Single-cell RNA sequencing identifies TGF-β as a key regenerative cue following LPS-induced lung injury

Kent A. Riemondy,[1] Nicole L. Jansing,[2] Peng Jiang,[3] Elizabeth F. Redente,[4,5] Austin E. Gillen,[1] Rui Fu,[1] Alyssa J. Miller,[6] Jason R. Spence,[6,7,8] Anthony N. Gerber,[2,5] Jay R. Hesselberth,[1] and Rachel L. Zemans[2,3,5,6]

[1]RNA Bioscience Initiative, University of Colorado School of Medicine, Aurora, Colorado, USA. [2]Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, National Jewish Health, Denver, Colorado, USA. [3]Division of Pulmonary Sciences and Critical Care Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA. [4]Program in Cell Biology, Department of Pediatrics, National Jewish Health, Denver, Colorado, USA. [5]Division of Pulmonary Sciences and Critical Care Medicine, University of Colorado Denver, Aurora, Colorado, USA. [6]Program in Cellular and Molecular Biology, University of Michigan, Ann Arbor, Michigan, USA. [7]Division of Gastroenterology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA. [8]Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, Michigan, USA.

Many lung diseases result from a failure of efficient regeneration of damaged alveolar epithelial cells (AECs) after lung injury. During regeneration, AEC2s proliferate to replace lost cells, after which proliferation halts and some AEC2s transdifferentiate into AEC1s to restore normal alveolar structure and function. Although the mechanisms underlying AEC2 proliferation have been studied, the mechanisms responsible for halting proliferation and inducing transdifferentiation are poorly understood. To identify candidate signaling pathways responsible for halting proliferation and inducing transdifferentiation, we performed single-cell RNA sequencing on AEC2s during regeneration in a murine model of lung injury induced by intratracheal LPS. Unsupervised clustering revealed distinct subpopulations of regenerating AEC2s: proliferating, cell cycle arrest, and transdifferentiating. Gene expression analysis of these transitional subpopulations revealed that TGF-β signaling was highly upregulated in the cell cycle arrest subpopulation and relatively downregulated in transdifferentiating cells. In cultured AEC2s, TGF-β was necessary for cell cycle arrest but impeded transdifferentiation. We conclude that during regeneration after LPS-induced lung injury, TGF-β is a critical signal halting AEC2 proliferation but must be inactivated to allow transdifferentiation. This study provides insight into the molecular mechanisms regulating alveolar regeneration and the pathogenesis of diseases resulting from a failure of regeneration.

Introduction

The alveolar epithelium consists of type 1 cells (AEC1s) and type 2 cells (AEC2s). AEC1s cover more than 95% of the alveolar surface and permit efficient gas exchange by virtue of their thin morphology. During lung injury, AECs, particularly AEC1s, die. Regeneration of the epithelium is required to restore normal lung structure and function; failure of efficient regeneration underlies the pathogenesis of many acute and chronic lung diseases. The principal progenitor responsible for regenerating the injured alveolar epithelium is the AEC2 (1). (Alternate progenitors are mobilized after severe injury with considerable AEC2 loss [refs. 1–6].) AEC2s replace lost cells via proliferation. However, once sufficient cell numbers are restored, AEC2s stop dividing and a subset transdifferentiate into AEC1s (1, 3, 7–9). Mechanisms underlying AEC2 proliferation after lung injury have been identified (10–12). For example, we and others have demonstrated a role for Wnt/β-catenin signaling (13–15). However, the molecular signals that induce proliferating AEC2s to exit the cell cycle and transdifferentiate into AEC1s are poorly understood.

To identify candidate signaling pathways implicated in cell cycle arrest and transdifferentiation during alveolar regeneration after injury, unbiased genome-wide expression analysis of regenerating AEC2s was indicated. However, our previous studies demonstrated that only a small fraction of AEC2s are mobilized

for regeneration (16, 17), precluding the use of bulk RNA sequencing. To discern rare subpopulations of AEC2s that are proliferating, exiting the cell cycle, or transdifferentiating, we exploited single-cell RNA sequencing (scRNAseq), which is ideally suited for the identification of subpopulations within a heterogeneous population and for providing insights into the regulation of transitions between cellular states (18). Lung injury was induced in *SftpcCreERT2;mTmG* mice, in which AEC2s and all of their progeny express GFP. AEC2-derived (GFP+) cells were isolated and subjected to scRNAseq. Unsupervised clustering revealed 3 distinct subpopulations of regenerating cells: proliferating, cell cycle arrest, and transdifferentiating. The gene expression profiles of these subpopulations were interrogated to identify candidate genes that may play a functional role in signaling proliferating cells to exit the cell cycle and transdifferentiate. TGF-β signaling was found to be highly activated in the cell cycle arrest subpopulation and relatively inactivated during transdifferentiation.

Although TGF-β is strongly implicated in the pathologic epithelial repair that characterizes pulmonary fibrosis (19), the role of TGF-β in physiologic epithelial repair remains undefined. TGF-β is known to inhibit epithelial cell proliferation (20–22), inducing cell cycle arrest via Smad3-dependent upregulation of cyclin-dependent kinase (CDK) inhibitors such as p15 (20–24). In animal models of lung injury, TGF-β levels nadir during AEC2 proliferation and then markedly increase at the end of the proliferation phase (25, 26). There are conflicting reports regarding the role of TGF-β signaling in AEC1 differentiation during alveologenesis (27–29) and from mature AEC2s (30–32). Since TGF-β signaling was activated in the cell cycle arrest subpopulation and relatively inactivated in transdifferentiating cells, we hypothesized that TGF-β is a critical signal inducing proliferating AEC2s to exit the cell cycle but must be inactivated to allow AEC2-to-AEC1 transdifferentiation, a hypothesis that we tested in cultured cells.

To our knowledge, this is the first reported scRNAseq study of regenerating AEC2s. We uncovered what we believe are novel regenerative transitional subpopulations, interrogated their gene expression profiles, confirmed the functional role of TGF-β in vitro, and generated a database of candidate pathways for future studies of physiologic and pathologic alveolar repair.

Results

scRNAseq of naive and regenerating AECs. Because we aimed to identify mechanisms of physiologic repair by AEC2s, we used a model of lung injury in which normal epithelial structure is restored primarily by AEC2s. In the LPS model, the proportion of lineage-labeled AEC2s remained constant during alveolar regeneration (Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.123637DS1). This suggested that AEC2s rather than an unlabeled cell type were the principal progenitor of nascent AEC2s and, unless other cell types directly differentiated into AEC1s, of nascent AEC1s. Our previous work revealed that at day 7 after LPS, some AEC2s are proliferating and some are transdifferentiating (16, 17). Accordingly, we selected day 7 as a time point to capture for scRNAseq cells that were proliferating, exiting the cell cycle, and transdifferentiating. *SftpcCreERT2;mTmG* mice were treated with LPS or left untreated. At day 7, "Naive AEC2s" (Tomato−GFP+) and "Naive Non-AEC2 Epithelial cells" (Tomato+GFP−CD45−EpCAM+T1α+) from control mice and "Injured AEC2-Derived cells" (Tomato−GFP+) from LPS-treated mice were sorted and subjected to scRNAseq (Supplemental Figures 2 and 3). Cells were projected into 2-dimensional space using t-distributed stochastic neighbor embedding (tSNE) (Figure 1A). The location of the sorted Naive AEC2, Naive Non-AEC2 Epithelial, and Injured AEC2-Derived cells within the tSNE plot is shown in Figure 1B. The tSNE plots derived from 2 separate scRNAseq experiments were similar (Supplemental Figure 4A). As anticipated, Naive AEC2s and Injured AEC2-Derived cells expressed high levels of GFP, whereas Naive Non-AEC2 Epithelial cells did not (Supplemental Figure 4B). Unsupervised clustering revealed 16 distinct clusters of cells (Figure 1C and Supplemental Figure 4C). The genes most differentially expressed by each cluster are shown in Supplemental Figure 4D and Supplemental Table 1.

Identification of naive AEC1 and AEC2 populations. We first examined the epithelial cells isolated from naive mice. Canonical club, basal, and ciliated cell markers were each expressed by a unique cluster of sorted Naive Non-AEC2 Epithelial cells (indicated by light orange, blue, and yellow circles, respectively, in Supplemental Figure 5, A–C) and were not expressed by the sorted Naive AEC2s. Canonical AEC1 markers were highly expressed by a distinct cluster of Naive Non-AEC2 Epithelial cells (indicated by green circles in Figure 1D and Supplemental Figure 5D). As anticipated, the sorted Naive AEC2s expressed canonical AEC2 markers (indicated by red circles in Figure 1E and Supplemental Figure 5E).

Sftpc+*Scgb1a1*+ cells (5) were identified (indicated by dark orange circles in Figure 1E and Supplemental Figure 5, A and E). Rare contaminating nonepithelial cells were identified: hematopoietic cells (indicated by pink circles in Supplemental Figure 6) and endothelial cells/fibroblasts (indicated by purple circles in Supplemental Figure 6). Based on expression of canonical markers, clusters representing naive cell types were identified (Figure 1F). Expression of canonical markers by each epithelial cell type is shown in Figure 1G.

The genes most differentially expressed by naive AEC1s and AEC2s are shown in Figure 1H and Supplemental Table 2. These included canonical AEC1 and AEC2 markers as well as additional markers specifically expressed by AEC1s or AEC2s. With the exception of *Hopx*, canonical AEC1 markers were less specific for AEC1s than the markers *Clic5* and *Akap5* (Figure 1D and Supplemental Figure 5D). *T1α* and *Aqp5* were expressed by airway epithelia (Figure 1D and Supplemental Figure 5D). *Aqp5* was also expressed by AEC2s, which was confirmed on the protein level (Supplemental Figure 5D and Supplemental Figure 7). The segregation of naive epithelial cells into clusters representative of known cell types demonstrated the robustness of the system and provided a framework for the identification of transitional regenerative cell types.

Identification of proliferating, cell cycle arrest, and transdifferentiating AEC2-derived cells. Airway epithelial cells and nonepithelial cells were excluded from further analysis. To identify proliferating AEC2s, expression of genes involved in cell proliferation was assessed. Cluster 10 was identified as the proliferating subpopulation (shown in blue circles in Figure 2A). Most proliferating AEC2s were Injured AEC2-Derived cells. However, rare Naive AEC2s clustered with the proliferating subpopulation and expressed markers of proliferation; these are likely the AEC2s responsible for homeostatic turnover (1, 9, 15). There were virtually no significant differences in gene expression between the Naive AEC2s and injured AEC2-Derived cells localized to this cluster (Supplemental Table 3). The proliferating subpopulation included cells in S, G_2, and M phase of the cell cycle (shown in blue circles in Supplemental Figure 8, A–C; ref. 33).

