Title

Expansion of hedgehog disrupts mesenchymal identity and induces emphysema phenotype

Journal

Journal of Clinical Investigation, 128(10)

ISSN

0021-9738

Authors

Wang, Chaoqun
de Mochel, Nabora S Reyes
Christenson, Stephanie A
et al.

Publication Date

2018-10-01

DOI

10.1172/jci99435

Peer reviewed

Expansion of hedgehog disrupts mesenchymal identity and induces emphysema phenotype

Chaoqun Wang,[1] Nabora S. Reyes de Mochel,[1] Stephanie A. Christenson,[1] Monica Cassandras,[1] Rebecca Moon,[1] Alexis N. Brumwell,[1] Lauren E. Byrnes,[2] Alfred Li,[3] Yasuyuki Yokosaki,[4] Peiying Shan,[5] Julie B. Sneddon,[2] David Jablons,[6] Patty J. Lee,[5] Michael A. Matthay,[1] Harold A. Chapman,[1] and Tien Peng[1,7]

[1]Department of Medicine, [2]Diabetes Center, and [3]Bone Imaging Research Core, (UCSF), San Francisco, California, USA. [4]Hiroshima University, Hiroshima, Japan. [5]Yale School of Medicine, New Haven, Connecticut, USA. [6]Department of Surgery and [7]Cardiovascular Research Institute, UCSF, San Francisco, California, USA.

GWAS have repeatedly mapped susceptibility loci for emphysema to genes that modify hedgehog signaling, but the functional relevance of hedgehog signaling to this morbid disease remains unclear. In the current study, we identified a broad population of mesenchymal cells in the adult murine lung receptive to hedgehog signaling, characterized by higher activation of hedgehog surrounding the proximal airway relative to the distal alveoli. Single-cell RNA-sequencing showed that the hedgehog-receptive mesenchyme is composed of mostly fibroblasts with distinct proximal and distal subsets with discrete identities. Ectopic hedgehog activation in the distal fibroblasts promoted expression of proximal fibroblast markers and loss of distal alveoli and airspace enlargement of over 20% compared with controls. We found that hedgehog suppressed mesenchymal-derived mitogens enriched in distal fibroblasts that regulate alveolar stem cell regeneration and airspace size. Finally, single-cell analysis of the human lung mesenchyme showed that segregated proximal-distal identity with preferential hedgehog activation in the proximal fibroblasts was conserved between mice and humans. In conclusion, we showed that differential hedgehog activation segregates mesenchymal identities of distinct fibroblast subsets and that disruption of fibroblast identity can alter the alveolar stem cell niche, leading to emphysematous changes in the murine lung.

Introduction

Emphysema is a major subtype of chronic obstructive pulmonary disease (COPD) linked to tobacco exposure and characterized by progressive loss of alveolar gas exchange surface and enlargement of distal airspace resulting from the loss of septated alveolar tissue (1). However, the observation that only a small subset of smokers eventually develop emphysema gave rationale to explore inherent genetic variations that confer susceptibility to this disease (2). Decades of research on emphysema pathogenesis have implicated diverse processes, such as inflammation, protease imbalance, apoptosis, autophagy, oxidative stress, etc. (3). However, recent large GWAS have identified numerous candidate genes that do not have clear roles in those processes previously identified (4), suggesting alternative mechanisms of pathogenesis. Hedgehog-interacting protein (HHIP) and patched 1 (PTCH1), both key modifiers of the hedgehog (Hh) pathway active during organ development, have been repeatedly implicated in multiple GWAS examining susceptibility to emphysema and lung-function decline (5–11), but the functional relevance of this association remains unclear.

The Hh ligand functions as a morphogen that produces asymmetric activation domains during tissue development to generate diverse transcriptional outputs that translate into cellular diversity, and disruption of these spatial domains produces downstream patterning defects related to the loss of specific cellular position or function (12–14). Interestingly, Hh activation continues to be present in adult organs during normal homeostasis, particularly within the mesenchyme, composed of cells such as fibroblasts, smooth muscle, pericytes, etc., that give rise to the connective tissues within the organs (15–19). In these organs, hedgehog is activated within a spatially restricted subset of the mesenchyme, where compartmental diversity is increasingly appreciated (20–22). Asymmetric Hh activation within the adult mesenchyme suggests that anatomical segregation of Hh-activation domains plays a role in the diversification of mesenchymal cell identity and maintenance of organ function and that disrupting the domain of Hh activation could lead to aberrant tissue remodeling.

The mammalian lung resembles an upside-down tree in which the trachea enters the chest cavity and bifurcates into multiple generations of increasingly narrow conducting airways that ultimately terminate into highly sacculated alveoli where gas exchange takes place (23). Within both the human and murine lungs, the bronchoalveolar ductal junction (BADJ) marks the abrupt transition between the terminal airway and the alveoli where the epithelial composition differs markedly between the 2 compartments (24, 25). Putative stem/progenitor cells have been characterized at the BADJ (26, 27), including the bronchioalveolar stem cells (BASCs), with the capacity to differentiate into bronchiolar (SCGB1A1$^+$) or alveolar (SFTPC$^+$) epithelial stem/progenitor cells (28). Recent studies have suggested that the mesenchymal cells underlying these compartment-specific stem/progenitor cells have a unique capacity to support the neighboring stem/progenitor cells, forming compartment-specific stem

Conflict of interest: The authors have declared that no conflict of interest exists.
Submitted: December 26, 2017; **Accepted:** July 5, 2018.
Reference information: *J Clin Invest*. 2018;128(10):4343–4358.
https://doi.org/10.1172/JCI99435.

cell niches in which epithelial-mesenchymal interactions dictate stem cell behavior (20, 22). We have previously shown that Hh is preferentially activated in the mesenchyme underlying the airway epithelium and less so in the alveoli, with the BADJ marking the transition point where the activation tapers off (29). This suggests that Hh activation may play a role in differentiating the capacities of the stem cell niches of the proximal airway and distal alveoli to support their respective stem/progenitor cell populations through epithelial-mesenchymal feedback. For the rest of this study, we will refer to the airway structures of the mainstem bronchi to the BADJ as the proximal compartment and the alveolated lung distal to the BADJ as the distal compartment.

In this study, we developed a mouse tool to capture and genetically modify Hh-receptive mesenchyme in the adult lung, which is characterized by distinct subsets defined by unique transcriptomes and segregated by proximal-distal localization and levels of Hh activation. Disrupting the asymmetry of Hh through distal expansion of the activation domain resulted in aberrant mesenchymal identity that led to loss of alveolar gas-exchange surface regulated by a mesenchymal feedback loop to the alveolar stem cells. Single-cell analysis of the human lung mesenchyme revealed a conserved segregation of proximal and distal mesenchymal transcriptomes characterized by asymmetric Hh activation. Our study sheds light on how asymmetric developmental pathway activation can serve to segregate cellular identity and maintain organ function in the adult and illustrates how disruption of mesenchymal identity can disrupt the stem cell niche to induce loss of structural integrity, as seen in chronic diseases such as emphysema.

Results

Asymmetric Hh activation in the adult lung mesenchyme. We first determined whether there is evidence for asymmetry of Hh activation within the adult lung, which is functionally and anatomically divided into the proximal conducting airway and the distal alveoli. We previously reported that Sonic hedgehog (*Shh*) ligand expression persists in the adult lung epithelium and activates the underlying mesenchyme (29). Here, we showed that *Shh* is highly expressed in airway proximal (SCGB1A1$^+$) epithelium, with some expression in distal alveolar (SFTPC$^+$) epithelium, by a combination of in situ hybridization and immunofluorescence staining (Supplemental Figure 1A; supplemental material available online with this article; https://doi.org/10.1172/JCI99435DS1). Consistently, when we examined transcriptional effectors of Hh activation, we found that *Gli1*, a transcriptional readout of highly active Hh signaling (30), was preferentially expressed in the mesenchyme surrounding the proximal conducting airway of the adult lung, but largely absent from the distal alveoli (Figure 1A). In contrast, *Gli2*, a mediator of Shh activity whose expression does not depend on Hh activation (31), was expressed both around the proximal airway and in the distal alveoli of the adult lung (Figure 1A). Colocalization with *Pdgfra* demonstrated that *Gli2* was widely expressed in the distal alveolar mesenchymal fibroblasts that lack *Gli1* expression (Figure 1B).

To better isolate and study GLI2$^+$ cells, we generated a *Gli2$^{creERT2\text{-}tdT}$* mouse allele that allows us to isolate GLI2$^+$ cells by endogenous *tdTomato* reporter expression and activate Cre-inducible alleles within GLI2$^+$ cells (Supplemental Figure 1, B and C). Lineage analysis demonstrated that GLI2 only marks mesenchymal cells (Figure 1D and Supplemental Figure 1D). The expression of *Ptch1*, the receptor for SHH ligand, significantly overlapped with the expression of *Gli2* in both the proximal airway and distal alveoli (Figure 1C), which suggests a broad domain of GLI2$^+$ mesenchyme that is receptive to Hh signaling in both the proximal and distal compartments of the adult lung. Of note, the expression of *Ptch1* is also transcriptionally activated by the presence of Hh activation (32) and X-gal staining of the *Ptch1$^{LacZ/+}$* reporter showed much higher intensity in the proximal compared with the distal mesenchyme (Supplemental Figure 1E), suggesting asymmetric activation of Hh within the broad GLI2$^+$ mesenchymal domain.

To determine the degree of overlap between the GLI1$^+$ and GLI2$^+$ mesenchymal populations, we generated a *Gli1:Gli2* dual-color reporter (*Gli2$^{creERT2\text{-}tdT}$:Gli1^{EGFP}*). Flow cytometry analysis of *Gli2$^{creERT2\text{-}tdT}$:Gli1^{EGFP}* lungs digested into single-cell suspension showed that the proximal GLI1$^+$ mesenchyme constituted a subset entirely within a broader GLI2$^+$ mesenchymal population in the adult lung (Figure 1D). Utilizing the *Gli1:Gli2* double reporter, we were able to separate and collect the proximal (GLI1$^+$GLI2$^+$) from the distal (GLI1$^-$GLI2$^+$) mesenchyme for RNA collection and downstream analysis. Quantitative PCR (qPCR) showed that the proximal mesenchyme was enriched in the expression of Hh target genes *Gli1*, *Ptch1*, and *Ptch2* relative to the distal mesenchyme (Figure 1E). These data demonstrated a field of Hh-competent GLI2$^+$ mesenchyme present in the adult lung, characterized by asymmetric Hh activation along the proximal-distal (Hh^{hi}-Hh^{lo}) axis (Figure 1F).

