Gli1+ mesenchymal stromal cells form a pathological niche to promote airway progenitor metaplasia in the fibrotic lung

Monica Cassandras 1, Chaoqun Wang 1, Jaymin Kathiriya 1, Tatsuya Tsukui 1, Peri Matatia 2, Michael Matthay 1, Paul Wolters 1, Ari Molofsky 2, Dean Sheppard 1, Hal Chapman 1 and Tien Peng 1✉

Aberrant epithelial reprogramming can induce metaplastic differentiation at sites of tissue injury that culminates in transformed barriers composed of scar and metaplastic epithelium. While the plasticity of epithelial stem cells is well characterized, the identity and role of the niche has not been delineated in metaplasia. Here, we show that Gli1+ mesenchymal stromal cells (MSCs), previously shown to contribute to myofibroblasts during scarring, promote metaplastic differentiation of airway progenitors into KRT5+ basal cells. During fibrotic repair, Gli1+ MSCs integrate hedgehog activation signalling to upregulate BMP antagonism in the progenitor niche that promotes metaplasia. Restoring the balance towards BMP activation attenuated metaplastic KRT5+ differentiation while promoting adaptive alveolar differentiation into SFTPC+ epithelium. Finally, fibrotic human lungs demonstrate altered BMP activation in the metaplastic epithelium. These findings show that Gli1+ MSCs integrate hedgehog signalling as a rheostat to control BMP activation in the progenitor niche to determine regenerative outcome in fibrosis.

A canonical feature of wound repair is the formation of scar tissue, which is composed of inflammatory cells and resident mesenchymal subsets that dynamically remodel the wound to close any barrier gaps. Scars are sometimes accompanied by epithelial metaplasia, which is defined as the interconversion of one mature cell type into another that is not typically present at the site of injury9. One of the cardinal features of idiopathic pulmonary fibrosis (IPF), the most prevalent and deadly subtype of progressive fibrotic lung diseases, is the appearance of “bronchiolization”9 histology. This pathological feature denotes the ectopic appearance of proximal bronchus/airway epithelium within the distal lung, which is characterized by metaplastic KRT5+ basal cells lining cystic air spaces in the alveoli alongside patches of fibroblastic scars8–11. While the functional relationship between metaplasia and scarring is unclear in IPF, the appearance of metaplastic KRT5+ cells is correlated with worsening disease severity and survival12. This suggests that epithelial metaplasia is a clinically relevant feature of organ fibrosis and a potential therapeutic target.

The lung is populated by unique resident epithelial stem/progenitor cells along the proximal–distal axis. KRT5+ basal cells reside in the trachea and large airways in the murine lung, and SFTPC+ type 2 cells reside in the distal alveolar sacs to generate functional alveolar epithelium13. In recent years, numerous studies have shown that KRT5+-progenitors from the airway can migrate and reconstitute epithelium in the alveoli in response to severe damage9–15. While the heterogeneity within this airway progenitor population (EpCAM+ cells expressing combinations of Sox2, Tp63, Scgb1a1 and Rgb4) has not been fully resolved, genetic lineage tracing and transplant experiments clearly demonstrate that KRT5+-airway progenitors can replete the damaged alveolar surface with endogenous (SFTPC+) or metaplastic (KRT5+) epithelium with functional consequences depending on the regenerative outcome.

Gli1, a transcriptional readout of hedgehog (Hh) activation, labels mesenchymal cells adjacent to the airway epithelium14. Gli1+ cells exhibit properties of mesenchymal stromal cells (MSCs) ex vivo12–14 and contribute to fibrotic scarring in the lung and other organs by differentiating into myofibroblasts and secreting collagen13. We have previously demonstrated that the ectopic activation of Hh in the alveolar mesenchyme disrupts SFTPC+ alveolar progenitor renewal during homeostasis14. In this study, we set out to define the functional interactions between Gli1+ niche cells and the metaplastic differentiation of airway epithelial progenitors in the fibrotic lung.

Results

Gli1+ mesenchymal cells promote metaplastic KRT5 differentiation from the airway. SOX2+ progenitors are normally restricted to the airway epithelium during homeostasis, but they migrate into the alveoli during severe alveolar damage from influenza14. To determine whether SOX2+ progenitors have a similar migratory capacity in a fibrotic model, we adapted a repetitive-dose bleomycin model of lung fibrosis that better captured features of IPF compared with the single-dose model, including the appearance of cystic air spaces and metastatic basal cells in the alveoli15–18 (Extended Data Fig. 1a,b). Our genetic lineage tracing confirmed that Sox2 lineage-traced (Sox2 Lin+) airway progenitors expand and differentiate into metaplastic KRT5+SOX2+ or endogenous SFTPC+SOX2− epithelium in the alveoli in this model of severe fibrotic injury (Fig. 1a,b and Extended Data Fig. 1c,d). The fidelity of the Sox2/SOFTZ2 driver was validated by anti-SOX2 antibody staining of tamoxifen and corn oil controls (Extended Data Fig. 1e–g). Interestingly, p63 Lin+ airway

1Department of Medicine, Division of Pulmonary and Critical Care Medicine, Cardiovascular Research Institute, San Francisco, CA, USA. 2Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, USA. ✉E-mail: tien.peng@ucsf.edu
progenitors, which were previously found to contribute to KRT5+ basal cells after influenza infection8, did not significantly contribute to alveolar KRT5+ basal cells after bleomycin treatment (Extended Data Fig. 2a–d). In contrast, Sgb1a1 Lin+ airway progenitors, which co-express Sox2 (ref. 7), contributed a significant proportion of alveolar basal cells after bleomycin treatment (Extended Data Fig. 2e,f).

Immunofluorescence analysis of Gli1+ lineage-labelled cells (Gli1 Lin+) demonstrated that they are localized immediately beneath the basement membrane in close proximity to SOX2+ airway epithelium (Fig. 1c,d) during normal homeostasis. Gli1 Lin+ mesenchymal cells also responded to fibrotic injury by migrating into the alveolar compartment (Fig. 1e and Extended Data Fig. 3d). The fidelity of the Gli1creERT2 driver was validated by corn oil controls and a comparable Gli1 reporter (Extended Data Fig. 3a–c). Cell-to-cell distance quantification showed that the ectopic SOX2+ cells in the alveoli maintain a close distance to Gli1 Lin+ cells that migrated into the alveoli (Fig. 1f), which suggests that there is a functional interaction between migratory Gli1 Lin+ mesenchyme and SOX2+ progenitors as niche partners during repair.

To determine whether Gli1+ mesenchymal cells are required for metaplastic KRT5 differentiation in the alveoli, we genetically deleted Gli1+ cells by inducing the expression of diphtheria toxin A (DTA) within Gli1+ expressing cells and evaluated airway progenitor differentiation as previously described5,14. Tamoxifen induction of Gli1creERT2;R26DTA/DTA (Gli1-deleted) animals alone without injury did not change the composition of the airway epithelium (Extended Data Fig. 3g). Tamoxifen induction of Gli1-deleted animals followed by repetitive bleomycin treatment resulted in significantly reduced areas of metaplastic KRT5+ pods on histology compared with control R26DTA/DTA animals. There was also a higher proportion of alveolar epithelium marked by the type 2 cell marker SFTPC (Fig. 1g,h), along with a trend towards a decrease in fibrotic markers (Extended Data Fig. 3j). This suggests that activated Gli1+ mesenchyme cells form a pathological niche to promote metaplastic differentiation after fibrotic injury.

To determine the ability of Gli1+ mesenchymal cells to modify airway progenitors in vitro, we co-cultured airway progenitors isolated from the lung (separated from the trachea) with or without Gli1+ mesenchymal cells in a heterotypic three-dimensional (3D) organoid assay5 (Fig. 2a and Extended Data Fig. 4c). Cytospin analysis confirmed that ITGB4-sorted lung epithelium largely overlapped with SOX2, with minimal KRT5- or SFTPC-expressing cells (Extended Data Fig. 4a,b). ITGB4+ progenitors cultured without mesenchyme failed to form organoids, whereas co-culture of ITGB4+ airway progenitors with Gli1+ mesenchyme resulted in significant growth in the 3D Matrigel (Fig. 2b,c). Histological analysis of the organoids co-cultured with Gli1+ cells demonstrated heterogeneous lineage differentiation that included organoids containing either KRT5+ or SFTPC+ cells, but never both in the same organoid (Fig. 2d). To ensure that the newly formed KRT5+ or SFTPC+ cells did not arise from pre-existing KRT5+ or SFTPC+ cells in the ITGB4-sorted samples, we isolated ITGB4+ progenitors of lungs of Krt5- or Sftpc-lineage-traced lungs for co-culture with Gli1+ mesenchyme. Immunophenotyping showed that <5% of the KRT5+ or SFTPC+ cells in the organoids were Krt5 or Sftpc Lin+, respectively (Extended Data Fig. 4d–j). We next isolated Gli1 Lin+ or Lin− mesenchyme from PBS- or bleomycin-treated lungs for co-culture with healthy ITGB4+ airway progenitors (Fig. 2e). This demonstrated that Gli1 Lin+ mesenchyme increased the colony forming efficiency (CFE) of airway progenitors in both uninjured and injured lungs compared with Gli1 Lin− mesenchyme (Fig. 2f,g). Last, immunophenotyping of the organoids containing Gli1 Lin+ mesenchyme showed that bleomycin-induced injury in the mesenchyme significantly increased the proportion of KRT5+ basal cell organoids, with a trend towards a reduction in SFTPC+ organoids compared with PBS-treated Gli1 Lin+ mesenchyme (Fig. 2h). This demonstrates that Gli1+ mesenchyme forms a niche to support airway progenitor expansion and differentiation in vitro.

Mesenchymal Hh activation promotes metaplastic KRT5 differentiation. Gli1+ mesenchyme integrates extracellular cues during normal homeostasis and repair, including the availability of the SHH ligand that modulates the level of Hh activation in the tissue microenvironment5. Analysis of Shh expression during bleomycin-induced injury showed expression in the ectopic SOX2+ airway progenitors that migrated into the alveoli (Extended Data Fig. 5a). To determine the role of mesenchymal Hh activation as a niche signal modulating airway progenitor differentiation, we induced the expression of a constitutively active form of the Hh effector Smo (SmoM2) in Gli1+ cells5. Overactivation of Hh in the Gli1+ mesenchyme through tamoxifen induction of Gli1creERT2;R26Smom2 (SmoM2) (referred to as Hh activated) animals did not induce metaplastic differentiation in the uninjured lung (Extended Data Fig. 5b,c). However, when Hh overactivation was followed by fibrotic injury with bleomycin, Hh-activated animals demonstrated a significant increase in areas of the lung with metaplastic KRT5+ basal cells along with a significantly reduced proportion of functional SFTPC+ alveolar epithelium compared with controls (Fig. 3a,b). We then performed genetic deletion of Smo using Gli1creERT2;R26Smom2Smyc2Min/+ (referred to as Smo deleted) animals did not induce metaplastic differentiation in the uninjured lung (Extended Data Fig. 5d,e). For control (Extended Data Fig. 5f–i). These data show that Hh activation within Gli1+ mesenchyme, in conjunction with a fibrotic stimulus, acts in trans to promote metaplastic differentiation in the airway progenitor niche.

