HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2021 September 12.

Published in final edited form as:

Science. 2021 March 12; 371(6534): . doi:10.1126/science.abc3172.

The genomic, epigenomic, and biophysical cues controlling the emergence of the lung alveolus

Jarod A. Zepp[1,2,3,7,*], **Michael P. Morley**[1,2,3], **Claudia Loebel**[5], **Madison M. Kremp**[1,2,3], **Fatima N. Chaudhry**[4], **Maria C. Basil**[1,2], **John P. Leach**[1,2,3], **Derek C. Liberti**[1,2,3,4], **Terren K. Niethamer**[1,2,3], **Yun Ying**[1,2,3], **Sowmya Jayachandran**[6], **Apoorva Babu**[1,2,3], **Su Zhou**[1,2,3], **David B. Frank**[2,3,6], **Jason A. Burdick**[5], **Edward E. Morrisey**[1,2,3,4,*]

[1]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

[2]Penn-CHOP Lung Biology Institute, Perelman School of Medicine, University of Pennsylvania, PA, USA.

[3]Penn Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

[4]Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

[5]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA.

[6]Division of Pediatric Cardiology, Department of Pediatrics, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, USA.

[7]Division of Pulmonary Medicine, Department of Pediatrics, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, USA.

Abstract

The lung alveolus is the functional unit of the respiratory system required for gas exchange. During the transition to air breathing at birth, biophysical forces are thought to shape the emerging tissue niche. However, the intercellular signaling that drives these processes remain poorly understood. Applying a multimodal approach, we identify alveolar type 1 (AT1) epithelial cells as a distinct signaling hub. Lineage tracing demonstrates that AT1 progenitors align with receptive, force-exerting, myofibroblasts in a spatial and temporal manner. Through single cell chromatin accessibility and pathway expression (SCAPE) analysis, we demonstrate that AT1-restricted ligands are required for myofibroblasts and alveolar formation. These studies show that the

[*]Co-Corresponding authors: zeppj@email.chop.edu, emorrise@pennmedicine.upenn.edu.

Author contributions:

E.E.M. and J.A.Z. were responsible for conceptualization, investigation, data curation and analysis, methodology, validation, visualization and writing and review. C. L. and J.A.B. performed experiments and assisted with data analysis. M.P.M. and A.B. performed biostatistical analysis of the single-cell sequencing datasets. M.C.B. and D.B.F. performed experiments and provided peripheral lung pediatric tissue for single-cell sequencing. M.M.K., F.N.C, J.P.L, D.C.L., T.K.N., Y.Y., S.J. performed experiments and assisted with data analysis.

Competing interests: None to declare.

alignment of cell fates, mediated by biophysical and AT1-derived paracrine signals, drives the extensive tissue remodeling required for postnatal respiration.

Abstract

Introduction: The alveolar region of the lung develops in a period spanning the late embryonic and early post-natal stages of life. An intricate series of events including cellular proliferation, surfactant production and morphogenesis of the alveolar structure occur during the transition to air breathing, a process known as alveologenesis. This critical period establishes the spatial arrangement of alveolar epithelium, capillary endothelium and fibroblasts to generate the gas-exchange niche.

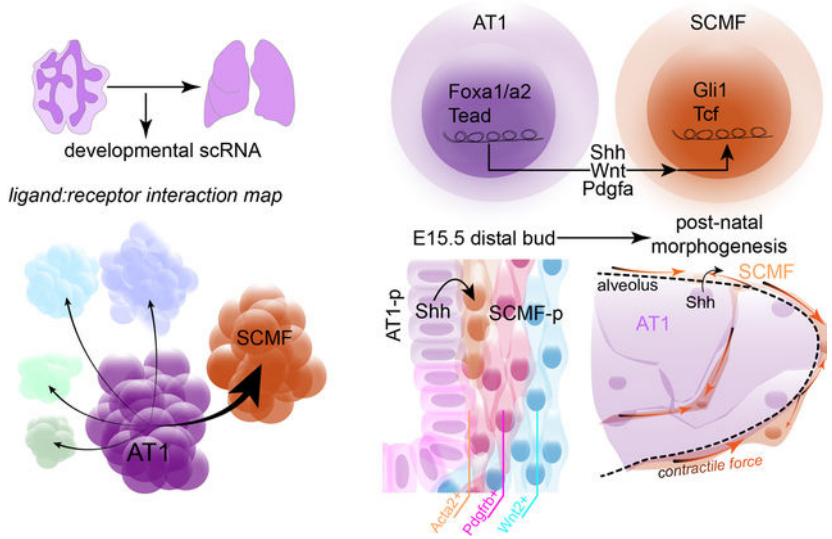
Rationale: The temporal and spatial alignment of cell compartments and the intercellular signaling that coordinates the development and maturation of the lung alveolus remain poorly characterized. Moreover, due to the extensive morphological changes that shape the alveolar niche, it is unclear the extent to which subsets of alveolar mesenchyme could exert mechanical force to promote correct alveolar architecture during early postnatal development. Single cell sequencing and new genetic lineage tracing tools have helped elucidate the cellular heterogeneity in all three cellular compartments in the lung, with extensive heterogeneity in the lung mesenchyme of particular note. We sought to integrate these new technologies and tools to assess the cell-cell communication that drives alveolar generation.

Results: We generated a single cell RNA-sequencing (scRNA-seq) atlas of the developing mouse lung which included epithelial, endothelial and mesenchymal compartments from seven timepoints spanning embryonic and post-natal stages. We then analyzed predicted ligand and receptor interactions and identified the alveolar type 1 epithelial cell (AT1) as a unique and highly enriched hub of ligands. The expression of cognate receptors for these ligands were restricted in subsets of developing mesenchymal cells. Mesenchymal progenitors are spatially and transcriptionally segregated into *Acta2*, *Pdgfrb* or *Wnt2* expressing subsets. By embryonic day E15.5 the mesenchymal progenitors are committed to generating distinct fibroblasts in the post-natal lung. We show by scRNA-seq and lineage tracing that the progenitors for the transient secondary crest myofibroblast (SCMF), which exists only during the early post-natal alveolarization period of lung development, is spatially and transcriptionally aligned with the AT1 cell progenitors. Compared to other alveolar fibroblasts we show that purified SCMF exert significantly more traction force *ex vivo*, indicating that they are a functionally specialized fibroblast lineage that can remodel the alveolus. Next, to identify intercellular signaling pathways that regulate cell lineage identity, we examined the single cell chromatin accessibility and pathway expression (SCAPE) of the AT1 and SCMF. We identify Foxa and Tead transcription factors as upstream regulators of several AT1-derived ligands including Shh and Wnt ligands. Conversely, SCMFs exhibit open chromatin with predicted Gli1 and Tcf target genes, implicating Shh and Wnt pathways in their development and function. To test these pathways *in vivo*, we generated AT1-cell specific conditional knockouts for Wnt-ligand secretion (Wls) and Shh. Conditional ablation of Shh from AT1 cells results in a loss of SCMF cells and subsequent alveolar simplification in the post-natal lung.

Conclusion: Our studies integrated single-cell genomics and genetic lineage tracing, to identify the spatial and temporal patterning of intercellular signaling pathways that are active during the development and maturation of the distal lung. A key finding from this work is that the AT1

epithelial cell, previously thought to primarily provide the thin gas-diffusible interface with the capillary endothelium, is also a crucial ligand expressing node required for proper lung development. These observations suggest that maintaining the viability of the AT1 cell is paramount to establish tissue homeostasis during lung development. Our whole-mount imaging and single cell biophysical measurement assays demonstrate that AT1 adjacent mesenchymal progenitors occupy anatomically discrete regions and are functionally specialized to mold the intricate architecture of the lung alveolus. These aspects of mesenchymal biology add to growing evidence underscoring the transcriptional and functional heterogeneity that underlies distinct fibroblast lineages.

Graphical Abstract



AT1 cells provide intercellular cues for alveolar development. A lung single cell developmental atlas reveals an enriched signature of ligand expression in alveolar type 1 (AT1) epithelial cells. Chromatin accessibility and whole-mount imaging identify the transcriptional and spatial alignment of AT1 and the secondary crest myofibroblast (SCMF) progenitors in the developing lung. The AT1-derived Shh signal is required for specification and outgrowth of the force-exerting SCMF.

Introduction

The development of the lung occurs during both embryonic and post-natal stages and is defined by multifarious processes including cell specification, proliferation and tissue morphogenesis. Embryonic development establishes the myriad of cell-types that make up the lung epithelium, endothelium and mesenchyme, including fibroblasts and smooth muscle cells. The processes of branching morphogenesis orchestrates the outgrowth and proximal-distal patterning of the airway and vascular tree. The late embryonic and early post-natal stages of lung development is when the alveolar region of the distal lung is refined to perform the central function of the lung, gas-exchange. Post-natal structural remodeling generates alveolar septa and arranges the alveolar cell lineages to create an efficient gas-exchange unit. Despite the importance of these final stages of organ maturation, there is a

paucity of information regarding how these steps are accomplished and what cellular, genomic, and biophysical drivers regulate them.

The lung alveolus facilitates the movement of air and exchange of gases between the circulatory system and the external environment. During branching morphogenesis, the distal lung endoderm commits to two functionally divergent alveolar epithelial cell fates, alveolar type 1 (AT1) or alveolar type 2 (AT2) cells (1, 2). While the AT2 cells are cuboidal and produce pulmonary surfactant that prevents alveolar collapse during breathing, the AT1 cells are thin and elongated to provide the thin diffusible surface for gas exchange. The alveolus also contains mesenchymal and endothelial cell lineages that are only partially defined in their development and functions in the mature organ. Thus, how the different cell types that occupy the alveoli cooperate to generate the expansive alveolar surface area remains poorly understood. Since the lung alveolus is a major site for damage caused by pathogens such as influenza and the novel coronavirus (SARS-CoV2 that causes COVID-19), a global understanding of how this tissue is formed and refined will provide insight for promoting its repair after acute injury.

We have utilized a multimodal approach combining single-cell genomic assays, cell type specific lineage tracing, single cell biophysical experiments and genetic pathway inactivation, to investigate the intercellular relationships that drive the emergence and formation of the alveolus. These studies have revealed that the AT1 epithelial cell is a signaling hub that coordinates mesenchymal cell development, especially during the transition to air breathing. Using lineage tracing, we define the emergence and genetic programming of multiple mesenchymal lineages including the transient secondary crest myofibroblast (SCMF), that promote formation of the alveolar niche in the perinatal period. Single-cell biophysical assays show that SCMFs preferentially drive the remodeling process, which is regulated in part by AT1-sourced Shh and Wnt ligands. Our study highlights the emerging paradigm that architectural refinement of tissue maturation and function occurs immediately before and after birth and is controlled by a combination of genomic, epigenomic, and biophysical cues.

Results

Transition to air scRNA-seq captures multi-compartment cell fate changes

To define cell-cell signaling that promotes the emergence of the alveolus we performed whole-lung single-cell RNA sequencing at multiple stages of mouse lung development (Fig. 1A). Concatenating this scRNA-seq dataset reveals segregation of transcriptionally distinct cell compartments (Fig. 1B, Suppl. Fig. 1–3) (3). Graph-based clustering identified the various alveolar cell types in the epithelial, endothelial and mesenchymal compartments (Fig. 1B). To further segregate inter-compartment cell subsets, we extracted the cells from each compartment and re-clustered the data (Fig. 1C–E). Dimension reduction was performed using the diffusion map (DM) algorithm, which computes gene expression transitions in an unsupervised manner (Fig. 1 f–h). These data show that most of the epithelial and endothelial cell types derived from single nodal points of early specification events that terminated with mature cell types (labeled). The mesenchymal DM revealed

numerous transitional cell nodes terminating in smooth-muscle cell types (ASM, VSM) and non-smooth-muscle Wnt2-Pα and MANC peripheral mesenchyme lineages (Fig. 1 h) (4).