Asymmetric Hh activation promotes anatomical segregation of mesenchymal transcriptomes. To determine whether Hh activation segregates mesenchymal identity within the GLI2$^+$ mesenchyme, we performed single-cell RNA-sequencing (scRNA-seq) on sorted GLI2$^+$ mesenchymal cells from the adult lung. We captured approximately 4,600 cells with a median of 2,246 genes detected per cell utilizing a droplet-based barcoding approach to capture single cells for transcript profiling (33). After filtering for sequencing quality and read depth (Supplemental Figure 2A and Methods), unsupervised graph-based clustering produced 4 distinct clusters (Figure 2A), with clusters 1 and 2 accounting for 27% and 66% of the cells, respectively, while clusters 3 and 4 combined to make up 7% of the total population. Differential expression analysis produced a list of signature genes for each subset (Supplemental Table 1). Immunofluorescence staining of cluster-specific signature genes confirmed that a small fraction of GLI2$^+$ cells contributed to mesothelial (cluster 3) and airway smooth muscle cells (cluster 4) (Supplemental Figure 2B). Clusters 1 and 2 were associated with each other, based on common expression of fibroblast markers, such as *Col1a1*, *Pdgfra*, and *Tcf21* (Supplemental Figure 2C), but they were segregated by markers that were anatomically distinct. Validation of cluster 1 markers demonstrated that they are predominantly expressed in the proximal fibroblasts surrounding the airway, whereas cluster 2 markers were expressed in the fibroblasts in the distal alveoli (Supplemental Figure 2D). While both the proximal and distal mesenchymal fibroblasts have been shown to participate in matrix production and paracrine signaling with surrounding cells, Gene Ontology (GO) analysis of

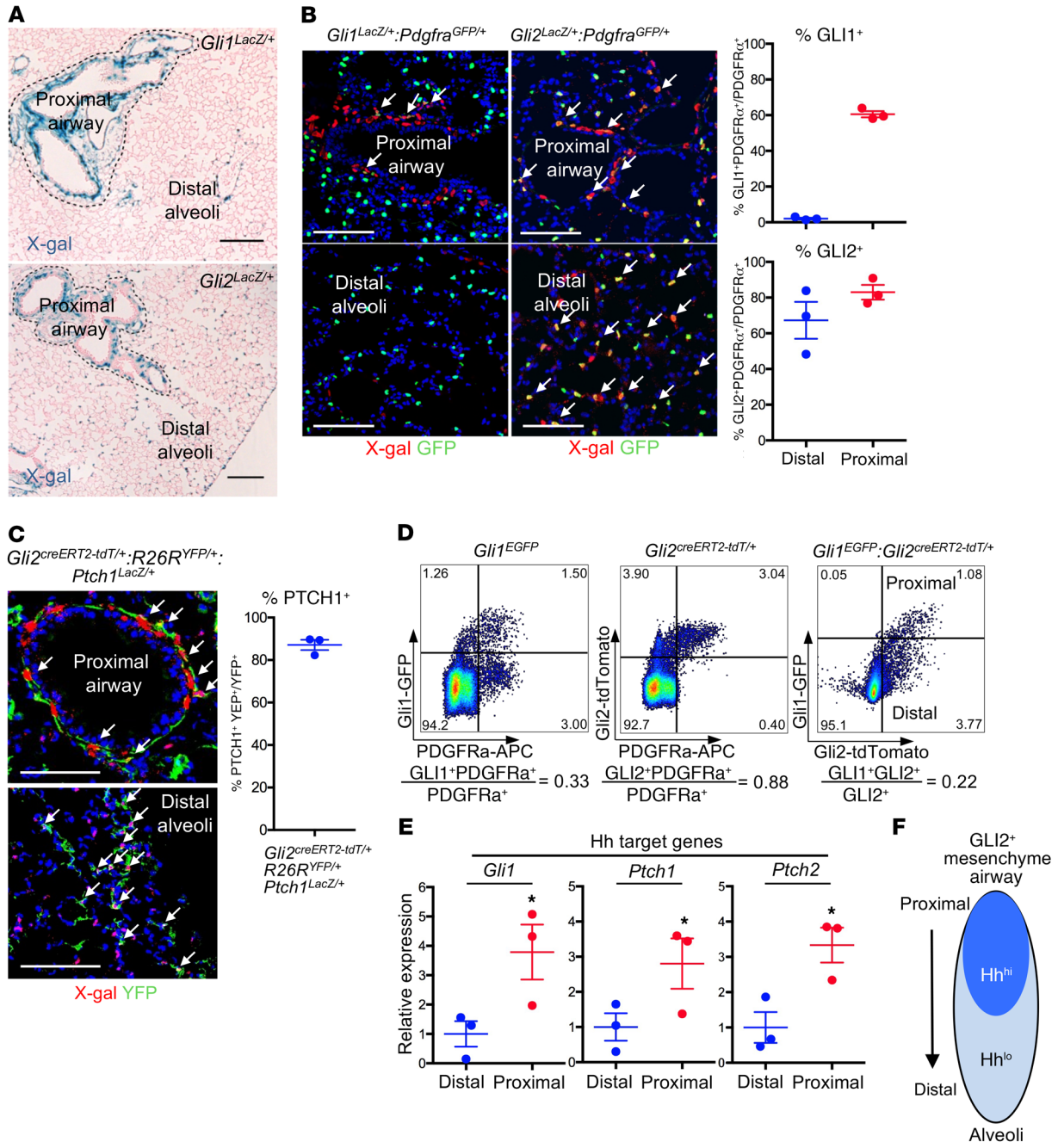


Figure 1. Proximal-distal asymmetry of Hh activation in the adult lung mesenchyme. (**A**) *Gli1*$^{LacZ/+}$ reporter demonstrates *Gli1* expression largely confined to the proximal airway, whereas *Gli2*$^{LacZ/+}$ reporter demonstrates *Gli2* expression both in the proximal airway and the distal alveoli. (**B**) Generation of double reporters (*Gli1*$^{LacZ/+}$:*Pdgfra*$^{GFP/+}$ and *Gli2*$^{LacZ/+}$:*Pdgfra*$^{GFP/+}$) demonstrates that GLI1$^+$ cells constitute a significant portion of the PDGFRα$^+$ proximal airway mesenchymal fibroblasts, but do not contribute to the distal mesenchymal fibroblast population. In contrast, GLI2$^+$ cells contribute to both the proximal and distal mesenchyme (arrows, overlap expression). (**C**) Generation of *Gli2*$^{creERT2\text{-}tdT/+}$:*R26R*$^{YFP/+}$:*Ptch1*$^{LacZ/+}$ followed by lineage labeling of GLI2$^+$ cells shows that *Gli2* and *Ptch1* expression largely overlaps in both the proximal and distal mesenchyme. (**D**) Left and center: utilizing fluorescent reporters of *Gli1* and *Gli2* (*Gli1*^{EGFP} and *Gli2*$^{creERT2\text{-}tdT/+}$), we confirmed that GLI1$^+$ and GLI2$^+$ cells contribute to a large proportion of the PDGFRα$^+$ mesenchyme in the lung. Right: crossing the 2 reporters generates a dual color *Gli1:Gli2* reporter (*Gli1*^{EGFP}:*Gli2*$^{creERT2\text{-}tdT/+}$), which shows that almost all of the GLI1$^+$ cells coexpress GLI2 and that the GLI2 domain can be divided into proximal GLI1$^+$ and distal GLI1$^-$ subpopulations. (**E**) qPCR analysis demonstrates significant enrichment of Hh target genes *Gli1*, *Ptch1*, and *Ptch2* in the sorted proximal mesenchyme (GLI1$^+$GLI2$^+$) relative to distal (GLI1$^-$GLI2$^+$). Statistical analysis was done using 1-tailed paired Student's *t* test. $^*P < 0.05$. (**F**) Schematic representing a broad GLI2$^+$ mesenchyme in the adult lung segregated by asymmetric activation of Hh along the proximal-distal axis. Scale bars: 100 μm. Data are represented as mean ± SEM, with $n = 3$ per group. Results were replicated ($n \geq 2$ experiments).

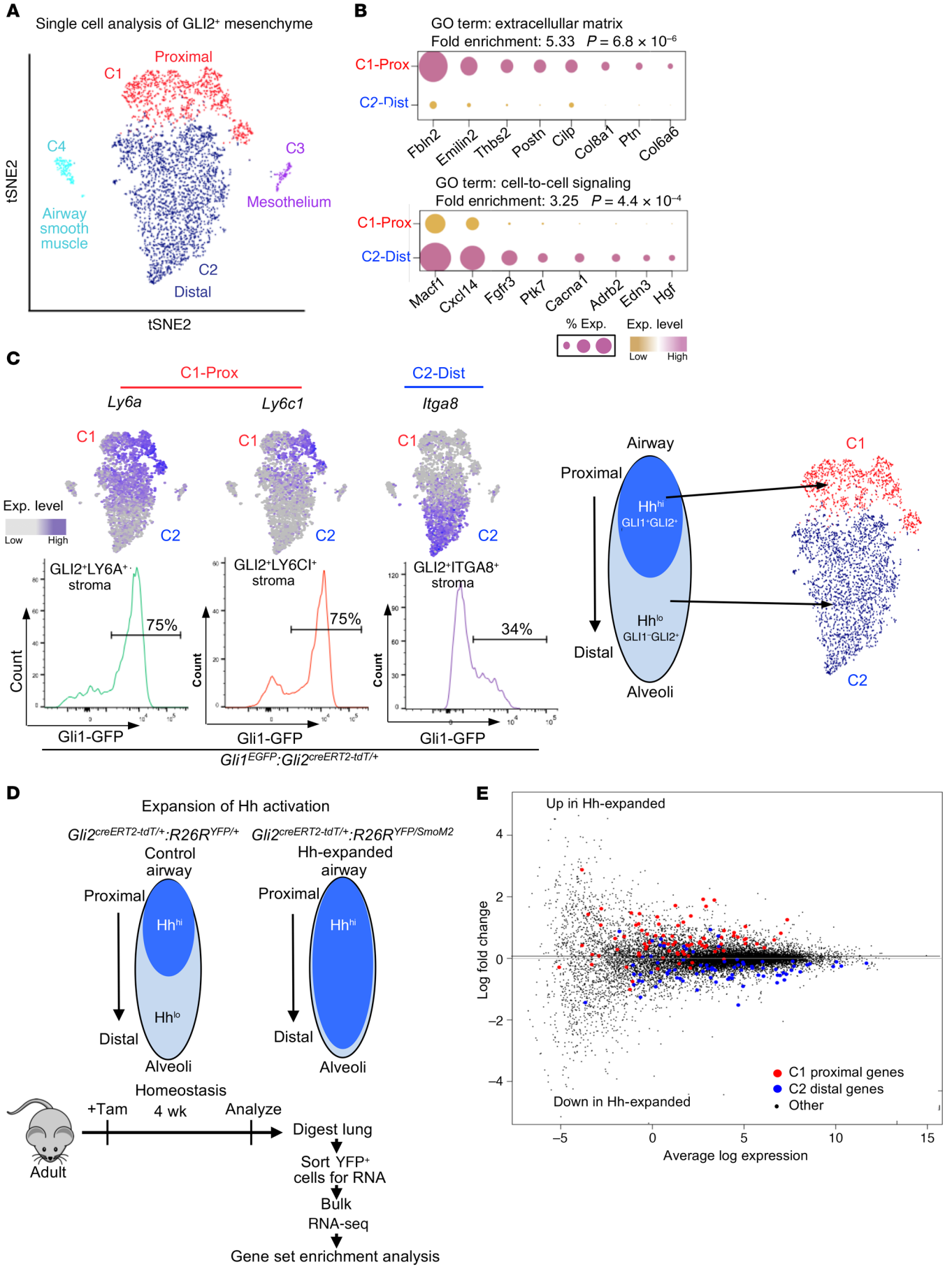

A
Single cell analysis of GLI2+ mesenchyme
Proximal
C1
C4
C3
Airway smooth muscle
Mesothelium
C2
Distal
tSNE2
tSNE2
B
GO term: extracellular matrix
Fold enrichment: 5.33 P = 6.8 × 10−6
C1-Prox
C2-Dist
Fbln2
Emilin2
Thbs2
Postn
Cilp
Col8a1
Ptn
Col6a6
GO term: cell-to-cell signaling
Fold enrichment: 3.25 P = 4.4 × 10−4
C1-Prox
C2-Dist
Macf1
Cxcl14
Fgfr3
Ptk7
Cacna1
Adrb2
Edn3
Hgf
% Exp.
Exp. level
Low
High
C
C1-Prox
C2-Dist
Ly6a
Ly6c1
Itga8
C1
C2
Exp. level
Low
High
GLI2+LY6A+ stroma
GLI2+LY6CI+ stroma
GLI2+ITGA8+ stroma
75%
75%
34%
Count
Gli1-GFP
Gli1EGFP:Gli2creERT2-tdT/+
Airway
Proximal
Distal
Alveoli
Hhhi
GLI1+GLI2+
Hhlo
GLI1−GLI2+
C1
C2
D
Expansion of Hh activation
Gli2creERT2-tdT/+:R26RYFP/+
Gli2creERT2-tdT/+:R26RYFP/SmoM2
Control airway
Hh-expanded airway
Proximal
Distal
Alveoli
Hhhi
Hhlo
Homeostasis
+Tam
4 wk
Analyze
Adult
Digest lung
Sort YFP+ cells for RNA
Bulk RNA-seq
Gene set enrichment analysis
E
Up in Hh-expanded
Down in Hh-expanded
Log fold change
Average log expression
C1 proximal genes
C2 distal genes
Other