To determine whether mesenchymal Hh activation in vivo promotes the differentiation of isolated airway progenitors in vitro, we isolated injured control or Hh-activated Gli1 Lin+ lungs mesenchyme with healthy ITGB4+ airway progenitors in our organoid assay.
Hh-activated Gli1 Lin+ mesenchyme increased the CFE of airway progenitors and increased KRT5+ basal cell organoids at the expense of SFTPC+ organoids compared with control mesenchyme (Fig. 3g–i). To activate Hh in the lung mesenchyme in vitro, we first generated a transgenic animal (UbccreERT2+/+;R26SmoM2+/+), whereby constitutively active SmoM2 could be induced in isolated lung mesenchymal cells in vitro utilizing a creERT2 allele driven by a ubiquitously expressed promoter as previously described11,15. We activated Hh in isolated lung mesenchyme by adding 4-hydroxytamoxifen (4OHT) followed by co-culture with wild-type ITGB4+ progenitors in our 3D organoid assay. Activation of Hh in the mesenchyme in vitro resulted in the appropriate induction of Gli1 in the separated mesenchyme (Extended Data Fig. 5h). Hh activation in the mesenchyme significantly increased the expression of Krt5 in the epithelial organoids and decreased the expression of Sftpc along with similar trends in immunophenotyping of KRT5+ and SFTPC+ organoids (Extended Data Fig. 5i–l). These results demonstrate that Hh activation of mesenchyme promotes the metaplastic
Single-cell transcriptome analysis of Gli1+ mesenchyme reveals upregulated BMP antagonism in the fibrotic niche. To determine the transcriptome alterations that could modify the ability of Gli1+ cells to alter airway progenitor behaviour, we performed single-cell RNA sequencing (scRNA-seq) on sorted Gli1+ mesenchymal cells before and after fibrotic injury (Fig. 4a). A total of 17,620 cells were captured from lineage-labelled Gli1+ cells from Gli1<sup>creERT<sup/>-R26<sup>/YFP</sup> lungs treated with PBS or bleomycin. PBS- and bleomycin-treated datasets were merged, and projection of transcriptomic variations of individual cells by uniform manifold approximation and projection (UMAP)<sup>20</sup> showed three main clusters that were defined by their anatomical localization and differentiation (Fig. 4b). Based on our prior study showing that mesenchyme surrounding the airway exhibits a distinct identity from those in the alveoli,<sup>17</sup>, cluster 1 and cluster 2 represent the airway and alveolar subsets, respectively, whereas cluster 3 represents a de novo population expressing myofibroblast markers that appears only during fibrotic injury in the alveoli (Fig. 4c). Bleomycin-injured Gli1+ cells demonstrated an increased proportion of cells in the alveolar subset (Fig. 4d), and gave rise to the entire myofibroblast subset that generates scarring in the alveoli along with portions of the alveolar subset (Fig. 4d, green box). RNA velocity analysis<sup>21</sup> on Gli1+ mesenchymal cells treated with bleomycin predicted two distinct trajectories, whereby airway mesenchymal cells gave rise to either alveolar mesenchymal cells or myofibroblasts during fibrotic repair (Fig. 4e). This result was supported by multicolour clonal analysis using the R26<sup>Rosa26<sup>-<sup>YFP</sup> reporter coupled with Gli1 lineage tracing after bleomycin-induced injury (Extended Data Fig. 6a). Differential gene expression analysis between Gli1+ cells treated with PBS versus bleomycin showed that Gli1+ cells retained Gli1 expression after injury even as they migrate into the alveoli (Extended Data Fig. 6b). Gene Ontology (GO) analysis of the upregulated genes in fibrotic Gli1+ mesenchyme classified them as “TGF-β signalling components” and “negative regulation of BMP signalling” (Fig. 4f,g). The upregulation of secreted BMP antagonists, particularly Grem2 and Fst, was most pronounced in the Gli1+<sup>−</sup> cells that appeared de novo in bleomycin, including the myofibroblast cluster (Fig. 4h). Secreted BMP antagonists are almost exclusively expressed in the mesenchyme of the lung in a publicly available murine lung cellular atlas (Tabula Muris)<sup>22</sup>. Indeed, numerous BMP ligands were also expressed in the Gli1+ mesenchyme, including Bmp3, Bmp4, Bmp5 and Bmp7 (Fig. 4i), with Bmp7 significantly downregulated in bleomycin treatment. Our single-cell analysis suggests that Gli1+ mesenchyme upregulates local BMP antagonism during fibrotic repair through the secretion of soluble BMP antagonists in the fibrotic niche.

Hh activation upregulates BMP antagonism in the fibrotic lung. During organ morphogenesis, Hh and BMP signalling are often mutually antagonistic, forming antiparallel gradients of activation within the same tissue to define segmental cell fate<sup>23</sup>. Activation of the BMP receptor leads to intracellular phosphorylation of SMAD1, SMAD5 and SMAD8 (SMAD1/5/8), which serves as a readout for BMP activation<sup>44</sup>. Quantification of phosphorylated SMAD1/5/8 (pSMAD1/5/8) staining in the Sox2+ cells in the alveoli after injury showed that the majority of BMP-activated Sox2+ cells were SCGB1A1<sup>+</sup> or SFTPC<sup>+</sup>, but rarely KRT5<sup>+</sup> (Extended Data Fig. 7a,b). To determine the effect of Hh activation on BMP activation, we quantified the pSMAD1/5/8 areas in the bleomycin-injured lungs, which showed a reduction of alveolar areas stained by pSMAD1/5/8 in the Hh-activated lungs compared with controls (Fig. 5a,b), which suggests that Hh activation antagonizes BMP activation in fibrotic repair.

Next, we examined the expression of secreted BMP antagonists in Hh-activated and control lungs treated with bleomycin. Genes encoding soluble BMP antagonists were upregulated in the fibrotic Hh-activated lungs compared with controls according to whole-lung quantitative PCR (qPCR) analysis (Fig. 5c). These data suggest that Hh activation antagonizes BMP activation during fibrotic repair by upregulating secreted antagonists of BMP signalling in the Gli1+ mesenchyme.

To determine whether Hh activation directly upregulates BMP antagonists in vitro, we collected Hh-inducible mesenchyme from Ubc<sup>creERT<sup/>-R26<sup>lox<sup>−<sup>YFP<sup> lungs and induced Hh activation with 4OHT (Fig. 5d and Extended Data Fig. 7c). Surprisingly, Hh-activated mesenchyme showed minimal induction of the BMP antagonist Grem2. Recombinant TGF-β1, a fibrotic stimulus upregulated in the fibrotic Gli1+ mesenchyme, also had minimal effect on Grem2 expression when added alone to control lung mesenchyme. However, TGF-β1 activation in conjunction with Hh activation in vitro significantly increased the expression of Grem2 (Fig. 5d), which was the most upregulated BMP antagonist in the Hh-activated lungs in vivo (Fig. 5c). This suggests a synergistic interaction between TGF-β and Hh in upregulating BMP antagonism in the fibrotic niche. Conversely, both Hh and TGF-β activation can independently suppress the expression of Bmp3, Bmp4, Bmp5 and Bmp7 from the mesenchyme, with the most profound effect on Bmp4 and Bmp5 when both Hh and TGF-β are activated (Fig. 5d). Hh activation in the mesenchyme reduced pSMAD1/5/8 staining in the epithelial organoids, along with expression of the BMP target genes Id1 and Id2 in the epithelium (Fig. 5e–h). These results demonstrate that Hh activation suppresses BMP activation by simultaneously upregulating soluble BMP antagonists in the mesenchyme and downregulating BMP ligands in the fibrotic niche (Fig. 5i).

BMP activation attenuates metaplastic airway progenitor differentiation. To test whether BMP activation can modify metaplastic KRT5 differentiation in vivo, we directly dosed recombinant BMP into the lung undergoing bleomycin treatment. Histologically, injured wild-type animals treated with recombinant human BMP4 (rhBMP4) demonstrated a significant reduction in the areas of lung occupied by KRT5<sup>+</sup> cystic air space concurrent with an increase in the proportion of SFTPC<sup>+</sup> cells compared to vehicle, along with an improvement in oximetry and fibrotic markers (Fig. 6a–c and 6d–f).
Extended Data Fig. 8a–c). One possible side effect of rhBMP4 is the induction of chondrogenic markers in the lung, as seen with qPCR results (Extended Data Fig. 8e). To determine the lineage origin of the SFTPC$^+$ cells seen in rhBMP4-treated lungs, we repeated the same experiment with Sox2-lineage-traced animals. Quantification of Sox2$^{Lin+}$ cells in the alveoli showed that rhBMP4 increased the proportion of SFTPC$^+$ cells with a concurrent reduction in KRT5$^+$ cells (Fig. 6d,e), with no difference in the total number of Sox2$^{Lin+}$ cells that migrated into the alveoli (Extended Data Fig. 8f).
Mesenchymal Hh activation promotes metaplastic KRT5 differentiation while suppressing SFTPC differentiation from airway. 