Individual timepoints across each compartment revealed that following the transition to air-breathing, multiple putative cell types arise with unique gene expression profiles (Suppl Fig. 1–3). Using this single cell transcriptome analysis, we generated fate-circuit tree schematics to illustrate these emerging and transitional cell types (Fig. 1i–k). In the epithelium, a cluster of *Sftpc*/*Spock2*/*Hopx*-expressing cells (AT1/AT2) arise at P3 and their gene expression signature suggests that they may be a transitional state of AT2 cells and may be similar to Spock2+/Axin2+ AT2 cells described previously (Suppl Fig. 1) (5). In the endothelium, the air transition datasets reveal a distinct cell subtype that is defined by *Car4* expression. These cells are restricted to the alveolar region exclusively after the transition to air and have recently been suggested to play a role in the maturation of the capillary plexus and in adult alveolar homeostasis (Suppl. Fig. 2 and Suppl. Fig. 4) (6–8).

The mesenchyme exhibits two new clusters of transcriptionally distinct cells beginning at P3, the secondary crest myofibroblast (SCMF) and the mesenchymal alveolar niche cell (MANC) (4, 9). *In situ* assays for RNA expression reveal anatomic restriction of mature mesenchymal subsets that are acquired as lung development progresses (Suppl. Fig. 3 and Suppl. Fig. 4). Of note, the integrated scRNA-seq time course dataset shows that the transcriptional signature of SCMFs disappears by the P15 and P42 timepoints (Suppl. Fig. 3). Numbering the transitions across each compartment reveals that the mesenchyme compartment undergoes at least seven cell transition states, compared to the epithelium or endothelium with four to five each (Fig. 1i–k). Together, our integrated time course dataset of the lung periphery reveals emergent cell types during the transition to air breathing and highlights the dynamic transcriptional states present in the mesenchymal compartment at this time.

AT1 cells are a node of paracrine signals communicating with the mesenchyme

In order to identify the intercellular signaling molecules that are temporally regulated during lung development, we performed differential gene expression in each cluster at all timepoints separately. After this we annotated transcripts for ligands and receptors and counted the number of predicted interacting pairs (Suppl. Fig. 5). These data showed two important trends: 1) the number of interactions across individual timepoints indicated ‘hot-spots’ of activity that increased over time, and 2) these data highlighted the alveolar epithelium as a source of ligands that signal to the mesenchyme (Fig. 2a and Suppl. Fig. 5). The number of predicted ligand-receptor interactions reveals that AT1 cells exhibit a greater number of total interactions compared to AT2 cells or airway epithelium (Fig. 2b–c). The ligands expressed by AT1 cells include Wnt ligands (*Wnt7a* and *Wnt3a*), *Shh*, *Pdgfa* and *Vegfa* (Fig. 2d). Using fluorescent mRNA *in situ* hybridization, we confirmed AT1-restricted (*Aqp5*+ cells) expression of several of these ligands by post-natal day 3 (Fig. 2d–g). AT1-restricted expression of these ligands increased by approximately 85% after the transition to air at P3 (Fig. 2h). Cognate receptor expression was detected in mesenchymal cell types at P3, including the SCMF cells, which express receptors for Wnts (*Fzd1*), Shh (*Smo*, *Ptch1*) and Pdgfs (*Pdgfra*, *Pdgfrb*) (Fig. 2d, right).

The AT1 signaling node is conserved in the human neonatal lung

These mouse data highlight the AT1 cell as a signaling node during the emergence of the alveolar gas exchange niche. We next assessed whether the AT1 signaling node was conserved in the human lung. scRNA-seq was performed on a 60 day old human patient-derived sample of peripheral lung. Previously known and some new putative cell-types within the epithelial, endothelial and mesenchymal compartments were identified (Fig. 2i). Notably, the transcriptional heterogeneity within the mesenchymal and epithelial compartments bears a high degree of similarity to that of the murine P3 data set based on clustering and marker gene expression (Suppl. Fig. 6). AT1 cells in the neonatal human lung are enriched for expression of the same ligands found in the mouse AT1 signaling node including *WNT7A*, *SHH*, *PDGFA* and *VEGFA* (Fig. 2j–k). Thus, our mouse and human lung developmental scRNA-seq datasets suggest that AT1 cells represent a signaling node in the developing alveolus and this function is evolutionarily conserved.

Temporal ontogeny of mesenchymal lineages including the SCMF

Since the AT1 cells express ligands that may impact development of the post-natal mesenchyme, we next used the scRNA-seq as a guide to infer mesenchymal differentiation trajectories. Diffusion map reduction was performed on E12.5 mesenchyme to define the temporal ontogeny of the various mesenchymal cell types (Fig. 3a). DM showed that *Wnt2* cells represented a single node, with *Acta2* or *Pdgfrb* expressing cells at the ends of the computed curves, suggesting that at E12.5 Wnt2-cells are multipotent (Fig. 3a–b). Lineage tracing of Wnt2+ cells starting at either E12.5 and examined at E17.5, shows that they are multipotent at E12.5, consistent with our previous work (10) (Fig. 3b). Epithelial and endothelial compartments were assessed at E12.5. The epithelium consisted of proximal-distal encoded epithelial progenitors (expressing *Sox2* and *Sox9*, respectively), consistent with findings from developmental studies (Suppl. Fig. 1) (2, 11–14). The endothelial compartment consisted of two distinct cell types, an *Aplnr*+/*Pecam1*+ and *Prox1*+/*Ccl21a*+ corresponding to a common vascular endothelial progenitor and a common lymphatic endothelial progenitor, respectively (Suppl. Fig. 2).

In situ hybridization showed that *Acta2* expression was anatomically localized around airways and blood vessels in the E12.5 lung tissue (Suppl. Fig. 4). *In situ* hybridization and lineage tracing revealed that *Acta2* expressing cells at E12.5 contributed almost exclusively to the ASM or VSM and were αSMA positive (Fig. 3c). Expression of *Pdgfrb* using i*n situ* hybridization revealed expression in a peri-bronchiolar or peri-vascular localization (Suppl. Fig. 4). Lineage tracing demonstrated that Pdgfrb+ cells developed into VSM or resided within close proximity to airways and blood vessels by E17.5 (Fig. 3d). These data reveal that embryonic Wnt2+ cells are multipotent whereas Acta2+ and Pdgfrb+ cells are committed to specific architectural niches during lung development (Fig. 3e).

We assessed the putative developmental ontogeny of the SCMF using the integrated scRNA-seq data and DM analysis (Fig. 3f–h). Using a combination of DM subsetting with lineage tracing, we tested whether Wnt2 expressing progenitors at E15.5 generate the SCMF lineage. Trajectory analysis from the DM embedding of Wnt2 to SCMF cell states predicts a low number of cells with a transitional state similar to the SCMF or ASM (Fig. 3i). Lineage

tracing at E15.5 with the *Wnt2*^{CreERT2} indicated that whereas Wnt2+ cells generated predominately Pdgfrα^{+} cells, they did not contribute to the SCMF lineage (Acta2+/Pdgfra+) (Fig. 3j). Flow cytometry revealed that lineage traced Wnt2+ cells contributed to the Pdgfrα$^{Dim(GFP)}$ (non-SCMF) rather than the Pdgfrα$^{Bright(GFP)}$ cell population (SCMFs) (Fig. 3k).

The next series of experiments tested the ability of myogenic cell types (Pdgfrb+, VSM and ASM) to generate SCMF. By DM analysis, Pdgfrb+ cells (AMP, VSM) showed that a limited number of cells have a transitional gene signature similar to SCMFs (Fig. 3l). This model was then validated by performing lineage tracing experiments using the *Pdgfrb*^{CreERT2} line, which labels VSM and Pdgfra^{-} fibroblasts (Fig. 3m). Pdgfrβ+ cells did not contribute significantly to the SCMF lineage (Fig. 3n). Next, DM analysis of ASM and SCMF cells revealed a continuity of gene expression that suggested a common *Acta2*+ progenitor (Fig. 3o). We confirmed that Acta2+ progenitors cells traced at E15.5 generated committed SCMF cells that were Pdgfra+ at P7 (Fig. 3p–q). These findings are consistent with and confirm previous lineage tracing studies showing that Acta2+ cells can generate SCMFs (15). By integrating the single-cell diffusion of gene expression with experimental validation through the use of lineage-tracing, we found that specification of postnatal mesenchymal lineages including the SCMF occurs around E15.5.

Spatial alignment of mesenchymal progenitors with the distal epithelium at E15.5

Based on the results from our lineage tracing studies and taking into consideration that AT1 cells express multiple ligands that could signal to adjacent mesenchymal cells, we next determined where Acta2+ progenitors spatially align with the developing epithelium. The scRNA-seq diffusion map revealed that from E15.5 to E17.5, AT1 and AT2 progenitors are progressively specified (Fig. 4a). Lineage tracing AT1 progenitors with the *Hopx*^{CreERT2} line (induction at E14.5, harvest at E15.5) showed anatomically distinct cells that bear the AT1 lineage mark versus cells expressing EGFP from the Sftpc^{EGFP} reporter (Fig. 4b). Consistent with these lineage tracing data, we reported previously that AT1 and AT2 cell fate is specified during this time period and in spatially distinct regions (1).

Next, we localized these progenitors and determined where Acta2 expression was observed at E15.5 in the mouse lung using an *Acta2*^{DsRed} reporter line (16). This reporter revealed anatomically restricted expression around the proximal and distal airways (Fig. 4c). Immunostaining confirmed that SCMF Acta2+ progenitors are anatomically located next to RAGE+ AT1 progenitors at E15.5 (Fig. 4c”’). We examined Hopx-expressing cells using the Hopx$^{3XFLAG-GFP}$ reporter crossed onto the Acta2^{DsRed} background. Consistent with the RAGE+ staining we found Hopx+ cells adjacent to the Acta2+ reporter positive cells in the stalk region of the developing epithelium just proximal to the distal tips (Fig. 4d). The *Acta2*^{DsRed} cells in the distal portion of the tissue were negative for Acta2 protein expression (αSMA) and appeared to be radiating out from the stalk region of the branching airway (Fig. 4e–h). Together, these data suggest that committed AT1 precursor cells align with SCMF progenitors. This spatial signaling node may stage the initiating signaling event to orchestrate the committed activity of SCMFs required for postnatal alveologenesis.

The SCMF lineage functions as a transitory architectural remodeler of the alveolus

To test whether SCMFs are a functionally distinct cell lineage, we generated a dual *Acta2*^{DsRed}:*Pdgfra*^{GFP} reporter line to purify these cells. Whole-mount imaging confirmed that this line faithfully reports Acta2+/Pdgfra+ SCMFs at P7 (Fig. 5a). Flow cytometry-based sorting (FACS) was conducted by gating for SCMF (DsRed^{Bright}-GFP^{Bright} double-positive) or Pdgfrα (DsRed^{Dim}-GFP^{Dim}) cells (Fig. 5b). Gene expression was assessed in sorted SCMFs versus Pdgfra^{Dim}+ mesenchyme (Fig. 5c). SCMFs expressed high levels of *Stc1* whereas Wnt2-Pα and MANC marker genes were enriched in the other Pdgfrα+, non-SCMFs, as denoted by expression of *Wnt2* and *Mfap5* (Fig. 5b–c). Next, to define the functional difference between these cell-types at P7, we tested whether these cells had differential capacity to exert mechanical force, which would allow them to drive architectural remodeling of the alveolar niche. Since *in vivo* modeling of this capacity is hampered by the lack of SCMF-specific Cre-driver, we purified SCMFs and Pdgfra+ (Pdgfra^{Dim}) cells, cultured them on a hydrogel (HA) embedded with fluorescent beads, and performed single cell traction force microscopy (TFM) (Fig. 5d) (17–19). TFM experiments showed that SCMFs exerted greater traction force and exhibited a lower cell spread area than the non-SCMF, Pdgfra+ cells in this in vitro assay (Fig. 5d–f). These data suggest that SCMFs are endowed with greater biophysical remodeling capabilities than other alveolar mesenchymal cells including non-SCMF Pdgfra+ cells based on the TFM *in vitro* assay.