Figure 2. Single-cell analysis of GLI2$^+$ mesenchyme in the adult lung. (**A**) Unbiased clustering of transcriptomes of individual GLI2$^+$ cells from the adult lung. Each cell is represented as a single dot that is colored by the clustering algorithm and plotted on the tSNE graph. Four distinct clusters emerge from a graph-based clustering algorithm with the majority of the cells in cluster 1 (C1) and cluster 2 (C2). (**B**) The top expressed genes in cluster 1 are enriched for the GO term extracellular matrix, while the cluster 2 gene signature is enriched for the GO term cell-to-cell signaling. Size of the dot plot represents proportion of cells within the cluster expressing the gene, and color denotes the level of expression (Exp.). Prox, proximal; Dist, distal. (**C**) Flow cytometry analysis of the *Gli1:Gli2* reporter (*Gli1^{EGFP}:Gli2$^{creERT2\text{-}tdT/+}$*) shows that the cluster 1 markers LY6A and LY6C1 detect cells that are predominantly in the Hh^{hi} (GFP$^+$, GLI1$^+$GLI2$^+$) fraction of the GLI2$^+$ mesenchyme, whereas the cluster 2 marker ITGA8 detects cells predominantly in the Hh^{lo} (GFP$^-$, GLI1$^-$GLI2$^+$) fraction. (**D**) Schematic of the experimental flow to isolate GLI2$^+$ cells where the Hh-activation domain is expanded to the distal alveoli for bulk RNA-seq, with $n = 4$ per group. Tam, tamoxifen. (**E**) The mean difference plot displays the log-fold differences (*y* axis) versus the mean counts for all genes in the bulk RNA-seq experiment, with each dot representing a gene detected in the GLI2$^+$ mesenchyme. The majority of the proximal genes are upregulated in the Hh-expanded Gli2$^+$ mesenchyme, while the majority of the distal genes are downregulated. Results were replicated ($n \geq 2$ experiments).

cluster 1 and 2 signature genes showed that cluster 1 (proximal) was enriched in extracellular matrix genes (fold enrichment: 5.33, $P = 6.8 \times 10^{-6}$), while cluster 2 (distal) was enriched in cell-to-cell signaling genes (fold enrichment: 3.25, $P = 4.4 \times 10^{-4}$, Figure 2B). To confirm that the proximally located cluster 1 was enriched for Hh-activated, mesenchyme-expressing *Gli1*, we performed flow cytometry analysis of signature genes expressed on the cell surface chosen from clusters 1 and 2 in *Gli1:Gli2* (*Gli2$^{creERT2\text{-}tdt}$:Gli1^{EGFP}*) dual-color reporter mouse lungs. Flow cytometry analysis demonstrated that cluster 1 markers *Ly6a* and *Ly6c1* were enriched in the GLI1$^+$ fraction of the GLI2$^+$ mesenchyme (Figure 2C). Conversely, the cluster 2 marker *Itga8* was enriched in the GLI1 fraction of the GLI2$^+$ mesenchyme (Figure 2C). This suggests that the GLI2$^+$ population is segregated into distinct mesenchymal subtypes along the proximal-distal axis of the lung characterized by asymmetric Hh activation.

Next, we tested to see whether Hh is a determinant of proximal identity where it is preferentially activated, rather than just a marker. We overexpressed a constitutively active form of the Hh effector, *Smo* (*SmoM2*) (34), throughout the entire GLI2$^+$ mesenchyme, expanding Hh activation into the distal mesenchyme (Figure 2D). To determine whether expansion of Hh activation promotes proximal versus distal identity, we performed bulk RNA-seq of the sorted GLI2$^+$ mesenchyme and performed gene set enrichment analysis of the genes upregulated and downregulated by Hh expansion with the proximal and distal mesenchymal fibroblast gene signature sets (top 100 genes) identified by scRNA-seq. The proximal signature genes were significantly enriched among genes upregulated in the Hh-expanded samples versus control (Camera, $P < 0.0001$, FDR < 0.0001; Figure 2E, Supplemental Figure 3A, and Supplemental Table 2). Conversely, distal signature genes were significantly enriched among genes downregulated in the Hh-expanded samples (Camera, $P < 0.0001$, FDR < 0.0001; Figure 2E, Supplemental Figure 3A, and Supplemental Table 3). qPCR analysis of the ITGA8$^+$GLI2$^+$ distal mesenchyme confirmed upregulation of proximal genes within the distal domain in Hh-expanded mutants versus controls (Supplemental Figure 3B). These data showed that Hh functions as a selector gene (35) within a field of Hh-competent mesenchyme in the adult lung, maintaining diversification by asymmetrically promoting proximal while suppressing distal mesenchymal identity in a spatially restricted manner.

Distal expansion of Hh activation induces emphysema. As recent GWAS have repeatedly implicated modifiers of Hh signaling in COPD/emphysema susceptibility and severity (5, 6, 10), we examined the transcriptome profiles of COPD/emphysema patients in the largest study to date that correlated gene expression within epithelial brushings with clinical disease severity (36). As the study only sampled the epithelium, we examined the expression of Hh ligands *SHH*, *IHH*, and *DHH*, which are expressed in the lung epithelium. Using a regression model that adjusts for sex, smoking status, age, and pack-years, we found that *SHH* expression in the epithelium significantly correlated with disease severity as measured by FEV1 (Supplemental Figure 4A, $P = 3.09 \times 10^{-8}$), the primary measure of lung function used to stratify COPD/emphysema (lower FEV1 = more severe disease). The expression levels of *IHH* and *DHH*, occurring in much lower abundance compared with those of *SHH*, were not significantly associated with lung function (Supplemental Figure 4A). Our transcript analysis of sorted epithelium in mouse and human showed that *SHH* was the predominant Hh ligand expressed in the lung (Supplemental Figure 4B), and that this expression was elevated with increasing severity of COPD/emphysema. Next, we examined *Shh* expression in a well-established murine model of emphysema utilizing cigarette exposure. Histological analysis of lungs exposed to 6 months of cigarette smoke demonstrated expansion of *Shh* expression in situ in the distal alveoli compared with room air controls (Supplemental Figure 4C).

Increased expression of *SHH* in human COPD/emphysema and expansion of *Shh* expression in murine models of COPD/emphysema suggest that Hh overactivation can induce emphysema, so we investigated the effect of ectopic activation of Hh on the distal airspace morphology of the Hh-expanded and control animals. Inducible expansion of Hh activation within the GLI2$^+$ mesenchyme induced over 20% enlargement of airspace compared with controls, as characterized by increase in mean linear intercept (MLI) (chord) length and alveolar size along with reduced density of alveoli in the distal lung, which are hallmarks of emphysema (Figure 3A). Furthermore, the mutants also displayed a more simplified alveolar structure in which secondary crests were absent from the enlarged alveoli that are characteristics of emphysema as well as alveolar developmental delay syndromes, such as bronchopulmonary dysplasia (37) (Supplemental Figure 5A). In addition to histological analysis of the airspace morphology, we also developed a radiographic method to quantify the degree of emphysema based on x-ray attenuation on computed tomography (CT) in explanted whole lungs. Human emphysematous lungs demonstrate increases in percentage of low attenuation areas (%LAA) in CT due to destruction of alveolar structures, resulting in large areas of air that appear as black space on CT (38). Similarly, we performed μCT of murine lungs of both control and Hh-expanded mutants, followed by 3D volume rendering

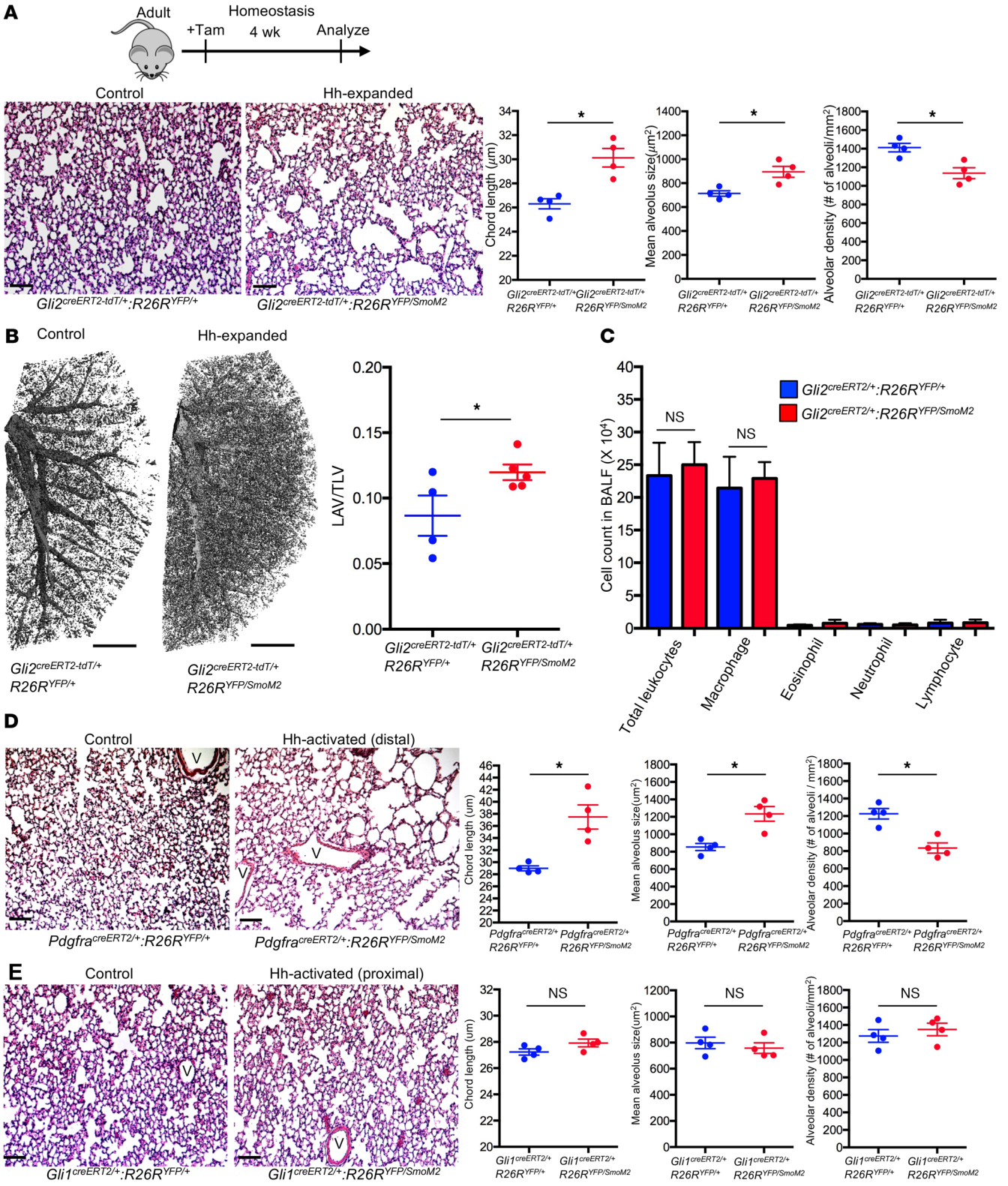


Figure 3. Expansion of Hh activation into the distal mesenchyme induces emphysematous changes. (**A**) Expansion of Hh activation (*Gli2*$^{creERT2\text{-}tdT/+}$:*R26R*$^{YFP/SmoM2}$) into the distal alveoli–induced emphysematous changes in the alveolar airspace characterized by increase in mean chord length, enlarged alveolus size, and reduced density of alveoli in the distal compartment. (**B**) μCT of murine lungs demonstrates increased percentage of LAV as a ratio of the TLV in Hh-expanded mutants. Dark dots indicate airspace. (**C**) Leukocyte counts in BALF show that macrophage is the major composition of leukocytes and there is no significant difference in the number of leukocytes and macrophages of the BALF from both control and Hh-expanded mutants. (**D**) Distal mesenchyme–specific activation of Hh utilizing the *Pdgfra*^{creERT2} allele phenocopied the Hh-expanded mutant (*Gli2*$^{creERT2\text{-}tdT/+}$:*R26R*$^{YFP/SmoM2}$). (**E**) Proximal mesenchyme-specific activation of Hh utilizing the *Gli1*^{creERT2} allele does not induce significant changes in the distal airspace morphology compared with controls. V, blood vessel. Data are represented as mean ± SEM, with $n \geq 4$ per group. Statistical analysis was done using the 1-tailed Student's *t* test. $^*P < 0.05$. Scale bars: 100 μm (**A**, **D**, and **E**); 1 mm (**B**). Results were replicated ($n \geq 2$ experiments).

to quantify the percentage of low attenuation volume (LAV) as a ratio of the total lung volume (TLV). Hh-expanded mutant lungs demonstrated a significantly elevated LAV/TLV ratio compared with controls (Figure 3B), consistent with the histological findings of airspace enlargement.

