, Col4a2, Nkd1, Napsa, Egfl6, Npnt* and *Rbpjl*) in the 170 identified Scgb1a1^{pos} Sftpc^{pos} cells, according to their position. The balloon size: percentage of positive cells. The colour intensity: $\log_2$(SCRINSHOT dots +1). Yellow: high, Dark blue: low. “n”: number of cells in the specified position.

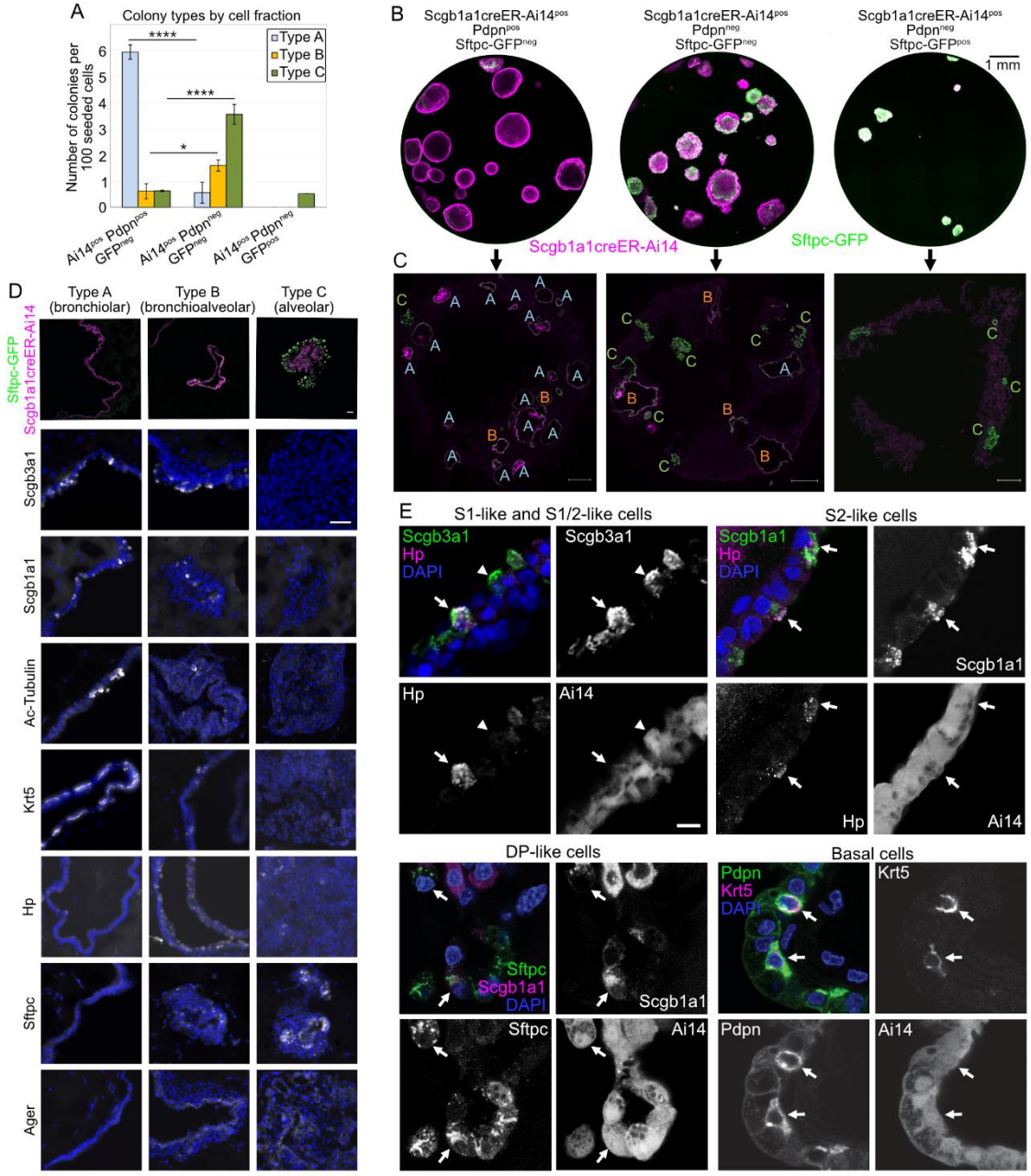

Figure 4. Clonogenic and differentiation potential of airway secretory cell states. (A) Bar plot of the numbers of derived colonies per 100 seeded cells (counted 2-3 whole wells per cell type from four animals, three independent experiments) with embedded percentage of colony types derived from each type of seeded cells, (counted one well of each cell type from three animals, three independent experiments). Type-A: bronchiolar colony, Type-B: bronchioalveolar colony, Type-C: alveolar colony. Error bars correspond to standard error from the mean, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Number of colonies counted from 4 different mice, type-A: 52, type-B: 16, type-C: 44. Two tailed equal variance Student's t-test (after a variance comparison test for all datasets) was performed. **(B)** Representative whole-culture images from Scgb1a1creER-Ai14^{pos} Pdpn^{pos} Sftpc-GFP^{neg}, Scgb1a1creER-Ai14^{pos}