The temporal scRNA-seq dataset suggests that the SCMF are a transcriptionally transient cell state or lineage. To interrogate the transient nature of these unique cells, we performed lineage tracing with the *Acta2*^{CreERT2}:*Pdgfra*^{GFP}:*R26R*^{tdT} reporter line starting at P7, confirming that this genetic tool labels SCMFs (Acta2+/Pdgfra+) during alveologenesis (Fig. 5g–h). Next, we traced SCMFs out to P21 and P60 and found that cells bearing the lineage trace were still present in the alveolar compartment but did not express ACTA2 protein (Fig. 5i–l). Induction of *Acta2*^{CreERT2} activity in the adult lung results in almost no cells within the alveolar niche that are traced (Fig. 5m). Quantification by flow cytometry indicates that SCMFs decrease in number significantly by P21 and P60 (Fig. 5n). Consistent with this, genes involved in apoptosis were found to be expressed at higher levels in SCMFs than non-SCMF Pdgfra^{Dim}+ cells (Fig. 5o). These data show that SCMFs are a unique transient, force-exerting cell lineage that reaches peak functional maturation during alveologenesis.

Single-Cell Chromatin Accessibility and Pathway Expression (SCAPE) analysis reveals ligand-receptor coupling critical for alveolar niche remodeling

The integration of scRNA-seq datasets with cell type specific lineage tracing suggests that SCMFs receive inductive cues from an AT1 signaling node to carry out their multi-faceted function in the post-natal lung. To better assess this cellular communication during tissue morphogenesis, we developed an approach that would allow us to identify transcription factors (TF) that are predicted to regulate ligand-receptor signaling between different cell types. Although scRNA-seq data from early post-natal lung (P3) showed enriched transcription factors (TF) genes for each cell type, these data lacked the genomic information to implicate TF target genes (Fig. 6a). Therefore, to identify putative transcriptional regulators of cell-cell signaling activity, we generated single-cell chromatin

accessibility or scATAC-seq from P3 mouse lungs and integrated these data with our P3 scRNA-seq dataset using the Seurat Signac extension (Fig. 6a–c). Using this method, we were able to directly map all cell types in scRNA-seq space with scATAC-seq space (Fig. 6b–c). SCMFs were observed in this integrated dataset, indicating that they are a distinct cell-type with a unique chromatin profile (Fig. 6c). Focusing on promoter enriched open chromatin, we searched for transcription factor motifs and identified transcription factors within each cell-type (Fig. 6d–e). Multiple AT1 enriched TFs were identified including Foxa1 and Tead3/4. Previous studies have shown that Foxa1/2 as well as Hippo pathway components upstream of Tead transcription factors regulate lung epithelial development and AT1 cell development, respectively (20–22). Gene ontology (GO) and pathways enrichment analysis of open chromatin states in promoters of AT1 or SCMFs indicates that the transcriptome of AT1 cells is enriched for genes associated with epithelial cell development while the SCMF transcriptome is enriched for genes associated with Wnt and MAPK pathways (Fig. 6f).

To align the transcriptional profiles of inter-cellular communication and cell-specific DNA accessibility, we developed Single-cell Chromatin Accessibility and Pathway Expression (SCAPE) analysis (flowchart in Suppl. Fig. 7). This bioinformatics pipeline takes the ligand expressing cell-type (in this case AT1s) then filters the data as follows: 1) identifies open and enriched peaks via scATAC-seq, 2) determine whether the putative downstream gene is expressed by scRNA-seq, 3) filters on predicted secreted ligands, and 4) defines the associated TF motifs in the promoters of these ligands and determines if the TF is expressed by scRNA-seq. This allows the user to connect transcriptional regulation to paracrine signaling factors expressed in each cell lineage (Fig. 6g). To define receiving cells for these ligands including the SCMF, cognate receptors are assessed based on scRNA-seq expression, and where transcriptional effectors for these pathways are known, putative target genes based on TF binding sites in open chromatin defined by scATAC-seq are identified in receiving cells (Fig. 6g). GO and Kegg pathway analysis can then be used to predict cellular responses in receiving cells (Fig. 6g, right). These data show that transcription factors, such as Foxa1/2, within the AT1 cells may serve as critical regulators of numerous ligand transcripts that are uniquely expressed in the early postnatal AT1 cell (Fig. 6h). Thus, SCAPE identifies critical pathways emanating from AT1 cells that direct the behavior of SCMFs during postnatal alveolar remodeling.

AT1-derived ligand signaling is required for alveologenesis *in vivo*—Based on the SCAPE analysis, we hypothesized that AT1 cells, through the transcriptional activity of Foxa1/2, Tead and/or Klf family members, express ligands such as *Wnt*, *Shh* and *Pdgfa* that mediate the development of mesenchymal lineages including the SCMF (Fig. 7a). We confirmed expression of these ligands and the responsive genes or receptors in the cognate pathway were expressed in the early postnatal lung (Suppl. Fig. 7).

To test whether AT1 cells are an essential source of ligands *in vivo*, we deleted Wnt-ligand secretion using the Wls^{Flox} conditional allele or Shh via the Shh^{Flox} conditional allele and the AT1 restricted driver $Hopx^{CreERT2}$. Since our imaging data suggested a physical association of AT1 and SCMF progenitors (Fig. 4), we hypothesized that the AT1-derived Wnt or Shh may impact the specification of the SCMF lineage and thus alveolar maturation.

We induced deletion of *Wls* or *Shh* at the time of SCMF specification, E15.5, and examined the accumulation of alveolar-restricted ACTA2+ SCMF cells at post-natal day 5 (Fig. 7b). Efficient deletion of *Shh* and *Wls* was verified by in situ hybridization in AT1 cells at P5 (Suppl Fig. 7). Deletion of *Wls* from AT1 cells resulted in only minor changes in morphology or alveolar structure (diameter) and a slight reduction in the Acta2 expression in the alveolus (Fig. 7c–h). However, deletion of *Shh* from AT1 cells resulted in a significant simplification of the alveoli and a reduction in the amount of Acta2 expression in the alveolus (Fig. 7i–n). These data suggest that AT1-derived Shh ligand secretion is critical for the specification and maintenance of the SCMF cell fate. The deletion of Shh did not change the number of Car4+ (endothelial) cells, and we did not observe any differences in elastin expression, caspase-3 positive cells or abnormal effects on ASM or VSM in the Shh-deficient tissue (Suppl. Fig. 7o–t). Thus, SCAPE allows for the identification of molecular regulators of intercellular communication in the developing alveolar niche.

Discussion

Mammalian adaptation to postnatal life requires the appropriate transcriptional, signaling, and cellular responses for refinement of tissue niches and function. This is particularly important in the lung alveolar niche that forms a functional interface with the external environment, but is also susceptible to pathogens that lead to respiratory infections. Thus, a global understanding of how the cellular constituents are assembled will lead to better therapeutic approaches to promote alveolar repair and regeneration after acute injury. Leveraging a multi-modal single cell, biophysical, and genetic analysis, we have identified the AT1 cell as a previously unappreciated node of intercellular signaling in the alveolar niche. Both mouse and human AT1 cells preferentially express high levels of critical signaling factors such as Shh, Pdgf, and Wnts. Using our integrated SCAPE analysis, we show that this inter-cellular communication node is crucial for the development of a transient mesenchymal lineage called the SCMF, which are also a unique driver of biophysical tissue remodeling in the alveolar niche. Together, our studies highlight the interconnected nature of cellular communications that promotes the emergence and early postnatal remodeling of the alveolar gas exchange niche in the lung and identify the AT1 lineage as a critical hub of this communications network.

During tissue development and remodeling, the spatial restriction of ligand bioavailability designates how heterogenous cell types respond to signals from neighboring cells. Our finding that the AT1 cell lineage acts as a signaling node during perinatal lung maturation suggests that many preconceived processes driving complex cellular remodeling to form tissue niches required for adult function need to be reassessed. AT1 cells are a distinct source of ligands for several important signaling pathways including Wnt, Shh, Pdgf and Vegf. The cognate receptors for these ligands are expressed on mesenchymal cells such as SCMF as well as endothelial cells. This AT1 signaling node is conserved in the neonatal human lung, along with conserved expression of cognate receptors in human SCMFs. The physical architecture of the AT1 cell may allow it to uniquely interact with the multifarious lineages within the lung alveolus. The extensive cellular projections made by AT1s allow it to form intimate associations with other mesenchymal cells, AT2 cells, and capillary

endothelium (6, 23). Thus, the ability to make physical contact with so many other cell types in the alveolar niche may allow AT1s to disperse signaling ligands across a vast area.

Mapping the developmental ontogeny of all mesenchymal, endothelial, and epithelial cells within the distal mouse lung also provided insight into the intimate association of AT1 progenitors with the developing SCMF lineage. Our data shows that Acta2+ progenitors emerging from the ASM at E15.5 lie in close proximity to early AT1 progenitors, which are located between the most distal branch bud and more dedicated airway epithelial progenitors (1). Later in lung development, mechanical signals derived from fetal breathing movements, may influence epithelial cell biology including specification and differentiation (24). During the transition to air breathing, AT1 cells continue to spread and fold to accommodate the increased surface area being generated, in part, by mechanical force exerted by mesenchymal cells (25, 26). The TFM experiments used in our study further refine these models to prove that SCMF are the force-exerting cell subset relative to other Pdgfra-positive fibroblasts. Our experiments show that deletion of *Shh* from the developing AT1 cells is critical for specification and outgrowth of the SCMF lineage. Although elastin expression was unchanged in our knock-out model, we did not evaluate whether the elastin network was disrupted as has been reported in other models (25). Since our *in situ* data reveal a gradient of *Shh* expression in the distal epithelium, we cannot exclude a possible role for AT2-expressed *Shh* in the embryonic stages. A prior study used a Gli1 lineage tracing strategy to monitor Shh-responsive lung mesoderm to show that SCMF precursors may be specified early in lung development (9). Thus, our data along with previous studies indicate that AT1-derived Shh is a crucial factor for mesenchymal identity, in particular the SCMF lineage and alveolar development. Although a recent study suggests that smooth muscle is not necessary for airway epithelial cell development (27), future studies could examine whether mechanical tension-dependent signaling is important for AT1 cell biology in the early postnatal lung.

The development of the SCAPE informatic pipeline allows the assessment of cell-cell signaling using integrated scRNA-seq and scATAC-seq datasets. SCAPE identified several TF-signaling axes that have previously been show to play important roles in lung development including Foxa1/2-Shh (10, 22, 28, 29), and Wnt-Lef/TCF (4, 30–36). These previous studies did not delineate the cell type specific expression of these ligands, receptors, or TFs. These studies also did not explore the integration of these paracrine pathways and the transcriptional readouts, which can affect cell lineage fate and differentiation. Our data suggest that loss of the AT1 signaling node in pediatric diseases could lead to aberrant cell-cell communication within the alveolus, contributing to respiratory insufficiency. Further work is also needed to study the role of AT1 cell derived signals, such as *Shh* identified in our study or *Vegfa* as has been recently reported elsewhere, in adult lung homeostasis and in the context of chronic pulmonary diseases (6, 37).