To determine whether this phenotype is due to Hh activation in the proximal or distal mesenchyme, we utilized a transgenic *Pdgfra*^{creERT2} allele in which lineage labeling only occurs in the distal mesenchyme at our operating tamoxifen dose while sparing the proximal mesenchyme (Supplemental Figure 5B). Activation of the *SmoM2* allele (*Pdgfra*^{creERT2}*:R26R*$^{YFP/SmoM2}$) within the distal mesenchyme phenocopied the airspace morphology changes seen in the Hh-expanded mutant using the *Gli2* allele, resulting in enlarged alveoli with reduced alveolar density (Figure 3D). In fact, the airspace enlargement phenotype seen with Hh expansion using the *Pdgfra*^{creERT2} driver was more severe compared with that using the *Gli2*$^{creERT2\text{-}tdT}$ driver, which was likely due to the difference in expression levels of the genes driving *creERT2* expression. In contrast, we could induce Hh activation only within the proximal mesenchyme while sparing the distal mesenchyme using the *Gli1*^{creERT2} driver (Supplemental Figure 5B), and *Gli1*^{SmoGOF} (*Gli1*^{creERT2}*:R26R*$^{YFP/SmoM2}$) mutants did not display significant distal airspace morphological changes compared with controls (Figure 3E). This showed that domain-specific activation of Hh in the distal mesenchyme induces airspace enlargement similar to that seen in emphysema, a lung disease linked to Hh signaling in multiple large-cohort GWAS.

Hh expansion disrupts the distal alveolar niche. While the prevailing notion of emphysema ascribes the loss of alveoli to excessive inflammation most commonly due to tobacco exposure (1), we did not observe evidence of increased inflammatory cells in the lungs of the Hh-expanded mutants when we examined the leukocytes present in the bronchoalveolar lavage fluid (BALF) as well as the composition of immune cells in the digested lungs by flow cytometry (Figure 3C and Supplemental Figure 6A). We also did not observe an increase in the number of apoptotic cells, as indicated by cleaved caspase-3 staining in the lungs of controls and mutants (Supplemental Figure 6B). However, we did notice the loss of SFTPC$^+$ type 2 pneumocytes, a putative resident stem cell in the distal alveoli (39), from the distal alveolar region of Hh-expanded mutants (Figure 4A). Distal expansion of Hh disrupted normal alveolar stem cell renewal, as Hh-expanded (*Gli2*^{SmoGOF}) mutants demonstrated significantly reduced fractions of SFTPC$^+$ alveolar stem cells that incorporated BrdU compared with controls during normal homeostasis as well as injury with bleomycin (Figure 4, B–D). These data showed that Hh activation in the distal mesenchyme could attenuate alveolar stem cell renewal in a non–cell autonomous manner, possibly altering mesenchymal feedback to the stem cells.

Among the genes identified in the distal mesenchymal fibroblast signature that was enriched in cell-to-cell signaling was hepatocyte growth factor (*Hgf*), a mesenchymal-derived mitogen implicated in stem cell renewal (19). We confirmed that *Hgf* was expressed in distal alveolar mesenchyme in situ (Figure 4E), and deficiency of HGF in the lung has been linked to emphysema (40, 41). Immunofluorescence staining of the lung showed that the HGF receptor *Met* was expressed almost exclusively in the SFTPC$^+$ distal alveolar stem cells that were adjacent to Gli2$^+$ distal mesenchyme in vivo (Figure 4F), suggesting that Hh activation in the distal alveolar mesenchyme might regulate HGF activation of MET in the alveolar stem cells.

To determine the effect of Hh activation on the distal alveolar niche in vitro, we first generated a transgenic animal (*Ubc*^{creERT2}*:R26R*$^{SmoM2/+}$) in which constitutively active SmoM2 could be induced in isolated lung mesenchymal cells in vitro utilizing a *creERT2* allele driven by a ubiquitously expressing promoter (42). Addition of 4-hydroxytamoxifen (4OHT) to the Hh-inducible lung mesenchymal fibroblasts activated *Gli1* expression, while *Hgf* expression was downregulated (Figure 5A). We then cocultured the Hh-inducible mesenchyme with isolated SFTPC$^+$ distal alveolar stem cells, and mesenchymal Hh activation with 4OHT resulted in significantly reduced distal alveolar organoid growth compared with vehicles, which was rescued with escalating doses of recombinant HGF (Figure 5B). HGF interaction with the MET receptor induces MET autophosphorylation to activate the tyrosine kinase activity of the receptor (43). To determine whether Hh expansion in the distal alveolar mesenchyme affected MET activation in vivo, we quantified MET phosphorylation in the Hh-expanded lungs and showed that ectopic Hh activation in the alveolar mesenchyme reduced the ratio of phosphorylated MET to total MET (Figure 5C; see complete unedited blots in the supplemental material).

Deletion of *Met* during lung development and maturation has previously been shown to disrupt normal alveolarization (44). To determine whether the HGF-MET axis plays a role in maintaining the distal alveolar airspace during adult homeostasis in vivo, we inducibly deleted *Met* from the SFTPC$^+$ alveolar stem cells by generating a *SFTPC*^{creERT2}*:Met*$^{fl/fl}$ (*SFTPC*^{MetCKO}) mutant. Tamoxifen induction of the adult *SFTPC*^{MetCKO} mutant resulted in an enlargement of alveolar airspace and loss of alveolar density (Figure 5D) similar to that previously seen in the Hh-expanded mutants utilizing both the *Gli2* and *Pdgfrα creERT2* alleles. These results showed that expansion of Hh activation disrupts distal mesenchymal feedback to the alveolar stem cells via HGF-MET signaling that maintains normal alveolar homeostasis and contributes to emphysematous changes in the distal lung.

In addition to HGF, Wnt ligands are enriched in the alveolar mesenchyme, which provides mitogenic cues to Wnt-receptive alveolar stem cells (45, 46). Here, we showed that activation of the Hh-inducible mesenchymal fibroblasts in vitro suppressed the expression of *Wnt1*, *Wnt2*, *Wnt3a*, and *Wnt7b* (Supplemental Figure 7A) and that the addition of the Wnt agonist Chiron (CHIR) also partially rescued the defect in SFTPC$^+$ distal alveolar stem cell proliferation resulting from Hh activation in the mesenchyme (Supplemental Figure 7B). This suggests that Hh activation could suppress a broad program of mitogenic feedback to the epithelial stem cell compartment in the distal alveoli.

Hedgehog asymmetry and mesenchymal segregation is conserved in the human lung. To determine whether the human lung mesenchyme exhibits segregation similar to that of transcriptomes based on anatomical location, we performed scRNA-seq on the human lung mesenchyme. We isolated lung specimens from both the proximal portion enriched in airways and the distal portion enriched in alveoli separately. The lung fragments were

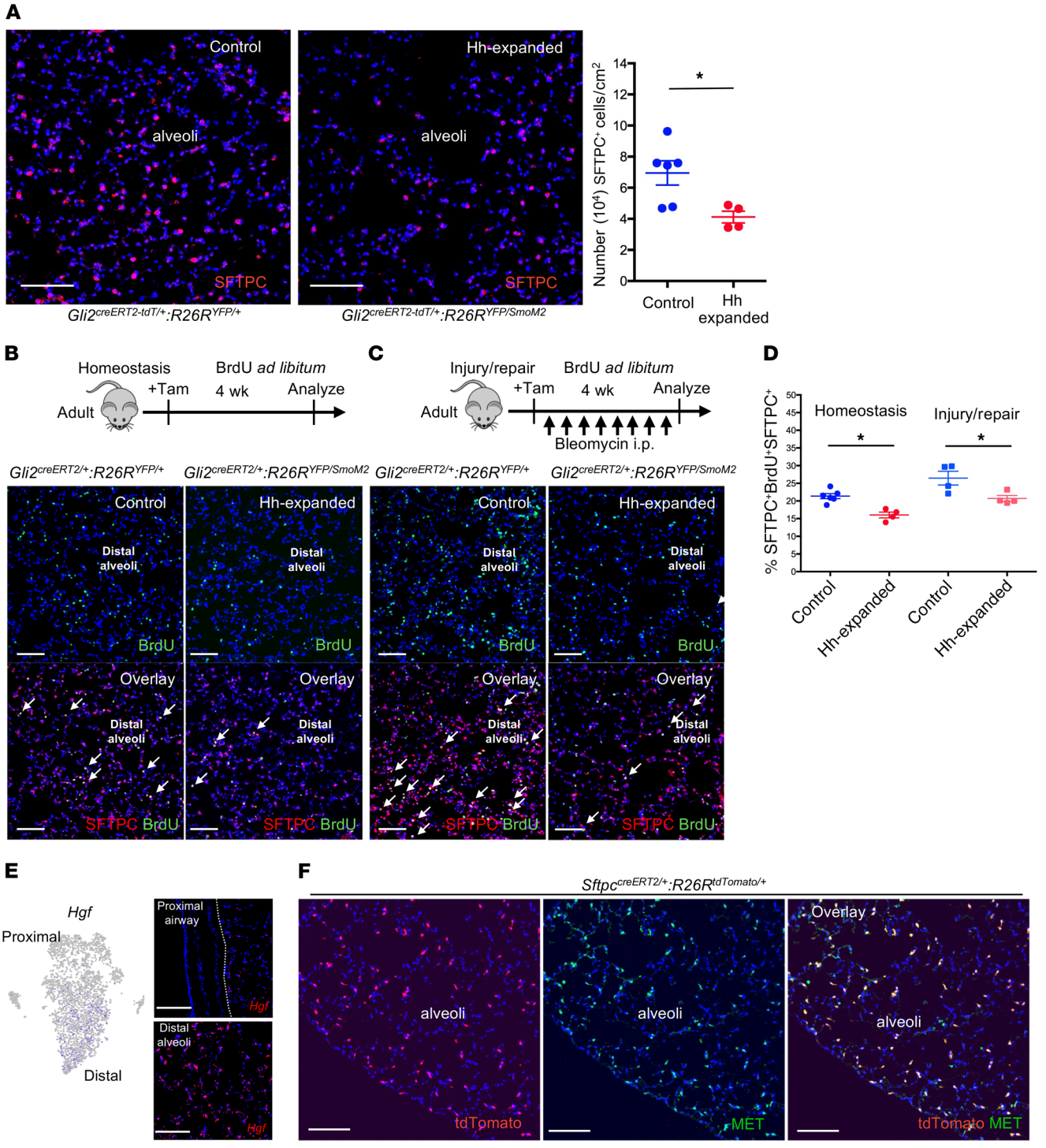


Figure 4. Hh activation in distal mesenchyme reduces proliferation of alveolar stem cells. (**A**) Expansion of Hh into the distal mesenchyme reduced the number of SFTPC⁺ stem cells in the alveoli. (**B**–**D**) Expansion of Hh activation into the distal mesenchyme reduced BrdU incorporation into SFTPC⁺ alveolar stem cells during normal homeostasis (**B**) as well as after bleomycin-induced injury (**C**). Arrows indicate overlapping staining of SFTPC and BrdU. Statistical analysis of the percentage of BrdU-incorporated SFTPC⁺ cells (**D**). (**E**) tSNE plot demonstrates that *Hgf* is enriched in cluster 2 (bottom cluster), which marks the distal mesenchyme. In situ confirms expression of *Hgf* in the distal mesenchyme. (**F**) SFTPC⁺ alveolar stem cells express the HGF receptor, MET. Data are represented as mean ± SEM, with $n \geq 4$ per group. Statistical analysis was done using 1-tailed Student's *t* test. $^*P < 0.05$. Scale bars: 100 μm. Results were replicated ($n \geq 2$ experiments).