Pdpn^{neg} Sftpc-GFP^{neg} and Scgb1a1creER-Ai14^{pos} Sftpc-GFP^{pos} seeded cells. Scale-bar: 1mm. **(C)** Images of 10 µm sections of representative cultures showing Ai14, and GFP transgene fluorescence. The letters indicate the colony annotations. **(D)** Representative images of the three types of colonies, showing the Ai14, and GFP transgene fluorescence and the immunofluorescence signal for bronchiolar (Scgb1a1), S1 (Scgb3a1), basal (Krt5), ciliated (acetylated Tubulin), and distal epithelial (Hp, Sftpc, Ager) markers. Nuclei-DAPI: blue. Scale bar: 50 µm. **(E)** Confocal microscopy images for the detection of S1 (Scgb3a1^{pos} Hp^{pos}), intermediate (Scgb3a1^{pos} Hp^{pos}), S2 (Scgb1a1^{pos} Hp^{pos}) and DP (Scgb1a1^{pos} Sftpc^{pos}) cells in analysed colonies. Arrows indicate positive cells for both analysed markers. Scale-bar 10 µm. Whole well culture section from each fraction of cells was stained and analysed for three biological replicates.

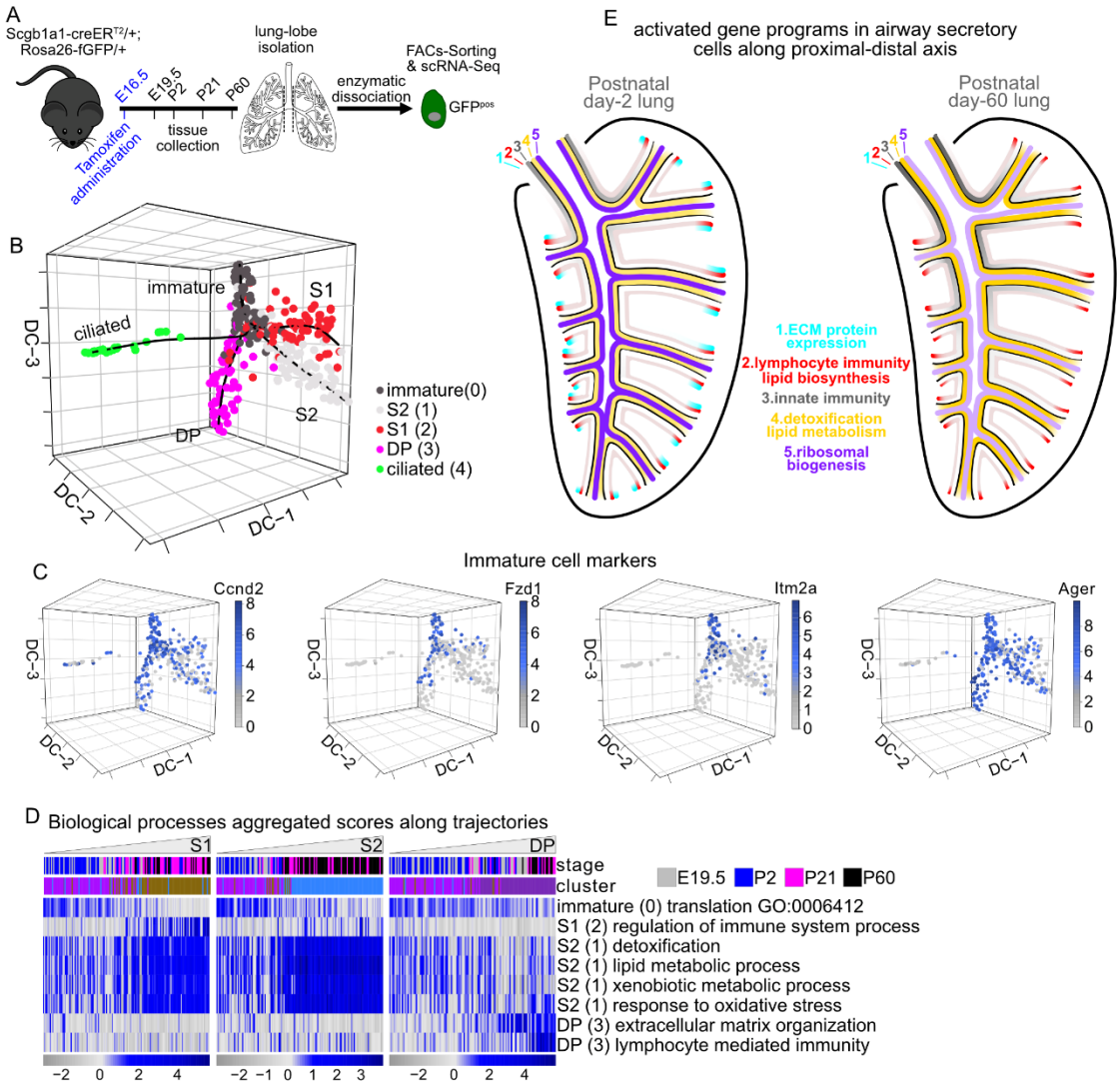

Figure 5. Lineage-tracing of airway secretory cell heterogeneity. **(A)** Experimental design for the isolation and single-cell RNA sequencing (scRNA-Seq) of the labelled cells from the Scgb1a1-CreER$^{T2\ pos/neg}$;Rosa26-fGFP$^{pos/neg}$ reporter mice. **(B)** Diffusion-map plot of 354 full-length, single-cell cDNA libraries. Colours: suggested clusters. Lines: four distinct lineage-trajectories, calculated by Slingshot. **(C)** Diffusion-map plots of the perinatally expressed genes *Ccnd2*, *Fzd1*, *Itm2a* and *Ager*. Expression levels: $\log_2$(normalized counts+1) (library size was normalized to 10^6). Blue: high, Gray: zero. **(D)** Heatmaps of the aggregated gene expression scores of the indicated biological processes (see Suppl. Table 6). The cells were ordered according to the pseudotime values of the trajectories in "B". Yellow: high, Purple: low. **(E)** Synopsis of the gene expression programs activation in airway secretory epithelial cells, along the proximal-distal axis, in the postnatal day-2 (left) and -60 (right) lungs. Colour intensity: activation level. Dark: high, Fade: low. ECM: extracellular matrix.

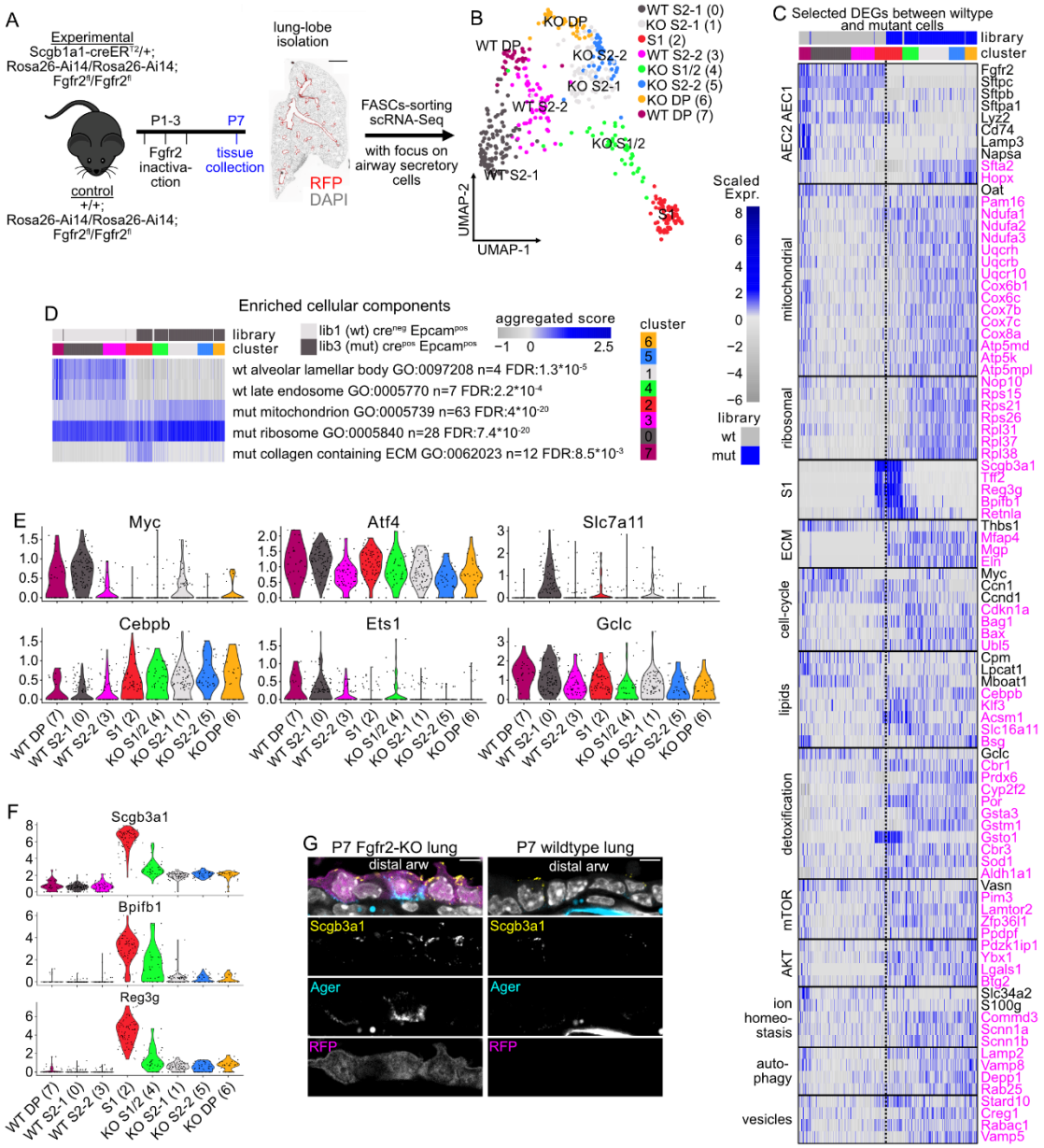

Figure 6. Fgfr2 inactivation in airway secretory cells causes extended gene-expression changes. (A) Experimental design for the perinatal inactivation of Fgfr2 in Scgb1a1^{pos} cells and analysis with single-cell RNA sequencing (scRNA-Seq) and histology. **(B)** UMAP-map plot of equal numbers of randomly-selected mutant and wildtype airway secretory cells of clusters -2 and -4 in Suppl. Fig. 4F. Colours: suggested clusters. **(C)** Heatmap of the aggregated scores of selected statistically significant, altered cellular components according to GO-analysis. The results are based on the statistically-significant, differentially expressed genes between the wildtype (library-1) and the mutant (library-3) airway secretory cells. The cells are ordered according to the clusters (colours as in "B"). Score: blue: high, grey: low. "FDR": false discovery rate, "n": number of genes. **(D)** Heatmap showing the expression of selected, differentially expressed genes between the wildtype (library-1) and the mutant (library-3) airway secretory cells. Genes are organized in distinct