To identify cells that had stopped proliferating and were exiting the cell cycle, we assessed expression of genes characteristic of cell cycle (G_1) arrest (Figure 2B and Supplemental Figure 8D) and CDKs not unique to G_1 (Supplemental Figure 8E and ref. 33). Cluster 8 was identified as the cell cycle arrest subpopulation (shown in green circles in Figure 2B and Supplemental Figure 8D). Based on upregulation of AEC1 markers and downregulation of AEC2 markers, cluster 6 was identified as the transdifferentiating subpopulation (shown in red circles in Figure 2, C and D, and Supplemental Figure 8, F and G). The AEC2-derived cells in the transdifferentiating subpopulation expressed many but not all genes highly expressed by mature (naive) AEC1s. For example, *Igfbp2*, which is a late marker of AEC1 differentiation during alveologenesis (34), was expressed by most naive AEC1s but not by transdifferentiating AEC2-derived cells (Supplemental Figure 8H). A few Injured AEC2-Derived (GFP+) cells clustered with the naive AEC1 population, suggesting that they had fully differentiated into mature AEC1s (Supplemental Figure 8I). A summary of gene expression by the 3 regenerative subpopulations — proliferating, cell cycle arrest, and transdifferentiating — is shown in Figure 2E. The number of cells in each regenerative subpopulation is shown in Supplemental Figure 8J. The presence of these 3 regenerative subpopulations was validated by immunostaining and/or in situ hybridization (Figure 3, Supplemental Figure 8K, and refs. 16, 17). The identification of the regenerative subpopulations — proliferating, cell cycle arrest, and transdifferentiating — was substantiated by pathway analysis (Supplemental Figure 8L). Aside from these 3 distinct regenerative subpopulations, which appear as peninsulas on the tSNE plot, all other clusters of Injured AEC2-Derived cells (clusters 0–4) had similar gene expression profiles (Supplemental Figure 4D and Supplemental Table 1), many of them converging with clustering at lower resolution (Supplemental Figure 9A). Therefore, we grouped these clusters as "Other Injured AEC2-Derived cells." To summarize, a tSNE indicating all identified naive and injured populations is shown in Figure 4A. Notably, there was minimal overlap on the tSNE between the Naive and Other Injured AEC2-Derived cells (Figure 1B and Figure 4A). Although canonical correlation analysis and hierarchical clustering suggested that the Other Injured AEC2-Derived cells were only mildly perturbed as compared with the Naive AEC2s (Supplemental Figure 9, B and C), their segregation on tSNE was due to differences in gene expression. The genes most differentially expressed between the Naive AEC2s and Other Injured AEC2-Derived cells are shown in Supplemental Figure 9D and Supplemental Table 4. The most striking difference is the activation of inflammatory pathways by the injured cells (Supplemental Figure 9E).

Gene expression profiles of regenerative subpopulations. To identify candidate signaling pathways that may regulate progression through the stages of alveolar regeneration, we examined the gene expression profiles of the
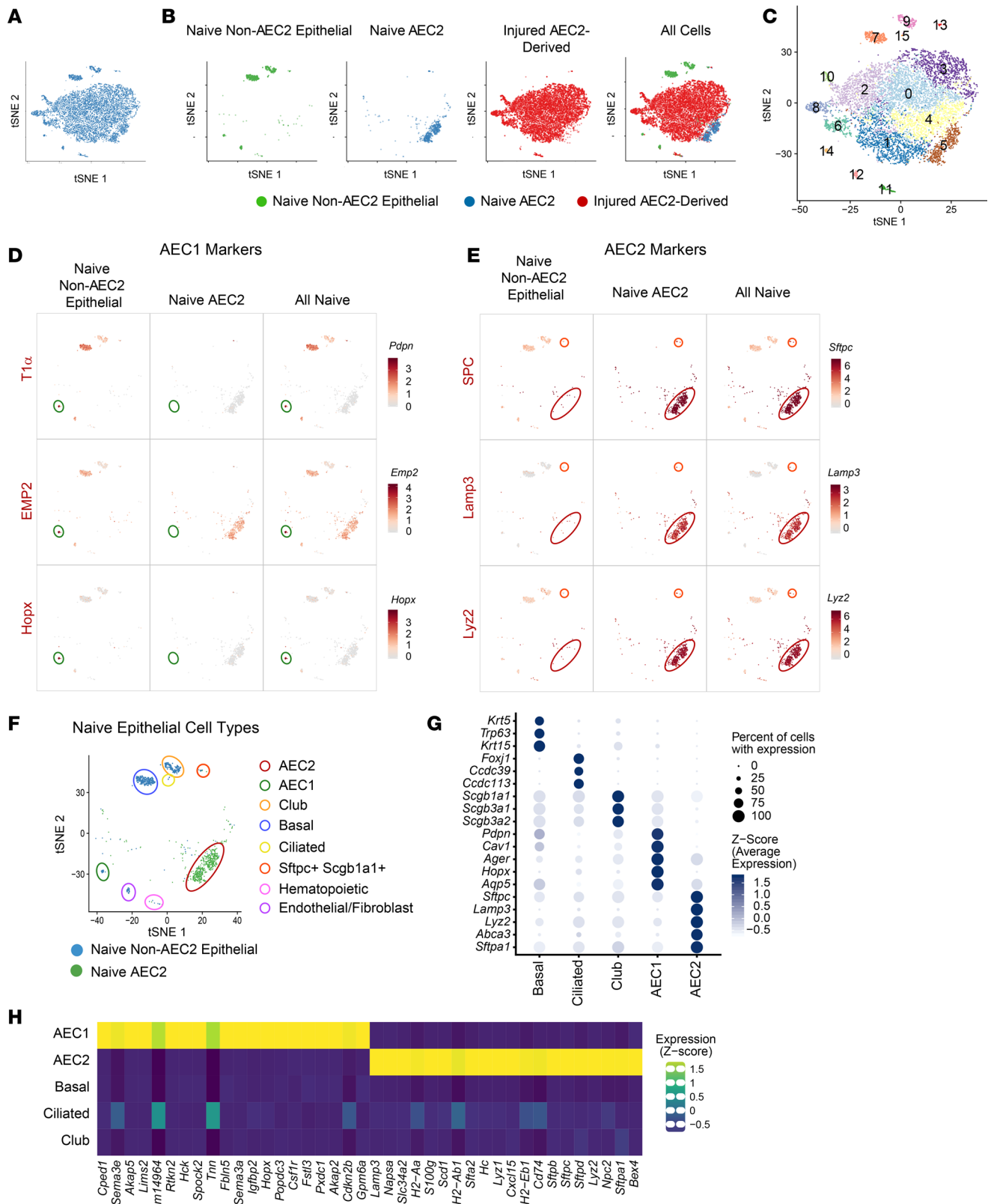
Figure 1. scRNAseq reveals naive lung epithelial cell types. *SftpcCreERT2;mTmG* mice were treated with or without LPS. At day 7, non-AEC2 epithelial cells and AEC2s from naive mice and AEC2-derived cells from LPS-treated mice were subjected to scRNAseq. (**A**) tSNE plot of all cells sequenced. (**B**) Locations within the tSNE plot of Naive Non-AEC2 Epithelial, Naive AEC2, and Injured AEC2-Derived cells. (**C**) Unsupervised clustering with a resolution of 0.6. (**D** and **E**) Gene expression of canonical AEC1 (**D**) and AEC2 (**E**) markers (natural log of normalized counts). (**F**) Based on gene expression patterns, specific epithelial and nonepithelial cell types were identified. (**G**) Expression levels of canonical cell markers by the identified cell types. (**H**) Heatmap of the 20 most differentially expressed genes in naive AEC2 and AEC1s compared with other naive epithelial cell types, ranked in order of Bonfer-

roni-corrected *P* value. Green circles (cluster 14), AEC1s; red circles (cluster 5), AEC2s; dark orange circles (cluster 13), *Scgb1a1+Sftpc+* cells; light orange circles (cluster 9), club cells; yellow circles (cluster 15), ciliated cells; blue circles (cluster 7), basal cells; pink circles (cluster 11), hematopoietic cells; purple circles (cluster 12), endothelial cells/fibroblasts. *n* = 2 mice per group.

proliferating, cell cycle arrest, and transdifferentiating subpopulations. The most differentially expressed genes in each subpopulation are shown in Figure 4B and Supplemental Table 5. The transcription factors and cell surface markers most differentially expressed by each of the regenerative subpopulations are shown in Supplemental Tables 6 and 7. Since transdifferentiation involves extensive cell spreading (16), we identified the most differentially expressed cell motility genes (Supplemental Table 8). Pathway analysis revealed that TGF-β was the most highly activated upstream regulator of gene expression by the 3 subpopulations of regenerating AEC2s taken together in comparison with all other Injured AEC2-Derived cells (Figure 4C). The next most highly activated pathways in the regenerating AEC2s were Myc (a Wnt/β-catenin target gene) and GM-CSF; these were most highly activated in the proliferating subpopulation, consistent with their established roles in AEC2 proliferation (13–15, 35, 36).

TGF-β signaling induces AEC2 cell cycle arrest. Although TGF-β was the most highly activated upstream regulator of gene expression by the regenerating AEC2s taken together, TGF-β signaling was largely inactive in naive and proliferating AEC2s, highly activated in the cell cycle arrest subpopulation, and relatively inactivated in the transdifferentiating subpopulation (Figure 4C). In fact, TGF-β signaling was the top upstream regulator of the cell cycle arrest subpopulation ($P = 5.17 \times 10^{-29}$). Several genes in the TGF-β pathway were among the most highly expressed genes in this subpopulation (Figure 4B), and additional TGF-β pathway genes were specifically upregulated during cell cycle arrest (Figure 5, A and B, and Supplemental Figure 10, A–C). These included *TGFβ2*, α_v and β_6 integrins, which are critical for activation of TGF-β during lung injury (37), and many TGF-β/Smad target genes. Expression of *TGFβ2* and β_6 integrin was confirmed by in situ hybridization (Figure 5C and Supplemental Figure 10, D and E). There was no evidence of epithelial-mesenchymal transition as determined by expression of epithelial and mesenchymal markers (Supplemental Figure 10, F and G).

Since TGF-β signaling was specifically activated in the cell cycle arrest subpopulation and TGF-β induces G_1 arrest in other cell types (21), we hypothesized that TGF-β induces cell cycle arrest in proliferating AEC2s. To test this hypothesis, we cultured primary AEC2s under 2-dimensional culture conditions in which they are known not to proliferate. To assess whether 2-dimensional culture induces AEC2s to recapitulate the cell cycle arrest phase observed during epithelial regeneration in vivo, we assessed expression of the G_1 arrest markers that were upregulated in the cell cycle arrest subpopulation in vivo. *p15*, *p53*, and *CDK4* were upregulated within the first day of culture (Figure 6A). The CDK inhibitors that were not differentially expressed in the cell cycle arrest subpopulation in vivo, *p21* and *p27*, were also not significantly upregulated in vitro (Rachel L. Zemans, unpublished observations), supporting the notion that gene expression by cultured AEC2s was similar to that of AEC2s in the cell cycle arrest state during alveolar regeneration in vivo. Cultured AEC2s failed to undergo S phase, as assessed by EdU incorporation, confirming that they were in cell cycle arrest (Figure 6B). We next assessed whether TGF-β signaling was activated during cell cycle arrest in vitro. TGF-β activation was demonstrated by increased levels of p-Smad3 (Figure 6C and Supplemental Figure 11) and upregulation of many of the same TGF-β pathway genes that were upregulated during cell cycle arrest in vivo, including *TGFβ2* and α_v and β_6 integrins (Figure 6D). To determine whether TGF-β signaling was necessary for cell cycle arrest, cultured AEC2s were treated with a TGF-β pharmacologic inhibitor or neutralizing antibody. TGF-β inhibition prevented cell cycle arrest (Figure 6E) and attenuated the upregulation of p15 (Figure 6F).