Fig. 3 | Schematic (a, upper), histology images (a, lower) and quantification (b) of areas of KRT5+ and percentage of SFTPC+ cells after constitutive Hh activation (activation with SmoM2) in Gli1+ cells followed by bleomycin injury (n = 9 per group). c-e, Schematic (c, lower), histology images (c, lower) and quantification of areas of KRT5+ and percentage of SFTPC+ cells followed by bleomycin injury (n = 7 for control, n = 9 for Hh inactivated). Data are expressed as the mean ± s.d. f-i, Schematic of the experimental protocol (f), images (g) CFE expressed as percentage (h, left), organoid size (h, right) and composition (i) of airway progenitors co-cultured with Gli1+ mesenchyme isolated from Hh-activated fibrotic lungs compared with Gli1+ mesenchyme from control fibrotic lung (n = 5 per group for h, n = 3 per group for i). j, Model of Hh activation in the Gli1+ mesenchyme acting in trans to promote SOX2+ progenitor differentiation into metaplastic KRT5+ basal cells. For b, d, e, h and i, each data point represents one animal (b, d and e) or one well (h and i); all data are expressed as the mean ± s.d., with one-tailed unpaired Student’s t-tests performed. Scale bars, 100 μm (a and c).
Fig. 4 | Single-cell transcriptome analysis of Gli1 Lin+ mesenchyme reveals upregulated BMP antagonism in the fibrotic niche. a, Schematic of the workflow used to isolate Gli1 Lin+ cells during homeostasis (PBS) and fibrosis (bleomycin) for single-cell analysis (one animal per treatment). b, Left: the heatmap of merged Gli1 Lin+ cells (PBS and bleomycin) segregates into three distinct clusters based on the transcriptome, with highest expressing cluster-defining genes labelled on the right (coloured bars denote relative expression). Right: UMAP plot of the three distinct clusters that appear segregated based on anatomical location. c, In situ (RNAsecope) and immunofluorescence analysis shows the spatial distribution of signature genes for alveolar, airway and myofibroblast (MF) clusters (coloured bars denote expression level). This experiment was repeated twice independently with similar results. Scale bars, 100 μm. d, UMAP plot with cells of origin based on treatment shows an increase in the proportion of cells in the alveolar and myofibroblast cluster after fibrotic injury. The proportion of the clusters (coloured bars denote expression level). This experiment was repeated twice independently with similar results. Scale bars, 100 μm.

f, GO term analysis of differentially expressed genes between PBS- and bleomycin-treated Gli1 Lin+ cells reveals enrichment for genes involved in the negative (Neg) regulation of bMP signalling during fibrotic repair (P adj. = 2.26 × 10^{-2}; MAST test). g, Gene correlation plot, with each dot representing a gene and genes significantly upregulated in fibrotic Gli1 Lin+ cells in red (P adj. < 0.1, log(f.c.) > 0.15) and downregulated (P adj. < 0.1, log(f.c.) < -0.15) in blue. Secreted BMP antagonists (GO term: negative regulation of BMP signalling) are labelled in green (Grem2 P adj. = 1.32 × 10^{-2}; Fst P adj. = 2.97 × 10^{-22}; Wnt5a P adj. = 2.18 × 10^{-3}; Sfrp1 P adj. = 2.99 × 10^{-22}; Fzd13 P adj. = 1.78 × 10^{-17}; Fzd7 P adj. = 2.36 × 10^{-18}; Bmp7 P adj. = 8.77 × 10^{-15}; MAST test). h, Gene feature plots (upper) and violin plots (lower) of secreted BMP antagonists show enrichment in myofibroblast and part of the distal cluster (green box). The violin bodies represent the distribution of the cells (n = 11,925 cells). i, Gene feature plots of BMP ligand expression in Gli1 Lin+ mesenchyme.
To test the effect of BMP modulation in the airway progenitor in vitro, we co-cultured ITGB4+ airway progenitors and Gli1+ mesenchyme with recombinant BMP4 or with a BMP antagonist, GREMLIN2, in our 3D organoid assay (Extended Data Fig. 8g). Addition of BMP4 increased the expression of Sftpc in the epithelial organoid and reduced the expression of Krt5, whereas GREMLIN2 had the opposite effects (Extended Data Fig. 8i). Furthermore, addition of BMP4 to airway progenitors co-cultured with Hh-activated mesenchyme attenuated KRT5 metaplasia in the organoids (Fig. 6f,g), thereby further supporting the idea that BMP antagonism occurs downstream of Hh activation to drive KRT5 differentiation. These data demonstrate that BMP activation promotes the differentiation of mobilized airway progenitors into adaptive SFTPC+ alveolar epithelium (Fig. 6h) and that the prevention of epithelial metaplasia is physiologically significant for improving gas exchange.

### IFP lungs display altered BMP activation in the epithelium.

To determine whether fibrotic human lung mesenchyme undergoes similar transcriptomic alterations as seen in the murine Gli1+ mesenchyme, we analysed scRNA-seq data of human lungs from patients with IFP and from healthy controls (n = 3 per group). The diseased lungs were collected from patients with an established diagnosis of IFP undergoing lung transplantation, and controls were from sex- and age-matched cadaveric donor lungs without prior history of lung disease. Mesenchymal cells were segregated based on the presence of PDGFRα as a mesenchymal marker, with PDGFRα+ cells from each lung were merged for comparison (Fig. 7a,b and Extended Data Fig. 9a). UMAP analysis showed the emergence of cellular subsets in the fibrotic lung that are distinct from PDGFRα− cells from healthy controls (Extended Data Figs. 7b and 9a), with cluster 3 derived almost entirely from IFP lungs and enriched for scar components such as genes associated with myofibroblast differentiation (ACTA2 and ASPN) and matrix deposition (COL1A1) (Fig. 7c,d). Differential gene expression analysis of the scRNA-seq dataset with GO analysis showed a significant enrichment of BMP antagonists in the fibrotic human lung mesenchyme compared with healthy controls (Fig. 7e,f). Furthermore, the upregulated BMP antagonists demonstrated particular enrichment in the scar-forming fibroblasts (cluster 3) (Fig. 7g).

Histological analysis of IFP lungs demonstrated areas of intense scarring in the alveoli, which were marked by the presence of SMA+ fibroblastic foci, which are often localized adjacent to metaplastic KRT5+ cells, neither of which were present in the healthy alveoli (Fig. 7h). Similar to mouse, metaplastic KRT5+ basal cells in the alveoli of IFP lungs co-labelled with SOX2, along with KRT14, NGFR and p63 (Extended Data Fig. 9b), which are also normally markers of airway differentiation in the human lung22. Cell-to-cell distance quantification of metaplastic SOX2+KRT5+ cells in the alveoli demonstrated that they are equally proximate to the SMA+ fibroblastic foci compared with the endogenous SFTPC+ epithelium (Extended Data Fig. 9c,d).

Unlike in the mouse lung, KRT5+ basal cells in healthy human airways extend more distally, lining the basal side of the pseudostratified epithelium of the distal airways5. Analysis of pSMAD1/5/8 showed intense staining in the airway epithelium of control lungs, including uniform expression in the KRT5+ basal cells (Fig. 7j). However, in the IFP lungs, the airway epithelium appeared dysplastic, with overgrowth of the SOX2+ epithelium and nonuniform distribution of pSMAD1/5/8 staining within the disorganized KRT5+ basal cell layer (Fig. 7j). Furthermore, metaplastic SOX2+KRT5+ basal cells in the alveoli demonstrated a significant reduction in pSMAD1/5/8 staining compared with their airway counterparts (Fig. 7j,k), similar to what we observed in the mouse data (Extended Data Fig. 7a,b). Finally, unsorted epithelial organoids were cultured from dissociated IFP lungs and treated with BMP4, which significantly attenuated the growth of KRT5+ organoids (Extended Data Fig. 9e,f). This suggests that BMP activation can either attenuate KRT5+ basal cell proliferation or reduce differentiation into KRT5+ basal cells from a precursor population not yet identified in human lung. These data show that the fibroblastic scars in IFP upregulate BMP antagonism in the local environment, concurrent with a downregulation of BMP activation in metaplastic KRT5+ cells that appears to be a conserved feature of epithelial metaplasia in lung fibrosis.

### Discussion

The ectopic appearance of mature epithelial cell types under pathological settings (that is, metaplasia) is a feature of chronic wounds that often appears at tissue transition zones3. Our study demonstrated how a putative MSC population (Gli1+ mesenchymal cells) that contributes to scarring can form a pathological niche to promote epithelial metaplasia, thereby providing a functional link between two commonly observed pathological phenomena in organ fibrosis. Utilizing temporally and spatially precise in vivo manipulation and ex vivo cellular modelling, we showed that Gli1+ mesenchyme integrates Hh activation to modify the extracellular BMP environment, which in turn regulates the metaplastic fate of adjacent airway progenitors in fibrotic repair (Extended Data Fig. 10).

Hh and BMP are key developmental morphogens that often form activation gradients in an antiparallel orientation to generate asymmetric patterning. Cells in the developing foregut and neural tube sense the local balance between Hh and BMP activation that specifies ventral (Hh predominant) versus dorsal (BMP predominant) fate23. While mutual antagonism between Hh and BMP is...
crucial for morphogenesis, questions remain as to whether and how the two pathways continue to interact in adult tissues during maintenance and repair. In the respiratory system, it has been reported that BMP antagonists are upregulated after injury to the tracheal epithelium, and inhibition of BMP drives basal cell proliferation\textsuperscript{26,27}. Our single-cell data show that \(Gli1^+\) mesenchyme is the key source of numerous soluble BMP antagonists, along with BMP ligands that are expressed in the mesenchyme and other cell types. This suggests that there is a cellular microenvironment in the adult lung where the balance between BMP ligands and soluble antagonists is tightly regulated, and Hh controls the secretion of BMP antagonists in \(Gli1^+\) mesenchymal cells to modulate the state of local BMP activation. We have previously shown that Hh maintains airway mesenchymal cell identity\textsuperscript{15}, but what was surprising about our new data is that Hh also appeared to promote airway epithelial cell fate \textit{in trans} by driving BMP antagonism in the fibrotic progenitor niche. This suggests a mechanism whereby Hh and BMP continue to compete in adult tissues to segregate compartmental identity during regeneration by acting through a mesenchymal effector that modifies epithelial cell fate.

Overactivation of Hh has been reported in patients with IPF\textsuperscript{28–30}, but pharmacological inhibition of the pathway has produced mixed...
results in experimental models of lung fibrosis. The overwhelming endpoints of these experimental fibrotic studies have been the reduction of scarring, as measured by collagen content in the lungs as a therapeutic endpoint. However, our study demonstrated that epithelial metaplasia is an equally important endpoint in lung fibrosis. We found that administration of rhBMP4 can promote the repletion of alveolar epithelium with endogenous SFTPC+ type 2 cells, which also improved gas exchange as seen in the improved...
oximetry results. While the cellular origin of the ectopic KRT5+ basal cells in human IPF remains unclear, the downregulation of BMP activation in the ectopic alveolar KRT5+ basal cells compared with airway KRT5+ basal cells suggests that BMP could regulate the ‘bronchiolization’ of alveoli, as supported by previous studies demonstrating an upregulation of BMP antagonists in lung fibrosis33–35. Our study highlights the need for other physiologically relevant endpoints in experimental fibrosis studies, and suggest that epithelial metaplasia presents both an attractive therapeutic target and a clinical endpoint for tissue fibrosis.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of
data and code availability are available at https://doi.org/10.1038/s41556-020-00591-9.