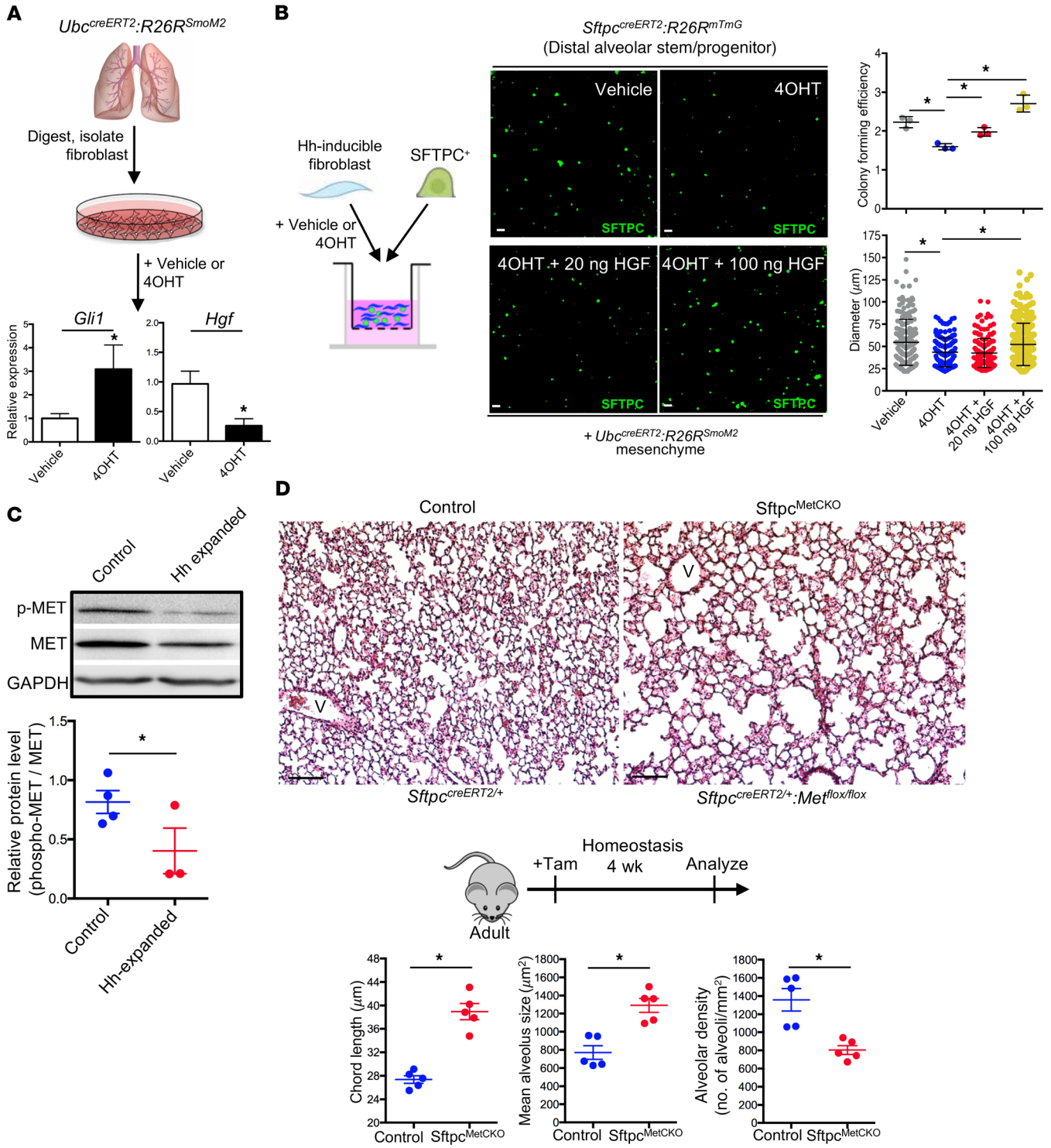
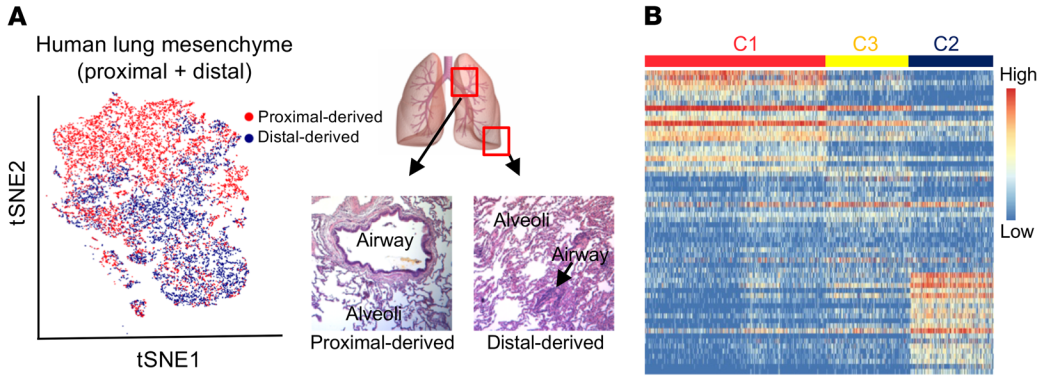
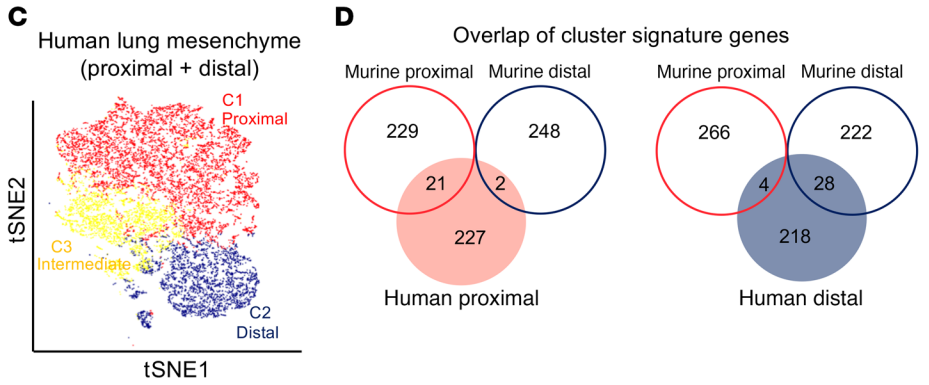
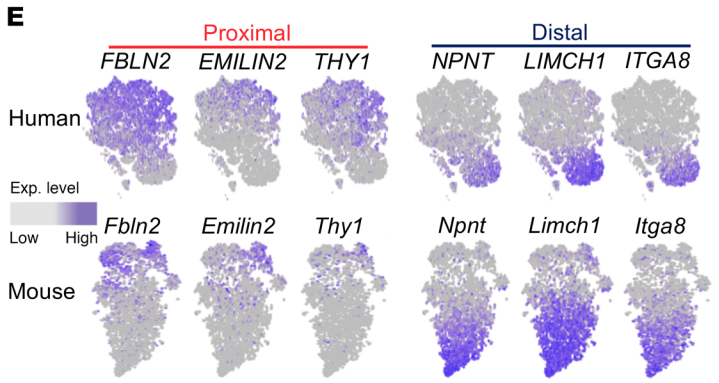
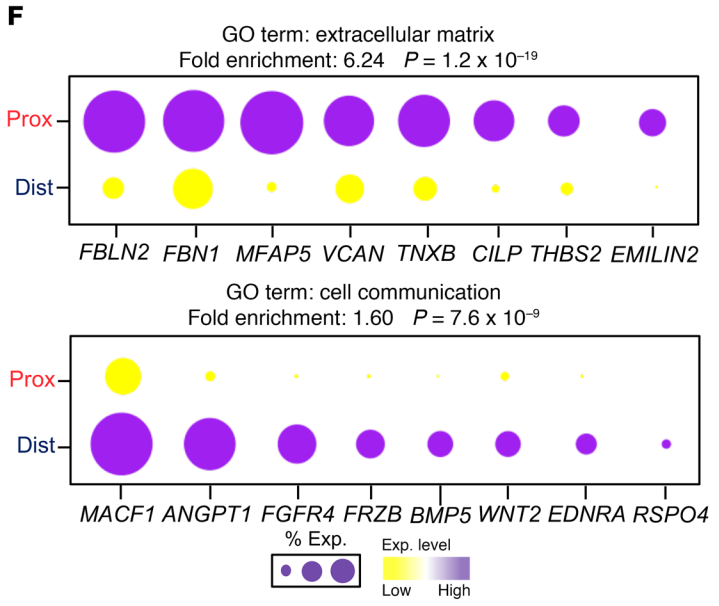


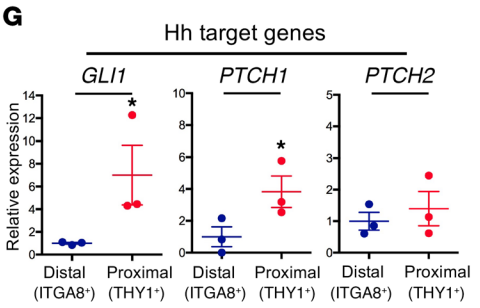
Figure 5. Hh activation disrupts the alveolar niche through suppression of mesenchymal feedback to stem/progenitor cells. (**A**) Isolation of lung mesenchymal fibroblasts from *Ubc*$^{creERT2/+}$*:R26R*$^{SmoM2/+}$ animals followed by induction of 4OHT-activated *Gli1* expression while *Hgf* expression was concurrently downregulated. Data are represented as mean ± SEM, with $n = 3$ per group. (**B**) SFTPC$^+$ distal alveolar stem/progenitor cells were cocultured with Hh-inducible (*Ubc*$^{creERT2/+}$*:R26R*$^{SmoM2/+}$) mesenchyme. 4OHT induction significantly reduced the formation of alveolar organoids, which was partially rescued with the addition of recombinant HGF. Colony-forming efficiency was defined as follows: (no. of colonies/no. added progenitors) × 100. Data are represented as mean ± SD, with $n = 3$ per group. (**C**) Western blotting analysis of phosphorylated MET (p-MET), MET, and GAPDH in whole-lung lysate from Hh-expanded and control mice. Relative protein level of p-MET was calculated by p-MET/MET. Data are represented as mean ± SEM, with $n \geq 3$ per group. (**D**) Inducible deletion of the HGF receptor, MET, from SFTPC$^+$ alveolar stem/progenitor- induced airspace enlargement in the distal compartment of the lung. Data are represented as mean ± SEM, with $n = 5$ per group. Scale bars: 200 μm (**B**); 100 μm (**D**). Each dot represents individual biological replicate. Statistical analysis was done using 1-tailed Student's *t* test (**A**, **C**, and **D**) and 1-way ANOVA with Fisher's LSD test (**B**). *$P < 0.05$. Results were replicated ($n \geq 2$ experiments).


A
Human lung mesenchyme
(proximal + distal)
Proximal-derived
Distal-derived
tSNE2
tSNE1
Airway
Alveoli
Proximal-derived
Alveoli
Airway
Distal-derived
B
C1
C3
C2
High
Low


C
Human lung mesenchyme
(proximal + distal)
C1
Proximal
C3
Intermediate
C2
Distal
tSNE2
tSNE1
D
Overlap of cluster signature genes
Murine proximal
Murine distal
229
248
21
2
227
Human proximal
Murine proximal
Murine distal
266
222
4
28
218
Human distal


E
Proximal
Distal
FBLN2
EMILIN2
THY1
NPNT
LIMCH1
ITGA8
Human
Exp. level
Low
High
Fbln2
Emilin2
Thy1
Npnt
Limch1
Itga8
Mouse


F
GO term: extracellular matrix
Fold enrichment: 6.24 P = 1.2 x 10−19
Prox
Dist
FBLN2 FBN1 MFAP5 VCAN TNXB CILP THBS2 EMILIN2
GO term: cell communication
Fold enrichment: 1.60 P = 7.6 x 10−9
Prox
Dist
MACF1 ANGPT1 FGFR4 FRZB BMP5 WNT2 EDNRA RSPO4
% Exp.
Exp. level
Low
High


G
Hh target genes
GLI1
PTCH1
PTCH2
Relative expression
Distal (ITGA8+)
Proximal (THY1+)
Distal (ITGA8+)
Proximal (THY1+)
Distal (ITGA8+)
Proximal (THY1+)