TGF-β inhibition promotes transdifferentiation. While TGF-β signaling was activated in the cell cycle arrest subpopulation, it was relatively inactivated in the transdifferentiating cells in vivo (Figure 4C, Figure 5, A and B, and Supplemental Figure 10, A and B). This suggests the possibility that while TGF-β induces cell cycle arrest, TGF-β signaling must be inactivated to allow AEC1 differentiation. To test this, we again used cultured AEC2s, which begin to express AEC1 markers by day 3 of culture (Figure 7, A–C), although they do not fully differentiate into AEC1s (15, 38–40). To determine whether TGF-β inactivation promotes transdifferentiation, cultured AEC2s were treated with the TGF-β inhibitor or neutralizing antibody. TGF-β inhibition promoted AEC1 differentiation (Figure 7, D and E). Similarly, AEC2s isolated from *Tgfbr2fl/fl* mice were treated with an adenovirus expressing Cre, which resulted in gene deletion (Supplemental Figure 12). Knockout of *Tgfbr2* promoted AEC1 differentiation
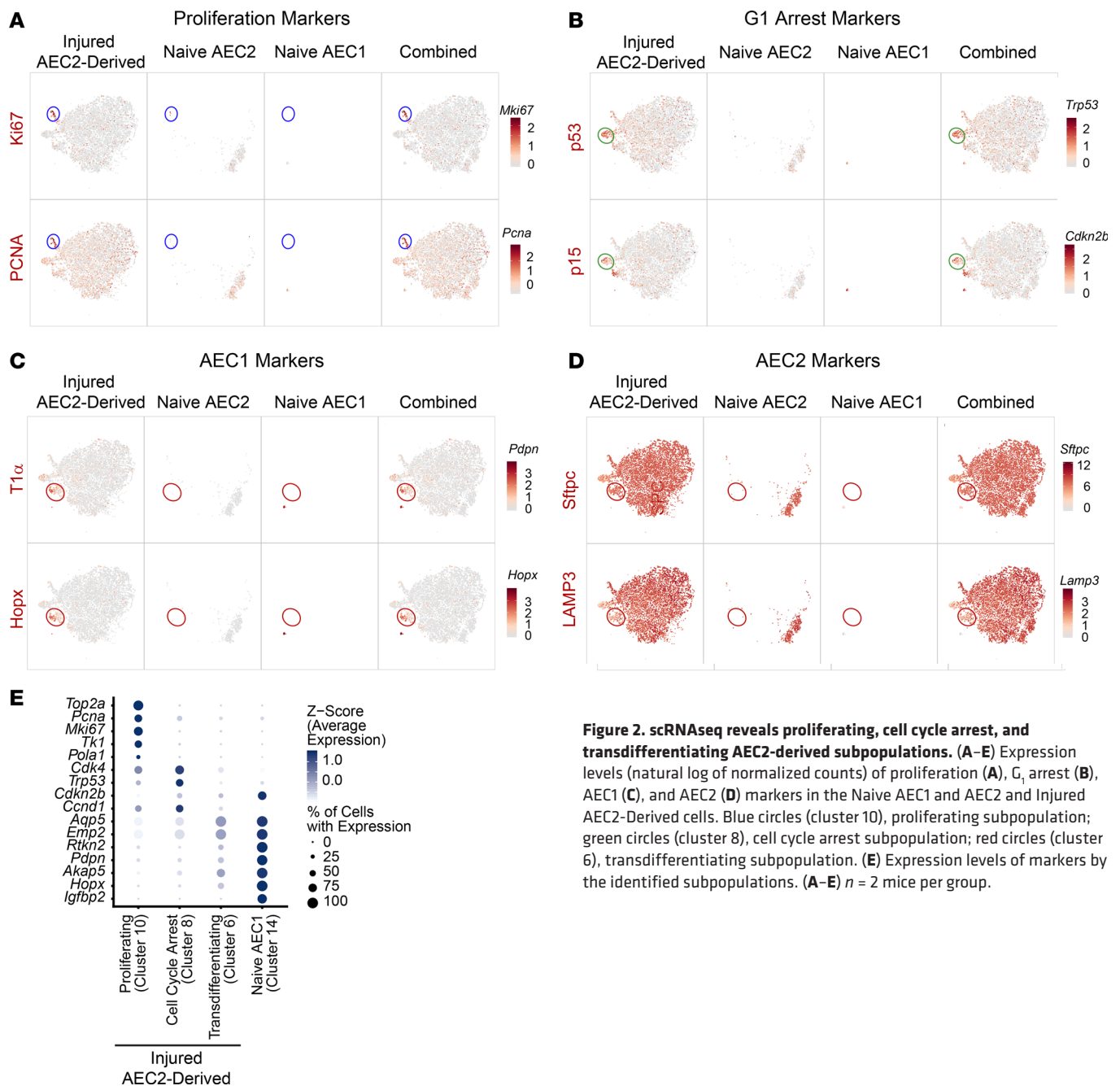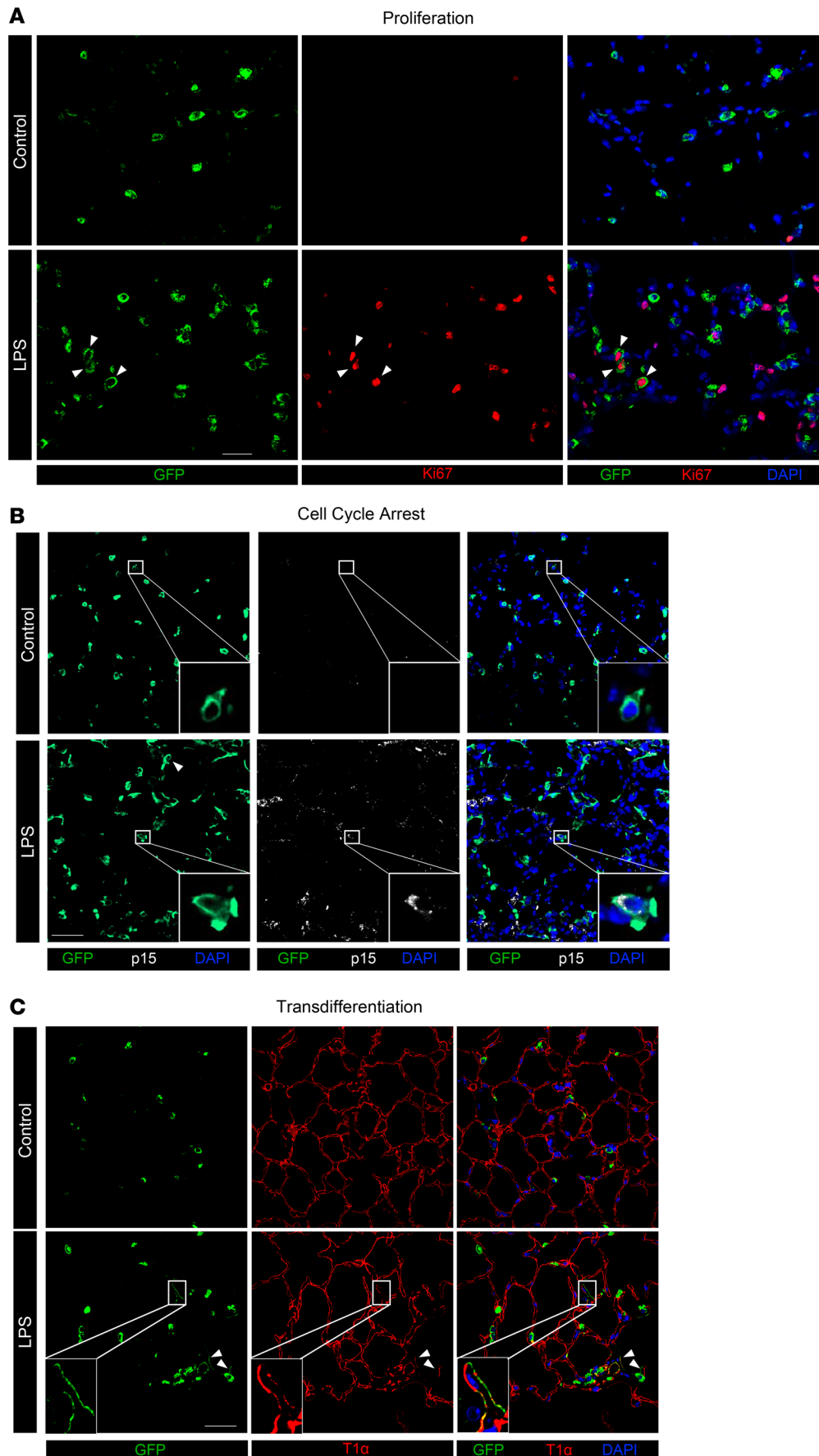
Figure 2. scRNAseq reveals proliferating, cell cycle arrest, and transdifferentiating AEC2-derived subpopulations. (A–E) Expression levels (natural log of normalized counts) of proliferation (A), G$_1$ arrest (B), AEC1 (C), and AEC2 (D) markers in the Naive AEC1 and AEC2 and Injured AEC2-Derived cells. Blue circles (cluster 10), proliferating subpopulation; green circles (cluster 8), cell cycle arrest subpopulation; red circles (cluster 6), transdifferentiating subpopulation. (E) Expression levels of markers by the identified subpopulations. (A–E) n = 2 mice per group.

(Figure 7, F and G). Taken together, these data suggest that TGF-β induces cell cycle arrest but inactivation of TGF-β promotes AEC1 differentiation.

Lineage relationship between regenerative intermediate subpopulations. To ascertain whether the proliferating, cell cycle arrest, and transdifferentiating subpopulations represent sequential regenerative stages in a linear trajectory or divergent fates, we performed pseudotime analysis. The results were variable depending on the input data and trajectory inference algorithm (Supplemental Figure 13, A–C). The cell cycle arrest and transdifferentiating subpopulations had similar gene expression (Supplemental Figure 9, A and C, and Supplemental Figure 13D). Therefore, our data are inconclusive as to cell fate trajectory.