categories according to previous knowledge (see Suppl. Table 7F-G). Colour: scaled expression. blue: high, grey: low. **(E)** Violin plots of the transcription factors *Myc*, *Atf4, Cebpb and Ets1* in addition to the involved in glutathione metabolism *Slc7a11* and *Gclc*. **(F)** Violin plots of the *Scgb3a1*, *Bpifb1* and *Reg3g* showing their up-regulation in Fgfr2-mutant cells. Expression levels: $\log_2$(normalized UMI-counts+1) (library size was normalized to 10.000). Colours as in "B". **(G)** Confocal microscopy, single-step images of distal airway epithelium from a postnatal day-7 (P7) Fgfr2-mutant lung (left) and a wildtype littermate (right). Immunofluorescence for Scgb3a1 (Yellow) and Ager (Cyan). Rosa26-RFP (magenta) indicates cells that underwent recombination. Nuclei-DAPI: grey. Scale-bar 5 µm. "arw": airway. Three lungs for each condition were analysed.

Literature

1. Morrisey, E.E. & Hogan, B.L. Preparing for the first breath: genetic and cellular mechanisms in lung development. *Dev Cell* **18**, 8-23 (2010).
2. Tata, P.R. & Rajagopal, J. Plasticity in the lung: making and breaking cell identity. *Development* **144**, 755-766 (2017).
3. Reynolds, S.D., Reynolds, P.R., Pryhuber, G.S., Finder, J.D. & Stripp, B.R. Secretoglobins SCGB3A1 and SCGB3A2 define secretory cell subsets in mouse and human airways. *Am J Respir Crit Care Med* **166**, 1498-1509 (2002).
4. Singh, G. & Katyal, S.L. Clara cell proteins. *Ann N Y Acad Sci* **923**, 43-58 (2000).
5. Davis, J.D. & Wypych, T.P. Cellular and functional heterogeneity of the airway epithelium. *Mucosal Immunology* **14**, 978-990 (2021).
6. Hewitt, R.J. & Lloyd, C.M. Regulation of immune responses by the airway epithelial cell landscape. *Nature Reviews Immunology* **21**, 347-362 (2021).
7. Fanucchi, M.V., Murphy, M.E., Buckpitt, A.R., Philpot, R.M. & Plopper, C.G. Pulmonary cytochrome P450 monooxygenase and Clara cell differentiation in mice. *Am J Respir Cell Mol Biol* **17**, 302-314 (1997).
8. Rawlins, E.L., Ostrowski, L.E., Randell, S.H. & Hogan, B.L. Lung development and repair: contribution of the ciliated lineage. *Proc Natl Acad Sci U S A* **104**, 410-417 (2007).
9. Tan, F.E. *et al.* Myb promotes centriole amplification and later steps of the multiciliogenesis program. *Development* **140**, 4277-4286 (2013).
10. Tilley, A.E., Walters, M.S., Shaykhiev, R. & Crystal, R.G. Cilia dysfunction in lung disease. *Annu Rev Physiol* **77**, 379-406 (2015).
11. Giangreco, A. *et al.* Stem cells are dispensable for lung homeostasis but restore airways after injury. *Proc Natl Acad Sci U S A* **106**, 9286-9291 (2009).
12. Rawlins, E.L. *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* **4**, 525-534 (2009).
13. Kathiriya, J.J., Brumwell, A.N., Jackson, J.R., Tang, X. & Chapman, H.A. Distinct Airway Epithelial Stem Cells Hide among Club Cells but Mobilize to Promote Alveolar Regeneration. *Cell Stem Cell* (2020).
14. Reynolds, S.D., Giangreco, A., Power, J.H.T. & Stripp, B.R. Neuroepithelial Bodies of Pulmonary Airways Serve as a Reservoir of Progenitor Cells Capable of Epithelial Regeneration. *The American Journal of Pathology* **156**, 269-278 (2000).
15. Kim, C.F. *et al.* Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* **121**, 823-835 (2005).
16. Guha, A., Deshpande, A., Jain, A., Sebastiani, P. & Cardoso, W.V. Uroplakin 3a(+) Cells Are a Distinctive Population of Epithelial Progenitors that Contribute to Airway Maintenance and Post-injury Repair. *Cell Rep* **19**, 246-254 (2017).
17. Liu, Q. *et al.* Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. *Nat Genet* (2019).
18. Salwig, I. *et al.* Bronchioalveolar stem cells are a main source for regeneration of distal lung epithelia in vivo. *EMBO J* (2019).
19. Strunz, M. *et al.* Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. *Nat Commun* **11**, 3559 (2020).
20. Choi, J. *et al.* Inflammatory Signals Induce AT2 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration. *Cell Stem Cell* **27**, 366-382 e367 (2020).
21. Kobayashi, Y. *et al.* Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis. *Nat Cell Biol* **22**, 934-946 (2020).
22. Metzger, R.J., Klein, O.D., Martin, G.R. & Krasnow, M.A. The branching programme of mouse lung development. *Nature* **453**, 745-750 (2008).
23. *Lung Stem Cells in Development, Health and Disease*. (2021).