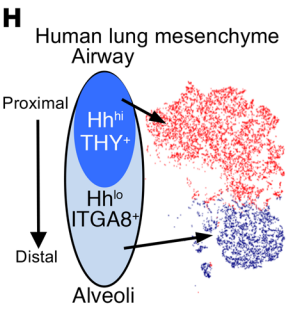

H
Human lung mesenchyme
Airway
Proximal
Hhhi
THY+
Hhlo
ITGA8+
Distal
Alveoli

Figure 6. Single-cell analysis of the human lung mesenchyme. (**A**) Proximal (enriched in airway) and distal (enriched in alveoli) fragments from the same lung were extracted for separate scRNA-seq, with the data merged onto the same tSNE projection demonstrating anatomical origin of the cell (each dot represents 1 cell; red and blue dots are from the proximal and distal fragments, respectively). (**B**) Unbiased clustering of individual transcriptomes of the human lung mesenchyme. (**C**) Three distinct subsets emerge from the tSNE projection labeled as proximal, distal, and intermediate. (**D**) The human proximal and distal mesenchymal subsets demonstrate significant overlap of homologous genes compared with their murine counterparts. (**E**) Feature plots showing human and murine mesenchymal subsets of comparable anatomic localization have similar enriched genes (feature dot plots of murine *Itga8*, *Thy1*, *Npnt*, and *Fbln2* appear in Figure 2C and Supplemental Figure 3B). (**F**) The top expressed genes in the proximal subset were enriched for the GO term extracellular matrix, while the distal gene signature was enriched for the GO term cell communication. Size of dot plot represents proportion of cells within the cluster expressing the gene, and color denotes level of expression. (**G**) qPCR analysis of the sorted proximal (THY1$^+$) and distal mesenchyme (ITGA8$^+$) demonstrates enrichment of Hh target genes *GLI1*, *PTCH1*, and *PTCH2* in the proximal mesenchyme relative to distal mesenchyme. Data are represented as mean ± SEM, with $n = 3$ per group. Statistical analysis was done using 1-tailed Student's *t* test. $^*P < 0.05$. (**H**) Schematic representing human lung mesenchyme segregated by asymmetric activation of Hh along the proximal-distal axis. Results were replicated ($n \geq 2$ experiments).

then digested and FACS sorted for mesenchymal cells based on negative expression of epithelial, hematopoietic, and endothelial markers (EPCAM$^-$CD45$^-$CD11b$^-$CD31$^-$). Library preparation was performed separately for the proximal- and distal-derived mesenchymal cells so that each subset had distinctive barcodes to determine anatomical origin (Figure 6A and Supplemental Figure 8A), and approximately 10,000 cells were captured from each specimen (Supplemental Figure 8B); these cells were then merged for analysis. Unsupervised clustering with hierarchical analysis of gene expression of the merged data sets revealed 3 distinct subsets of mesenchyme based on transcriptome identity (Figure 6, B and C), with clusters 1 and 2 demonstrating distinct segregation and enrichment of cells from the proximal- and distal-derived fragments, respectively (Figure 6A). Immunofluorescence staining confirmed localization of cluster signature genes to the proximal airway or distal alveoli (Supplemental Figure 8C). Cluster 3 shared some gene expression patterns with both proximal and distal subsets with a smaller number of unique genes, and we labeled this cluster intermediate (Figure 6, B and C). Applying the same pipeline to identify subset-specific signature genes (Supplemental Table 4), we found that the proximal and distal human lung mesenchymal subsets demonstrated significant overlap of homologous genes with their murine counterparts (human and murine proximal, odds ratio, 8.2, $P = 2.4 \times 10^{-12}$; human and murine distal, odds ratio, 11.8, $P = 2.2 \times 10^{-14}$; Figure 6, D and E), while the human intermediate subset did not resemble either murine subset (Supplemental Figure 8, D and E). Similarly to what was seen in the mouse lung, GO analysis of the human lung mesenchymal signature genes also found enrichment in extracellular matrix (fold enrichment: 6.24, $P = 1.2 \times 10^{-19}$) and cell communication (fold enrichment: 1.6, $P = 7.6 \times 10^{-9}$) genes in the proximal and distal subsets, respectively, suggesting a segregation of mesenchymal function based on anatomical location that is conserved across species (Figure 6F). The intermediate mesenchymal subset, which we believe was novel compared with the murine mesenchymal subsets, was enriched for genes involved in cholesterol metabolism (fold enrichment: 14.9, $P = 1.7 \times 10^{-4}$; Supplemental Figure 8E), which suggests that its members might function as putative lipofibroblasts that promote surfactant biosynthesis in the alveolar epithelium (47). Furthermore, we were able to separate the proximal and distal mesenchyme by FACS sorting with antibodies against cell-surface markers (THY1 and ITGA8, markers of proximal and distal mesenchymal identity, respectively, in both mouse and human lungs) enriched in each subset (Supplemental Figure 8F). qPCR analysis of the sorted mesenchymal cells demonstrated that proximal (THY1$^+$) mesenchyme was enriched in Hh target genes relative to the distal mesenchyme (ITGA8$^+$) (Figure 6G), reflecting a asymmetry of Hh activation similar to that seen in the murine lung. This demonstrated that segregation of mesenchymal identity based on asymmetric Hh activation is a conserved organizational feature of the adult lung (Figure 6H).

Discussion

While region-specific diversity of the mesenchymal compartment is coming into focus in adult tissue, the mechanism to generate and maintain this diversity during normal homeostasis has not been defined. Our study demonstrates that adult organs continue to utilize a developmental morphogen whereby asymmetric Hh activation diversifies the mesenchyme into proximal and distal populations with distinct identity, which might exhibit a level of plasticity modulated by the level of Hh activation (Figure 7A). Single-cell transcriptome analysis identified distinct proximal and distal mesenchymal gene signatures that are in part maintained by asymmetric Hh activation in the mesenchyme in both the murine and human lung. In addition, we generated a model to disrupt the asymmetry of Hh activation in the lung mesenchyme by expanding the domain of Hh activation into the distal compartment and utilized both single-cell and bulk population RNA-seq to demonstrate the disruption of distal mesenchymal identity, leading to loss of tissue integrity in the adult lung. While mesenchymal plasticity in the context of transdifferentiation between specific cell types (e.g., differentiation of lipofibroblast to myofibroblasts) during fibrotic injury has recently been described (48, 49), our study suggests that mesenchymal plasticity between region-specific subsets (e.g., distal to proximal mesenchyme and vice versa) may play a role in tissue response and outcomes to injury repair.

Beyond its role of providing the structural scaffold of adult organs, the mesenchyme has been recognized as a vital component of the stem cell niche through reciprocal interactions with epithelial stem cells. We demonstrate that the distal mesenchyme is enriched in paracrine factors known to be important in stem cell maintenance and renewal. Recent work has shown that the presence of GLI2 in the mesenchyme is required for the expression of mammary epithelial stem cell niche signals that regulate ductal morphogenesis and maintenance (50), suggesting that GLI2 is a marker of specialized mesenchyme to support the stem cell compartment. In this study, we generated a tool to isolate and dissect the GLI2$^+$ mesenchyme, demonstrating that the GLI2$^+$ mesenchyme participated in feedback to alveolar stem cells. Furthermore, we showed that the proper niche function is mediated by asymmetric Hh activation, as expansion of Hh activation into

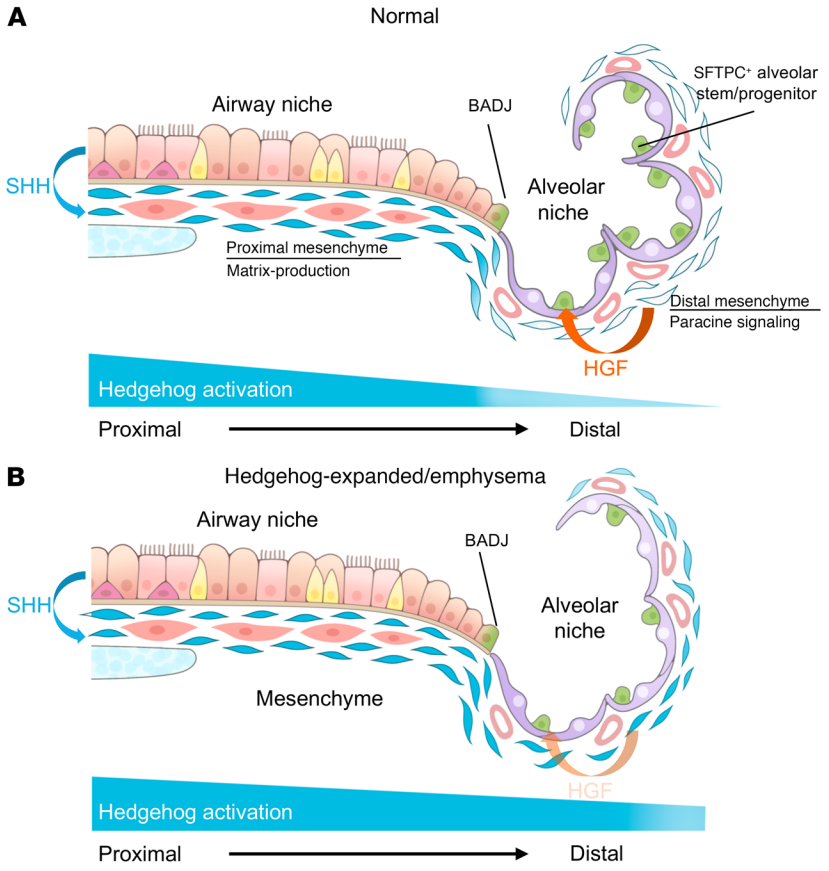


Figure 7. Model of Hh asymmetry regulating discrete mesenchymal identities. (**A**) Model of Hh asymmetry maintaining distinct proximal-distal mesenchymal identities and (**B**) ectopic Hh activation disrupting distal mesenchymal identity and the alveolar stem cell niche.

the distal mesenchyme reduced its capacity to support the alveolar stem cells in the distal compartment of the lung, resulting in a loss of alveoli that resembles emphysema.

Multiple large-cohort GWAS have identified Hh signaling components *HHIP* and *PTCH1* as genetic determinants of emphysema and lung function, respectively, in humans (5, 7), but the mechanism underlying this association remains unclear. Both HHIP and PTCH1 negatively regulate the Hh pathway (51, 52), and enhancer analysis has shown that the polymorphisms of HHIP that confers susceptibility to emphysema are loss-of-function (53), implying a derepressed Hh pathway in those with genetic susceptibility to emphysema. This study showed that disruption of the Hh gradient through ectopic activation in the distal mesenchyme induces emphysematous changes in the murine lung, in part by disrupting the mesenchymal feedback of HGF to the distal epithelial stem cells in the alveoli (Figure 7B).

A limitation of our study is that we have not established a direct link between loss of *HHIP*, the Hh-signaling component most commonly associated with human emphysema, and the gain-of-function Hh phenotype that we have utilized to induce an emphysematous phenotype in mouse models. Future studies addressing the role of HHIP in modulating Hh signaling in the adult lung are necessary to address the potential mechanistic role of Hh in human emphysema. Another caveat of our study is that we could not find alterations in inflammation or apoptosis in our murine model of emphysema, which does not exactly recapitulate the human disease. Finally, although we provide evidence that the SHH ligand is elevated in COPD/emphysema patients along with single-cell transcriptome studies showing differential expression of Hh targets in human fibroblast subsets, further studies in human emphysema tissues are required to determine whether region-specific identities are altered in the disease. While we acknowledge that the pathogenesis of human emphysema is multifactorial, we do believe that our study combined with the GWAS offers a paradigm where ectopic Hh activation can disrupt mesenchymal feedback to the stem cell compartment to cause alveolar loss to contribute to the emphysema phenotype.

Methods

Animal studies. Generation and genotyping of the *Gli2*^{lacz} (54), *Gli1*^{lacz} (31), *Gli1*^{EGFP} (55), *Gli1*^{creERT2} (30), Ptch1^{lacz} 56), *Pdgfra*^{creERT} (57), *Pdgfra*^{EGFP} (58), *Ubc*^{creERT2} (42), *c-Met*$^{fl/fl}$ (59), and *R26R*^{SmoM2} (34), lines have been previously described. Animals were obtained from the Jackson Laboratory. The *Gli1*^{EGFP} strain was obtained through GENSAT (55). *SFTPC*^{creERT2} mice were generated as previously described (60). To generate the *Gli2*$^{creERT2\text{-}tdT}$ mouse, a *CreERT2-2A-tdTomato* expression cassette was inserted into the ATG site in exon 2 of the *Gli2* locus (Supplemental Figure 1B). Targeting vector construction, embryonic stem (ES) cell selection, and injection were performed by the Ingenious Targeting Laboratory. The presence of the *Gli2-Cre-ERT2-2A-tdTomato* allele was determined by genomic DNA PCR with a WT band of 230 bp and a mutant band of 500 bp (genotyping primer for *Gli2*$^{creERT2\text{-}tdT}$, common forward: CCTGGGGTCAGAAGACTGAG, WT reverse: CTGCTGTCCTCCAAGAGACC, mutant reverse: CCAGTGAAACAGCATTGCTGT). Animals between the ages of 8 and 12 weeks old were used for the experiments, and littermates were used as controls, with a balance of sexes between the groups. Unless otherwise indicated, all mice were maintained on a mixed background. Tamoxifen (MilliporeSigma) was dissolved in corn oil and administered i.p. at 200 mg/kg per day for 3 days for lineage-tracing studies, with the exception of *Gli2*$^{creERT2\text{-}tdT}$, which was given at 200 mg/kg per day for 5 days. For injury experiments, bleomycin (Hospira) was dissolved in PBS and injected i.p. at 50 U/kg animal twice a week for 4 weeks. For BrdU-tracking experiments, mice were supplied water with 0.5 mg/ml BrdU (MilliporeSigma) and 1% sucrose continuously for 4 weeks. For cigarette smoke exposure, 8-week-old C57BL/6J mice were exposed to room air or the smoke from nonfiltered research cigarettes (2R4; University of Kentucky, Lexington, Kentucky, USA), as previously described (61). During the first week, mice received half a cigarette twice a day to allow for acclimation and then 2 cigarettes per day (1 cigarette per session, 2 sessions per day) for up to 6 months.