Received: 15 January 2020; Accepted: 9 September 2020; Published online: 12 October 2020

References
1. Giroux, V. & Rustgi, A. K. Metaplasia: tissue injury adaptation and a precursor to the dysplasia-cancer sequence. Nat. Rev. Cancer 17, 594–604 (2017).
2. Chilosi, M. et al. Abnormal re-epithelialization and lung remodeling in idiopathic pulmonary fibrosis: the role of ΔN-p63. Lab. Invest. 82, 1335–1345 (2002).
3. Seibold, M. A. et al. The idiopathic pulmonary fibrosis honeycomb cyst contains a mucociliary pseudostratified epithelium. PLoS ONE 8, e56658 (2013).
4. Xu, Y. et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight 1, e90558 (2016).
5. Prasse, A. et al. BAL cell gene expression is indicative of outcome and airway basal cell involvement in idiopathic pulmonary fibrosis. Am. J. Respir. Crit. Care Med. 199, 622–630 (2019).
6. Hogan, B. L. et al. Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. Cell Stem Cell 15, 123–138 (2014).
7. Ray, S. et al. Rare SOX2+ airway progenitor cells generate KRT5+ cells that repopulate damaged alveolar parenchyma following influenza virus infection. Stem Cell Rep. 7, 817–825 (2016).
8. Vaughan, A. E. et al. Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature 517, 621–625 (2015).
9. Yang, Y. et al. Spatial–temporal lineage restrictions of embryonic p63+ progenitors establish distinct stem cell pools in adult airways. Dev. Cell 44, e794 (2018).
10. Yee, M. et al. Alternative progenitor lineages regenerate the adult lung depleted of alveolar epithelial type 2 cells. Am. J. Respir. Cell Mol. Biol. 50, 453–464 (2017).
11. Peng, T. et al. Hogan's directed differentiation strategy maintains alveolar quiescence and regulates repair and regeneration. Nature 526, 578–582 (2015).
12. Zhao, H. et al. The sputum provides a niche for mesenchymal stem cells of craniofacial bones. Nat. Cell Biol. 17, 386–396 (2015).
13. Zhao, H. et al. Secretion of Shh by a neurovascular bundle niche supports mesenchymal stem cell homeostasis in the adult mouse incisor. Cell Stem Cell 14, 160–173 (2014).
14. Kramann, R. et al. Perivascular Gli1+ progenitors are key contributors to injury-induced organ fibrosis. Cell Stem Cell 16, 51–66 (2015).
15. Wang, C. et al. Expansion of hoganoid disrupts mesenchymal identity and induces emphysema phenotype. J. Clin. Invest. 128, 3453–3458 (2018).
16. Xi, Y. et al. Local lung hypoxia determines epithelial fate decisions during alveolar regeneration. Nat. Cell Biol. 19, 904–914 (2017).
17. Degryse, A. L. et al. Replicative intratracheal bleomycin models several features of idiopathic pulmonary fibrosis. Am. J. Physiol. Lung Cell Mol. Physiol. 299, L442–L452 (2010).
18. Kurche, J. S. et al. Muc5b enhances murine honeycomb-like cyst formation. Am. J. Respir. Cell Mol. Biol. 61, 544–546 (2019).
19. Jeong, J., Mao, J., Tenzen, T., Kottmann, A. H. & McMahon, A. P. Hedgehog signaling in the neural crest cells regulates the patterning and growth of facial primordia. Gene Dev. 18, 937–951 (2004).
20. Recht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. https://doi.org/10.1038/nbt.4314 (2018).
21. La Manno, G. et al. RNA velocity of single cells. Nature 560, 494–498 (2018).
22. The Tabula Muris Consortium et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 562, 367–372 (2018).
23. Zagoski, M. et al. Decoding of position in the developing neural tube from antiparallel morphogen gradients. Science 356, 1379–1383 (2017).
24. Massague, J., Seoane, J. & Wotton, D. Smad transcription factors. Genes Dev. 19, 2783–2810 (2005).
25. Plantier, L. et al. Ectopic respiratory epithelial cell differentiation in bronchiolised distal airspaces in idiopathic pulmonary fibrosis. Thorax 66, 651–657 (2011).
26. Tadokoro, T., Gao, X., Hong, C. C., Hotten, D. & Hogan, B. L. BMP signaling and cellular dynamics during regeneration of airway epithelium from basal progenitors. Development 143, 764–773 (2016).
27. Mou, H. et al. Dual SMAD signaling inhibition enables long-term expansion of diverse epithelial basal cells. Cell Stem Cell 19, 217–231 (2016).
28. Bolanos, A. L. et al. Role of sonic hedgehog in idiopathic pulmonary fibrosis. Am. J. Physiol. Lung Cell Mol. Physiol. 303, L978–L990 (2012).
29. Jia, G. et al. CXCL14 is a candidate biomarker for Hedgehog signalling in idiopathic pulmonary fibrosis. Thorax 72, 780–787 (2017).
30. Lee, J. et al. Increased primary cilia in idiopathic pulmonary fibrosis. Mol. Cells 41, 224–233 (2018).
31. Liu, L. et al. Hedgehog signaling in neonatal and adult lung. Am. J. Respir. Cell Mol. Biol. 48, 703–710 (2013).
32. Moshai, E. et al. Targeting the hedgehog-gloma-associated oncogene homolog pathway inhibits bleomycin-induced lung fibrosis in mice. Am. J. Respir. Cell Mol. Biol. 51, 11–25 (2014).
33. Koli, K. et al. Bone morphogenetic protein-4 inhibitor gremlin is overexpressed in idiopathic pulmonary fibrosis. Am. J. Pathol. 169, 61–71 (2006).
34. Myllariemi, M. et al. Gremlin-mediated decrease in bone morphogenetic protein signaling promotes pulmonary fibrosis. Am. J. Respir. Crit. Care Med. 177, 321–329 (2008).
35. Dong, Y. et al. Blocking follistatin-like 1 attenuates bleomycin-induced pulmonary fibrosis in mice. J. Exp. Med. 212, 235–252 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020
Methods

Animal studies. All animals were housed and treated in compliance with the Institutional Animal Care and Use Committee (IACUC) protocol approved at the University of California San Francisco (UCSF). Mice were housed under pathogen-free conditions at room temperature, maintained under a standard 12 h light–dark cycle and monitored daily by caretakers or researchers. Cage, bedding and food were changed twice a week. C57BL/6 mice were obtained from Charles River Laboratories. The generation and genotyping of the Gli1creERT, Gli1LacZ, Sox2creERT, SggBuTA, p63creERT, SPCC60, KRT5creERT, Smoo2, Gli1cre, R26RsfGFP, R26RtdTomato, R26Ror, SmoFlox and R26Flox lines were performed as previously described by The Jackson Laboratory. For all animal experiments, 6–12-week-old littermate male and female mice were gender balanced and randomly assigned to experimental groups. Mice used in experiments were not involved in previous procedures.

Animal treatment. For lineage-tracing studies, tamoxifen (T5648, Sigma-Aldrich) was dissolved in corn oil and administered intraperitoneally at 200 mg per kg body weight per day for five consecutive days, followed by at least 7 days before bleomycin was administered. For bleomycin-induced injury, mice were given pharmaceutical-grade bleomycin (Hospira) dissolved in PBS via intranasal installation once a week for 4 weeks. Mixed-background mice were given a dose of 1 U per kg and C57BL/6-background mice were given 0.75 U per kg per dose. Mice were killed 7 days after the final treatment. For BMP4 treatment, mice were treated with 5 µg/ml of hrBMP4 protein in 50 µl (314-BP-010, R&D Systems) via intranasal installation twice a week for 4 weeks, concurrently with bleomycin treatment. Mice were killed 7 days after the final bleomycin treatment and 2 days after the final BMP4 treatment.

Human lung tissue. Studies involving human tissue were approved and performed in compliance with the UCSF Institutional Review Board. IPF lungs were obtained from patients independently diagnosed with interstitial pneumonia or scleroderma at the time of lung transplant. Sex- and age-matched healthy human lungs were provided by the Northern California Donor Network from brain-dead donors in whom the lung was independently rejected for transplantation. All subjects provided written informed consent. Samples were collected from all ethnic groups, both male and female individuals, and all samples were from individuals over 50 years old. Investigators were blind to the patient information besides the relevant history of lung disease. No patient recruitment was performed for this study.

Histology and immunofluorescence. The right ventricles of mice were perfused with 1–3 ml PBS and the lungs were inflated with 1–3 ml 4% paraformaldehyde (PFA) in PBS, and then fixed in 4% PFA overnight at 4 °C. After fixation, the lungs were washed with cold PBS 4 times for 30 min each at 4 °C and then dehydrated in a series of increasing ethanol concentration washes (30%, 50%, 70%, 95% and 100%) for a minimum of 2 h per wash. The dehydrated lungs were incubated with xylene for 1 h at room temperature and then with paraffin at 65 °C for 90 min twice, and then embedded in paraffin. The lungs were sectioned at 8 µm on a microtome.

For histological analysis of organoid assays, Transwells were fixed in 4% PFA overnight at 4 °C, then washed in PBS overnight at least three times. Transwells were cut and embedded in OCT, then 8-µm sections were cut on a microtome.

For immunofluorescence staining, paraffin sections were incubated in xylene for 10 min twice, then rehydrated in ethanol washes (100% twice, 70%, 50% and 50% ethanol for 10 min each) at room temperature. OCT-embedded slides were fixed in 4% PFA overnight at 4 °C, then washed in PBS overnight at least three times. Transwells were cut and embedded in OCT, then 8-µm sections were cut on a cryostat.

For single-cell capture and sequencing. For mouse scRNA-seq, all live Lin+ cells were sorted from two Gli1creERT2/R26RsfGFP adult lungs. Lin+ cells were isolated based on forward and side scatter, DAPI exclusion and GFP fluorescence. One mouse was challenged with repetitive-dose bleomycin and one was treated with vehicle at the same time points as a control. The sort was performed on the Cyte 8. The single animal was cut with scissors and resuspended in a digestion cocktail of 0.25% collagenase A (SC136, Millipore Sigma), dispase II (17105041, Thermo Fisher; 15 U/ml) and DNase I (DN25, Sigma-Aldrich; 50 U/ml) and removed from the chest. The lung was further digested with a razor blade and incubated in digestion cocktail for 45 min at 37 °C with continuous shaking. The mixture was then washed with sorting buffer (2% fetal bovine serum (FBS) and 1% penicillin–streptomycin in DMEM medium (31053028, Thermo Fisher)). The mixture was passed through a 70-µm cell strainer and resuspended in RBC lysis buffer (NC906751, Thermo Fisher), then passed through a 40-µm cell strainer. Cell suspensions were incubated with the appropriate conjugated antibodies in sorting buffer for 30 min at 4 °C and then washed with sorting buffer.

For human samples, multiple pieces were collected from whole lungs and separated into smaller pieces for a total weight of approximately 1 g. Tissue was cut with scissors and resuspended in a digestion cocktail of 0.25% collagenase A (SC136, Millipore Sigma), dispase II (17105041, Thermo Fisher; 15 U/ml) and DNase I (DN25, Sigma-Aldrich; 50 U/ml) in RPMI medium (R7388, Millipore Sigma). The tissue was digested at 37 °C for 1 h with intermittent resuspension. Cells were passed through a 100-µm cell strainer washed with PBS and more EDTA. Cells were then mounted with TrueBlack to reduce background and with Fluorosave to maintain fluorescence.