Discussion

Harnessing the power of scRNAseq to discern cellular heterogeneity, we identified what we believe are novel transitional subpopulations of regenerating AEC2s — proliferating, cell cycle arrest, and transdifferentiating. Interrogation of their gene expression profiles revealed candidate signaling pathways that may

Figure 3. Immunofluorescence/in situ hybridization validates proliferating, cell cycle arrest, and transdifferentiating AEC2-derived subpopulations. Immunofluorescence staining of lung sections for GFP (**A**–**C**), Ki67 (**A**), and T1α (**C**) and in situ hybridization for p15 (**B**). Insets and arrowheads indicate double-positive cells. Scale bars: 25 μm (**A**), 50 μm (**B** and **C**). $n \geq 4$ mice per group.
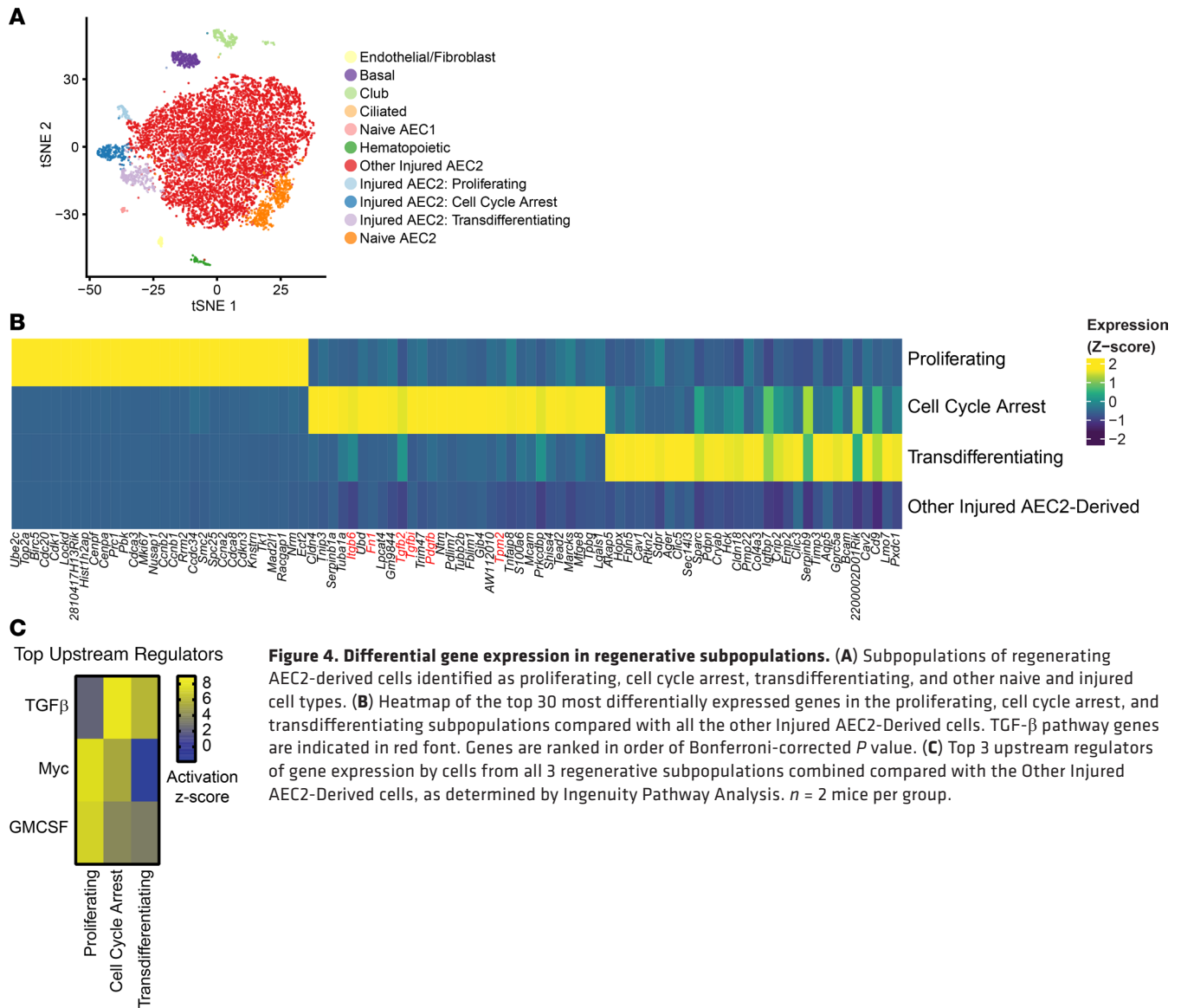
Figure 4. Differential gene expression in regenerative subpopulations. (**A**) Subpopulations of regenerating AEC2-derived cells identified as proliferating, cell cycle arrest, transdifferentiating, and other naive and injured cell types. (**B**) Heatmap of the top 30 most differentially expressed genes in the proliferating, cell cycle arrest, and transdifferentiating subpopulations compared with all the other Injured AEC2-Derived cells. TGF-β pathway genes are indicated in red font. Genes are ranked in order of Bonferroni-corrected P value. (**C**) Top 3 upstream regulators of gene expression by cells from all 3 regenerative subpopulations combined compared with the Other Injured AEC2-Derived cells, as determined by Ingenuity Pathway Analysis. n = 2 mice per group.

regulate alveolar regeneration after LPS-induced lung injury. One such pathway, TGF-β, was upregulated in the cell cycle arrest subpopulation and relatively downregulated in the transdifferentiating cells. Functional experiments in cultured cells confirmed that TGF-β induced cell cycle arrest, while deactivation promoted transdifferentiation. These data suggest that in the LPS model, TGF-β is a critical checkpoint in alveolar regeneration, halting the proliferation phase of alveolar regeneration but impeding transdifferentiation. While the mechanisms of AEC2 proliferation have been investigated, this study addresses a largely unanswered question in the field of lung regeneration: the mechanisms that halt proliferation and promote transdifferentiation, thus restoring normal alveolar structure and function.

While the antimitogenic effects of TGF-β in epithelia are well established (20, 21), we demonstrate here that in the context of alveolar regeneration TGF-β was activated in AEC2s at the end of the proliferation phase, signaling its termination. Building on previous studies demonstrating increased total lung TGF-β levels at the end of the proliferation phase in several models of lung injury (25, 26), this scRNAseq analysis now uncovers a specific subpopulation of AEC2s that respond to TGF-β by exiting the cell cycle. Here, TGF-β induced G_1 arrest via upregulation of $p15$, which, in contrast to p21 and p27, is known to specifically inhibit the G_1-to-S transition (23, 33). Notably, the cell cycle arrest subpopulation also expressed lower levels of AEC2 markers than the naive and proliferating AEC2s. Since TGF-β is known to downregulate surfactant protein expression (31, 41–44), we propose that TGF-β triggers coordinated $p15$ upregulation and surfactant protein downregulation, in turn inducing cell cycle exit and dedifferentiation, perhaps in anticipation of transdifferentiation.
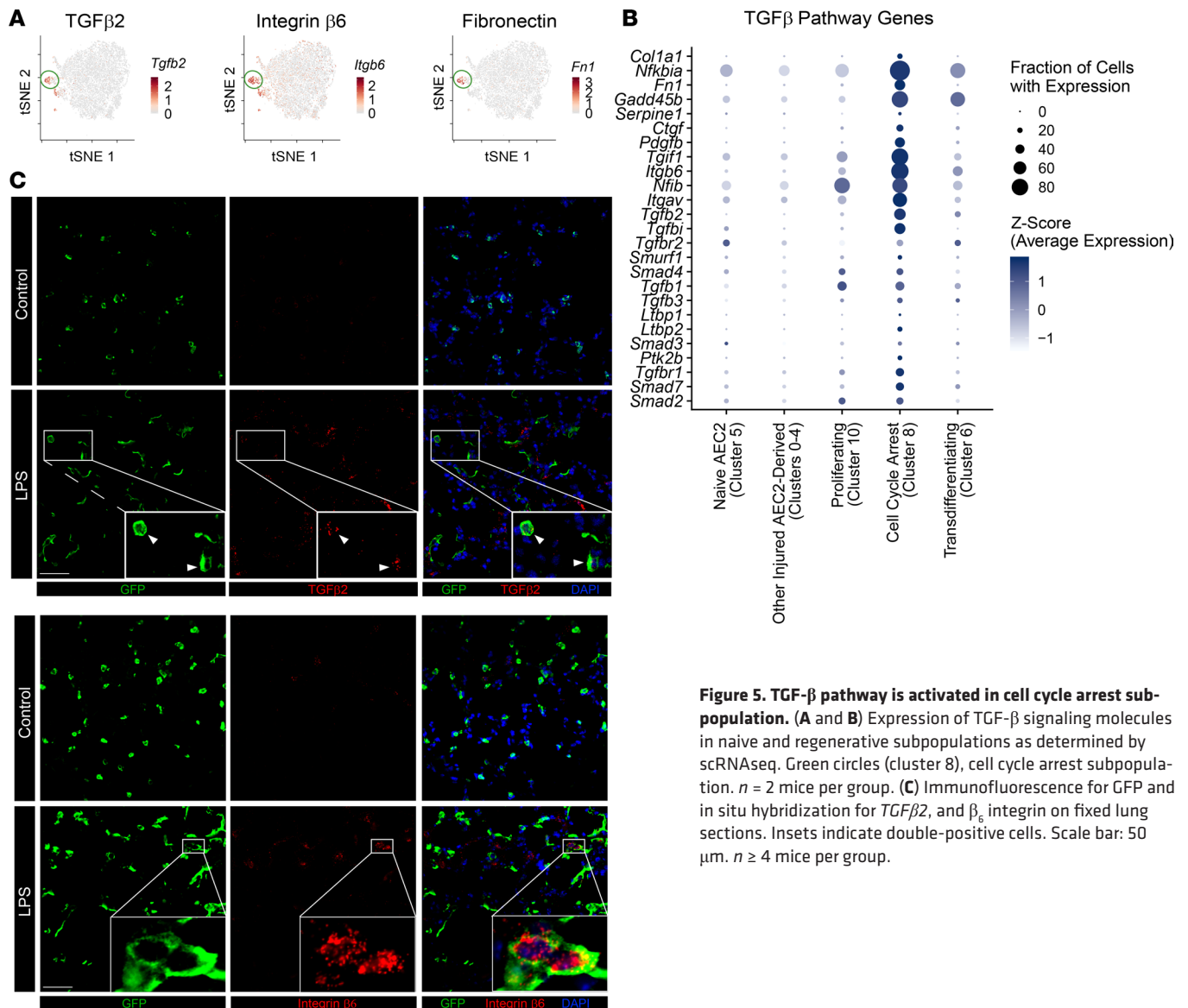
Figure 5. TGF-β pathway is activated in cell cycle arrest subpopulation. (**A** and **B**) Expression of TGF-β signaling molecules in naive and regenerative subpopulations as determined by scRNAseq. Green circles (cluster 8), cell cycle arrest subpopulation. *n* = 2 mice per group. (**C**) Immunofluorescence for GFP and in situ hybridization for *TGFβ2*, and β_6 integrin on fixed lung sections. Insets indicate double-positive cells. Scale bar: 50 μm. *n* ≥ 4 mice per group.

The data presented here suggest that inhibition of TGF-β promotes transdifferentiation in cultured cells. Our study is more definitive than those suggesting that TGF-β promotes transdifferentiation (30, 31), as we used multiple methods of TGF-β gain and loss of function in 2 species and assessed validated AEC1 markers on both the mRNA and protein levels. Moreover, our findings were corroborated by in vivo expression data. A recent report demonstrated that bone morphogenetic protein 4 (BMP4) induces transdifferentiation (32). Accordingly, in our study, *BMP4* and its target genes *Id1* and *Id3* were highly differentially expressed by the transdifferentiating cells (Supplemental Table 5). The inhibitory role of TGF-β demonstrated here suggests that transdifferentiation may be tightly regulated by a balance of opposing TGF-β/BMP stimuli, an example of classic TGF-β/BMP antagonism.