Our delineation of the progressive formation of the lung alveolus reveals the complexity in the resident cell lineages and their extensive communications with each other. The definition of the AT1 lineage as a signaling node should spur new and more focused attention towards this poorly understood cell type. Moreover, the extensive heterogeneity within the distal lung mesenchymal lineages and the significant alterations in their fate and state across time,

indicate that cross-talk between the various mesenchymal, epithelial, and endothelial lineages will be equally as complex and will require additional work to decipher. Given the ongoing scourge of the COVID-19 pandemic, increased appreciation of the active cell-cell communication that occurs to both form the delicate alveolar structure, and maintain its function during homeostasis or in the face of acute injury may provide new strategies for developing therapies to repair and regenerate this critical tissue niche.

Materials and methods summary

All animal procedures were approved by the University of Pennsylvania Institutional Animal Care and Use Committee. We preformed all lineage tracing, genetic reporter imaging, FACS based cell sorting and conditional knockout experiments in embryonic and early post-natal mice of both sexes. Lineage tracing was initiated by administration of tamoxifen in corn oil via oral gavage (pregnant dams) or intra-peritoneal injection (pups). Single cell suspension of lung tissue was generated through mechanical and enzymatic processing. Murine single-cell RNA sequencing atlas was generated from CD-1 female mice. Single-cell RNA and ATAC libraries were performed using 10x genomics Chromium device and protocols (RNA, V2 chemistry). Processing of the single cell data was conducted using the following software packages; Cell Ranger, Seurat V3 and Signac. All imaging was performed on a Leica SP8 confocal system and image files were processed using FIJI. Tissue from lineage-tracing and reporter were thick-cut (150 μm) sections of lung tissue and images presented are max-projections of Z-series. RNA in situ probes and staining were performed according to the manufacturer's protocol.

Methods

Mice

All mouse experiments were conducted according to the guidelines and adherence to protocols reviewed and approved by the University of Pennsylvania Institutional Animal Care and Use Committee. All mice were housed in normal conditions with regular light cycles as previously described. The following reporter strains used in this study are commercially available from Jackson Laboratories (MA, USA). Rosa26-EYFP (Stock No. 007903), Rosa26-tdTomato (Stock No. 007914), $Pdgfra^{GFP}$ (listed as $Pdgfra^{H2B\text{-}GFP}$ line, Stock No. 007669), $Pdgfrb^{CreERT2}$ (Stock No. 029684), Wls^{Flox} (Stock No. 012888), Shh^{Flox} (Stock no. 004293), $Hopx^{GFP}$ (listed as $Hopx^{3FlagGFP}$, Stock No. 029271), $Hopx^{CreERT2}$ (Stock No. 017606), $Sftpc^{EGFP}$ (Stock No. 028356). The $Wnt2^{CreERT2}$ and $Axin2^{CreERT2}$ lines were generated as described previously (5, 10). The $Acta2^{CreERT2}$ was generously provided by Dr. Pierre Chambon and Dr. Daniel Metzger (38). The $Acta2^{DsRED}$ transgenic reporter line was generously provided by Dr. David Brenner (16). All mouse lines were maintained on an outbred CD-1 strain (Charles River Laboratories).

Tamoxifen induction of cell-lineage tracing

Tamoxifen (Sigma-Aldrich) was dissolved in corn-oil (Sigma-Aldrich) for a stock concentration of 10–20 mg/ml. For embryonic lineage tracing, pregnant dams received a single dose of tamoxifen (100 mg/kg) via oral gavage. For early postnatal induction

timepoints P3 or P7 intraperitoneal administration of tamoxifen was used at a concentration of 100 mg/kg, light pressure was applied at the injection site for at least 15 seconds to prevent tamoxifen leak from the peritoneal cavity.

Histology and Whole-mount tissue processing

Tissues were harvest from euthanized mice, the chest cavity was exposed and the lung was cleared of blood by perfusing PBS into the right ventricle. Lungs were then inflated with 2% or 4% fresh paraformaldehyde under constant pressure of 15–30 cm and allowed to fix overnight. For paraffin embedding and RNA *in situ* analysis, the tissue was washed with PBS and dehydrated to 100 ethanol. For *in situ*, slide processing and probe hybridization was conducted according to the manufacture's protocol on 4% PFA fixed tissue (ACD-Bio). For whole-mount imaging, after PFA fixation the tissue was washed at least three times with PBS. Lung lobes were then embedded in 4% low-melting agarose. After the agarose solidified the embedded tissue was cut at 150 μm using a Vibratome (Leica). Tissue slices were permeabilized using perm-buffer (PBS + 1% Triton-X100) (Sigma-Aldrich) for at least one hour at room temperature. Tissue slices were stained with primary antibody for two days at 4°C in block buffer (PBS + 1% Triton-X100 + 1% BSA). After washing three times in perm-buffer, secondary antibody was incubated with tissue overnight in block buffer. After this, a subsequent wash in perm buffer was performed. This was followed by tissue clearing with ScaleA2 and ScaleB4 (39) plus DAPI for at least one week each. Tissue was mounted onto a coverslip for imaging by confocal microscopy.

Single Cell Suspension and Staining

The protocol for generating a single-cell suspension from murine lung tissue has been detailed previously (4). Briefly, lung tissue was minced with razor blade and brought into a suspension with PBS, Collagenase Type 1, DNase and Dispase, then incubated in a 37°C water bath for 15–45 minutes depending on the age of the tissue (less time was required for younger aged tissue). Flow cytometry was performed on a single cell suspension after ACK lysis to remove red blood cells. Human tissue was collected via an approved IRB#12–009788 to D.B.F. and described previously in (40).

Flow Cytometry and Cell Sorting

Single suspensions were stained with fluorophore conjugated antibodies. Antibodies used in this study are Epcam-APCef780, CD31-PeCy7, CD45-PeCy7 (all from eBiosciences). DAPI (Sigma-Aldrich) stain was used to exclude dead cells. For cell sorting, single cells were gated using doublet-discrimination parameters and cells were collected in FACS buffer (1x HBSS, 2% FBS, 1 mM EDTA, 25 mM HEPES).

Single cell RNA sequencing (scRNA-seq)

For mouse samples a single cell suspension prep was generated as outlined above. Sorting was carried out by gating on Live (DAPI negative), CD45-PeCy7 negative cells. Cells were collected in FACS buffer. Cells were pelleted and counted by trypan blue exclusion. The cell pellet was resuspended in FACS buffer according to the recommendations provide by 10X Genomics, for optimal cell recovery, aiming for 10,000 cells. The cells were then loaded

onto the 10x Chromium (10X Genomics) and libraries were prepared according to the manufacturer's protocol using 10X Single Cell 3' v2 chemistry. Libraries were sequenced on the Illumina HiSeq2500 instrument. Data processing was conducted by aligning reads and obtaining unique molecular identifier (UMI) counts via the Cell Ranger pipeline (10X Genomics). For further processing, integration and downstream analysis, the Seurat (V3, http://satijalab.org/seurat/) package was used (3). Cells with less than 200 genes and greater than 2 Median absolute deviation above the median were removed. Cells with potential stress signals were removed if the percent mitochondrial reads were greater that 5%. Feature (gene) data was scaled in order to remove unwanted sources of variation using the Seurat ScaleData function for percent mitochondrial reads, number of genes detected and predicted cell cycle phase. Non-linear dimension reduction was performed using uniform manifold projection (UMAP) and diffusion map using the Destiny R package (41). Graph-based clustering was performed using the Louvain algorithm and parameters were set empirically by testing top 10 differentially expressed genes between the clustering results. Marker genes were used to aid in the identification of putative cell-types. Trajectory analysis on diffusion map reduction was performed using the R Slingshot package (41, 42). Subsetting was performed by assessing marker gene expression across clusters at each time-point. The subsetted clusters were then selected and then UMAP and clustering was performed. Clusters associated with the following cell-types and marker genes were excluded from our analysis: erythroid cells (*Alas2*+), immune cells (*Ptprc*+), tracheal mesenchyme (*Sox9*+, *Col2a1*+, *Foxg1*+) and mesothelium (*Wt1*+, *Msln*+). Human single cell suspensions for sequencing were prepared as described previously (40).

Single Cell ATAC-seq (scATAC-seq)

Post-natal day 3 lung was harvested as outline above for the scRNA-seq. Nuclei isolation and counting were conducted according to the manufacturer's protocol (10X Genomics, Single Cell ATAC v1 reagents). Data processing such as read filtering and alignment, transposase cut site identification and peak accessibility was conducted using Cell Ranger ATAC pipeline. Data was further processed using the Seurat extension Signac (https://satijalab.org/signac/). Cells were removed if peak fragments were less than 3000, percent reads in peaks less than 15%, a genome black list ratio greater than 0.025 and a mononucleosomal/nucleosome-free ratio greater than 10. Normalization was performed using term frequency-inverse document frequency (TF-IDF) and dimension reduction by running singular value decomposition (SVD) on the TD-IDF normalized matrix. Non-linear reduction is performed similar to scRNA-seq with UMAP and graph-based clustering with the Louvain algorithm. Gene activity score was computed using the chromatin accessibility reads associated with each gene. scATAC-seq and scRNA-seq were integrated using a cross-modality integration and label transfer method to identify shared correlation patterns in the gene activity matrix and scRNA-seq dataset using the FindIntegrationAnchors and IntegrateData functions. Integrated scRNA-seq samples were processed as previously described except were scaled and normalized using the Seurat SCTransform function. Motif activity scores were calculated using R package chromVAR as implemented in Signac (43). Cluster motif enrichment tests were performed in Signac using the FindMarkers function and a logistic regression framework as the test procedure.

Single-Cell Chromatin Accessibility and Pathway Expression (SCAPE)

We have constructed a scale-free directed graph using scRNA-seq and scATAC-seq to create a predicted Cell to Cell communication pathway.

The graph consists of 5 different types of nodes:

1. **Celltype**: Celltype define from scRNA-seq clusters and defined by cell types based on expression of cell specific gene markers.
2. **Transcription factor binding site (TFBS) motifs**: Cell type enriched motifs obtain from scATAC-seq analysis as described previously.
3. **Ligand**: A gene defined as a Ligand in the Famtom5 ligand receptor database and expressed in 20% of the cells in a celltype (44).
4. **Receptor**: A gene defined as a Ligand in the Fantom5 ligand receptor data base and expressed in 20% of the cells in a celltype.
5. **Transcription factor genes**: Genes defined as transcription factors by the AnimalDB and expressed in 20% of the cells in a celltype.

To simplify the graph we only considered the following edges.

1. **Celltype to motif**: relationship between cell type and which motifs are enriched in open chromatin sites.
2. **Motif to ligand**: This relationship defines which transcription factors may be regulating a ligand. This is defined if a ligand is expressed in celltype and the motif is enriched in the promoter region of the gene: The promoter region being defined by 2.5kb upstream and 500 bp downstream.
3. **Ligand to receptor**: This relation is defined using the Fantom5 ligand-rector database.
4. **Receptor to Transcription factor:** relationship between a receptor and Transcription factor using known signal transduction pathway data.

A custom R script utilizing the iGraph R package was used to filter and construct the network data object. The iGraph object was exported and visualized using the Cystoscape application (https://cytoscape.org/)

RNA isolation and quantitative PCR

Cells collected by FACS into FACS buffer were pelleted then lysed with Trizol-LS Reagent (Life Technologies) and RNA was isolated according to manufacturer's protocol. The First Strand IV synthesis system (Life Technologies) was used to make cDNA. Quantitative PCR was performed using 2x Power SYBR green reagents on the QuantStudio 7 Thermocycler (Life Technologies).