Lung digestion and fluorescence-activated cell sorting. For mouse, whole lung was dissected from adult animals and tracheally perfused with a digestion cocktail of collagenase type I (17100017, Theromo Fisher; 225 U/ml), dispase II (17105041, Theromo Fisher; 15 U/ml) and DNase I (DN25, Sigma-Aldrich; 50 U/ml) and removed from the chest. The lung was further digested with a razor blade and incubated in digestion cocktail for 45 min at 37 °C with continuous shaking. The mixture was then washed with sorting buffer (2% fetal bovine serum (FBS) and 1% penicillin–streptomycin in DMEM medium (31053028, Thermo Fisher)). The mixture was passed through a 70-µm cell strainer and resuspended in RBC lysis buffer (NC906751, Thermo Fisher), then passed through a 40-µm cell strainer. Cell suspensions were incubated with the appropriate conjugated antibodies in sorting buffer for 30 min at 4 °C and then washed with sorting buffer.

Fluorescence-activated cell sorting (FACS) was performed on a BD FACSAria II using FACSDiva software. Doublets and dead cells were excluded based on forward and side scatter. Cells and DRAQ7 (7406, Cell Signaling Technology) for DNA content, respectively. Immune and endothelial cells were excluded using CD45 (Alexa Fluor 700; 565010, BioLegend; 1:200; clone 30-F11) and CD31 (APC/Fire550; 102528, BioLegend; 1:200; clone ME13.3), respectively. For Gli1LacZ and Gli1creERT2/R26RsfGFP sorting, epithelial cells were also excluded using CD326 (B2421; 562314, BD; 1:200; clone G8.8), then Gli1 cells were sorted using endogenous GFP fluorescence. For organoid sorting, cells were sorted for CD326 and then CD104 (PE; 123610, BioLegend; 1:100; or Alexa Fluor 647; 123608, BioLegend; 1:75; clone 346–11A). Cells were sorted into sorting buffer. Analysis was performed using FlowJo v10 software.

For human samples, multiple pieces were collected from whole lungs and separated into smaller pieces for a total weight of approximately 1 g. Tissue was cut with scissors and resuspended in a digestion cocktail of 0.25% collagenase A (SC136, Millipore Sigma), dispase II (17105041, Theromo Fisher; 15 U/ml) and DNase I (DN25, Sigma-Aldrich; 50 U/ml) in RPMI medium (R7388, Millipore Sigma). The tissue was digested at 37 °C for 1 h with intermittent resuspension. Cells were passed through a 100-µm cell strainer washed with PBS and then removed from the chest. Cells were then mounted with TrueBlack to reduce background and with Fluorosave to maintain fluorescence.

Single-cell capture and sequencing. For mouse scRNA-seq, all live Lin+ cells were sorted from two Gli1creERT2/R26RsfGFP adult lungs. Lin+ cells were isolated based on forward and side scatter, DAPI exclusion and GFP fluorescence. One mouse was challenged with repetitive-dose bleomycin and one was treated with vehicle at the same time points as a control. The sort was performed on the Cyte 8. The single animal was cut with scissors and resuspended in a digestion cocktail of 0.25% collagenase A (SC136, Millipore Sigma), dispase II (17105041, Theromo Fisher; 15 U/ml) and DNase I (DN25, Sigma-Aldrich; 50 U/ml) in RPMI medium (R7388, Millipore Sigma). The tissue was digested at 37 °C for 1 h with intermittent resuspension. Cells were passed through a 100-µm cell strainer washed with PBS and more EDTA. Cells were then mounted with TrueBlack to reduce background and with Fluorosave to maintain fluorescence.

Organoid assay. Adult mouse lungs underwent FACs for ITGB4+ epithelial progenitor cells and Gli1+ or Gli1lin+ mesenchymal cells. The trachea was dissected and removed from the lung before lung digestion and FACs. ITGB4+ epithelial cells and Gli1+ mesenchymal cells were sorted (1.5 × 10⁶ epithelial cells x 10⁶ mesenchymal cells per well) in modified MTEC medium diluted 1:1 in growth-factor-reduced Matrigel (C-40230A, Theromo Fisher). Modified MTEC culture medium (CC-3118, Lonza) comprises small airway basal media (SABM) with selected components from a SABM bullet kit including insulin, transferrin,
bovine pituitary extract, retinoic acid and human epidermal growth factor. CTA (0.1 μg ml⁻¹; C8052, Sigma Aldrich), 5% FBS and 1% penicillin–streptomycin were also added to the medium. The cell suspension–Matrigel mixture was placed in a Transwell and incubated in growth medium with 10 μg ROCK inhibitor (72252, Stellenc Technologies) in a 24-well plate for 48 h, after which the medium was replenished every other day (lacking ROCK inhibitor). Each experimental condition was performed in triplicate. Where applicable, rhBMP4 (314-30-B, R&D Systems; 50 ng ml⁻¹) and GREM2 (2069-PR-050, R&D Systems; 1.5 μg ml⁻¹) were added to the medium after 48 h and replenished in every medium change. Colonies were assayed after 12–14 days.

To extract RNA from the organoid assays, cell suspension–Matrigel mixtures in the Transwells were washed with PBS and incubated in a digestion cocktail of collagenase type I (17100017, Thermo Fisher; 225 μl⁻¹), dispase II (17105041, Thermo Fisher; 15 μl ml⁻¹) and DNase I (DN25, Sigma-Aldrich; 50 μg ml⁻¹) for 1 h at 37 °C with intermittent resuspension. The mixture was removed from the Transwell and resuspended in Tryple (12563011, Thermo Fisher) and shaken at 37 °C for 20 min. Cells were resuspended in sorting buffer (2% FBS and 1% penicillin–streptomycin in DMEM) and blocked with rat serum (19860, Stellenc Technologies; 1:50) for 10 min at 4 °C. Cells were then stained with biotin anti-mouse CD326 (118294, BioLegend; 1:250; clone G8.8) for 30 min at 4 °C. Streptavidin beads (19860, Stellenc Technologies; 1:50) were added to isolate the epithelial cells.

qPCR with reverse transcription. Total RNA was obtained from cells isolated from organoid assays or cultured primary lung fibroblasts using a PicoPure RNA Isolation kit (KIT0204, Applied Biosystems) or a RNeasy kit (74004, Qiagen) following the manufacturers’ protocols. RNA from mouse lung tissue was obtained by removing the entire left lobe, homogenizing in TRIzol (15956018, Thermo Fisher), and extracting RNA using an E.Z.N.A. Total RNA kit (R683-01, Omega Bio-Tek) following the manufacturer’s instructions. cDNA was synthesized from total RNA using the SuperScript Strand Synthesis System (18080544, 100060840, Thermo Fisher), qPCR was performed using the SYBR Green system (F451L, Thermo Fisher). Primers are listed in Supplementary Table 1. Relative gene expression levels after qPCR with reverse transcription were defined using the ΔΔCT method and normalizing to GAPDH.

RNAseq in situ hybridization. Paraffin-embedded lung sections were processed for RNA in situ detection of Fbn2 (447931, Advanced Cell Diagnostics), Npnt (316771, Advanced Cell Diagnostics) and Sih (600951, Advanced Cell Diagnostics) using a RNAseq Multiplex Fluorescent Reagent kit v.2 (323100, Advanced Cell Diagnostics) according to the manufacturer’s instructions.

Pulse oximetry. The MouseOx Pulse Oximeter system (Starr Life Sciences) was used to measure arterial oxygen saturation from awake mice on the day of euthanasia. Mice were shaved before measurement. Mice were measured at five measurements per second for at least 5 min, and at least ten successful readings were obtained. All successful measurements (error code = 0) were averaged for each mouse.

Immunofluorescence image quantification. Sections were imaged for quantification on a Zeiss Lumar V12 microscope. At least three samples per genotype per condition were imaged, and at least three randomly selected regions were analysed for each sample. Cell counts for SFTPC cells were performed on Fiji using the Cell Counter plug-in. Five high-magnification images of randomly selected alveolar regions were counted per sample. SFTPC quantification is presented as a percentage of total organoids analysed. For all analyses, the performer was blind to the specimen genotype per condition during data collection and analyses. Results were averaged between each specimen, and standard deviations were calculated per genotype per condition.

scRNA-seq analysis. To build transcript profiles of individual cells, the CellRanger v.2.1.1 software with default settings was used for de-multplexing, aligning reads with STAR software to GRCh38 for the mouse genome and Hg19 for the human genome, and counting unique molecular identifiers. We used the Seurat R package along with a gene–barcode matrix provided by CellRanger for downstream analysis36. In total, we filtered the data in two different steps. We first filtered the dataset by only accepting cells that expressed a minimum of 200 genes and genes that were expressed in at least 3 cells. Our second filter was set to accept cells with less than 6,000 unique gene counts and 5% mitochondrial counts. After removing unwanted cells, we used LogNormalize to normalize the gene expression measurements for each cell. We calculated 2,000 features that exhibited high cell-to-cell variation, which were used in principle component analysis after scaling the data. We used the JackStrawPlot function in the Seurat package to create Scree plots and compare P values (significance) for each principle component. We selected ten different principle components for clustering of both mouse and human cells. Clustering results were visualized using the UMAP algorithm in the Seurat package. For mouse datasets, we further integrated the PDGFRA⁺ clusters from both PBS and bleomycin samples to compare the datasets37. Differentially expressed genes between PBS and bleomycin datasets were identified using a MAST test in the DESeq package in R37. RNA velocity was calculated on the integrated object using the Velocytto package and RNA velocity in the SeuratWrappers package in R8. For human scRNA-seq analysis, we similarly integrated the PDGFRA⁺ clusters from both three healthy controls and three patients with IPF, and the differentially expressed genes were identified with a MAST test. For cluster visualization and individual gene visualization on all clusters we used the UMAP algorithm. GO enrichment analysis was performed using the PANTHER Overrepresentation test and entering the top differentially expressed genes with adjusted P (P adj.) values of <0.1 and fold-change (log(f.c.)) values of >0.15 or <0.15.