The extent to which alveolar regeneration recapitulates alveologenesis remains a fundamental question in the field. Expression of *Igfbp2* by the mature AEC1s but not the regenerative transdifferentiating cells is consistent with *Igfbp2* expression by differentiating AEC1s only late in alveologenesis (34). The specific role of TGF-β in alveologenesis remains controversial (27–29). However, our finding that inhibition of TGF-β promoted AEC1 differentiation during regeneration is consistent with previous reports that inhibition of transcriptional pathways promotes AEC1 differentiation during alveologenesis and homeostasis. During alveologenesis, nascent AEC1s express no new transcription factors, suggesting that AEC1 differentiation may be due to withdrawal of inhibitory transcriptional pathways (18). Indeed, downregulation of transcriptional pathways such as
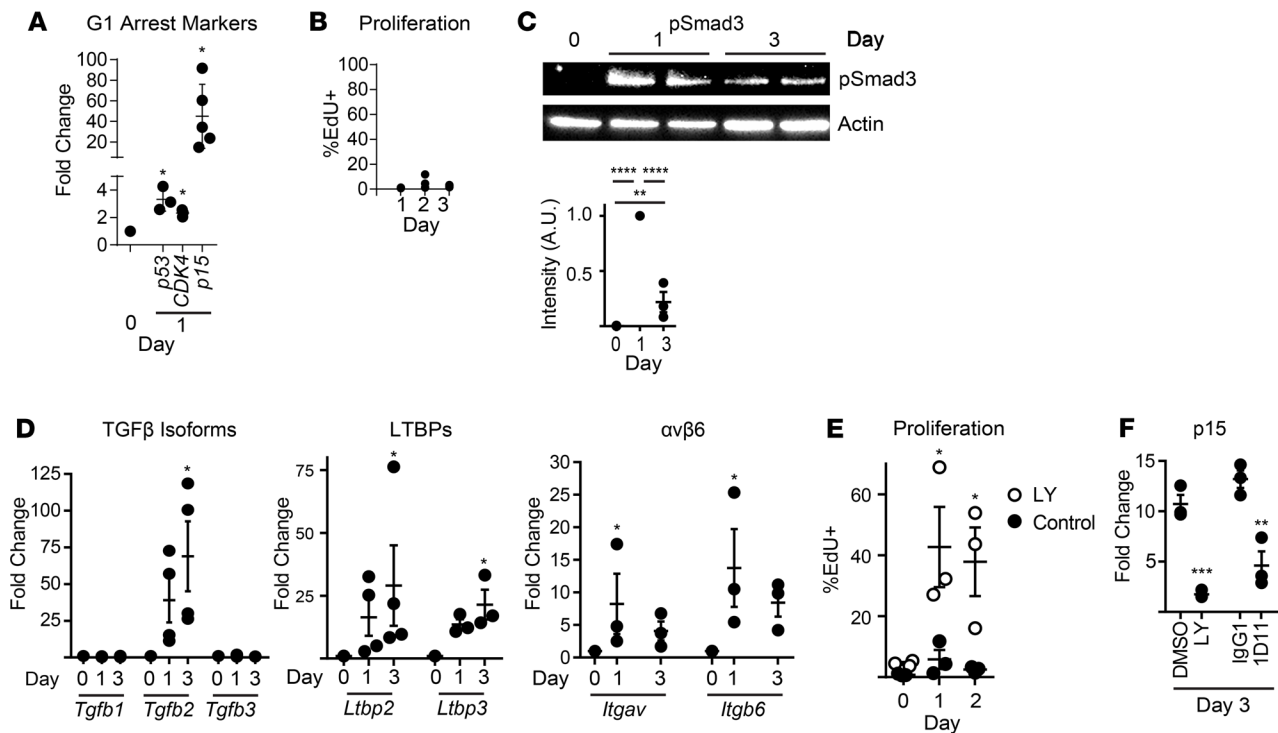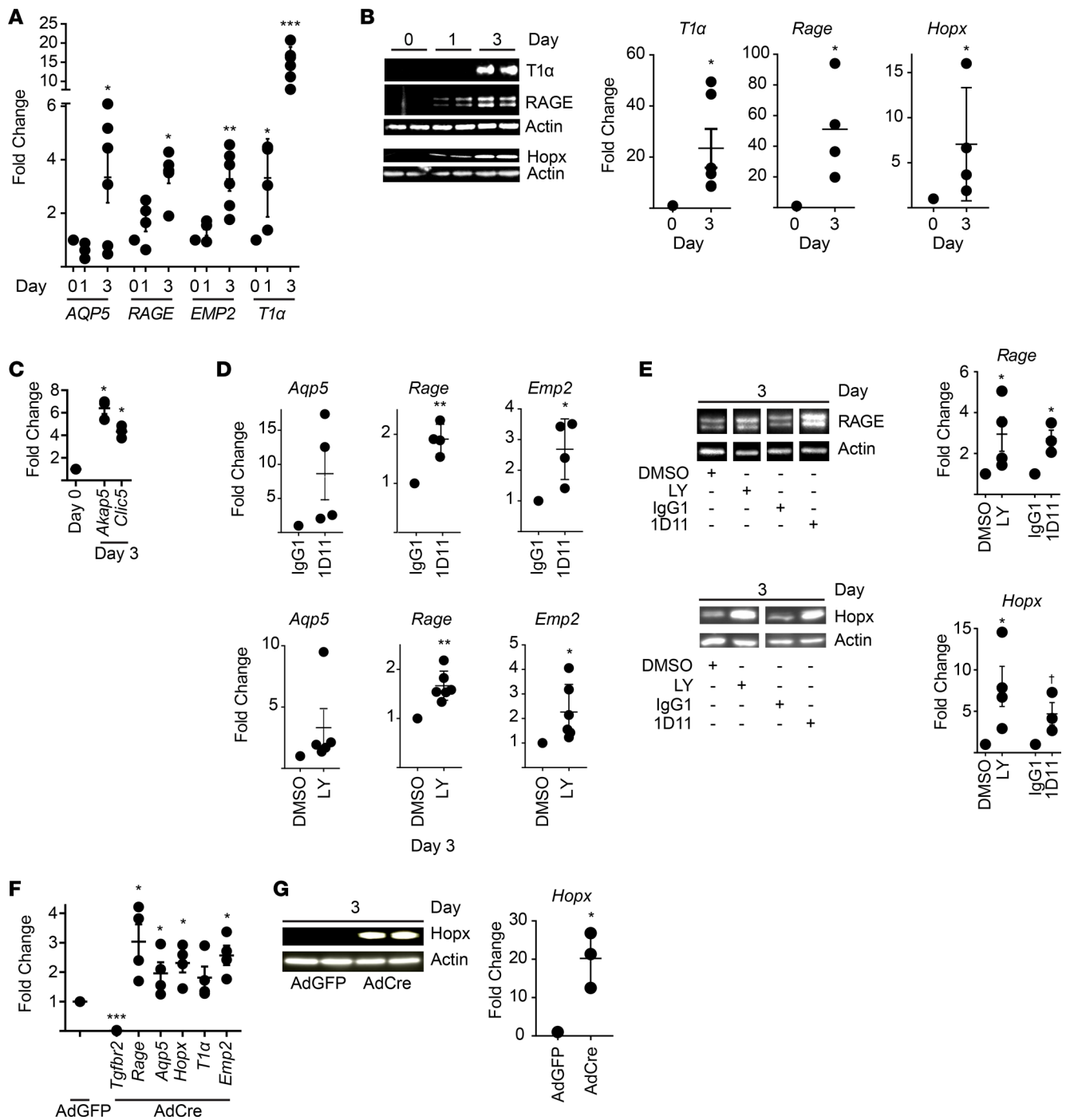
Figure 6. TGF-β induces AEC2 cell cycle arrest in cultured AEC2s. Primary rat AEC2s were cultured in the presence or absence of the TGFβRI inhibitor LY364947 (LY) or the TGF-β–neutralizing antibody 1D11 with or without EdU. (**A**, **D**, and **F**) Quantitative PCR was performed. Data are presented as fold change relative to day 0. (**B** and **E**) EdU incorporation is shown. (**C**) Western blotting of cell lysates for p-Smad3 and actin. Densitometry of p-Smad3 corrected for actin is shown. Experiments were performed at least 3 times, each with 2 technical replicates. Two-tailed t test or 1- or 2-way ANOVA with post hoc analysis for multiple comparisons was performed and corrected for repeated measures. Mean ± SEM is shown. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$ compared with day 0 (**A** and **D**) or day 1 (**C**) or LY364947 vs. DMSO or 1D11 vs. IgG1 (**E** and **F**).

Wnt/β-catenin is necessary for AEC1 differentiation during both alveologenesis and homeostasis (14, 15, 45, 46). Since Wnt/β-catenin signaling promotes proliferation (13–15) but inhibits transdifferentiation, we propose the following construct: During alveolar regeneration, Wnt/β-catenin stimulates AEC2 proliferation. Once lost cells have been replaced, TGF-β induces cell cycle arrest, halting proliferation; withdrawal of Wnt/β-catenin and TGF-β promotes efficient transdifferentiation. Future studies exploring the mechanisms of Wnt/TGF-β crosstalk in the context of alveolar regeneration are warranted.

This study also provided insights into the homeostatic and inflammatory functions of the alveolar epithelium. First, the rare proliferating naive AEC2s and those mobilized to proliferate after injury had similar gene expression profiles. This suggests that the mechanisms underlying homeostatic turnover and regeneration are conserved and the rare AEC2s responsible for homeostatic turnover are likely not inherently different from the AEC2s mobilized to proliferate after injury but rather are specified by their location within the niche (1, 9, 15). Second, whereas the widely accepted AEC1 markers *T1α* and *AQP5* were expressed by other lung epithelial cell types, as has been previously reported (47–49), we identified highly specific AEC1 markers that may be exploited to drive Cre expression or isolate AEC1s, invaluable techniques that remain somewhat elusive (2, 34, 40, 48, 50–52). Finally, we identified immune pathways activated in the Injured AEC2-Derived cells, which could be investigated to elucidate the inflammatory and host defense functions of the alveolar epithelium.