Microscopy and Image Processing

All images were acquired on a Lecia TCS SP8 laser-scanning confocal system. Image processing was performed in ImageJ (FIJI). Data presented from whole-mount imaging is

from z-stack max intensity composites. Cell counting was performed using the colocalization and spots detection in Imaris (Bitplane). Alveolar diameter and SMA-volume from whole-mount (fluorescence intensity in 3D) measurements were performed manually in FIJI and Imaris, respectively. For the latter, surface rendering was performed (complexity set at 0.8 μm), enabling split touching of objects around 20μm, airway and vascular smooth muscle was identified and removed manually for the volume statistics.

Single Cell Traction force microscopy (TFM)

Traction force microscopy (TFM) was performed as previously described. For the preparation of hydrogels, the tetrabutylammonium salt of sodium hyaluronate (Lifecore, 64 kDa) was modified by esterification with 5-norbornene-2-carboxylic acid, 4-(dimethylamino)pyridine and ditert-butyl decarbonate, followed by dialysis. Hydrogel films (4wt% polymer, 10.0 kPa ± 1.1 kPa, 100 μm thickness) were ultraviolet-photopolymerized on thiolated coverslips with fluorescent beads (500 nm diameter at 2% vol/vol (Bangs Laboratories FSPP003)) and in the presence of 1 mM thiolated RGD. SCMF and Pdgfrα cells were seeded at 5×10^4 cells/ per well in a 6-well plate and cultured in DMEM F12 media for 48 hours before TFM analysis. Cells and hydrogel-embedded fluorescent beads were captured in fluorescence using an Observer 7 Widefield Microscope (Carl Zeiss). Images were acquired prior and after cell lysis with a buffer containing 10% sodium dodecyl sulfate/1% Triton X-100. ImageJ was used to analyze data, including stack alignment, particle image velocimetry and Fourier transform traction cytometry (Poisson's ratio: 0.45, regularization parameter: 1×10^{-9})

Statistical Analysis and Software

Statistical analysis and graphing was performed in Graphpad Prism 8 for Mac. A two-tailed non-parametric Mann-Whitney U-test was used for comparison between two experimental groups. Data were considered significant if $p < 0.05$. Flow cytometry plots and relative percentages cells were created and analyzed in FlowJo v10 software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

The authors would like to acknowledge the assistance provided by J. Murray and F. Tuluc at the Children's Hospital of Philadelphia (CHOP) Flow Cytometry Core and R. Pellegrino da Silva, F. Mafra, M. Gonzales and J. Smiler of the CHOP Center for Applied Genomics (CAG) Core for assistance with scRNA-seq library preparation and sequencing as well as technical assistance from E. Lu and J. Rurik at Penn. We thank A. Stout and the Cell and Developmental Biology Microscopy Core at Penn. The authors would also like to acknowledge The Ayla Gunner Prushansky Research Fund at CHOP.

Funding:

This work was supported by NIH grants R01-HL087825, U01-HL134745-01, R01-HL132999 and R01-HL132349 to E.E.M. NIH grants 4T32-HL007843 and K99-HL141684 to J.A.Z. and K08-HL140129 to D.B.F.

Data and materials availability:

Single cell sequencing (scRNA-seq and scATAC-seq) are deposited in the Gene Expression Omnibus (GEO) under super series accession number GSE149563. Requests for reagents and animal lines can be sent to E.E.M.

References

1. Frank DB et al., Early lineage specification defines alveolar epithelial ontogeny in the murine lung. Proc Natl Acad Sci U S A, (2019).
2. Rawlins EL, Clark CP, Xue Y, Hogan BL, The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development 136, 3741–3745 (2009). [PubMed: 19855016]
3. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R, Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420 (2018). [PubMed: 29608179]
4. Zepp JA et al., Distinct Mesenchymal Lineages and Niches Promote Epithelial Self-Renewal and Myofibrogenesis in the Lung. Cell 170, 1134–1148 e1110 (2017). [PubMed: 28886382]
5. Frank DB et al., Emergence of a Wave of Wnt Signaling that Regulates Lung Alveologenesis by Controlling Epithelial Self-Renewal and Differentiation. Cell Rep 17, 2312–2325 (2016). [PubMed: 27880906]
6. Vila Ellis L et al., Epithelial Vegfa Specifies a Distinct Endothelial Population in the Mouse Lung. Dev Cell 52, 617–630 e616 (2020). [PubMed: 32059772]
7. Niethamer TK et al., Defining the role of pulmonary endothelial cell heterogeneity in the response to acute lung injury. Elife 9, (2020).
8. Gillich A et al., Capillary cell-type specialization in the alveolus. Nature 586, 785–789 (2020). [PubMed: 33057196]
9. Li C et al., Progenitors of secondary crest myofibroblasts are developmentally committed in early lung mesoderm. Stem Cells 33, 999–1012 (2015). [PubMed: 25448080]
10. Peng T et al., Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature 500, 589–592 (2013). [PubMed: 23873040]
11. Alanis DM, Chang DR, Akiyama H, Krasnow MA, Chen J, Two nested developmental waves demarcate a compartment boundary in the mouse lung. Nat Commun 5, 3923 (2014). [PubMed: 24879355]
12. Rockich BE et al., Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proc Natl Acad Sci U S A 110, E4456–4464 (2013). [PubMed: 24191021]
13. Chang DR et al., Lung epithelial branching program antagonizes alveolar differentiation. Proc Natl Acad Sci U S A 110, 18042–18051 (2013). [PubMed: 24058167]
14. Perl AK, Kist R, Shan Z, Scherer G, Whitsett JA, Normal lung development and function after Sox9 inactivation in the respiratory epithelium. Genesis 41, 23–32 (2005). [PubMed: 15645446]
15. Moiseenko A et al., Origin and characterization of alpha smooth muscle actin-positive cells during murine lung development. Stem Cells 35, 1566–1578 (2017). [PubMed: 28370670]
16. Magness ST, Bataller R, Yang L, Brenner DA, A dual reporter gene transgenic mouse demonstrates heterogeneity in hepatic fibrogenic cell populations. Hepatology 40, 1151–1159 (2004). [PubMed: 15389867]
17. Dembo M, Wang YL, Stresses at the cell-to-substrate interface during locomotion of fibroblasts. Biophys J 76, 2307–2316 (1999). [PubMed: 10096925]
18. Loebel C et al., Cross-Linking Chemistry of Tyramine-Modified Hyaluronan Hydrogels Alters Mesenchymal Stem Cell Early Attachment and Behavior. Biomacromolecules 18, 855–864 (2017). [PubMed: 28146630]
19. Tseng Q et al., Spatial organization of the extracellular matrix regulates cell-cell junction positioning. Proc Natl Acad Sci U S A 109, 1506–1511 (2012). [PubMed: 22307605]

20. Volckaert T et al., Hippo signaling promotes lung epithelial lineage commitment by curbing Fgf10 and beta-catenin signaling. Development 146, (2019).
21. Nantie LB et al., Lats1/2 inactivation reveals Hippo function in alveolar type I cell differentiation during lung transition to air breathing. Development 145, (2018).
22. Wan H et al., Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. J Biol Chem 280, 13809–13816 (2005). [PubMed: 15668254]
23. Yang J et al., The development and plasticity of alveolar type 1 cells. Development 143, 54–65 (2016). [PubMed: 26586225]
24. Li J et al., The Strength of Mechanical Forces Determines the Differentiation of Alveolar Epithelial Cells. Dev Cell 44, 297–312 e295 (2018). [PubMed: 29408236]
25. Branchfield K et al., A three-dimensional study of alveologenesis in mouse lung. Dev Biol 409, 429–441 (2016). [PubMed: 26632490]
26. Li R, Li X, Hagood J, Zhu MS, Sun X, Myofibroblast contraction is essential for generating and regenerating the gas-exchange surface. J Clin Invest 130, 2859–2871 (2020). [PubMed: 32338642]
27. Young RE et al., Smooth Muscle Differentiation Is Essential for Airway Size, Tracheal Cartilage Segmentation, but Dispensable for Epithelial Branching. Dev Cell 53, 73–85 e75 (2020). [PubMed: 32142630]
28. Peng T et al., Hedgehog actively maintains adult lung quiescence and regulates repair and regeneration. Nature 526, 578–582 (2015). [PubMed: 26436454]
29. Wan H et al., Foxa2 regulates alveolarization and goblet cell hyperplasia. Development 131, 953–964 (2004). [PubMed: 14757645]
30. Zacharias WJ et al., Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. Nature 555, 251–255 (2018). [PubMed: 29489752]
31. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ, Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. Science 359, 1118–1123 (2018). [PubMed: 29420258]
32. Kadzik RS et al., Wnt ligand/Frizzled 2 receptor signaling regulates tube shape and branch-point formation in the lung through control of epithelial cell shape. Proc Natl Acad Sci U S A 111, 12444–12449 (2014). [PubMed: 25114215]
33. Kneidinger N et al., Activation of the WNT/beta-catenin pathway attenuates experimental emphysema. Am J Respir Crit Care Med 183, 723–733 (2011). [PubMed: 20889911]
34. Goss AM et al., Wnt2 signaling is necessary and sufficient to activate the airway smooth muscle program in the lung by regulating myocardin/Mrtf-B and Fgf10 expression. Dev Biol 356, 541–552 (2011). [PubMed: 21704027]
35. Goss AM et al., Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev Cell 17, 290–298 (2009). [PubMed: 19686689]
36. Shu W, Jiang YQ, Lu MM, Morrisey EE, Wnt7b regulates mesenchymal proliferation and vascular development in the lung. Development 129, 4831–4842 (2002). [PubMed: 12361974]
37. Raredon MSB et al., Single-cell connectomic analysis of adult mammalian lungs. Sci Adv 5, eaaw3851 (2019). [PubMed: 31840053]
38. Wendling O, Bornert JM, Chambon P, Metzger D, Efficient temporally-controlled targeted mutagenesis in smooth muscle cells of the adult mouse. Genesis 47, 14–18 (2009). [PubMed: 18942088]
39. Hama H et al., Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. Nat Neurosci 14, 1481–1488 (2011). [PubMed: 21878933]
40. Swarr DT et al., Novel Molecular and Phenotypic Insights into Congenital Lung Malformations. Am J Respir Crit Care Med 197, 1328–1339 (2018). [PubMed: 29328793]
41. Angerer P et al., destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics 32, 1241–1243 (2016). [PubMed: 26668002]
42. Street K et al., Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics 19, 477 (2018). [PubMed: 29914354]

43. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ, chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods 14, 975–978 (2017). [PubMed: 28825706]

44. Ramilowski JA et al., A draft network of ligand-receptor-mediated multicellular signalling in human. Nat Commun 6, 7866 (2015). [PubMed: 26198319]