Statistics and reproducibility. All statistical analyses were performed in GraphPad Prism 6.0 using t-test (unpaired, one-tailed) T-tests to determine the P value. Specific P values are labelled in the plots or figure legends, where significant values are P<0.05. Ordinary one-way analysis of variance (ANOVA) was used for multiple comparison to a control group. All data in graphs are presented as the mean ± s.d. No data were excluded from the analyses presented in this study. All animal studies were independently performed twice with successful replication, except for the scRNA-seq. All in vitro experiments were performed at least three times with successful replication. Three or more biological replicates were used for every study and specific n values are listed in the figure legends.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The mouse RNA-seq data reported in this paper (Fig. 1) are deposited in the NCBI Gene Expression Omnibus (GEO) under the accession number GSE140032. The human RNA-seq data reported in this paper (Fig. 7) are deposited in the NCBI GEO under the accession number GSE132771. Previously published sequencing data from the Tabula Muris cellular database were referenced here and are publicly available at https://tabula-muris.ds.cribiohubb.org/. There are no restrictions on data availability. All other data supporting the findings of this study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability
This study does not include any custom code.

References
36. Butler, A. et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).
37. Stuart, T. et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902 (2019).
38. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 16, 278 (2015).

Acknowledgements
We thank K. N. Cheong, A. Bruinwell, A. Baldwin and N. Reyes de Mochel for providing technical assistance; C. Gralapp for assistance with model illustration; M. Looney for critical review of manuscript; the Paranasus Flow Cytometry Core for assistance with cell sorting for bulk and single-cell RNA analysis (P30DK063720); members of the Biological Imaging and Development Core (P30DK063720); and E. Wu and the Institute for Human Genetics Core for the processing of single-cell RNA samples and high-throughput sequencing. This work is supported by NIH grants DP2AG056034, K08HL121146 and R01HL142552 to T.P., along with a Tobacco Related Disease Research Program New Investigator Award and a Pulmonary Hypertension Association award to T.P. and a Nina Ireland Program Award to M.M. for human lung collection.
Author contributions
M.C. and T.P. conceived the experiments. M.C., C.W., J.K., T.T., P.M., M.M. and P.W. performed the experiments and data analyses. D.S., A.M. and H.C. provided expertise and feedback. M.C. and T.P. wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41556-020-00591-9.
Supplementary information is available for this paper at https://doi.org/10.1038/s41556-020-00591-9.
Correspondence and requests for materials should be addressed to T.P.
Reprints and permissions information is available at www.nature.com/reprints.
**Extended Data Fig. 1** | Epithelial progenitors expand distally in fibrotic repair. (**a, b**) Time course showing appearance of metaplastic airway epithelium in the alveoli after 1, 2, 3 or 4 doses of weekly bleomycin injury. (n = 3 per group; each data point represents one animal). Data are expressed as mean ± SD. (**c, d**) Contribution of Sox2 Lin+ cells to the various epithelial lineages in the alveoli after bleomycin injury. (n = 6 per group; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (**e**) Lack of Sox2 Lin+ cells in corn oil-treated SoxcreERT2;R26mTmG lung. (2 animals). (**f**) Co-localization of Sox2 Lin+ cells and SOX2 staining. (n = 3 animals). Data are expressed as mean ± SD. (**g**) Extended washout after tamoxifen (>4wks) followed by bleomycin demonstrates similar expansion of Sox2 Lin+ cells into the alveoli (3 animals). AW = airway. Scale bars, 100 μm.
Extended Data Fig. 2 | Lineage tracing of diverse airway progenitors after bleomycin injury. (a, b) p63 Lin+ cells also expanded into the alveoli after bleomycin but did not contribute significantly to KRT5+ basal cells in the alveoli. (n = 3 animals). Data are expressed as mean ± SD. (c, d) Comparison of p63 Lin+ and Sox2 Lin+ cells that expand into the alveoli after bleomycin injury. (n = 3 for p63 Lin, n = 6 for Sox2 Lin; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (e, f) Contribution of Scgb1a1 Lin+ cells to the various epithelial lineages in the alveoli after bleomycin injury. (n = 3 per group; each data point represents one animal). Data are expressed as mean ± SD. AW = airway. Scale bars, 100 μm.
Extended Data Fig. 3 | Gli1 Lin+ mesenchyme expands distally in fibrotic repair. (a) Lack of Gli1 Lin+ cells in corn oil-treated Gli1<sup>creERT2;R26<sup>YFP</sup></sup> lung (2 animals). (b, c) Comparison of Gli1 Lin+ cell localization with Gli1+ cells in Gli1<sup>LacZ/+</sup> lung. These experiments were each repeated three times independently with similar results. (d, e) Extended washout after tamoxifen (>4wks) followed by bleomycin demonstrates similar expansion of Gli1 Lin+ cells into the alveoli. (n = 3 per group; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (f, g) Deletion of Gli1+ cells does not alter airway epithelial composition during normal homeostasis. (n = 3 per group; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (h–j) Gli1-deleted lungs treated with bleomycin demonstrate reduced expression of Gli1 along with changes in fibrotic markers (n = 6 for control, n = 5 for Gli1- deleted; each data point represents one animal; for (i), Col1α1 P = 0.0862; Snai1 P = 0.3126; Cdh2 P = 0.0310; one-tailed unpaired Student’s t-test for (h–j)). Data are expressed as mean ± SD. (k) Weight loss after bleomycin injury in control and Gli1- deleted animals (n = 6 for control, n = 5 for Gli1-deleted; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. AW = airway. Scale bars, 100 μm.
Extended Data Fig. 4 | See next page for caption.
Extended Data Fig. 4 | ITGB4 + airway progenitors give rise to KRT5 + and SFTPC + cells that are not from pre-existing Krt5 + and Sftpc + cells.

(a) Gating strategy for isolation of ITGB4 + airway progenitors and Gli1 + mesenchymal cells by FACS, presented in Fig. 2a, e; 3f; 5e; 6f and Extended Data Fig. 4d, h; 5h; 8g. (b) Cytospin of freshly sorted ITGB4 + confirms that majority of cells are SOX2 +, SCGB1A1 +, and TubIV +, with cells rarely positive for SFTPC, KRT5, p63, or NGFR (n = 1 for NGFR, n = 3 for rest; each data point represents one biological replicate). Data are expressed as mean ± SD. (b) Z-axis scan shows that Gli1 Lin + mesenchyme settle to bottom of the Matrigel well. (d–g) Lineage trace of pre-existing Krt5 Lin + (tdTomato +) cells followed by isolation of lung for ITGB4 + airway progenitor organoid culture with Gli1 + cells. Tracheal Krt5 Lin + cells demonstrate expression of basal cell markers KRT5 and p63 (e), while few Krt5 Lin + cells appear in ITGB4 + cell-derived organoid isolated from the lung of the same animal (f, g). The overwhelming majority of KRT5 + cells in the ITGB4-derived organoids are Krt5 Lin-negative. (n = 3 wells for (f, g); n = 1 sample for (e)). Data are expressed as mean ± SD. (h–j) Lineage trace of pre-existing Sftpc Lin + (membrane GFP +) cells followed by isolation of the lung epithelium for organoid culture with Gli1 + cells. ITGB4 + (red box) or Sftpc Lin + (green box) cells were sorted from the same lung. The overwhelming majority of SFTPC + organoids in the ITGB4 - derived organoids are Sftpc Lin-negative, while most SFTPC + cells in the Sftpc Lin + derived organoids are Lin + (n = 3; each data point represents one well). Data are expressed as mean ± SD. Scale bars, 100 μm.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Characterization of Hh- activation during homeostasis and in fibrotic repair. (a) RNAscope in situ of Shh shows expression in ectopic SOX2+ cells in the alveoli after bleomycin injury. This experiment was repeated independently twice with similar results. (b, c) Histology quantification shows no evidence of KRT5+ metaplasia in Hh-activated animals without injury when compared to controls (n = 3 per group; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (d-g) After bleomycin injury, Hh-activated lungs show similar expansion of Gli1+ cells in distal alveoli compared to controls, a trend towards increased myofibroblasts (SMA+) differentiation of the Gli1 Lin+ cells in the alveoli, and no difference in ectopic SCGB1A1+ cells in the alveoli (for e, n = 4 for control, n = 6 for Hh-activated; for f, g, n = 9 per group; each data point represents one animal; one-tailed unpaired Student’s t-test for e-g). Data are expressed as mean ± SD. (h) Model of 3D airway organoid assay using Hh-inducible mesenchyme whereby pretreatment of 4-hydroxytamoxifen (4OHT) induces Hh activation in the mesenchyme as shown by upregulation of Gli1 transcript compared to vehicle (ethanol) (n = 3 per group; each data point represents one biological replicate; one-tailed unpaired Student’s t-test). (i-l) Mesenchymal Hh activation in vitro reduces CFE and organoid size, but increases the expression of Krt5 and number of KRT5+ organoids while reducing Sftpc expression and number of SFTPC+ organoids (i-k n = 3 per group; l n = 3 for Hh-activated, n = 5 for control; each data point represents one well; one-tailed unpaired Student’s t-test for j-l). Data are expressed as mean ± SD. Hh = hedgehog, 4OHT = 4-hydroxytamoxifen, AW = airway. Scale bars, 100 μm.
Extended Data Fig. 6 | Expansion of Gli1 Lin+ mesenchyme into the alveoli. (a) Multi-colour stochastic recombination of Gli1 Lin+ cells tagged with one of four possible fluorescent proteins (GFP, YFP, RFP, or CFP) shows clones of single color clustering around the airway in PBS treated lungs, and clones of single color spanning the airway and alveoli following bleomycin injury (4 animals per group) (b) Violin plots showing retention of Gli1 expression in Gli1 Lin+ cells following bleomycin injury as they migrate into the alveoli. Each black dot represents one cell (n values listed in source data for Fig. 4d). AW = airway. Scale bars, 100 μm.
Extended Data Fig. 7 | Spatial regulation of BMP activation. (a, b) pSMAD1/5/8 staining of Sox2 Lin+ cells in the alveoli after bleomycin injury demonstrates that the majority of pSMAD1/5/8+/Sox2 Lin+ cells are SCGB1A1+ or SFTPC+, and rarely KRT5+ (n = 3 for KRT5 quantification, n = 4 for SCGB1A1 and SFTPC quantification; each data point represents one animal). Data are expressed as mean ± SD. (c) qPCR of Hh-inducible cultured mesenchyme isolated from UbcCreERT2+/R26SmoM2/ mice lungs treated with 4OHT+/−TGF-β, showing effect on various BMP antagonists expression (n = 3; each data point represents one biological replicate). Wnt5a: Control v TGF-β P = 0.0094, Control v Hh-activated P = 0.1066, Control v TGF-β+Hh-activated P = 0.002; Fst: Control v TGF-β P = 0.0598, Control v Hh-activated P = 0.0389, Control v TGF-β+Hh-activated P = 0.3833; Fstl1: Control v TGF-β P = 0.2051, Control v Hh-activated P = 0.4620, Control v TGF-β+Hh-activated P = 0.028; Fstl3: Control v TGF-β P = 0.0005, Control v Hh-activated P = 0.0704, Control v TGF-β+Hh-activated P = 0.0003; Grem1: Control v TGF-β P = 0.4845, Control v Hh-activated P = 0.2422, Control v TGF-β+Hh-activated P = 0.1417; Fzd1: Control v TGF-β P = 0.3341, Control v Hh-activated P = 0.2689, Control v TGF-β+Hh-activated P = 0.3320; Sfrp1: Control v TGF-β P = 0.4103, Control v Hh-activated P = 0.4614, Control v TGF-β+Hh-activated P = 0.1093; ordinary one-way ANOVA test. Data are expressed as mean ± SD. Scale bars, 100 µm.
Extended Data Fig. 8 | Effect of rhBMP4 in vivo and in vitro. (a–d) Effect of rhBMP4 treatment on fibrotic marker expression, myofibroblasts in the alveoli, weight change, and presence of ectopic SCGB1A1+ cells, in bleomycin injured C57BL/6 animals (n = 8 per group; each data point represents one animal). For (a), Col1a1 P = 0.0015; Snai1 P = 0.0769; Cdh2 P = 0.1771; Acta2 P = 0.0934; Tagln P = 0.3782; one-tailed unpaired Student’s t-test for (a–d). Data are expressed as mean ± SD. (e) Effect of rhBMP4 on the expression of markers of chondrogenesis, adipogenesis, and osteogenesis in the lung (n = 8 per group; each data point represents one animal). Acan P = 0.0021; Col2a1 P = 0.2363; Col10a1 P = 0.4201; Plin2 P = 0.0578; Pparg P = 0.4343; Fabp4 P = 0.0006; Spp1 P = 0.3117; Alpl P = 0.0474; Ibsp P = 0.3371; Runx2 P < 0.0001; one-tailed unpaired Student’s t-test. Data are expressed as mean ± SD. (f) Effect of rhBMP4 on the total number of Sox2 Lin+ cells in the alveoli after bleomycin injury (n = 4 per group; each data point represents one animal; one-tailed unpaired Student’s t-test). Data are expressed as mean ± SD. (g–i) Airway progenitor organoid co-cultured with Gli1+ mesenchyme treated with BMP4 demonstrates increased CFE and enhanced Sftpc expression with reduced Krt5 (n = 4 for solvent, n = 3 for BMP4; each data point represents one well; one-tailed unpaired Student’s t-test for (h, i)). Conversely, the BMP antagonist GREM2 acts to suppress CFE and Sftpc expression while enhancing Krt5 (n = 4 per group; each datapoint represents one well; one-tailed unpaired Student’s t-test for (hi)). Data are expressed as mean ± SD.
Extended Data Fig. 9 | Metaplastic KRT5 + basal cells and SMA + fibroblastic foci in the alveoli of IPF lungs. (a) UMAP projection of mesenchyme from each individual donor that was used in single cell RNA-seq analysis with contributing cell number. (b) Expression of basal cell markers NGFR, p63, and KRT14, in KRT5 + basal cells in human IPF, along with mouse bleomycin injured lungs and murine airway progenitor-derived organoids. This experiment was repeated three times independently with similar results. (c) Ectopic expression of KRT5 and SMA in the alveoli of IPF lungs form honeycomb cysts and fibroblast foci, respectively. This experiment was repeated independently twice with similar results. (d) Average cell-to-cell distance shows equivalent proximity of metaplastic KRT5 and endogenous SFTPC cells to SMA + fibroblastic foci (n = 3,280 KRT5 cells, n = 2,215 SFTPC cells; each datapoint represents an individual distance measurement with all measurements plotted for 3 samples; one-tailed unpaired Student’s t-test). (e, f) Unsorted epithelial organoids derived from dissociated IPF lungs grew KRT5 + organoids, and BMP4 attenuated the number and size of IPF-derived organoids. (n = 5; each data point represents one well; one-tailed unpaired Student’s t-test) Data are expressed as mean ± SD. Scale bars, 100 μm.
Extended Data Fig. 10 | Gli1 + mesenchyme modulates SOX2 + progenitor fate. Model of Gli1 + mesenchyme integrating Hh activation as a rheostat that controls local BMP activation to determine SOX2 + airway epithelial progenitor fate.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed
---|---
☐ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐ | The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
☐ | A description of all covariates tested
☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐ | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.
☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐ | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about: availability of computer code