24. Boers, J.E., Ambergen, A.W. & Thunnissen, F.B. Number and proliferation of basal and parabasal cells in normal human airway epithelium. *Am J Respir Crit Care Med* **157**, 2000-2006 (1998).
25. Montoro, D.T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**, 319-324 (2018).
26. Plasschaert, L.W. *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377-381 (2018).
27. Kadur Lakshminarasimha Murthy, P. *et al.* Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. *Nature* **604**, 111-119 (2022).
28. Vanderbilt, J.N. *et al.* High-Efficiency Type II Cell–Enhanced Green Fluorescent Protein Expression Facilitates Cellular Identification, Tracking, and Isolation. *American Journal of Respiratory Cell and Molecular Biology* **53**, 14-21 (2015).
29. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neuroscience* **13**, 133-140 (2010).
30. Sountoulidis, A. *et al.* SCRINSHOT, a spatial method for single-cell resolution mapping of cell states in tissue sections. *bioRxiv*, 2020.2002.2007.938571 (2020).
31. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv 2018. *arXiv preprint arXiv:1802.03426* (1802).
32. Jiang, M. *et al.* VEGF receptor 2 (KDR) protects airways from mucus metaplasia through a Sox9-dependent pathway. *Dev Cell* **56**, 1646-1660 e1645 (2021).
33. Hurskainen, M. *et al.* Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat Commun* **12**, 1565 (2021).
34. Angerer, P. *et al.* destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241-1243 (2015).
35. Guha, A. *et al.* Analysis of Notch signaling-dependent gene expression in developing airways reveals diversity of Clara cells. *PLoS One* **9**, e88848 (2014).
36. Bingle, L. *et al.* BPIFB1 (LPLUNC1) is upregulated in cystic fibrosis lung disease. *Histochem Cell Biol* **138**, 749-758 (2012).
37. Winkelmann, V.E. *et al.* Inflammation-induced upregulation of P2X(4) expression augments mucin secretion in airway epithelia. *Am J Physiol Lung Cell Mol Physiol* **316**, L58-L70 (2019).
38. Rothenberg, M.E. *et al.* IL-13 receptor alpha1 differentially regulates aeroallergen-induced lung responses. *J Immunol* **187**, 4873-4880 (2011).
39. Khalifeh-Soltani, A. *et al.* The Mfge8-alpha8beta1-PTEN pathway regulates airway smooth muscle contraction in allergic inflammation. *FASEB J*, fj201800109R (2018).
40. Yang, Z.C. *et al.* Targeted inhibition of Six1 attenuates allergic airway inflammation and remodeling in asthmatic mice. *Biomed Pharmacother* **84**, 1820-1825 (2016).
41. Park, K.S. *et al.* SPDEF regulates goblet cell hyperplasia in the airway epithelium. *J Clin Invest* **117**, 978-988 (2007).
42. Tian, B. *et al.* Central Role of the NF-kappaB Pathway in the Scgb1a1-Expressing Epithelium in Mediating Respiratory Syncytial Virus-Induced Airway Inflammation. *J Virol* **92** (2018).
43. Kugler, M.C., Joyner, A.L., Loomis, C.A. & Munger, J.S. Sonic hedgehog signaling in the lung. From development to disease. *Am J Respir Cell Mol Biol* **52**, 1-13 (2015).
44. Peng, T. *et al.* Hedgehog actively maintains adult lung quiescence and regulates repair and regeneration. *Nature* **526**, 578-582 (2015).
45. Pariollaud, M. *et al.* Circadian clock component REV-ERBalpha controls homeostatic regulation of pulmonary inflammation. *J Clin Invest* **128**, 2281-2296 (2018).
46. Movassagh, H. *et al.* Semaphorin 3E Alleviates Hallmarks of House Dust Mite-Induced Allergic Airway Disease. *Am J Pathol* **187**, 1566-1576 (2017).
47. Toubi, E. & Vadasz, Z. Semaphorin3A is a promising therapeutic tool for bronchial asthma. *Allergy* **75**, 481-483 (2020).