Although scRNAseq is a powerful technique, there are several caveats. First, the validity of scRNAseq studies rests on the assumption that the cell isolation yields representative cells without significant changes in gene expression. Therefore, we performed in situ hybridization (ISH) and immunofluorescence staining of key markers to validate the regenerative subpopulations. The percentage of cells in the transdifferentiating subpopulation as determined by scRNAseq (Supplemental Figure 8J) was similar to our previous rigorous quantification of transdifferentiation in situ (16). The proliferating subpopulation was underestimated by scRNAseq as compared with our previous estimates (17), perhaps because of impaired recoverability of cells from the lung digest. The percentage of cells in the cell cycle arrest subpopulation as determined

Figure 7. TGF-β inhibits AEC2-to-AEC1 transdifferentiation in cultured AEC2s. (A–E) Primary rat AEC2s were cultured in the presence or absence of the TGFβRI inhibitor LY364947 (LY) or the TGF-β–neutralizing antibody 1D11 for the indicated time periods. (A, C, and D) Quantitative PCR (qPCR) for AEC1 markers. (B and E) Western blotting for AEC1 markers with densitometry. Vertical white lines indicate that the lanes were run on the same gel but were noncontiguous. Data are shown as fold change compared with day 0 (A–C) or fold change compared with IgG or DMSO controls (D and E). (F and G) AEC2s were isolated from *Tgfbr2fl/fl* mice, transduced with AdGFP or AdCre, and cultured for 3 days. qPCR (F) and Western blotting with densi- tometry (G) shown as fold change in AdCre samples compared with AdGFP samples. Each experiment was performed at least 3 times, except the 1D11 experiments with 2 technical replicates each. Two-tailed *t* test or 1- or 2-way ANOVA with post hoc analysis for multiple comparisons was performed and corrected for repeated measures. Mean ± SEM is shown. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, †$P = 0.07$ calculated for fold change relative to day 0 (A–C) or DMSO vs. LY or IgG1 vs. 1D11 (D and E) or AdGFP vs. AdCre (F and G).

by scRNAseq was less than that determined by ISH for p15, TGF-β2, or β$_6$ integrin, probably because some transdifferentiating cells express these genes also. However, while coexpression of these markers by scRNAseq was quantified, coexpression by ISH was not. Another minor limitation of our study was the limited number of naive AEC1s sequenced, although they were sufficient to provide a reference point for identifying transdifferentiating cells.

This study sets the stage for future investigation of remaining unknowns. First, although our data imply autocrine signaling by TGF-β2, other cellular sources and isoforms of TGF-β are not excluded; the mechanisms by which adequate replacement of lost cells is sensed, thus triggering TGF-β expression, remain to be determined. The mechanism by which TGF-β signaling is inhibited to allow transdifferentiation is also unknown. In addition, our study examined only the LPS model, a model in which epithelial injury is neutrophil dependent likely via oxidants and proteases, though the precise underlying molecular mechanisms are not entirely understood (53, 54). Based on studies in the bleomycin and hyperoxia models showing that TGF-β levels nadir during AEC2 proliferation and then increase at the end of the proliferation phase (25, 26), we suspect that the regenerative function of TGF-β demonstrated here is also operative after direct lung injury; however, this has not been shown. In fact, the function of TGF-β in alveolar regeneration in the LPS model has not yet been validated in vivo. Precisely timing TGF-β loss of function to the cell cycle arrest or transdifferentiation phase without effects on other regenerative phases or on the initial injury (55) will be necessary. Finally, the lineage trajectory of the 3 regenerative subpopulations remains unknown. Our data do not definitely establish whether the cell cycle arrest subpopulation represents an intermediate state between proliferation and transdifferentiation or whether the cell cycle arrest and transdifferentiating subpopulations represent divergent fates. Similarity in gene expression (Supplemental Figure 13D), nearness on hierarchical clustering (Supplemental Figure 9C), and convergence on low-resolution tSNE plots (Supplemental Figure 9A) demonstrate that the cell cycle arrest and transdifferentiating subpopulations are closely related. These findings taken together with expression of AEC1 markers by the cell cycle arrest subpopulation (Figure 2C and Supplemental Figure 8F) suggest that its fate is to transdifferentiate. However, pseudotime analysis raised the possibility that some cells in the cell cycle arrest subpopulation may not ultimately transdifferentiate (Supplemental Figure 13, A–C). It is quite possible that the cell cycle arrest subpopulation is heterogeneous with respect to fate, with some cells transdifferentiating, some maintaining an AEC2 phenotype, and, if proliferation yields an excessive number of AEC2s, some undergoing apoptosis. Pseudotime analysis also suggested the possibility that some AEC2s may transdifferentiate without first undergoing a round of proliferation, though previous studies have clearly demonstrated that at least some cells proliferate prior to transdifferentiation (7, 56). Ultimately, these questions of cell fate will need to be addressed by lineage tracing experiments.

The termination of proliferation and initiation of transdifferentiation are of critical physiologic and clinical relevance. These processes are necessary to prevent AEC2 hyperplasia and restore the exquisitely thin AEC1s that cover more than 95% of the alveolar surface and permit efficient gas exchange. Studies of lung regeneration are relevant to diverse lung diseases that result from a failure of normal repair. However, the mechanisms underlying the termination of proliferation and initiation of transdifferentiation have specific relevance to the pathogenesis of pulmonary fibrosis, which is driven by excessive TGF-β and is characterized by AEC2 hyperplasia with a paucity of AEC1s (57). Based on the findings presented here, we speculate that supraphysiologic levels of TGF-β may prevent transdifferentiation, suspending postproliferative AEC2s in a transitional state reminiscent of the cell cycle arrest subpopulation, with the failure to regenerate AEC1s in turn stimulating fibrogenesis (19, 58).

Alveolar regeneration is undoubtedly regulated by a complex interaction of multiple signaling pathways. The full impact of our study will be realized by future mechanistic studies built on the foundation established here. The transcription factors, signaling molecules, and cell motility genes differentially expressed by the regenerative subpopulations are promising candidates for regulating the transitions of regenerating AEC2s from quiescence to proliferation to cell cycle arrest and transdifferentiation. The cell surface markers identified may be used to purify large numbers of cells from a specific subpopulation for deeper transcriptomic or proteomic analysis or for in vitro functional experiments.

In summary, this study uncovered potentially novel transitional subpopulations of regenerating AECs and their gene expression profiles, revealed some similarities between alveolar regeneration, homeostasis, and development, and established TGF-β as a critical checkpoint in alveolar regeneration in the LPS model. Our findings will serve as a foundation for additional studies investigating the mechanisms of

physiologic alveolar regeneration, the extent to which they recapitulate development, and the manner in which they are dysregulated in disease.

Methods

Animals. Animals were maintained in a pathogen-free environment on a 12-hour light/12-hour dark cycle with full access to food and water. *SftpcCreERT2* mice (59), obtained from Harold Chapman (UCSF, San Francisco, California, USA), were crossed to *Rosa26-mTmG* (abbreviated *mTmG*) mice (The Jackson Laboratory). Male *SftpcCreERT2*[+/–]*;mTmG*[+/–] mice were given 20 mg/ml tamoxifen in corn oil at a dose of 0.25 mg/g body weight i.p. every other day for 3 doses at age 6 weeks. Recombination efficiency is shown in Supplemental Figure 1. At 12 weeks of age, mice were treated with 40 μg LPS (*E. coli* 0111:B4, List Biological Laboratories) in 50 μl 0.9% saline by intratracheal instillation or left untreated. Mice were randomly assigned to treatment groups. Mice were euthanized 7 or 27 days later. Sprague-Dawley rats aged 2–4 months were purchased from Charles River or obtained from the University of Michigan Rodent Recycling Program. *Tgfbr2*[fl/fl] mice were purchased from The Jackson Laboratory.

Fluorescence-activated cell sorting. Lungs were perfused with 10 ml PBS at 20 cmH_2O, lavaged twice with 1 ml PBS plus 5 mM EDTA plus 5 mM EGTA, twice with RPMI 1640 plus 25 mM HEPES, and once with 4.5 U/ml elastase (Worthington LS00229) in RPMI plus 25 mM HEPES. Next, 2 ml elastase 4.5 U/ml was instilled into the lungs, followed immediately with 500 μl 1% low–melting point agarose. Lungs were submerged into 2 ml elastase 4.5 U/ml and incubated for 45 minutes at 37°C. The trachea and bronchial tree were dissected away, and remaining lung tissue was submerged in 5 ml ice-cold FBS with 0.2 mg/ml DNAse (Sigma-Aldrich DN25). Tissue was minced to approximately 0.5- to 1-mm^3 fragments with scissors, shaken vigorously for 2 minutes, and filtered over 100-, 40-, and 20-μm filters. Cell suspension was centrifuged at 350 *g* for 12 minutes at 4°C onto a Percoll cushion. The cell pellet was resuspended in 30 μl DNAse (4 mg/ml), and 300 μl RPMI 1640 plus 25 mM HEPES plus 20% FBS was added. Cells were stained with anti-T1α–PE/Cy7 (eBioscience 25-5381), anti-EpCAM–647 (BioLegend 118212), and anti-CD45–A700 (eBioscience 56-0451-80) antibodies with FcBlock (BD Pharmingen 553142) for 60 minutes at 4°C. DAPI (Thermo Fisher Scientific 62248) 1 μg/ml was added. Cells were sorted on a Moflo XDP (Beckman). Gates were drawn based on fluorescence-minus-one controls. DAPI[–]Tomato[–]GFP[+] cells and DAPI[–]Tomato[+]GFP[–]CD45[–]EpCAM[+]T1α[+] cells were collected. To confirm purity and viability of the sorted populations, an aliquot of the sorted cells was run through the cytometer a second time, and cells were found to be approximately 95% live and 92%–98% pure, as defined by predicted expression of markers.

Single-cell RNA sequencing. Five thousand to fifteen thousand Tomato[–]GFP[+] cells from the LPS-treated mice ("Injured AEC2-Derived"), 2,000 Tomato[–]GFP[+] cells from the control mice ("Naive AEC2"), and 30 μl of undiluted Tomato[+]GFP[–]CD45[–]EpCAM[+]T1α[+] cells from the control mice ("Naive Non-AEC2 Epithelial") were loaded into a Chromium Single Cell 3′ Solution Capture System (10X Genomics). The processed libraries were sequenced on an Illumina HiSEQ4000 instrument using 26 × 8 × 101 sequencing. The scRNAseq experiment was performed twice, with 1 control and 1 LPS mouse in each experiment. Sequencing metrics, including depth and saturation, are shown in Supplemental Table 9.

Immunofluorescence and ISH. Lung sections were immunostained as previously described (16). Briefly, lungs were fixed with phosphate-buffered formalin, processed, embedded in paraffin, and cut into 4-μm sections. Lung sections were deparaffinized, hydrated, boiled in Target Retrieval Solution (Dako S1699), blocked in 5% goat or donkey serum in Tris-buffered saline with 0.05% Tween (TTBS), and incubated with antibodies against T1α (University of Iowa Developmental Studies Hybridoma Bank 8.1.1), GFP (Abcam ab13970), AQP5 (Millipore AB15858), and/or proSPC (Millipore AB3786) overnight at 4°C. Secondary antibodies, including Alexa Fluor 488–conjugated anti-rabbit (Invitrogen A11034), Cy3-conjugated anti-chicken (Jackson ImmunoResearch Laboratories 703-165-155), and Alexa Fluor 647–conjugated anti-hamster (Jackson ImmunoResearch Laboratories 107-605-142), were applied for 1 hour at room temperature. In situ hybridization (ISH) was performed using the RNAScope manual fluorescent multiplex kit (ACDbio) according to the manufacturer's instructions. The mouse *Tgfb2* probe was generated by ACDbio targeting 1491–2468 of the *Tgfb2* gene (accession number NM_009367.3, ACDbio 406181). The mouse *Itgb6* probe was generated by ACDbio targeting 533–1496 of the *Itgb6* gene (accession NM_001159564.1, ACDbio catalog 312501). The mouse *p15* probe was generated by ACDbio targeting 119–1302 of the *p15* gene (accession NM_007670, ACDbio 458341-C2). At the end of the RNAScope protocol, slides were

stained with the anti-GFP antibody as described above. Images were acquired using the ×20 air or ×63 oil objective of an Axiovert 200 microscope (Zeiss) or an A1 confocal microscope (Nikon). The percentage of GFP⁺ cells that stained positive for the *p15* RNA probe was counted manually. The percentage of GFP⁺ cells that stained positive for at least 3 *Tgfb2* or *Itgb6* RNA molecules was counted manually. The percentage of proSPC⁺ cells that were GFP⁺ was quantified using NIH ImageJ software.