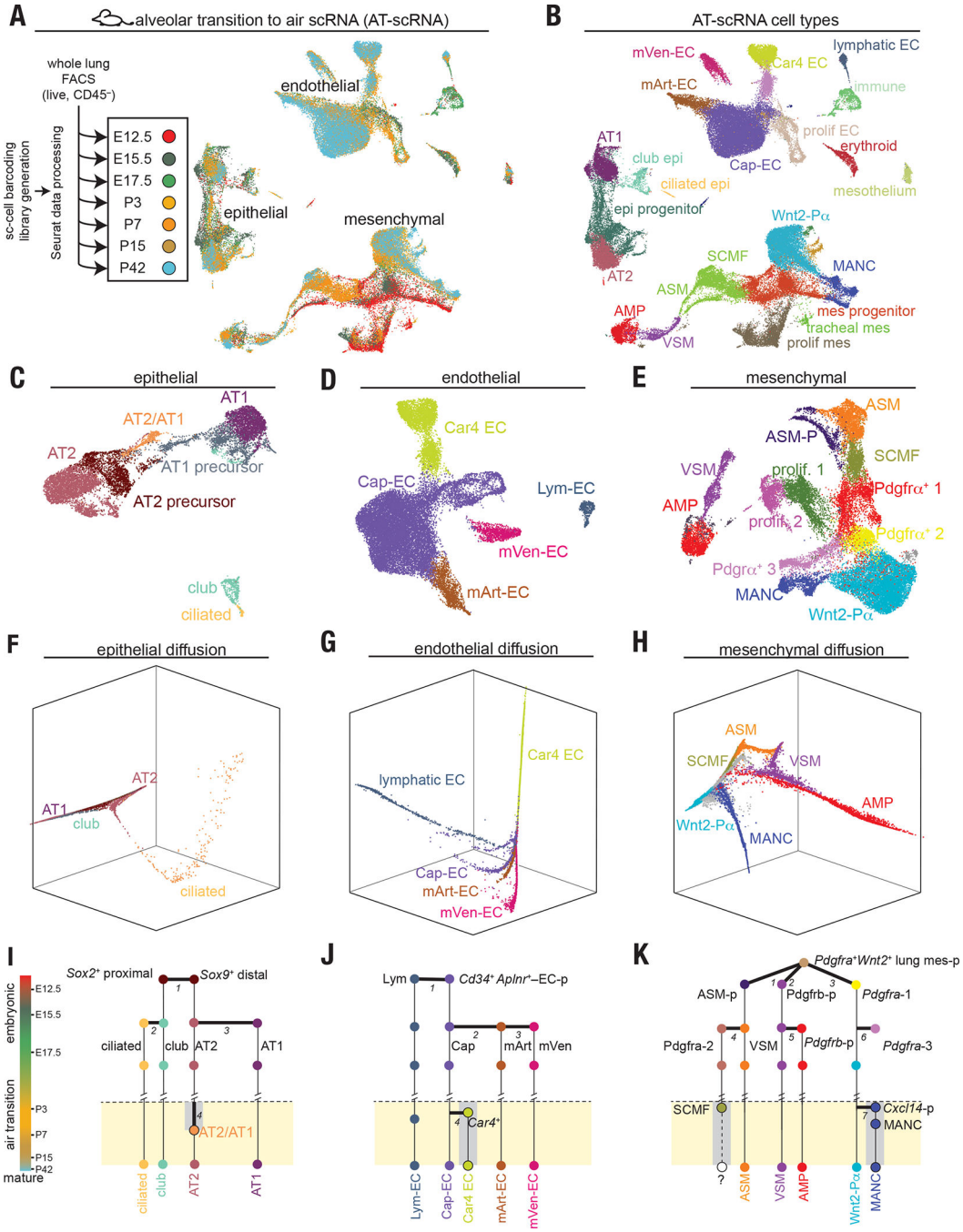


Figure 1. Longitudinal single-cell transcriptomics of the murine lung during the transition to air breathing.

(a) Experimental outline for alveolar transition to air (AT–) whole-lung single-cell isolation and concatenation of seven developmental time-points. (b) Graph-based clustering and cell-type annotation of integrated time-points using Seurat V3 and based on enriched gene expression profiles. (c) Subsetting and clustering of epithelium, (d) endothelium, (e) mesenchyme. (f) Epithelial diffusion map reduction, shown in three-dimensional space. (g) Same method as in (f) but for endothelial compartment. (h) Mesenchymal diffusion map

embedding. (i–k) Model diagrams of cell lineage circuits based on individual time-point UMAP to identify transcriptionally distinct clusters.

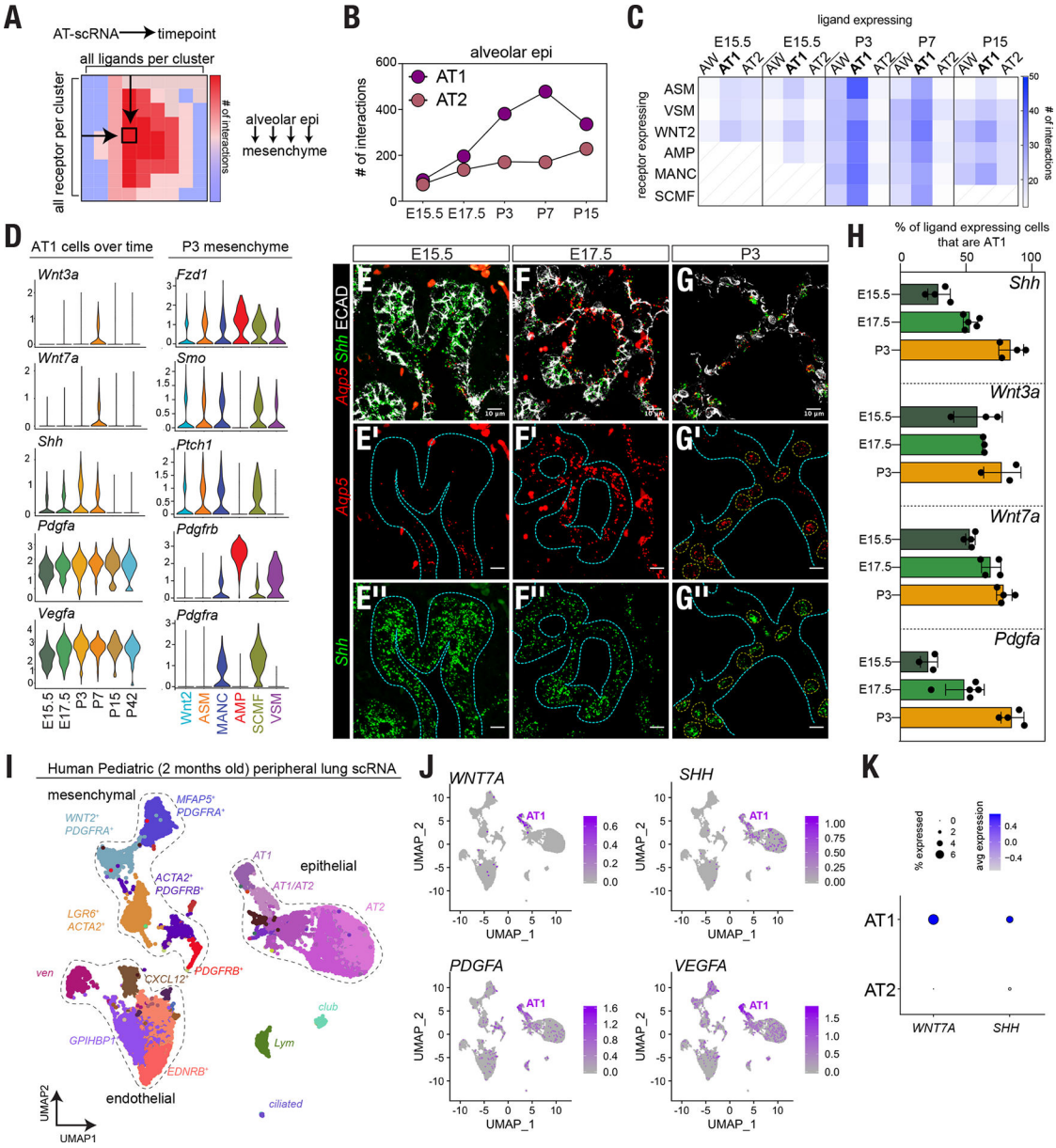


Figure 2. Alveolar type 1 cells are a signaling hub in the mouse and human neonatal lung.

(a) Model of intersectional analysis of each cell-type (cluster) from every timepoint in the AT-scRNA datasets. Number of interactions were plotted to identify enriched alveolar niche derived paracrine and autocrine ligands and receptors. (b) Number of interactions changing over developmental timepoints between AT1 and AT2 cells. (c) Total number of interactions between ligand-expressing epithelium AT1, AT2 and AW (airway, including ciliated/club cells) and receptor-expressing mesenchyme, diagonal lines represent an absence of a cell-type/cluster at the associated timepoint. (d) Expression levels of the indicated ligands in AT1 cells. (right) Expression of cognate receptors for AT1-derived ligands in P3 mesenchyme. (e–g) Fluorescent RNA in situ (FISH) for *Shh* and *Aqp5* together with IHC for E-Cadherin (ECAD) to mark epithelial cell boundaries, showing an AT1 restricted expression profile by P3. (h) Quantification of FISH/IHC for ligands expressed in the AT1 cell lineage. (i) UMAP

of scRNA derived from peripheral lung tissue from a 2-month-old human pediatric sample. Shown are putative epithelial cell-types and identifying marker genes for mesenchymal and endothelial cell compartments. (j) Marker gene expression showing enrichment and expression of ligand transcripts in AT1 cells in the pediatric tissue. (k) DotPlot from subsetted AT1 and AT2 cells derived from the human scRNA. Quantification data in (h) is averaged from at least 3 images derived from 3–5 individual samples, error bars are mean ± SEM.

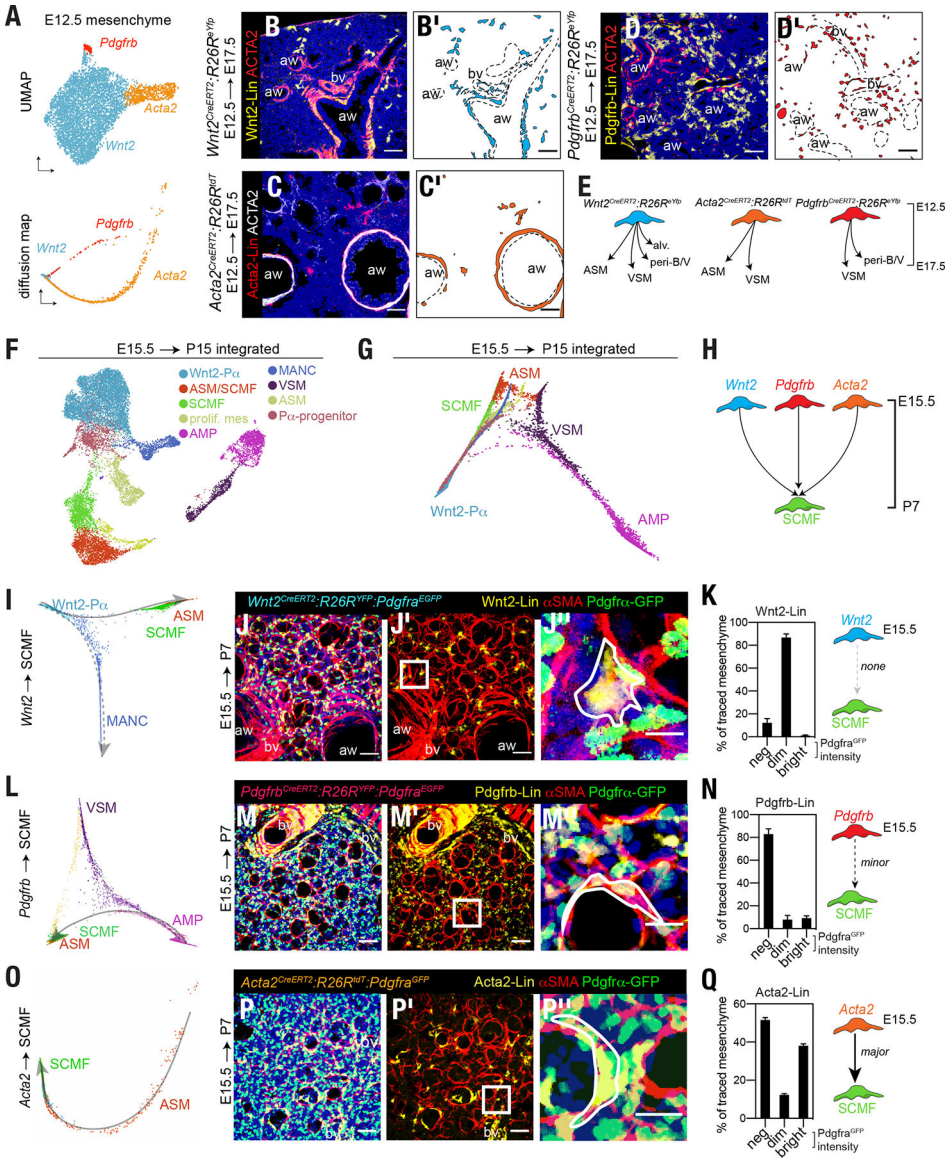


Figure 3. Transcriptional lineage inference and genetic lineage tracing identifies mesenchymal progenitors.