48. Ahmadvand, N. *et al.* Fgfr2b signaling is essential for the maintenance of the alveolar epithelial type 2 lineage during lung homeostasis in mice. *Cell Mol Life Sci* **79**, 302 (2022).
49. Brownfield, D.G. *et al.* Alveolar cell fate selection and lifelong maintenance of AT2 cells by FGF signaling. *Nat Commun* **13**, 7137 (2022).
50. Liberti, D.C. *et al.* Alveolar epithelial cell fate is maintained in a spatially restricted manner to promote lung regeneration after acute injury. *Cell Reports* **35**, 109092 (2021).
51. Tian, Y. *et al.* LRRK2 plays essential roles in maintaining lung homeostasis and preventing the development of pulmonary fibrosis. *Proc Natl Acad Sci U S A* **118** (2021).
52. Zhang, Z. *et al.* Transcription factor Etv5 is essential for the maintenance of alveolar type II cells. *Proc Natl Acad Sci U S A* **114**, 3903-3908 (2017).
53. Little, D.R. *et al.* Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1. *Proc Natl Acad Sci U S A* **116**, 20545-20555 (2019).
54. Little, D.R. *et al.* Differential chromatin binding of the lung lineage transcription factor NKX2-1 resolves opposing murine alveolar cell fates in vivo. *Nat Commun* **12**, 2509 (2021).
55. Sountoulidis, A. *et al.* SCRINSHOT enables spatial mapping of cell states in tissue sections with single-cell resolution. *PLOS Biology* **18**, e3000675 (2020).
56. Lee, J.-H. *et al.* Surfactant Protein–C Chromatin-Bound Green Fluorescence Protein Reporter Mice Reveal Heterogeneity of Surfactant Protein C–Expressing Lung Cells. *American Journal of Respiratory Cell and Molecular Biology* **48**, 288-298 (2013).
57. Stupnikov, M.R., Yang, Y., Mori, M., Lu, J. & Cardoso, W.V. Jagged and Delta-like ligands control distinct events during airway progenitor cell differentiation. *eLife* **8**, e50487 (2019).
58. Chen, H. *et al.* Airway epithelial progenitors are region specific and show differential responses to bleomycin-induced lung injury. *Stem Cells* **30**, 1948-1960 (2012).
59. Teisanu, R.M. *et al.* Functional analysis of two distinct bronchiolar progenitors during lung injury and repair. *Am J Respir Cell Mol Biol* **44**, 794-803 (2011).
60. McQualter, J.L., Yuen, K., Williams, B. & Bertoncello, I. Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. *Proc Natl Acad Sci U S A* **107**, 1414-1419 (2010).
61. Gomperts, B.N., Gong-Cooper, X. & Hackett, B.P. Foxj1 regulates basal body anchoring to the cytoskeleton of ciliated pulmonary epithelial cells. *J Cell Sci* **117**, 1329-1337 (2004).
62. Balasooriya, G.I., Goschorska, M., Piddini, E. & Rawlins, E.L. FGFR2 is required for airway basal cell self-renewal and terminal differentiation. *Development* **144**, 1600-1606 (2017).
63. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**, 171-181 (2014).
64. Pogach, M.S., Cao, Y., Millien, G., Ramirez, M.I. & Williams, M.C. Key developmental regulators change during hyperoxia-induced injury and recovery in adult mouse lung. *J Cell Biochem* **100**, 1415-1429 (2007).
65. Ahmadvand, N. *et al.* Identification of a novel subset of alveolar type 2 cells enriched in PD-L1 and expanded following pneumonectomy. *Eur Respir J* **58** (2021).
66. Namkoong, S. *et al.* The integral membrane protein ITM2A, a transcriptional target of PKA-CREB, regulates autophagic flux via interaction with the vacuolar ATPase. *Autophagy* **11**, 756-768 (2015).
67. Chung, M.-I. & Hogan, B.L.M. Ager-CreERT2: A New Genetic Tool for Studying Lung Alveolar Development, Homeostasis, and Repair. *American Journal of Respiratory Cell and Molecular Biology* **59**, 706-712 (2018).

68. Bellusci, S., Grindley, J., Emoto, H., Itoh, N. & Hogan, B.L.M. Fibroblast Growth Factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung. *Development* **124**, 4867-4878 (1997).
69. Jones, M.R. *et al.* FGFR2b signalling restricts lineage-flexible alveolar progenitors during mouse lung development and converges in mature alveolar type 2 cells. *Cell Mol Life Sci* **79**, 609 (2022).
70. Herriges, J.C. *et al.* FGF-Regulated ETV Transcription Factors Control FGF-SHH Feedback Loop in Lung Branching. *Dev Cell* **35**, 322-332 (2015).
71. Dorry, S.J., Ansbro, B.O., Ornitz, D.M., Mutlu, G.M. & Guzy, R.D. FGFR2 Is Required for AEC2 Homeostasis and Survival after Bleomycin-induced Lung Injury. *Am J Respir Cell Mol Biol* **62**, 608-621 (2020).
72. Moerlooze, L.D. *et al.* An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis. *Development* **127**, 483-492 (2000).
73. Jain, R. *et al.* Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. *Nat Commun* **6**, 6727 (2015).
74. Yang, J. *et al.* The development and plasticity of alveolar type 1 cells. *Development* **143**, 54-65 (2016).
75. Liu, J. *et al.* CREG1 promotes lysosomal biogenesis and function. *Autophagy* **17**, 4249-4265 (2021).
76. Ikezawa, M. *et al.* Loss of VAMP5 in mice results in duplication of the ureter and insufficient expansion of the lung. *Dev Dyn* **247**, 754-762 (2018).
77. Zepp, J.A. *et al.* Genomic, epigenomic, and biophysical cues controlling the emergence of the lung alveolus. *Science* **371** (2021).
78. Zaidi, A.R.S., Dresman, S., Burt, C., Rule, S. & McCallum, L. Molecular signatures for CCN1, p21 and p27 in progressive mantle cell lymphoma. *J Cell Commun Signal* **13**, 421-434 (2019).
79. Negretti, N.M. *et al.* A single-cell atlas of mouse lung development. *Development* **148** (2021).
80. Basu, A. & Haldar, S. The relationship between BcI2, Bax and p53: consequences for cell cycle progression and cell death. *Mol Hum Reprod* **4**, 1099-1109 (1998).
81. Gijón, M.A., Riekhof, W.R., Zarini, S., Murphy, R.C. & Voelker, D.R. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. *J Biol Chem* **283**, 30235-30245 (2008).
82. Vessey, D.A., Kelley, M. & Warren, R.S. Characterization of the CoA ligases of human liver mitochondria catalyzing the activation of short- and medium-chain fatty acids and xenobiotic carboxylic acids. *Biochim Biophys Acta* **1428**, 455-462 (1999).
83. Prokesch, A. *et al.* Arxes: retrotransposed genes required for adipogenesis. *Nucleic Acids Res* **39**, 3224-3239 (2011).
84. Maeda, K. *et al.* Role of the fatty acid binding protein mal1 in obesity and insulin resistance. *Diabetes* **52**, 300-307 (2003).
85. Lin, S., Ikegami, M., Moon, C., Naren, A.P. & Shannon, J.M. Lysophosphatidylcholine Acyltransferase 1 (LPCAT1) Specifically Interacts with Phospholipid Transfer Protein StarD10 to Facilitate Surfactant Phospholipid Trafficking in Alveolar Type II Cells. *J Biol Chem* **290**, 18559-18574 (2015).
86. Guo, L. *et al.* Transactivation of Atg4b by C/EBPβ promotes autophagy to facilitate adipogenesis. *Mol Cell Biol* **33**, 3180-3190 (2013).
87. Guntur, A.R. *et al.* An essential role for the circadian-regulated gene nocturnin in osteogenesis: the importance of local timekeeping in skeletal homeostasis. *Ann N Y Acad Sci* **1237**, 58-63 (2011).
88. Sue, N. *et al.* Targeted disruption of the basic Krüppel-like factor gene (Klf3) reveals a role in adipogenesis. *Mol Cell Biol* **28**, 3967-3978 (2008).
89. Kim, Y. *et al.* Function of the pentose phosphate pathway and its key enzyme, transketolase, in the regulation of the meiotic cell cycle in oocytes. *Clin Exp Reprod Med* **39**, 58-67 (2012).