Human lung collection. Human lungs were received from brain-dead donors from the Northern California Transplant Donor Network. Screening criteria for selection of healthy lung for specimen collection were previously described (62).

Murine cell culture and organoid assay. For $Ubc^{creERT2}$:$R26R^{SmoM2}$ mesenchymal cells, lung was digested as above and then cells were cultured on gelatin-treated tissue culture plates with DMEM-F12 plus 10% FBS. Medium was refreshed every other day, and primary lung mesenchymal cells were maintained for no more than 3 passages. The $Ubc^{creERT2}$:$R26R^{SmoM2}$ mesenchymal cells were pretreated with vehicle or 1 μg/ml 4OHT for 72 hours before coculturing with epithelium. For organoid assay, GFP^+ alveolar epithelia were FACS sorted from tamoxifen-induced SPC^{creERT}:$R26R^{mTmG}$ lungs and cocultured with $Ubc^{creERT2}$:$R26R^{SmoM2}$ mesenchymal cells (5×10^3 epithelial cells, 3×10^4 mesenchymal cells/well) in a modified small airway epithelial cell growth medium (SAGM) diluted 1:1 in growth factor-reduced Matrigel (Corning). The modified SAGM comprised SAGM (PromoCell) with selected components from the SAGM Bullet Kit (PromoCell), including 5 μg/ml insulin, 10 μg/ml transferrin, 0.004 ml/ml bovine pituitary extract, 0.1 ng/ml retinoic acid, and 10 ng/ml EGF. Additional components included 0.1 μg/ml cholera toxin (MilliporeSigma) and 5% FBS (Life Technologies). Cell suspension–Matrigel mixture was placed in a Transwell and incubated in growth medium with 10 μM ROCK inhibitor (MilliporeSigma) in a 24-well plate for 48 hours, after which the medium was replenished every 48 hours (lacking ROCK inhibitor). Recombinant HGF (R&D Systems) was added at 20 ng/ml and 100 ng/ml, and CHIR (DMSO was used as control) was added at 20 nM after 48 hours and replenished every 2 to 3 days. Colonies were assayed after 14 days. Each experimental condition was performed in triplicate and counted by an investigator blinded to the experimental conditions. Colony-forming efficiency was equal to the following: number of GFP^+ colonies divided by number of GFP^+ epithelial cells cultured per well. Areas of individual colonies were assayed on Fiji, and over 140 colonies were randomly sized per experimental condition.

Single-cell capture and sequencing. $Gli2^+$ murine cells and human lung mesenchymal cells ($EPCAM^-CD45^-CD31^-CD11b^-$) were sorted as described above and resuspended in 50 μl PBS with 0.04% BSA at 1,000 cells/μl and loaded onto a single lane in the chromium controller to produce gel bead-in emulsions (GEMs). GEMs underwent reverse transcription for RNA barcoding and cDNA amplification, with the library prepped with the Chromium Single Cell 3′ Reagent Version 2 Kit (10X Genomics). Sample was sequenced on the HiSeq2500 (Illumina) in rapid run mode.

RNAscope in situ hybridization. Mouse *Shh*, *Hgf*, *Fbln2*, *Thy1*, and *Npnt* probes for RNA in situ detection were purchased from Advanced Cell Diagnostics. Paraffin-embedded lung sections were processed for RNA in situ detection using the RNAscope Multiplex Fluorescent Reagent Kit, version 2 (Advanced Cell Diagnostics), according to the manufacturer's instructions.

Expression overlay image analysis. Sections included in cell-count analysis were acquired using confocal microscopy. At least 3 animals per genotype were used, with at least 5 randomly selected sections chosen for each animal. Cell counts were performed on Fiji using the cell counter plug-in, and the investigator was blinded to the specimen genotype and condition. Results were averaged between specimens, and SDs were calculated per genotype. One tailed Student's *t* test was used to determine *P* values.

MLI, airspace size, and alveolar density analysis. For alveolar morphometric analysis, mouse lungs were processed according to the above protocol for paraffin-embedded samples, with the exception of inflation with 4% PFA with a constant pressure of 25 cm H_2O. The paraffin-embedded lung sections were stained by H&E in order to analyze the alveolar morphology metric. At least 5 randomly selected sections from each genotype were selected for analysis. The MLI was calculated as the linear sum of the lengths of all lines randomly drawn across the images divided by the number of intersections between alveolar walls. A minimum of 1,000 intercepts from 60 lines drawn across the lung in a randomized fashion were obtained for each lung, and the analysis was carried out using Fiji with the cell counter plug-in. The airspace size was measured with Fiji using the analyze particles tool. Images were first converted to 8 bit and inverted, and the analyze particle function was used with a set minimum of 50 μm^2 and maximum of infinity to identify and quantify alveoli in the image. The average airspace size for each lung was quantified by dividing the total airspace by the number of alveoli. At least 3,500 alveoli were measured for each lung. The alveolar density was the reciprocal of the airspace size. The airspace of airway and pulmonary vessels was excluded. Statistical analysis was done using 1-tailed Student's *t* test. Data are presented as mean ± SEM.

Contrast-enhanced μCT imaging of lung airspace. Lung specimens were harvested and fixed in 4% PFA at 4°C for 24 hours. Specimens were then washed 4 times in PBS for 30 minutes each time. Lungs were stained in a 25% Lugol's iodine solution at 4°C for 48 hours. Specimens were then embedded in a scanning medium of 3% agarose. Following agarose embedding, lung specimens were scanned on a Scanco Medical μCT 50 high-resolution specimen scanner. X-ray energy was set at 45 kVp and 88 μA. Specimens were scanned at a resolution of 10 μm voxel size and using an integration time of 500 ms. Quantitative analyses were performed using the Scanco μCT Evaluation Program. The region of interest was manually delineated slice by slice to include the entire lung volume, which we called TLV. An upper threshold of 160 grayscale units, out of 1,000, was applied to segment airspace and excluded tissue. A Gaussian filter with a σ of 0.8 and a support of 1 was applied to remove background noise, and the airspace volume, which we called LAV, was then quantified and reported as a ratio of LAV/TLV for each specimen.

scRNA-seq analysis. We built transcript profiles of individual cells by using Cell Ranger v1.1.0 software with default settings for demultiplexing, aligning reads with STAR software to mouse genome GRCm38 and Hg19 for humans, and counting unique molecular identifiers (UMIs). We used the Seurat v2.0 R package along with a gene-barcode matrix provided by CellRanger for downstream analysis (63). In total, we filtered the data in 2 different steps. We first filtered the data set by only accepting cells that contained about 200 genes/cell and genes that were expressed at a minimum of 3 cells. The UMI were log normalized genes with high expression, and variable genes were identified by the mean variance relationship method. Our second filter was done by using linear regression. We removed cells that had above 4,500 UMIs and a 2×10^{-5} percent.mito in the mouse data and 5,000 UMIs with a percent.mito of 0.1 in the human data. Using the RegressOut function, we took out cells that contained high mitochondrial genes. Next we used principle component analysis (PCA) to identify components that could be found within our data set for unsupervised clustering. JackStrawPlot function in Seurat was used to find significant principle components (PC) for each data set. We selected 10 different PCAs for unsupervised clustering of both data sets. Clustering results were visualize using the t-distributed

stochastic neighbor embedding (tSNE) algorithm in the Seurat package. For mouse data, 7 clusters were identified, but 2 clusters were not analyzed because they contained fewer than 10 cells and 1 cluster was suppressed because it contained more than 10% mitochondrial genes in the top 50 genes at an adjusted *P* value, leaving 4 main clusters. In the human data set, we found 4 clusters that encompassed the mesenchyme, which was divided into proximal, intermediate, and distal clusters. Differentially expressed genes between clusters were identified using the Cell Ranger method, which applies the Shrinkage estimation of dispersion (sSeq), employing a negative binomial exact test. The algorithm runs for each cluster against the rest of the clusters through a fast asymptotic β test used in edgeR. Cell Ranger computes relative library size as the total UMI counts for each cell divided by the mean UMI counts per cell as a way to normalize by the per cell library size, which is a parameter incorporated in calculations of differentially expressed genes. Dot plots were created using the Seurat DotPlot function to visualize the percentage of cells expressing a gene and the average expression level between our clusters. For cluster visualization and individual gene visualization on all clusters, we used the tSNE function. GO enrichment analysis was performed using the PANTHER overrepresentation test and entering the top 250 overexpressed genes in each cluster. Odds ratios and *P* values of overlap between mouse and human mesenchymal subsets were calculated using Fisher's exact test for enrichment using a base value of 21,000 genes in the mouse and human genome, respectively.

Bulk RNA-seq data analysis. Sequencing was done using Sanger/Illumina 1.9, and there were an average of 45 million reads per sample with a total of 4 biological replicates per condition. Quality control of reads was conducted by using FastQC (Babraham Bioinformatics). Ligation adaptors were removed using the Cutadapt and Sickle. Sequencing reads were aligned using HISAT and assembled with Stringtie software to the reference genome *Mus musculus*, UCSC version mm10. All gene counts of the biological replicates were concatenated while running DEseq for differential gene expression.

Camera gene set enrichment analysis. We used Camera (competitive gene set test accounting for intergene correlation) to compare the single-cell sequencing data sets with the whole-tissue data set (64). Camera determines whether a group of genes (gene set) is significantly overrepresented (enriched) at extremes of a ranked list, in which case the gene set is considered correlated with the phenotype. This is done using a rank-based test, an extension of the Wilcoxon Mann-Whitney that adjusts for intergene correlation among the elements of the test set. The analysis was carried out using the limma package in the R statistical environment. Raw gene counts from the whole-tissue experiment were first filtered to include only those genes indexed in the HGNC database and with an Ensembl gene biotype of protein_coding. Genes with low or zero counts were removed by filtering out genes with less than 50 total read counts across samples. Gene counts were then normalized using the limma voom function, and this data set was used as the test set in the Camera analysis. Four gene sets from the single-cell sequencing experiments were generated to analyze against this test set in the Camera analysis. These 4 gene sets were derived from the top 100 genes significantly upregulated (fold-change > 1.5, FDR < 0.05) in association with (a) the proximal phenotype and (b) the distal phenotype.

Human lung epithelial brushing transcriptome analysis. Bronchial epithelial brushings obtained from sixth through eighth generation bronchi of former and current smokers with ($n = 85$) and without COPD ($n = 152$) were previously profiled by Affymetrix HG 1.0 ST Arrays (36). Spirometry was done in all participants. Raw microarray files were deposited in the NCBI's Gene Expression Omnibus database (GEO GSE37147). Inclusion/exclusion criteria were previously published (65). The microarray data set underwent quantile normalization and probe summarization using the RMA algorithm (affy package, Bioconductor, R). Entrez Gene custom chip definition files were used for annotation (available at http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp). Batch effect was minimized using Combat. Regression analyses were performed using the limma package in R between $\log_2$ gene expression forced expiratory volume in 1 second (FEV1). Analyses were done before and after adjustment for age, sex, smoking status, and pack-years.

For more information, see Supplemental Methods.

Data access. scRNA-seq and bulk RNA-seq data sets were deposited in the NCBI's Gene Expression Omnibus database (GEO GSE102592).

Statistics. Statistical analysis was carried out using GraphPad Prism software. One-tailed Student's *t* tests were used to generate *P* values. One-way ANOVA was used to determine whether there were statistical differences among 3 groups, followed by Fisher's least significant difference (LSD) test for pairwise comparisons if the overall test was statistically significant. A *P* value of less than 0.05 was considered significant.