Bioinformatics. FASTQs were processed using Cell Ranger (version 2.1.1) with the mm10 genome assembly to generate unique molecular identifier (UMI) gene count matrices per sample. An aggregated matrix was also generated by downsampling of the mapped reads in each sample to the same depth as the sample with the lowest read count using cellranger aggr. The aggregated matrix was next processed using Seurat (version 2.3.0) to perform initial quality control filtering, normalization, and clustering (60). Cells were removed if there were fewer than 100 genes detected, a UMI count less than 1024, or greater than 3 standard deviations above the mean of each sample, or if the proportion of UMIs mapped to mitochondrial genes was greater than 5%. Genes were excluded if they were detectable in fewer than 3 cells. Following filtering, the UMI counts were normalized to library size (total number of UMIs detected), scaled by 10,000, and log-transformed. Principal component analysis was performed on the *Z* scores of the normalized expression values, and 16 dimensions were selected for tSNE projection using a perplexity of 30. Graph-based clustering was performed using the top 15 principal components, with the 30 nearest neighbors, and a resolution of 0.6 or other resolution. Following clustering and tSNE projection, the original non-downsampled UMI matrices for each sample were normalized as described above, and the resulting gene expression values were used for all downstream analyses. Reads that did not align to the mouse genome were aligned to the eGFP coding sequence (GenBank U55761.1) using bowtie2 (version 2.1.0) in very-sensitive-local mode (61). UMIs were counted using umi-tools (version 0.5.3) with the directional algorithm (62). Genes differentially expressed in each cluster compared with other clusters in each tested comparison were determined using FindAllMarkers() in Seurat. Genes were included if they were detectable in a minimum of 25% of cells in the cluster of interest, there was at least a 0.25-fold-change (natural log scale) difference in expression between the cluster of interest and the other clusters, and the *P* value was less than 0.01. Notably, because of the relatively low efficiency of current scRNAseq library preparation methods, many mRNAs, particularly those of low to moderate expression levels, are not detected in some cells in which they are present (63, 64). Heatmaps were generated using ComplexHeatmap. Genes were ordered using hierarchical clustering of euclidean distances using the complete-linkage method (65). Cells were ordered in pseudotime using Monocle2 with the DDRTree method for dimensionality reduction (version 2.10.0) (66). RNA velocity and partition-based graph abstraction were computed using the velocyto and scanpy Python packages (67, 68). Canonical correlation analysis was performed using the RunCCA and AlignSubspace Seurat commands. Dotplots were generated using the DotPlot function in the Seurat package. Hierarchical clustering of the average gene expression values of marker genes of nonepithelial lung cell types (69) was used to examine the relationships between clusters.

The transcription factor list was obtained from the TF-Checkpoint database (70). Cell surface markers were obtained from a database derived from mass spectrometry data (71). Additionally, the UniProt database was queried to extract putative cell surface proteins (locations:(location:"cell membrane") annotation:(type:transmem) AND organism:"Mus musculus (Mouse) [10090]"). Genes related to cell motility were those classified by the following Gene Ontology (GO) terms: GO:0048870 cell motility, GO:0016477 cell migration, GO:0005925 focal adhesion, GO:0030054 cell junction, and GO:0031589 cell substrate adhesion. The list of TGF-β pathway genes was obtained from Ingenuity Pathway Analysis. The list of mesenchymal markers were the top 80 markers differentially expressed by fibroblasts in normal mouse lung (69).

The list of differentially expressed genes in each cluster, containing gene identifiers, fold change, and adjusted *P* values, was uploaded into Ingenuity Pathway Analysis software (Qiagen). The "core analysis" function was used to find significant upstream transcriptional regulators. Significance was assessed by Fisher's exact test, and activation *Z* scores were computed. The list of differentially expressed genes was uploaded into iPathwayGuide software (Advaita Bioinformatics) and the GO biological processes enriched for the differentially expressed genes in each cluster were determined. To identify biological processes enriched in Injured AEC2s compared with Naive AEC2s, genes differentially expressed between Injured AEC2 and Naive AEC2 cells were uploaded to DAVID (https://david.ncifcrf.gov/).

Data/code availability. All data were deposited in the NCBI's Gene Expression Omnibus database (GEO GSE113049). Analysis and processing scripts are provided at a GitHub repository (https://github.com/rnabioco/lung-scrna).

AEC culture. AEC2s were isolated from mice and rats as previously described (72). Briefly, mouse lungs were perfused with PBS, instilled with dispase followed by agarose plug, then incubated in dispase for 30 minutes at room temperature. Tissue was minced and filtered. Single-cell suspension was subjected to magnetic cell sorting using negative selection with CD45 antibody and positive selection with EpCAM antibody. Rat lungs were perfused and instilled with fluorocarbon for 20 minutes at 37°C followed by elastase digestion for 15 minutes at 37°C. Tissue was minced and filtered, and cells were purified by density gradient centrifugation. The rat cell isolation yields 85%–93% purity and the mouse cell isolation yields 90%–92% purity as determined by staining cytospins for proSPC. Freshly isolated rat AEC2s were cultured in DMEM (Invitrogen) containing 44 mM NaHCO$_3$, 1 mM sodium pyruvate, 4 mM L-alanyl-glutamine, 90 μg/ml streptomycin, 40 μg/ml penicillin, and 10% FBS (HyClone) on tissue culture plastic or Transwell inserts (Millipore MCHT12H48) coated with 20% Matrigel (BD Biosciences 356234) and 80% rat tail collagen (Corning 354236) or glass coverslips. Cells were treated with 2.5 μM LY364947 (Cayman Chemicals 13341), 40 μg anti–TGF-β neutralizing antibody 1D11 (R&D MAB1835), or DMSO or mIgG1 (R&D MAB002) as controls. Freshly isolated mouse AEC2s were cultured on Transwell inserts (Corning 3401) coated with mouse laminin (Millipore Sigma CC095) in DMEM/F12 with 2% FBS, 90 μg/ml streptomycin, and 40 μg/ml penicillin. Cells were transduced with Ad5CMVeGFP (AdGFP) or Ad5CMVCre-eGFP (AdCre) (University of Iowa Viral Vector Core VVC-U of Iowa-4 or VVC-U of Iowa-1174), MOI 10, for 48 hours, after which cells were washed and cultured in DMEM/F12 with 10% FBS. For RNA extraction, cells were lysed in Buffer RLT and RNA was purified using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA was reverse-transcribed into cDNA using qScript XLT cDNA SuperMix (QuantaBio) or Quantitect (Qiagen) according to the manufacturer's instructions. cDNA was analyzed by quantitative PCR using primers listed in Supplemental Table 10. Quantitative PCR was performed for 40 cycles on the CFX96 (Bio-Rad) or StepOnePlus (Applied Biosystems) using PerfeCTa SYBR Green FastMix (QuantaBio) or iQ SYBR Green Supermix (Bio-Rad). Relative mRNA expression was calculated using the $2^{-\Delta\Delta Ct}$ method. For Figure 6D, the values listed for day 3 are from day 2 or 3. For genotyping, cells or tails were incubated in 25 mM NaOH, 0.2 mM disodium EDTA, pH 12.0, at 95°C for 30 minutes, followed by neutralization with 40 mM Tris, pH 5.0. PCR was performed on extracted genomic DNA using GoTaq DNA polymerase according to the manufacturer's instructions using the primers listed in Supplemental Table 10. For Western blotting, cells were lysed in 10 mM Tris, pH 8.0, 50 mM NaCl, 0.5% sodium deoxycholate, 0.2% SDS, and 1% IGEPAL, supplemented with Protease Inhibitor Cocktail (Thermo Fisher Scientific 1860932). Lysates were boiled in Laemmli buffer and analyzed by SDS-PAGE and immunoblotting with antibodies against p-Smad2/3 (Abcam ab51451), Smad3 (Abcam ab40854), T1α (gift of Mary Williams, Boston University, Boston, Massachusetts, USA), Hopx (Santa Cruz sc-30216), RAGE (Life Technologies PA1-84173), or actin (Abcam ab3280). Densitometry was performed using ImageLab (Bio-Rad) or NIH ImageJ software. For indicated experiments, 10 μM EdU (Thermo Fisher Scientific C10340) was added to the culture medium, and 24 hours later cells were fixed in 4% paraformaldehyde and stained according to the manufacturer's instructions. Data more than 2 SDs from the mean were excluded, a criterion that was established prior to experimentation.

Statistics. The scRNAseq experiment, with 1 control and 1 LPS-treated mouse, was performed twice. Genes differentially expressed between clusters were determined using the Wilcoxon rank sum test using expression values in each cluster compared with all other clusters. Differential gene expression is reported using adjusted *P* values with Bonferroni correction for multiple hypothesis testing. The lower limit of detection for the *P* values was $2.225074 \times 10^{-308}$. The overlap *P* values were calculated using Fisher's exact test. The *P* values for GO biological processes in iPathwayGuide were corrected for Elim pruning. For scRNAseq differential gene expression experiments, $P \leq 0.01$ was considered significant. For in vitro experiments, 2-tailed *t* test or 1- or 2-way ANOVA with post hoc analysis for multiple comparisons was performed and corrected for repeated measures. *P* less than 0.05 was considered significant. Mean ± SEM is shown in graphs.

Study approval. Animal protocols were approved by the Animal Care and Use Committees at National Jewish Health and the University of Michigan.

Author contributions

RLZ, KAR, ANG, and JRH conceived and designed the study. KAR, NLJ, PJ, EFR, AEG, RF, AJM, JRS, and RLZ performed data analysis and acquisition. KAR, ANG, JRH, and RLZ interpreted data. RLZ drafted and revised the manuscript. All authors approved the final manuscript.

Acknowledgments

Address correspondence to: Rachel L. Zemans, Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, University of Michigan, 4062 BSRB/SPC 2200, 109 Zina Pitcher Place, Ann Arbor, Michigan 48109-2200, USA. Phone: 734.936.9372; Email: zemansr@med.umich.edu.