(a, upper) Model of lung embryonic development showing the anatomical localization of multipotent Wnt2-expressing mesenchymal progenitors at E8.5. (a, lower) Sub-selected mesenchyme from E12.5 mouse lung tissue displayed in UMAP or DM as labeled. (b) Whole-mount imaging of Wnt2-lineage analysis on mice induced at E12.5 and harvested at E17.5. (c and d) Similar experimental outline as in (b) lineage tracing *Pdgfrb*– or *Acta2*–expressing cells. (e) Summary of mesenchymal lineage tracing at E12.5. All representative images and shown are from three mice and lineage tracing experiments were performed at least three times. (f) Integrated scRNA from mesenchymal cells derived from the indicated timepoints spanning the transition to air. (g) Diffusion map reduction of the E15.5 to P15 mesenchyme dataset revealing at least three progenitor nodes for the SCMF cell type. (h) Model of the possible progenitors derived from the DM analysis that may contribute to the SCMF lineage that peaks in number at P7. (i) DM trajectory analysis on sub-selected Wnt2/

MANC/SCMF/ASM clusters. (j) Whole-mount images from the indicated mice harvest at P7 show little to no overlap of Wnt2-lineage marked cells producing αSMA protein. (k) Quantification data from flow cytometry for Wnt2-lineage traced $Pdgfra^{GFP}$ expressing cells. (l) DM trajectory analysis on the ASM/VSM/AMP/SCMF clusters. (m–n) Whole mount imaging from the indicated mice harvested at P7 show a limited contribution to the SCMF lineage and quantification by flow cytometry. (o) DM from sub-selected ASM and SCMF shows cells aligned continuously on the trajectory curve. (p–q) Whole-mount imaging and quantification of SCMF progenitors labeled with Acta2 are $Pdgfra^{GFP}$ expressing and αSMA positive. All imaging and quantification data (error bars, mean ± SEM) shown are representative of three independently performed experiments.

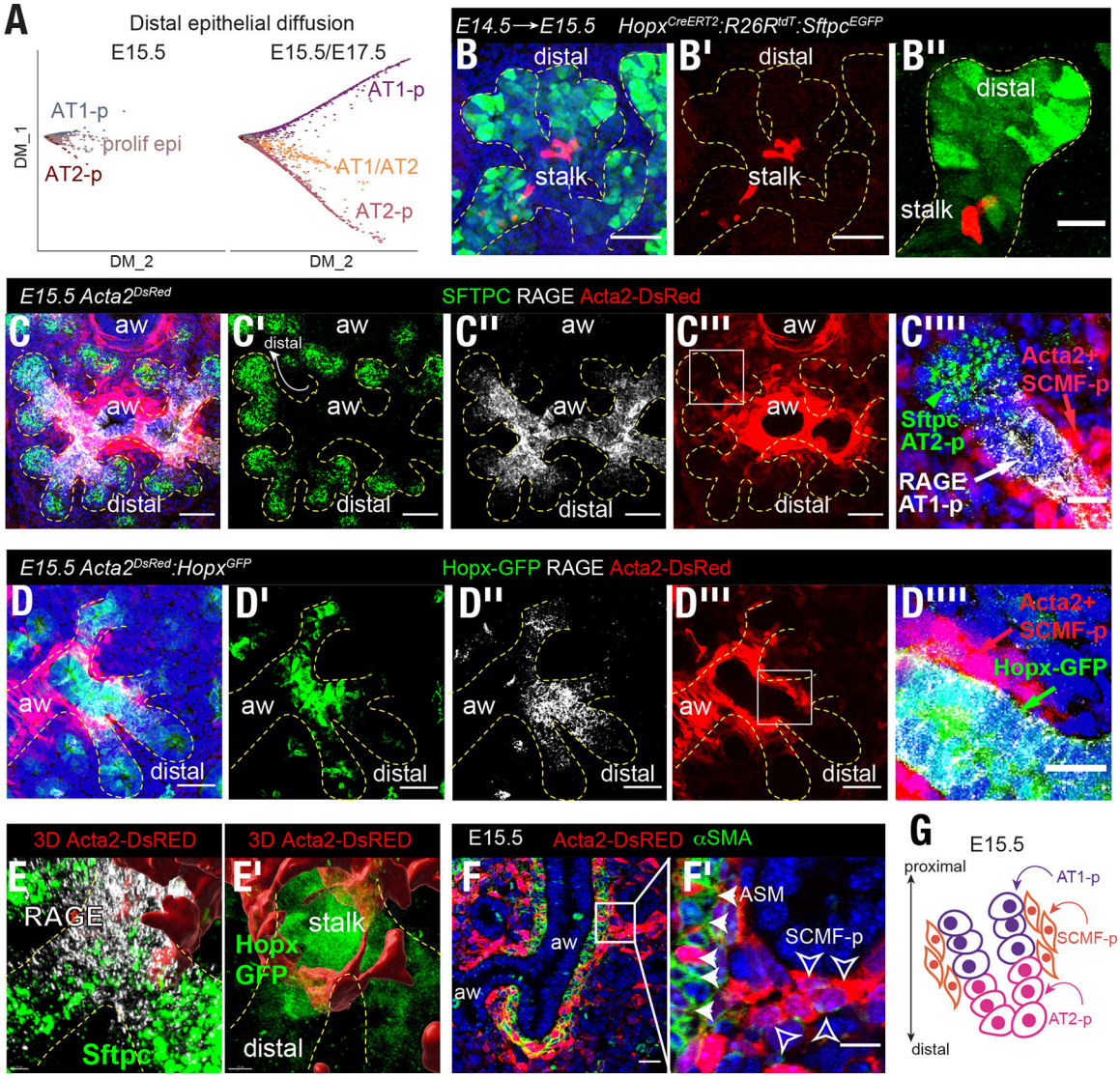


Figure 4. Embryonic AT1-precursors align with Acta2-expressing mesenchymal progenitors.

(a) ScRNA diffusion map of subsetted distal epithelium. Data shown are integrated E15.5 and E17.5 epithelial cells displaying two distinct lineage paths, either AT1 or AT2 cells. (b) Whole-mount lungs from lineage traced AT1-p cells (tdTomato+) showing distinct localization just proximal to the distal Spc-reporter positive (GFP) distal tips. Imaging data are representative max projections of z-series from three embryonic lungs. (c) Localization of Acta2-expressing cells at E15.5 using an *Acta2*^{DsRed} reporter line relative to RAGE or SFTPC protein-expressing AT1-p and AT2-p cells, respectively. (c"') High magnification of Acta2-expressing cells in close proximity to AT1-p cells at E15.5. (d) Localization of Hopx-expressing cells (GFP+) at E15.5 and relation to Acta2-expressing cells. Shown are representative max projection of z-series from 3 embryonic lungs. (e–f) Three-dimensional renderings of Acta2-expressing cells (Red) and SPC–, Rage– and Hopx-Gfp –positive cells. (g) Acta2-expressing progenitors at E15.5 are aSMA-protein negative. (h) Model summarizing the alignment of Acta2– and AT1– progenitors in the distal lung at E15.5.

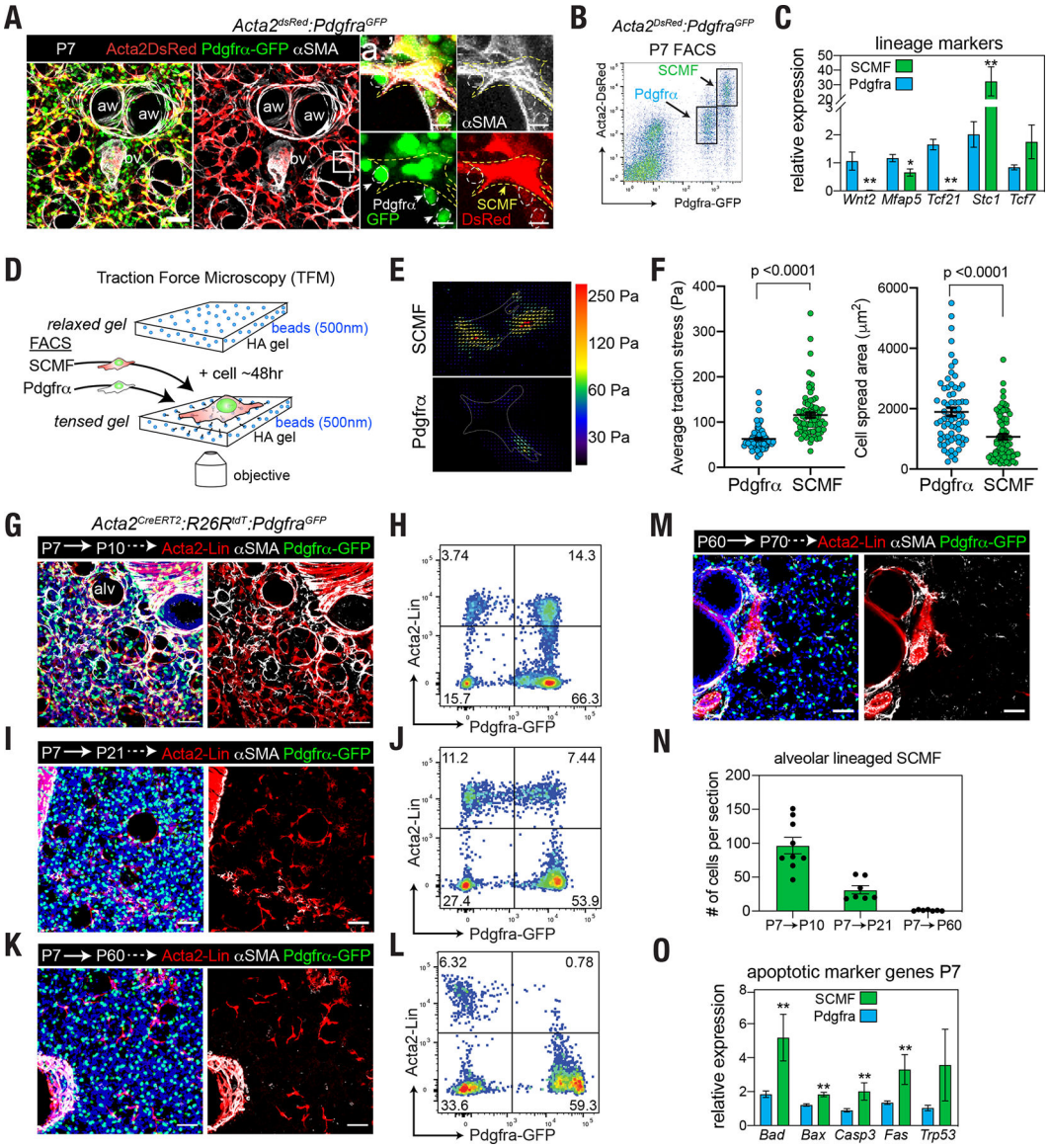


Figure 5. Biophysical characterization of pure SCMFs reveals enhanced force generation during alveologenesis.