Study approval. All mice were maintained under specific pathogen–free conditions at UCSF according to IACUC protocol AN111302. Human lung specimens were derived from brain-dead donors and deidentified from the deceased donor; therefore, the requirement for informed consent for human subject research is not applicable under US Department of Health and Human Services regulations.

Author contributions

CW and TP conceived the experiments. CW, RM, MC, PS, and ANB performed the experiments. CW and RM performed flow cytometry and histological analysis. NSRDM and LEB performed analysis of scRNA-seq data sets. NSRDM and SAC performed analysis of bulk RNA-seq data sets. MAM and DJ procured human lung specimens. AL performed imaging acquisition and analysis of lung μCT. JBS, YY, HAC, and PJL provided expertise and feedback. TP and CW wrote the manuscript.

Acknowledgments

We thank Christopher Law, Zimu Deng, Julia Jackson, Andrew Vaughan, Ian Driver, and David Frank for providing technical assistance; Ophir Klein for sharing in situ hybridization resources; Ed Morrisey for providing the CreERT2-2A-tdTomato targeting cassette; Ari Molofsky for flow cytometry resources; Chris Gralapp for assistance with model illustration; Mark Looney, Dean Sheppard, Ophir Klein, Prescott Woodruff, John Fahy, and Rushika Perera for critical review of the manuscript; the Parnassus Flow Cytometry Core for assistance with cell sorting for bulk and single-cell RNA analysis; Eunice Wan and the Institute for Human Genetics Core for processing of single-cell RNA samples and high-throughput sequencing. TP is an NIH New Innovator, and this work was supported by NIH grants DP2AG056034 and K08HL121146 as well as the Marcus Precision Medicine and Pulmonary Hypertension Association Award (to TP). Access to flow

cyotmeters was supported by Diabetes Research Center Grant NIH P30 KD063720 and NIH S10 1S10OD021822-01. Human lung tissue procurement was supported by the Nina Ireland Program Award (to MAM).

Address correspondence to: Tien Peng, University of California, San Francisco, 513 Parnassus Avenue, HSE Building, Room 1312, Box 0130, San Francisco, California 94143, USA. Phone: 415.514.4180; Email: tien.peng@ucsf.edu.

1. Tuder RM, Yoshida T, Arap W, Pasqualini R, Petrache I. State of the art. Cellular and molecular mechanisms of alveolar destruction in emphysema: an evolutionary perspective. *Proc Am Thorac Soc*. 2006;3(6):503–510.
2. Wan ES, Silverman EK. Genetics of COPD and emphysema. *Chest*. 2009;136(3):859–866.
3. Tuder RM, Petrache I. Pathogenesis of chronic obstructive pulmonary disease. *J Clin Invest*. 2012;122(8):2749–2755.
4. Berndt A, Leme AS, Shapiro SD. Emerging genetics of COPD. *EMBO Mol Med*. 2012;4(11):1144–1155.
5. Pillai SG, et al. Loci identified by genome-wide association studies influence different disease-related phenotypes in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2010;182(12):1498–1505.
6. Castaldi PJ, et al. Genome-wide association identifies regulatory Loci associated with distinct local histogram emphysema patterns. *Am J Respir Crit Care Med*. 2014;190(4):399–409.
7. Hancock DB, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet*. 2010;42(1):45–52.
8. Li X, et al. Importance of hedgehog interacting protein and other lung function genes in asthma. *J Allergy Clin Immunol*. 2011;127(6):1457–1465.
9. Castaldi PJ, et al. Genetic control of gene expression at novel and established chronic obstructive pulmonary disease loci. *Hum Mol Genet*. 2015;24(4):1200–1210.
10. Pillai SG, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet*. 2009;5(3):e1000421.
11. Loth DW, et al. Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nat Genet*. 2014;46(7):669–677.
12. Gurdon JB, Bourillot PY. Morphogen gradient interpretation. *Nature*. 2001;413(6858):797–803.
13. Briscoe J, Small S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development*. 2015;142(23):3996–4009.
14. Wassef M, Joyner AL. Early mesencephalon/metencephalon patterning and development of the cerebellum. *Perspect Dev Neurobiol*. 1997;5(1):3–16.
15. Fabian SL, et al. Hedgehog-Gli pathway activation during kidney fibrosis. *Am J Pathol*. 2012;180(4):1441–1453.
16. Liu L, et al. Hedgehog signaling in neonatal and adult lung. *Am J Respir Cell Mol Biol*. 2013;48(6):703–710.
17. Peng YC, Levine CM, Zahid S, Wilson EL, Joyner AL. Sonic hedgehog signals to multiple prostate stromal stem cells that replenish distinct stromal subtypes during regeneration. *Proc Natl Acad Sci U S A*. 2013;110(51):20611–20616.
18. Shin K, et al. Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder. *Nature*. 2011;472(7341):110–114.
19. Lim A, Shin K, Zhao C, Kawano S, Beachy PA. Spatially restricted Hedgehog signalling regulates HGF-induced branching of the adult prostate. *Nat Cell Biol*. 2014;16(12):1135–1145.
20. Zepp JA, et al. Distinct mesenchymal lineages and niches promote epithelial self-renewal and myofibrogenesis in the lung. *Cell*. 2017;170(6):1134–1148.e10.
21. Yang H, Adam RC, Ge Y, Hua ZL, Fuchs E. Epithelial-Mesenchymal Micro-niches Govern Stem Cell Lineage Choices. *Cell*. 2017;169(3):483–496.e13.
22. Lee JH, et al. Anatomically and Functionally Distinct Lung Mesenchymal Populations Marked by Lgr5 and Lgr6. *Cell*. 2017;170(6):1149–1163.e12.
23. Robinson RJ, Russo J, Doolittle RL. 3D airway reconstruction using visible human data set and human casts with comparison to morphometric data. *Anat Rec (Hoboken)*. 2009;292(7):1028–1044.
24. Thiesse J, et al. Lung structure phenotype variation in inbred mouse strains revealed through in vivo micro-CT imaging. *J Appl Physiol*. 2010;109(6):1960–1968.
25. Miller FJ, Mercer RR, Crapo JD. Lower respiratory tract structure of laboratory animals and humans: dosimetry implications. *Aerosol Sci Technol*. 1993;18(3):257–271.
26. Hong KU, Reynolds SD, Giangreco A, Hurley CM, Stripp BR. Clara cell secretory protein-expressing cells of the airway neuroepithelial body microenvironment include a label-retaining subset and are critical for epithelial renewal after progenitor cell depletion. *Am J Respir Cell Mol Biol*. 2001;24(6):671–681.
27. Giangreco A, Reynolds SD, Stripp BR. Terminal bronchioles harbor a unique airway stem cell population that localizes to the bronchoalveolar duct junction. *Am J Pathol*. 2002;161(1):173–182.
28. Kim CF, et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell*. 2005;121(6):823–835.
29. Peng T, et al. Hedgehog actively maintains adult lung quiescence and regulates repair and regeneration. *Nature*. 2015;526(7574):578–582.
30. Ahn S, Joyner AL. In vivo analysis of quiescent adult neural stem cells responding to Sonic hedgehog. *Nature*. 2005;437(7060):894–897.
31. Bai CB, Auerbach W, Lee JS, Stephen D, Joyner AL. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development*. 2002;129(20):4753–4761.
32. Agren M, Kogerman P, Kleman MI, Wessling M, Toftgård R. Expression of the PTCH1 tumor suppressor gene is regulated by alternative promoters and a single functional Gli-binding site. *Gene*. 2004;330:101–114.
33. Macosko EZ, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202–1214.
34. Jeong J, Mao J, Tenzen T, Kottmann AH, McMahon AP. Hedgehog signaling in the neural crest cells regulates the patterning and growth of facial primordia. *Genes Dev*. 2004;18(8):937–951.
35. Lawrence PA, Struhl G. Morphogens, compartments, and pattern: lessons from drosophila? *Cell*. 1996;85(7):951–961.
36. Steiling K, et al. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. *Am J Respir Crit Care Med*. 2013;187(9):933–942.
37. Bourbon JR, Boucherat O, Boczkowski J, Crestani B, Delacourt C. Bronchopulmonary dysplasia and emphysema: in search of common therapeutic targets. *Trends Mol Med*. 2009;15(4):169–179.
38. Wang Z, et al. Optimal threshold in CT quantification of emphysema. *Eur Radiol*. 2013;23(4):975–984.
39. Barkauskas CE, et al. Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest*. 2013;123(7):3025–3036.
40. Plantier L, et al. Defect of hepatocyte growth factor production by fibroblasts in human pulmonary emphysema. *Am J Physiol Lung Cell Mol Physiol*. 2005;288(4):L641–L647.
41. Kennelly H, Mahon BP, English K. Human mesenchymal stromal cells exert HGF dependent cytoprotective effects in a human relevant preclinical model of COPD. *Sci Rep*. 2016;6:38207.
42. Ruzankina Y, et al. Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell*. 2007;1(1):113–126.
43. Naldini L, et al. Hepatocyte growth factor (HGF) stimulates the tyrosine kinase activity of the receptor encoded by the proto-oncogene c-MET. *Oncogene*. 1991;6(4):501–504.
44. Calvi C, et al. Hepatocyte growth factor, a determinant of airspace homeostasis in the murine lung. *PLoS Genet*. 2013;9(2):e1003228.
45. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ. Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science*. 2018;359(6380):1118–1123.
46. Zacharias WJ, et al. Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature*. 2018;555(7695):251–255.
47. Torday J, Rehan V. Neutral lipid trafficking regulates alveolar type II cell surfactant phospholipid and surfactant protein expression. *Exp Lung Res*. 2011;37(6):376–386.
48. Plikus MV, et al. Regeneration of fat cells from myofibroblasts during wound healing. *Science*. 2017;355(6326):748–752.
49. El Agha E, et al. Two-way conversion between lipogenic and myogenic fibroblastic phenotypes marks the progression and resolution of lung fibrosis. *Cell Stem Cell*. 2017;20(2):261–273.e3.
50. Zhao C, et al. Stromal Gli2 activity coordinates a niche signaling program for mammary epithelial stem cells. *Science*. 2017;356(6335):eaal3485.
51. Chuang PT, McMahon AP. Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature*.

1999;397(6720):617–621.
52. Chen Y, Struhl G. Dual roles for patched in sequestering and transducing Hedgehog. *Cell*. 1996;87(3):553–563.
53. Zhou X, et al. Identification of a chronic obstructive pulmonary disease genetic determinant that regulates HHIP. *Hum Mol Genet*. 2012;21(6):1325–1335.
54. Bai CB, Joyner AL. Gli1 can rescue the in vivo function of Gli2. *Development*. 2001;128(24):5161–5172.
55. Gong S, et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature*. 2003;425(6961):917–925.
56. Goodrich LV, Milenkovic L, Higgins KM, Scott MP. Altered neural cell fates and medulloblastoma in mouse ptached mutants. *Science*. 1997;277(5329):1109–111.
57. Kang SH, Fukaya M, Yang JK, Rothstein JD, Bergles DE. NG2+ CNS glial progenitors remain committed to the oligodendrocyte lineage in postnatal life and following neurodegeneration. *Neuron*. 2010;68(4):668–681.
58. Hamilton TG, Klinghoffer RA, Corrin PD, Soriano P. Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. *Mol Cell Biol*. 2003;23(11):4013–4025.
59. Huh CG, Factor VM, Sanchez A, Uchida K, Conner EA, Thorgeirsson SS. Hepatocyte growth factor/c-met signaling pathway is required for efficeint liver regeneration and repair. *Proc Natl Acad Sci U S A*. 2004;101(13):4477–4482.
60. Chapman HA, et al. Integrin α6β4 identifies an adult distal lung epithelial population with regenerative potential in mice. *J Clin Invest*. 2011;121(7):2855–2862.
61. Zhang X, et al. Cathepsin E promotes pulmonary emphysema via mitochondrial fission. *Am J Pathol*. 2014;184(10):2730–2741.
62. Lee JW, Fang X, Gupta N, Serikov V, Matthay MA. Allogeneic human mesenchymal stem cells for treatment of E. coli endotoxin-induced acute lung injury in the ex vivo perfused human lung. *Proc Natl Acad Sci U S A*. 2009;106(38):16357–16362.
63. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33(5):495–502.
64. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res*. 2012;40(17):e133.
65. Steiling K, et al. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. *Am J Respir Crit Care Med*. 2013;187(9):933–942.