1. Barkauskas CE, et al. Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest*. 2013;123(7):3025–3036.
2. Jain R, et al. Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. *Nat Commun*. 2015;6:6727.
3. Rock JR, et al. Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proc Natl Acad Sci U S A*. 2011;108(52):E1475–E1483.
4. Vaughan AE, et al. Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. *Nature*. 2015;517(7536):621–625.
5. Kim CF, et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell*. 2005;121(6):823–835.
6. Ray S, et al. Rare SOX2[+] airway progenitor cells generate KRT5[+] cells that repopulate damaged alveolar parenchyma following influenza virus infection. *Stem Cell Reports*. 2016;7(5):817–825.
7. Evans MJ, Cabral LJ, Stephens RJ, Freeman G. Renewal of alveolar epithelium in the rat following exposure to NO2. *Am J Pathol*. 1973;70(2):175–198.
8. Adamson IY, Bowden DH. The type 2 cell as progenitor of alveolar epithelial regeneration. A cytodynamic study in mice after exposure to oxygen. *Lab Invest*. 1974;30(1):35–42.
9. Desai TJ, Brownfield DG, Krasnow MA. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature*. 2014;507(7491):190–194.
10. Liu Y, et al. FoxM1 mediates the progenitor function of type II epithelial cells in repairing alveolar injury induced by Pseudomonas aeruginosa. *J Exp Med*. 2011;208(7):1473–1484.
11. Dial CF, Tune MK, Doerschuk CM, Mock JR. Foxp3[+] regulatory T cell expression of keratinocyte growth factor enhances lung epithelial proliferation. *Am J Respir Cell Mol Biol*. 2017;57(2):162–173.
12. Rafii S, et al. Platelet-derived SDF-1 primes the pulmonary capillary vascular niche to drive lung alveolar regeneration. *Nat Cell Biol*. 2015;17(2):123–136.
13. Zemans RL, et al. Neutrophil transmigration triggers repair of the lung epithelium via β-catenin signaling. *Proc Natl Acad Sci U S A*. 2011;108(38):15990–15995.
14. Zacharias WJ, et al. Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature*. 2018;555(7695):251–255.
15. 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.
16. Jansing NL, McClendon J, Henson PM, Tuder RM, Hyde DM, Zemans RL. Unbiased quantitation of alveolar type II to alveolar type I cell transdifferentiation during repair after lung injury in mice. *Am J Respir Cell Mol Biol*. 2017;57(5):519–526.
17. Jansing NL, et al. Flow cytometry underestimates and planimetry overestimates alveolar epithelial type 2 cell expansion after lung injury. *Am J Respir Crit Care Med*. 2018;198(3):390–392.
18. Treutlein B, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*. 2014;509(7500):371–375.
19. Barkauskas CE, Noble PW. Cellular mechanisms of tissue fibrosis. 7. New insights into the cellular mechanisms of pulmonary fibrosis. *Am J Physiol Cell Physiol*. 2014;306(11):C987–C996.
20. Massagué J, Wotton D. Transcriptional control by the TGF-beta/Smad signaling system. *EMBO J*. 2000;19(8):1745–1754.
21. Massagué J, Blain SW, Lo RS. TGFβ signaling in growth control, cancer, and heritable disorders. *Cell*. 2000;103(2):295–309.
22. Massagué J. TGFβ in cancer. *Cell*. 2008;134(2):215–230.
23. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature*. 1994;371(6494):257–261.
24. Ashcroft GS, et al. Mice lacking Smad3 show accelerated wound healing and an impaired local inflammatory response. *Nat Cell Biol*. 1999;1(5):260–266.
25. Khalil N, O'Connor RN, Flanders KC, Shing W, Whitman CI. Regulation of type II alveolar epithelial cell proliferation by TGF-beta during bleomycin-induced lung injury in rats. *Am J Physiol*. 1994;267(5 pt 1):L498–L507.
26. Buckley S, Bui KC, Hussain M, Warburton D. Dynamics of TGF-beta 3 peptide activity during rat alveolar epithelial cell proliferative recovery from acute hyperoxia. *Am J Physiol*. 1996;271(1 pt 1):L54–L60.
27. Chen H, et al. TGF-beta receptor II in epithelia versus mesenchyme plays distinct roles in the developing lung. *Eur Respir J*. 2008;32(2):285–295.
28. Wang Y, et al. HDAC3-dependent epigenetic pathway controls lung alveolar epithelial cell remodeling and spreading via miR-17-92 and TGF-β signaling regulation. *Dev Cell*. 2016;36(3):303–315.
29. Li M, et al. Epithelium-specific deletion of TGF-β receptor type II protects mice from bleomycin-induced pulmonary fibrosis.

J Clin Invest. 2011;121(1):277–287.

30. Bhaskaran M, Kolliputi N, Wang Y, Gou D, Chintagari NR, Liu L. Trans-differentiation of alveolar epithelial type II cells to type I cells involves autocrine signaling by transforming growth factor β 1 through the Smad pathway. *J Biol Chem*. 2007;282(6):3968–3976.

31. Zhao L, Yee M, O'Reilly MA. Transdifferentiation of alveolar epithelial type II to type I cells is controlled by opposing TGF-β and BMP signaling. *Am J Physiol Lung Cell Mol Physiol*. 2013;305(6):L409–L418.

32. Chung MI, Bujnis M, Barkauskas CE, Kobayashi Y, Hogan BLM. Niche-mediated BMP/SMAD signaling regulates lung alveolar stem cell proliferation and differentiation. *Development*. 2018;145(9):dev163014.

33. Harper JV, Brooks G. The mammalian cell cycle: an overview. *Methods Mol Biol*. 2005;296:113–153.

34. Wang Y, et al. Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. *Proc Natl Acad Sci U S A*. 2018;115(10):2407–2412.

35. Cakarova L, et al. Macrophage tumor necrosis factor-alpha induces epithelial expression of granulocyte-macrophage colony-stimulating factor: impact on alveolar epithelial repair. *Am J Respir Crit Care Med*. 2009;180(6):521–532.

36. Huffman JA, Hull WM, Dranoff G, Mulligan RC, Whitsett JA. Pulmonary epithelial cell expression of GM-CSF corrects the alveolar proteinosis in GM-CSF-deficient mice. *J Clin Invest*. 1996;97(3):649–655.

37. Munger JS, et al. The integrin αvβ6 binds and activates latent TGF-β1: a mechanism for regulating pulmonary inflammation and fibrosis. *Cell*. 1999;96(3):319–328.

38. Gonzalez R, Yang YH, Griffin C, Allen L, Tigue Z, Dobbs L. Freshly isolated rat alveolar type I cells, type II cells, and cultured type II cells have distinct molecular phenotypes. *Am J Physiol Lung Cell Mol Physiol*. 2005;288(1):L179–L189.

39. Danto SI, Shannon JM, Borok Z, Zabski SM, Crandall ED. Reversible transdifferentiation of alveolar epithelial cells. *Am J Respir Cell Mol Biol*. 1995;12(5):497–502.

40. Marconett CN, et al. Cross-species transcriptome profiling identifies new alveolar epithelial type I cell-specific genes. *Am J Respir Cell Mol Biol*. 2017;56(3):310–321.

41. Correll KA, Edeen KE, Zemans RL, Redente EF, Mikels-Vigdal A, Mason RJ. TGF beta inhibits expression of SP-A, SP-B, SP-C, but not SP-D in human alveolar type II cells. *Biochem Biophys Res Commun*. 2018;499(4):843–848.

42. Beers MF, et al. TGF-β1 inhibits surfactant component expression and epithelial cell maturation in cultured human fetal lung. *Am J Physiol*. 1998;275(5):L950–L960.

43. Whitsett JA, Budden A, Hull WM, Clark JC, O'Reilly MA. Transforming growth factor-β inhibits surfactant protein A expression in vitro. *Biochim Biophys Acta*. 1992;1123(3):257–262.

44. Whitsett JA, Weaver TE, Lieberman MA, Clark JC, Daugherty C. Differential effects of epidermal growth factor and transforming growth factor-β on synthesis of Mr = 35,000 surfactant-associated protein in fetal lung. *J Biol Chem*. 1987;262(16):7908–7913.

45. Frank DB, et al. Emergence of a wave of Wnt signaling that regulates lung alveologenesis by controlling epithelial self-renewal and differentiation. *Cell Rep*. 2016;17(9):2312–2325.

46. Zhang Z, et al. Transcription factor Etv5 is essential for the maintenance of alveolar type II cells. *Proc Natl Acad Sci U S A*. 2017;114(15):3903–3908.

47. Rock JR, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc Natl Acad Sci U S A*. 2009;106(31):12771–12775.

48. Flodby P, et al. Directed expression of Cre in alveolar epithelial type 1 cells. *Am J Respir Cell Mol Biol*. 2010;43(2):173–178.

49. Krane CM, et al. Aquaporin 5-deficient mouse lungs are hyperresponsive to cholinergic stimulation. *Proc Natl Acad Sci U S A*. 2001;98(24):14114–14119.

50. O'Reilly MA. Giving new identities to alveolar epithelial type I cells. *Am J Respir Cell Mol Biol*. 2017;56(3):277–278.

51. Yang J, et al. The development and plasticity of alveolar type 1 cells. *Development*. 2016;143(1):54–65.

52. Gonzalez RF, Dobbs LG. Isolation and culture of alveolar epithelial Type I and Type II cells from rat lungs. *Methods Mol Biol*. 2013;945:145–159.

53. Rittirsch D, et al. Acute lung injury induced by lipopolysaccharide is independent of complement activation. *J Immunol*. 2008;180(11):7664–7672.

54. Grommes J, Soehnlein O. Contribution of neutrophils to acute lung injury. *Mol Med*. 2011;17(3-4):293–307.

55. Pittet JF, et al. TGF-β is a critical mediator of acute lung injury. *J Clin Invest*. 2001;107(12):1537–1544.

56. Evans MJ, Cabral LJ, Stephens RJ, Freeman G. Transformation of alveolar type 2 cells to type 1 cells following exposure to NO2. *Exp Mol Pathol*. 1975;22(1):142–150.

57. Katzenstein AL, Myers JL. Idiopathic pulmonary fibrosis: clinical relevance of pathologic classification. *Am J Respir Crit Care Med*. 1998;157(4 pt 1):1301–1315.

58. Kasper M, Barth K. Potential contribution of alveolar epithelial type I cells to pulmonary fibrosis. *Biosci Rep*. 2017;37(6):BSR20171301.

59. 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.

60. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–420.

61. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–359.

62. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res*. 2017;27(3):491–499.

63. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nat Methods*. 2014;11(7):740–742.

64. Zheng GX, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8:14049.

65. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849.

66. Trapnell C, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381–386.

67. La Manno G, et al. RNA velocity of single cells. *Nature*. 2018;560(7719):494–498.

68. Wolf FA, et al. Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *bioRxiv* website. https://doi.org/10.1101/208819. Posted October 25, 2017. Accessed March 12, 2019.

69. Reyfman PA, et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis [published online ahead of print December 15, 2018]. *Am J Respir Crit Care Med*. https://doi.org/10.1164/rccm.201712-2410OC.

70. Chawla K, Tripathi S, Thommesen L, Lægreid A, Kuiper M. TFcheckpoint: a curated compendium of specific DNA-binding RNA polymerase II transcription factors. *Bioinformatics*. 2013;29(19):2519–2520.

71. Bausch-Fluck D, et al. A mass spectrometric-derived cell surface protein atlas. *PLoS ONE*. 2015;10(3):e0121314.

72. Jansing NL, et al. Isolation of rat and mouse alveolar type II epithelial cells. *Methods Mol Biol*. 2018;1809:69–82.