(a) Whole-mount imaging from P7 lung tissue from an *Acta2*^{DsRed}:*Pdgfra*^{GFP} mouse to monitor SCMF and other non-SCMF, Pdgfra-expressing lineages as outlined in (a.'). (b) Flow cytometry plot gated on mesenchyme, showing the intensity of Acta2^{DsRed} and Pdgfra^{GFP} expression at P7. Boxes show the cells that were sorted for subsequent experiments. (c) Quantitative PCR was performed from the indicated cells and marker genes for either the SCMF and Pdgfra+ cells. (d) Experimental outline for traction force microscopy (TFM). (e) A representative force generation plot from SCMF and Pdgfra+ cells. (f) Quantification from individual cells in this assay. Data presented are means ± SEM from n=3 independently performed experiments. (g–k) Whole-mount imaging from lineage traced Acta2-expressing SCMF at the indicated times. To the right of the images is a representative flow cytometry plot showing the decrease of Acta2-lineaged Pdgfra+ cells (SCMF) over time. (m) Whole-mount imaging of lineage traced adult mouse lung

demonstrating no alveolar traced cells. (n) Cell counts from the whole mount tissue derived from enumerating lineage traced SCMF. (o) Quantitative PCR for the indicated apoptosis pathway related genes. Data presented in panel c. and o. are means ± SEM from n=6 mice per group. Two-tailed t-test, * $p \leq 0.05$, ** $p \leq 0.01$.

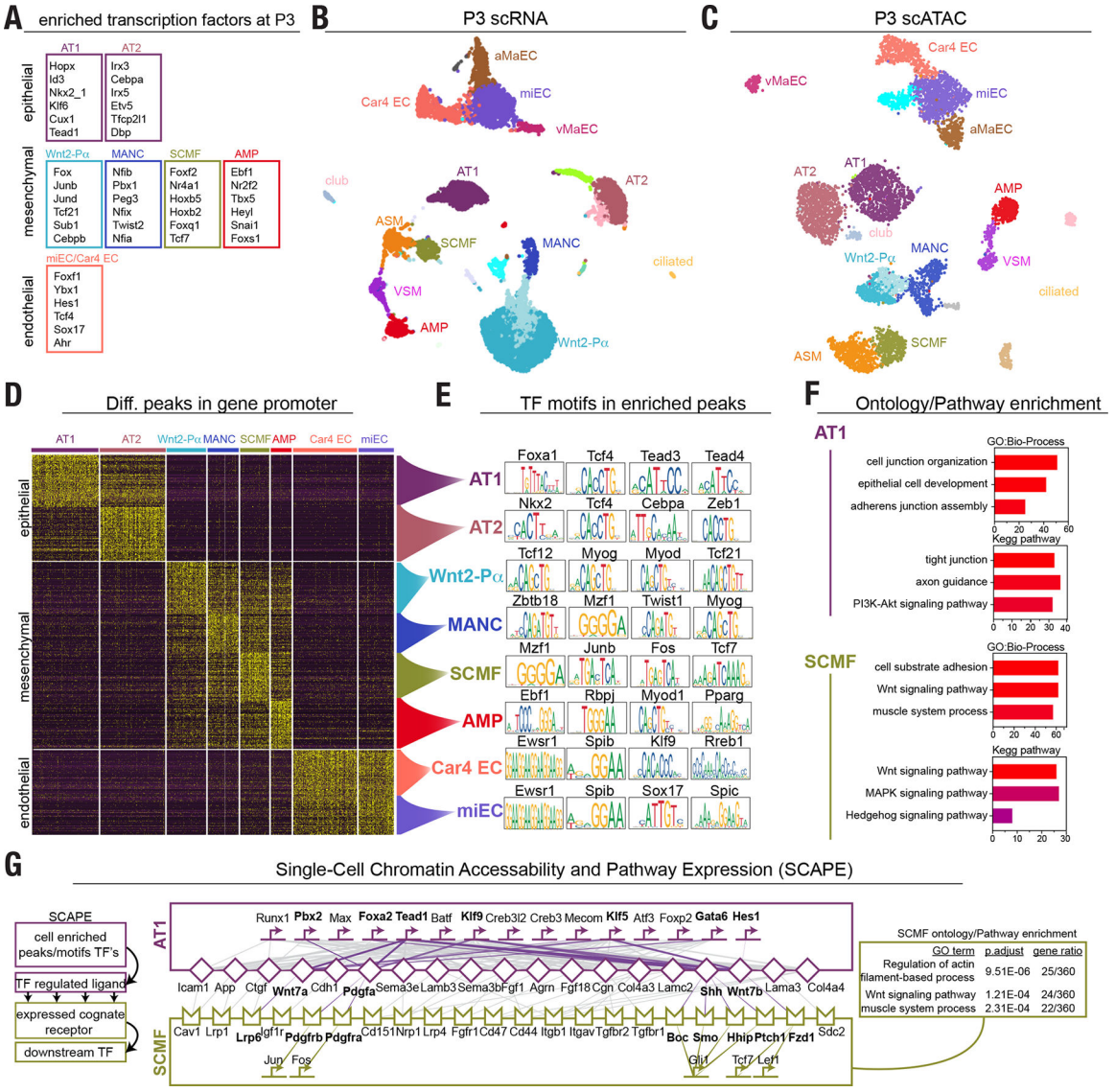


Figure 6. SCAPE analysis reveals cell-intrinsic modes of intercellular communication.

(a) Analysis of enriched transcription factors in cell comprising the alveolar niche that are expressed in the P3 scRNA. (b–c) UMAP of scRNA-seq (b) and scATAC-seq (c) from P3 mouse lung annotated by cell-type. Clusters identified in scRNA-seq were transferred to the scATAC-seq. (d) Heatmap showing peak enrichment based on cell-type, identified peaks in the promoters of genes. (e) Summary of the top transcription factor motifs appearing in the enriched promoter peaks shown in (d). (f) Gene ontology and pathway enrichment from all of the enriched TF's identified in the AT1 and SCMF cells showing canonical functions and pathways. (g) Left, a brief overview of the SCAPE pipeline that links ligand-expressing cell to a receiver cell. Center, cytoscape network showing the connecting nodes derived from AT1 TF to ligand to cognate receptor and informed TF in the receptive SCMF. Right, gene ontology performed on the downstream genes enriched in the scRNA from the expected TF that are shown.

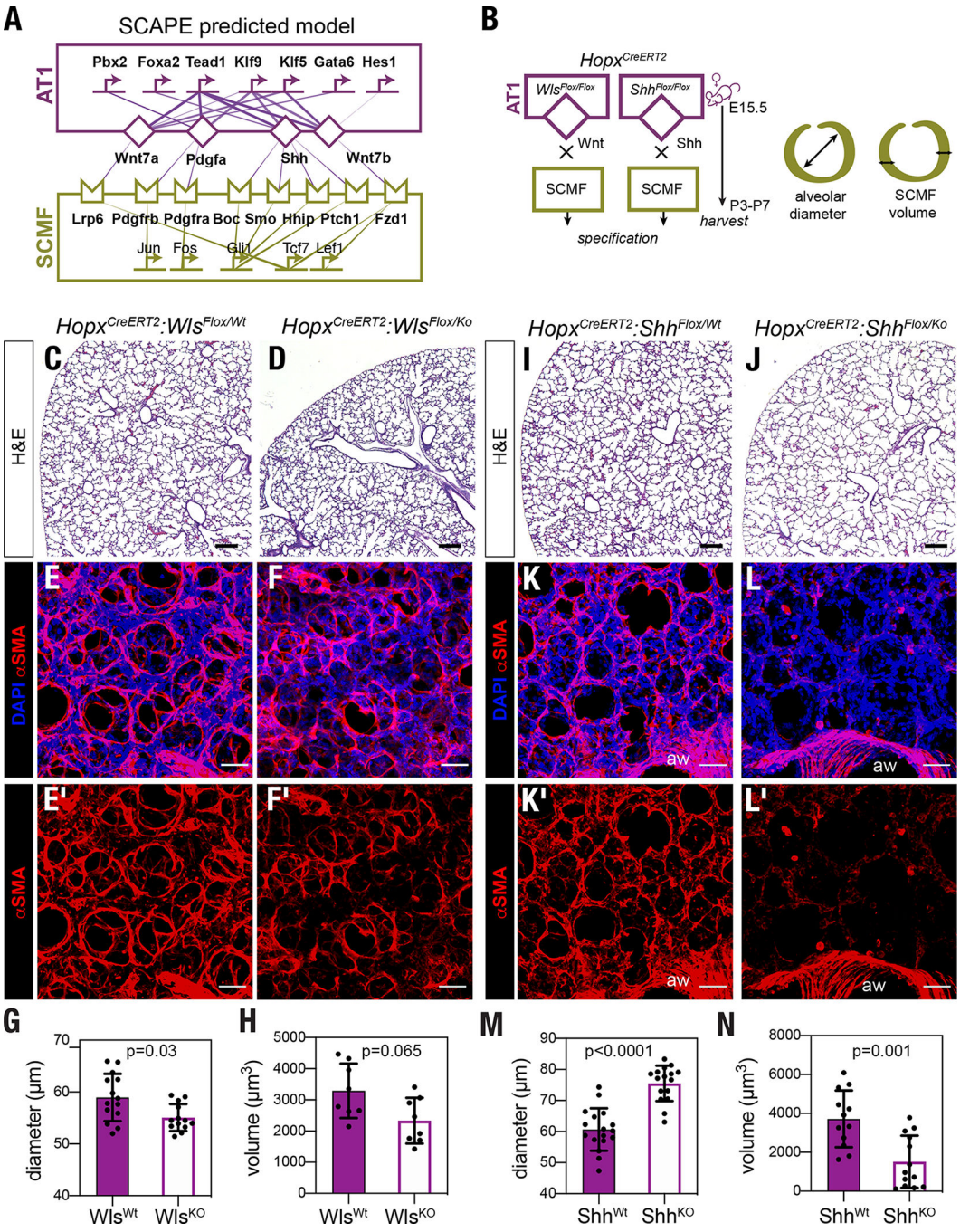


Figure 7. AT1 cells are a critical signaling hub *in vivo* during alveologenesis.

(a) SCAPE prediction model of showing the top TF regulated ligands that are expressed and have cognate receptors expressed in the SCMF. (b) Experimental outline to evaluate whether AT1-derived Wnt or Shh are required for the SCMF cell fate. The AT1-restricted inducible cre ($Hopx^{creERT2}$) is crossed to Wls^{Flox} or the Shh^{Flox} line. Pregnant dams are induced with tamoxifen at E15.5 and the tissue harvested at postnatal day 5 to assess alveolar diameter and volumetric analysis on SCMF-specific ACTA2 staining. (c–d) Representative H&E images of conditional Wls-KO and littermate control lungs. (e–f) Whole-mount imaging

derived from Wls-knockout in AT1 cells. (g–h) Diameter and volumetric analysis of the alveoli and SCMF cells. (i–j) Representative H&E images of conditional Shh-KO and littermate control lungs. (k–l) Whole-mount imaging from AT1-specific ablation of Shh shows enlarged alveoli in the mutant mice. (m–n) Alveolar diameter is increased and the ACTA2+ SCMF is reduced. The data shown from Wls-KO or Shh-KO are derived from four and six knockouts and littermate controls, respectively from three different experiments. The alveolar diameter measurements are averages from 2–3 wholemounts from each mouse and SMA volume measurements are averages from 2–3 wholemounts from 2 different lung lobes from either n=4 or n=6 mutants, data presented as means ± SD, two-tailed t-test.