Data collection

| Data analysis |
|---------------|
| STAR v2.1.3; CellRanger v2.1.1; MouseOx Pulse Oximeter v2; BD FACSDiva v8.0.1; Zeiss Lunar V12; Invitrogen EVOS MS5000 |

Data analysis

For single cell RNA-seq data analysis the CellRanger v2.1.1 software with default settings was used for de-multiplexing, aligning reads with STAR software v2.1.3 to mouse genome GRCm38, and counting unique molecular identifiers (UMIs). Downstream data analysis was conducted with Seurat v3.0 in R. RNA velocity was calculated on integrated objects using Velocyto package and RNA velocity v1.1 in SeuratWrappers package in R. Differential gene expression was conducted using DESeq in R. Gene ontology enrichment analysis was performed using the PANTHER Overrepresentation test on the top differentially expressed genes. For mouse arterial oxygen saturation analysis, the MouseOx Pulse Oximeter system (v2) was used for data collection and analysis. Image analysis and quantification was performed using Fiji 1.0 and Imaris 8.3.1 (which includes Big Sortomato v1 tool). Howo v1.0 was used for flow cytometry analysis. GraphPad Prism 6.0 was used for all statistical testing.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about: availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The mouse sequencing data reported in this paper (Figure 4) is deposited in NCBI Gene Expression Omnibus (GEO) under the accession number: GSE140032. The human sequencing data reported in this paper (Figure 7) is deposited in NCBI Gene Expression Omnibus (GEO) under the accession number: GSE132771. Gene
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Sample size**
No statistical test was performed to determine sample size. All experiments used a minimum of 3 independent samples per group in order to perform necessary statistical tests. For animal studies, sample size was determined based on the number of biological replicates available, which were strictly controlled for sex and age between experimental groups to minimize confounding factors in our analysis. Sample sizes were based on our prior experience and previous published studies using lineage tracing and lung injury models, including Ray et al., Stem Cell Reports, 2016; Vaughan et al., Nature, 2015; Kramann et al., Cell Stem Cell, 2015; Wang et al, JCI, 2018 [all cited in text]. For non-animal studies, sample size was determined following standard practices in similar fields of study, our prior experience, and ability to perform statistical tests (i.e a minimum of 3 independent samples per group).

**Data exclusions**
No data was excluded from the analysis presented in this paper.

**Replication**
All animal studies were independently performed twice with successful replication, with the exception of single cell RNA-sequencing which was performed once. All non-animal experiments were performed at least three times with successful replication. Three or more biological replicates were used for every study.

**Randomization**
For animal studies, age/sex/litter-matched mice were randomized into groups based on the presence of specific genetic alleles. For studies using animals of identical genotypes, age/sex/litter-matched mice were randomly assigned to experimental groups using publicly available online "list randomizer" tool. Animals of both experimental groups were housed in the same cages when possible to minimize confounding factors. All non-animal experiments were allocated into experimental groups randomly using similar randomizer tool.

**Blinding**
For RNA and histologic quantification, performer was blinded to specimen/genotype during data collection and analysis. All other forms of data collection and analysis were performed in an automated/operator-independent manner and therefore blinding was not necessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

**Materials & experimental systems**

- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants
- Clinical data

**Methods**

- Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

The following primary antibodies were used for immunofluorescence on mouse tissue:

- Chicken anti-GFP (Cat#K03-1020; Aves Labs; used 1:250; Lot#GFP8/798A)
- Goat anti-GFP (Cat#ab6673, Abcam; used 1:250; Lot#GR3272452-2)
- Rabbit anti-SMA (Cat#ab5694, Abcam; used 1:200; Lot#GR3183259-34)
- Rabbit anti-DES (Cat#ab15200, Abcam; used 1:200; Lot#GR303635-1)
- Goat anti-SCGB1a1 [Cat#AB-7972; Santa Cruz; used 1:500; Lot#A2616]
- Mouse anti-SCGB1a1 (Cat#sc-30313; Santa Cruz; used 1:200)
- Rabbit anti-SFTPC (Cat#ab3786; Millipore Sigma; used 1:2000)
- Goat anti-SFTPC (Cat#AB-706; Santa Cruz; used at 1:200)
- Rat anti-RAGE (Cat#ab81179; R&D Systems; used at 1:200)
- Chicken anti-KRT5 (Cat#R05901; Biolegend; used at 1:500; Lot#B271562)
- Goat anti-SOX2 (Cat#G115098; Neurosciences; used at 1:200; Lot#02343)
Chicken anti-RFP (Cat#600-901-379); Rockland; used: 1:200; Lot#42717
Rabbit anti-RFP (Cat#600-401-379); Rockland; used: 1:100
Mouse anti-TubV (Cat#111; Abcam; used: 1:200)
Rabbit anti-p-e36 (Cat#13109T; Cell Signaling Technology; used: 1:200)
Rabbit anti-NGFR (Cat#8875; Abcam; used: 1:200)
Mouse anti-KRT14 (Cat#C81; Sigma; used: 1:100)
Rabbit anti-p-smad1/5/8 (Cat#AB8848; Millipore; Sigma; used: 1:100; Lot#3258711)

The following primary antibodies were used for human tissue:
Rabbit anti-SMA (Cat#B56594; Abcam; used: 1:200; Lot#G3183259-34)
Rabbit anti-p-smad1/5/8 (Cat#AB8848-I; Millipore; Sigma; used: 1:100; Lot#3258711)
Rabbit anti-p-e36 (Cat#13109T; Cell Signaling Technology; used: 1:200)
Rabbit anti-NGFR (Cat#8875; Abcam; used: 1:200)
Mouse anti-KRT14 (Cat#C81; Sigma; used: 1:100)
Chicken anti-KRT5 (Cat#RS5901; Biolegend; used: 1:500; Lot#B271562)
Goat anti-SOX2 (Cat#ST5098; Neuronics; used: 1:200; Lot#02343)
Goat anti-SFTPC (Cat#sc-7705; Santa Cruz; used at 1:200).

The following secondary antibodies were used at 1:250 for both mouse and human tissue:
Donkey anti-Chicken IgG (H+L) Alexa Flour 488 (Cat#AB7357; Thermo Fisher; Lot#20678)
Donkey anti-Goat Alexa Flour 647 (Cat#A-2447; Thermo Fisher; Lot#203532)
Donkey anti-Mouse IgG Alexa Flour 555 (Cat# A-31570; Thermo Fisher)
Donkey anti-Rat Dylight 488 (Cat#AAS-10026; Thermo Fisher).