90. Luo, C. *et al.* Increased CYR61 expression activates CCND1/c-Myc pathway to promote nasal epithelial cells proliferation in chronic rhinosinusitis with nasal polyps. *Clinical Immunology* **247**, 109235 (2023).
91. Tameire, F. *et al.* ATF4 couples MYC-dependent translational activity to bioenergetic demands during tumour progression. *Nat Cell Biol* **21**, 889-899 (2019).
92. Berberich-Siebelt, F. *et al.* SUMOylation interferes with CCAAT/enhancer-binding protein beta-mediated c-myc repression, but not IL-4 activation in T cells. *J Immunol* **176**, 4843-4851 (2006).
93. Lim, J.K.M. *et al.* Cystine/glutamate antiporter xCT (SLC7A11) facilitates oncogenic RAS transformation by preserving intracellular redox balance. *Proc Natl Acad Sci U S A* **116**, 9433-9442 (2019).
94. Misra, I. & Griffith, O.W. Expression and purification of human gamma-glutamylcysteine synthetase. *Protein Expr Purif* **13**, 268-276 (1998).
95. Basil, M.C. *et al.* Human distal airways contain a multipotent secretory cell that can regenerate alveoli. *Nature* **604**, 120-126 (2022).
96. Sikkema, L. *et al.* An integrated cell atlas of the lung in health and disease. *Nat Med* **29**, 1563-1577 (2023).
97. Lyu, H. *et al.* Niche-mediated repair of airways is directed in an occupant-dependent manner. *Cell Rep* **41**, 111863 (2022).
98. Sauler, M. *et al.* Characterization of the COPD alveolar niche using single-cell RNA sequencing. *Nature Communications* **13**, 494 (2022).
99. Mori, M., Bjermer, L., Erjefalt, J.S., Stampfli, M.R. & Roos, A.B. Small airway epithelial-C/EBPbeta is increased in patients with advanced COPD. *Respir Res* **16**, 133 (2015).
100. 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* **12**, 1041-1055 (2019).
101. Elhamamsy, A.R., Metge, B.J., Alsheikh, H.A., Shevde, L.A. & Samant, R.S. Ribosome Biogenesis: A Central Player in Cancer Metastasis and Therapeutic Resistance. *Cancer Res* **82**, 2344-2353 (2022).
102. Travaglini, K.J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619-625 (2020).
103. Vieira Braga, F.A. *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* **25**, 1153-1163 (2019).
104. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nature Methods* **14**, 1083-1086 (2017).
105. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 296 (2019).
106. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
107. Finak, G. *et al.* MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* **16**, 278 (2015).
108. Li, Z., Nagai, J.S., Kuppe, C., Kramann, R. & Costa, I.G. scMEGA: single-cell multi-omic enhancer-based gene regulatory network inference. *Bioinformatics Advances* **3** (2023).
109. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
110. Van den Berge, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun* **11**, 1201 (2020).
111. Müllner, D. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. *Journal of Statistical Software* **53**, 1 - 18 (2013).
112. Hennig, C. & Imports, M. Package ‘fpc’. *Flexible Procedures for Clustering* (2015).
113. Kolde, R. & Kolde, M.R. Package ‘pheatmap’. *R package* **1**, 790 (2015).

114. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods* **10**, 1096-1098 (2013).
115. Ligges, U. & Maechler, M. scatterplot3d - An R Package for Visualizing Multivariate Data. *Journal of Statistical Software* **8**, 1 - 20 (2003).
116. McGinnis, C.S., Murrow, L.M. & Gartner, Z.J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337.e324 (2019).
117. Sountoulidis, A. *et al.* Activation of the canonical bone morphogenetic protein (BMP) pathway during lung morphogenesis and adult lung tissue repair. *PLoS One* **7**, e41460 (2012).
118. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682 (2012).
119. Lamprecht, M.R., Sabatini, D.M. & Carpenter, A.E. CellProfiler™: free, versatile software for automated biological image analysis. *BioTechniques* **42**, 71-75 (2007).
120. Royce, S.G. *et al.* Trefoil factor-2 reverses airway remodeling changes in allergic airways disease. *Am J Respir Cell Mol Biol* **48**, 135-144 (2013).
121. Rose, M.C. & Voynow, J.A. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* **86**, 245-278 (2006).
122. Lee, J.H. *et al.* Anatomically and Functionally Distinct Lung Mesenchymal Populations Marked by Lgr5 and Lgr6. *Cell* **170**, 1149-1163 e1112 (2017).
123. Whitsett, J.A., Wert, S.E. & Weaver, T.E. Alveolar surfactant homeostasis and the pathogenesis of pulmonary disease. *Annu Rev Med* **61**, 105-119 (2010).
124. Soupene, E., Kemaladewi, D.U. & Kuypers, F.A. ATP8A1 activity and phosphatidylserine transbilayer movement. *J Receptor Ligand Channel Res* **1**, 1-10 (2008).
125. Bach, L.A. Recent insights into the actions of IGFBP-6. *J Cell Commun Signal* **9**, 189-200 (2015).
126. Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371-375 (2014).
127. Markart, P. *et al.* Comparison of the microbicidal and muramidase activities of mouse lysozyme M and P. *Biochemical Journal* **380**, 385-392 (2004).
128. Senechal, K.R., Thaller, C. & Noebels, J.L. ADPEAF mutations reduce levels of secreted LGI1, a putative tumor suppressor protein linked to epilepsy. *Human Molecular Genetics* **14**, 1613-1620 (2005).
129. Fujiwara, H. *et al.* The Basement Membrane of Hair Follicle Stem Cells Is a Muscle Cell Niche. *Cell* **144**, 577-589 (2011).
130. Loscertales, M. *et al.* Type IV collagen drives alveolar epithelial–endothelial association and the morphogenetic movements of septation. *BMC Biology* **14**, 59 (2016).
131. Cheng, C.-C. *et al.* Hair follicle epidermal stem cells define a niche for tactile sensation. *eLife* **7**, e38883 (2018).
132. Brasch, F. *et al.* Involvement of Napsin A in the C- and N-terminal Processing of Surfactant Protein B in Type-II Pneumocytes of the Human Lung*. *Journal of Biological Chemistry* **278**, 49006-49014 (2003).
133. Niimi, T. *et al.* claudin-18, a Novel Downstream Target Gene for the T/EBP/NKX2.1 Homeodomain Transcription Factor, Encodes Lung- and Stomach-Specific Isoforms through Alternative Splicing. *Molecular and Cellular Biology* **21**, 7380-7390 (2001).
134. Zhang, Z. *et al.* Transcription factor Etv5 is essential for the maintenance of alveolar type II cells. *Proceedings of the National Academy of Sciences* **114**, 3903-3908 (2017).
135. Masui, T. *et al.* Replacement of Rbpj With Rbpjl in the PTF1 Complex Controls the Final Maturation of Pancreatic Acinar Cells. *Gastroenterology* **139**, 270-280 (2010).
136. Larraguibel, J. *et al.* Wnt ligand–dependent activation of the negative feedback regulator Nkd1. *Molecular Biology of the Cell* **26**, 2375-2384 (2015).

137. Borges, M. *et al.* An achaete-scute homologue essential for neuroendocrine differentiation in the lung. *Nature* **386**, 852-855 (1997).
138. McGovern, S., Pan, J., Oliver, G., Cutz, E. & Yeger, H. The role of hypoxia and neurogenic genes (Mash-1 and Prox-1) in the developmental programming and maturation of pulmonary neuroendocrine cells in fetal mouse lung. *Laboratory Investigation* **90**, 180-195 (2010).
139. Keith, I.M., Pelto-Huikko, M., Schalling, M. & Hökfelt, T. Calcitonin gene-related peptide and its mRNA in pulmonary neuroendocrine cells and ganglia. *Histochemistry* **96**, 311-315 (1991).
140. Kassambara, A. & Kassambara, M.A. Package 'ggpubr'. *R package version 0.1* **6** (2020).
141. Nikolic, M.Z. *et al.* Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *Elife* **6** (2017).
142. Bologna-Molina, R., Mosqueda-Taylor, A., Molina-Frechero, N., Mori-Estevez, A.D. & Sanchez-Acuna, G. Comparison of the value of PCNA and Ki-67 as markers of cell proliferation in ameloblastic tumors. *Med Oral Patol Oral Cir Bucal* **18**, e174-179 (2013).