For flow cytometry/cell separation on mouse cells, the following antibodies were used:
Rat anti-Mouse CD45-Alexa Flour 700 (BD; Cat#65010; Clone 30-F11; used: 1:200; Lot#338884)
Rat anti-Mouse CD31-APC-Fire 780 (BioLegend; Cat#102528; Clone MECL-3; used: 1:200; Lot#8259777)
Rat anti-Mouse CD26- BV421 (BD; Cat#563214; Clone G8.8; used: 1:200; Lot#8228560)
Rat anti-Mouse CD104-AF647 (Cat#136068; BioLegend; Clone 346-11A; used: 1:75)
Rat anti-Mouse CD104-PE (BD; Cat#12610; Biolegend; Clone 346-11A; used: 1:100; Lot#8268478)
Biot anti-Mouse CD326 (Cat#119204; BioLegend; used: 1:250; Clone G8.8).

For flow cytometry on human cells, the following antibodies were used:
Mouse anti-Human CD31: PE/Cy7 (Cat#301118; BioLegend; clone WM59; used: 1:200)
Mouse anti-Human CD45: APC (Cat#304037; BioLegend; clone HI30; used: 1:200)
Mouse anti-Human CD26: FITC (Cat#34204; BioLegend; clone 9C4; used: 1:200)
Mouse anti-Human CD235a: APC-Cy7 (Cat#349116; BioLegend; clone HI264; used: 1:200).

Validation statements are available from manufacturer's websites:
Chicken anti-GFP: https://www.aveslabs.com/products/green-fluorescent-protein-gfp-antibody/varient=25144111202404
Goat anti-GFP: https://www.abcam.com/gfp-antibody-ab6653.html
Rabbit anti-SMA: https://www.abcam.com/alpha-smooth-muscle-actin-antibody-ab5694.html
Rabbit anti-DES: https://www.abcam.com/desmin-antibody-cytoskeleton-marker-ab15200.html
Goat anti-SCGB1a1: https://www.scbt.com/p/cc10-antibody-t-18
Mouse anti-SCGB1a1: https://www.scbt.com/p/foi10-antibody-b-6
Goat anti-SFTPC: https://www.emdmillipore.com/EN/en/product/anti-prosurfactant-protein-c-prosp-c-antibody-MM-NF-AB3786
Goat anti-SFTPC: https://www.scbt.com/p/sp-c-antibody-m-20
Rat anti-RAGE: https://www.ndsystems.com/products/mouse-rat-rage-antibody-175410-mab1179
Chicken anti-KRT5: https://www.biolegend.com/en-us/products/keratin-5-s-polyclonal-chicken-antibody-purified-10957
Goat anti-SOX2: https://www.neuromics.com/pt-15098
Mouse anti-RFP: https://rockland-inc.com/Product.aspx?pid=35267
Rabbit anti-RFP: https://rockland-inc.com/store/Antibodies-to-GFP-and-Antibodies-to-RFP-600-401-379-04L_2429.napx
Mouse anti-Tub: https://www.abcam.com/pa-lv-tubulin-antibody-ons16-ab-13315.html
Rabbit anti-p63: https://www.cellsigna.com/products/primary-antibodies/p63-a-d2k8x-1-xp-rabbit-mab-13109
Rabbit anti-NGFR: https://www.abcam.com/pj7-n-5-gfr-receptor-antibody-ab8875.html
Mouse anti-KRT14: https://www.emdmillipore.com/EN/en/product/Anti-Cytokeratin-14-Antibody-clione-LU002-MM-NF-CBL197?ReferrerURL=https%3A%2F%2Fwww.google.com%2F
Rabbit anti-p-smad1/5/8: https://www.emdmillipore.com/EN/en/product/Anti-phospho-Smad1-Smad5-Smad8-Ser463-465-MM-NF-AB8848-I?ReferrerURL=https%3A%2F%2Fwww.google.com%2F
Goat anti-SFTPC: https://www.citeab.com/antibodies/827842-sc-7705-sp-c-antibody-c-19
Donkey anti-Chicken IgG (H+L) Alexa Flour 488: https://www.jacksonimmuno.com/catalog/products/703-545-155
Donkey anti-Rabbit IgG Alexa Flour 555: https://www.thermoscientific.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly
Cross-Adsorbed-Secondary-Antibody-Polygonal/A-21447
Donkey anti-Mouse IgG Alexa Flour 555: https://www.thermoscientific.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly
Cross-Adsorbed-Secondary-Antibody-Polygonal/A-31570
Donkey anti-Rat Dylight 488: https://www.thermoscientific.com/antibody/product/Donkey-anti-Rat-IgG-H-L-Highly
Cross-Adsorbed-Secondary-Antibody-Polygonal/A-30106
Rat anti-Mouse CD45-Alexa Flour 700: https://wwwbdbiosciences.com/US/applications/research/stem-cell-research/cancer-research/mouse/alexa-fluor-700-rat-anti-mouse-cd45-30-fl1j/p/566530
Rat anti-Mouse CD31-APC-Fire 750: https://www.biolegend.com/fr-ch/products/apc-fire-750-anti-mouse-cd31-antibody-15033
Rat anti-Mouse CD326-BV421: https://wwwbdbiosciences.com/US/applications/research/stem-cell-research/cancer-research/
Animals and other organisms

Policy information about studies involving animals. ARRIVE guidelines recommended for reporting animal research

Laboratory animals
Female and male adult transgenic mice in mixed or C57BL/6 background were used at 5-12 weeks of age. Mice were gender balanced between experimental groups. The laboratory mouse strains used in this study are: C57BL/6 (Charles River Laboratories Strain; No: 027), Gli1-EGFP (GENSAT; Stock No: DM1971), Gli1-creERT2 (The Jackson Laboratory; Stock No: 007913), Ubc-creERT2 (The Jackson Laboratory; Stock No: 007001), Sox2-creERT2 (The Jackson Laboratory, Stock No: 017593), R26-SmoM2 (The Jackson Laboratory, Stock No: 005130), Smo-flox (The Jackson Laboratory, Stock No: 004576), R26-EYFP (The Jackson Laboratory, Stock No: 006148), R26-tdTomato (The Jackson Laboratory, Stock No: 007950), R26-mTmG (The Jackson Laboratory, Stock No: 007576), R26-OVA (The Jackson Laboratory; Stock No: 008669), Scgb1a1-creERT2 (The Jackson Laboratory; Stock No: 016225), R26-Confetti (The Jackson Laboratory; Stock No: 017492), Gli1-LacZ (The Jackson Laboratory; Stock No: 008211), p63-creERT2 (The Jackson Laboratory; Stock No: 024564), KRT5-creERT2 (The Jackson Laboratory; Stock No: 029155), SFTPC-creERT2 (5fptcm1.1creERT2,rtTA)Hap. Both female and male mice were used for all strains. Mice were housed under pathogen-free conditions at room temperature, maintained under standard 12 hour light/dark cycle, and monitored daily by caretakers or researchers. Cage, bedding, and food was changed biweekly.

Wild animals
No wild animals were involved in this study.

Field-collected samples
No field-collected samples were involved in this study.

Ethics oversight
All animals were housed and treated in accordance with the Institutional Animal Care and Use Committee (IACUC) protocol approved at the University of California, San Francisco.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics
IPF lungs were collected from patients with established diagnoses of IPF undergoing lung transplantation, and the control healthy lungs were from sex and age matched cadaveric donors without prior history of lung disease obtained from brain-dead donors that were rejected for lung transplantation. Three healthy and three IPF lungs were used for RNA-seq analysis. For histologic analysis, three independent IPF samples were used. Samples were collected from all ethnic groups, both male and female patients, and all samples were from patients over 50 years old. Further patient information (such as genotypic information) was not available to researchers.

Recruitment
No patient recruitment was performed for this study; all samples were collected based on patient’s health condition. Healthy human lungs were provided by the Northern California Donor Network from deceased donors in which the lung was independently rejected for transplantation. IPF lungs were obtained from end-stage patients receiving lung transplantation. Investigators were blind to patient information besides relevant history of lung disease.

Ethics oversight
Studies involving human tissue were approved by the UCSF Institutional Review Board. Informed consent was received from each donor.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:
- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.
Methodology

Sample preparation
For mouse, whole lung was dissected from adult animals and tracheally perfused with a digestion cocktail of Collagenase Type I (Cat#17100017; Thermo Fisher; used 225 U/mL), Dispase II (Cat#17105041; Thermo Fisher; used 15 U/mL) and DNase I (Cat#DN25; Sigma-Aldrich; used 50 U/mL) and removed from the chest. The lung was further diced with a razor blade and incubated in digestion cocktail for 45 mins at 37 °C with continuous shaking. The mixture was then washed with sorting buffer (2% FBS and 1% Penicillin-Streptomycin in DMEM [Cat#31053028; Thermo Fisher]). The mixture was passed through a 70 μm cell strainer and resuspended in RBC lysis buffer (Cat#NC9067514; Thermo Fisher), then passed through a 40 μm cell strainer. Cell suspensions were incubated with the appropriate conjugated antibodies in sorting buffer for 30 min at 4 °C and washed with sorting buffer. For human, multiple pieces were collected from whole lungs and separated into smaller pieces for a total of approximately 1 gram. Tissue was cut with scissors and resuspended in a digestion cocktail of 0.25% Collagenase A (Cat#SCR136; Millipore Sigma), Dispase II (Cat#17105041; Thermo Fisher; used 15 U/mL) and DNase I (Cat#DN25; Sigma-Aldrich; used 50 U/mL) in RPMI (Cat#R7888; Millipore Sigma). The tissue was digested at 37 °C for 1 hour with intermittent resuspension. Cells were passed through a 100 μm cell strainer, washed with PBS, and resuspended in 0.5% BSA in PBS.

Instrument
All data was collected on a BD FACSaria II.

Software
Data was collected using BD FACSdiva software. Data was analyzed using FlowJo v10.

Cell population abundance
The sorted ITGB4+ airway progenitor population was approximately 1.5% of total cells and 18-22% of all EpCAM+ cells. Gli1 mesenchymal cells were approximately 1.5-2.5% of total cells and 18-23% of all mesenchymal cells. Small portions of sorted cells were run through the machine to confirm purity of over 85%. Immunofluorescent staining of sorted cells were also performed to confirm expression of expected markers.

Gating strategy
Cells were selected using FSC/SSC scatter profile to remove debris and doublets were excluded using FSC A/FSC H to select for single cells. Dead cells were removed using DRAQ7 or DAPI exclusion. ITGB4 airway progenitors were isolated from wild type or Gli1-EGFP reporter mice using the gating strategy: CD45-, CD31-, CD326+, CD104+. For mouse Gli1 cell sorting, the following gating strategy was used using Gli1-EGFP reporter mouse: CD45-, C031, CD326, GFP+. These gating strategies are included in extended data figure 4a.

☑️ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.