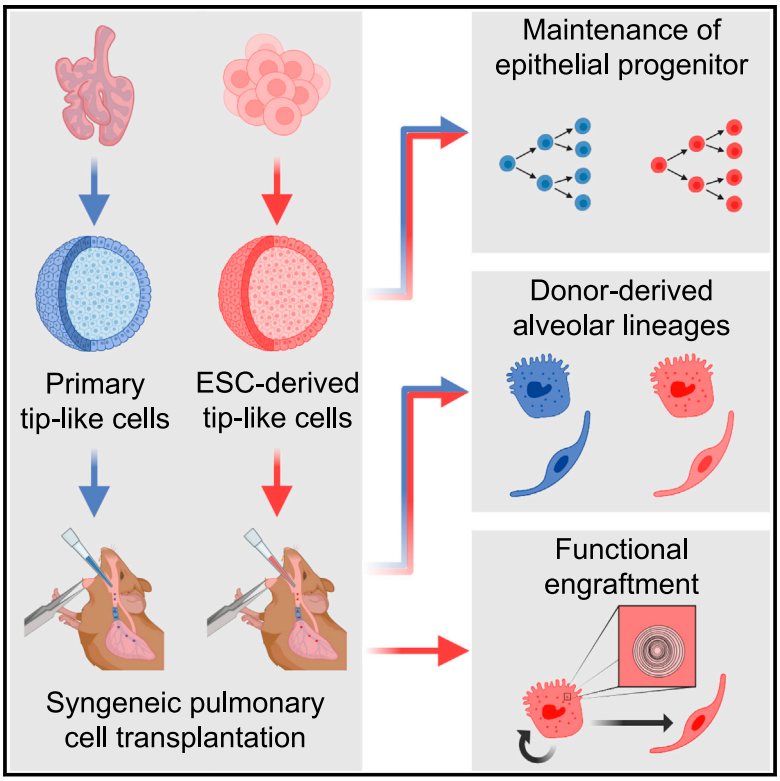
Cell Stem Cell

Article

Durable alveolar engraftment of PSC-derived lung epithelial cells into immunocompetent mice

Graphical abstract



Authors

Michael J. Herriges, Maria Yampolskaya, Bibek R. Thapa, ..., Carlos Villacorta-Martin, Pankaj Mehta, Darrell N. Kotton

Correspondence

dkotton@bu.edu

In brief

Kotton and colleagues differentiate embryonic stem cells (ESCs) into lung epithelial progenitors that can be transplanted into immunocompetent mouse lungs. These cells durably engraft into the recipient's lung, giving rise to mature alveolar type 1 (AT1)-like and AT2-like cells that are transcriptomically and functionally similar to endogenous lineages.

Highlights

- Directed differentiation of mouse ESCs into tip-like lung epithelial progenitors
- Primary and ESC-derived transplants produce similar AT1-like and AT2-like cells
- ESC-derived tip-like cells persist for at least 6 months in immunocompetent mice
- Donor-derived cells have the lamellar bodies and progenitor capacity of AT2 cells

Herriges et al., 2023, Cell Stem Cell *30*, 1217–1234
September 7, 2023
https://doi.org/10.1016/j.stem.2023.07.016

CellPress

Cell Stem Cell

CellPress
OPEN ACCESS

Article

Durable alveolar engraftment of PSC-derived lung epithelial cells into immunocompetent mice

Michael J. Herriges,[1,2] Maria Yampolskaya,[3] Bibek R. Thapa,[1,2] Jonathan Lindstrom-Vautrin,[1] Feiya Wang,[1] Jessie Huang,[1,2] Cheng-Lun Na,[4] Liang Ma,[1,2] McKenna M. Montminy,[1,2] Pushpinder Bawa,[1] Carlos Villacorta-Martin,[1] Pankaj Mehta,[3] and Darrell N. Kotton[1,2,5,*]

[1]Center for Regenerative Medicine, Boston University and Boston Medical Center, Boston, MA 02118, USA
[2]The Pulmonary Center and Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA
[3]Department of Physics, Boston University, Boston, MA 02215, USA
[4]Division of Pulmonary Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA
[5]Lead contact
*Correspondence: dkotton@bu.edu
https://doi.org/10.1016/j.stem.2023.07.016

SUMMARY

Durable reconstitution of the distal lung epithelium with pluripotent stem cell (PSC) derivatives, if realized, would represent a promising therapy for diseases that result from alveolar damage. Here, we differentiate murine PSCs into self-renewing lung epithelial progenitors able to engraft into the injured distal lung epithelium of immunocompetent, syngeneic mouse recipients. After transplantation, these progenitors mature in the distal lung, assuming the molecular phenotypes of alveolar type 2 (AT2) and type 1 (AT1) cells. After months *in vivo*, donor-derived cells retain their mature phenotypes, as characterized by single-cell RNA sequencing (scRNA-seq), histologic profiling, and functional assessment that demonstrates continued capacity of the engrafted cells to proliferate and differentiate. These results indicate durable reconstitution of the distal lung's facultative progenitor and differentiated epithelial cell compartments with PSC-derived cells, thus establishing a novel model for pulmonary cell therapy that can be utilized to better understand the mechanisms and utility of engraftment.

INTRODUCTION

Acute injuries to the distal lung epithelium, such as those resulting from COVID-19, and chronic lung diseases, such as pulmonary fibrosis or emphysema, represent leading causes of morbidity and mortality worldwide. Common to each of these illnesses is progressive destruction of the distal lung epithelium that can lead to a lethal reduction in respiratory function. Although lung transplants can be used to alleviate symptoms, this solution is severely limited by the insufficient supply of donor lungs and the continual risk of immune rejection of donor tissue despite life-long immunosuppressing drug regimens. One conceivable alternative to full organ transplantation is reconstitution of the injured epithelium through cell therapy, in which donor cells are engrafted directly into a patient to functionally replace lost endogenous cells. Although cell therapy has successfully been used to replace multiple cell types in patients,[1–7] lung epithelial reconstitution in humans has not yet been accomplished. Recent work has shown that cell transplantation is possible in injured mouse lungs with donor-derived cells surviving *in vivo* and expressing markers of mature epithelial lineages.[8–20] However, most of these studies followed the surviving cells for only brief periods and utilized either primary lung epithelial cells or immunocompromised recipients, limiting their potential for clinical application as a cell-based treatment. Furthermore, in many of these studies, it is still unclear, with two notable exceptions,[18,20] how donor-derived cells compare with endogenous cells on a wider transcriptional or functional level, which is a critical step toward developing truly therapeutic cell engraftment.

Within the hematopoietic system, similar clinical hurdles and biological questions were iteratively solved through mouse models of blood repopulation based on transplantation of mouse hematopoietic progenitor cells into immunocompetent syngeneic recipients, leading to human bone marrow transplant and peripheral blood stem cell transplant therapies that are now standard of care for a variety of blood diseases worldwide.[21,22] Development of a similar syngeneic murine transplantation assay for the lung epithelium has the potential to provide insight into the treatment and regeneration of this organ and can inform future work in human pre-clinical studies. However, a clinically relevant source of engraftable progenitors for the distal alveolar lung epithelium is not readily apparent since alveolar type 2 (AT2) cells, the endogenous progenitors of this tissue, are difficult to access and are not easily expanded *in vitro* for autologous therapy.[23,24] Cultured primary cells isolated from the murine distal fetal lung bud tip (hereafter referred to as primary tip-like cells) have been transplanted previously into alveoli,[10] making these

progenitors a compelling source of donor cells, but a similar human population of autologous embryonic tip cells would be difficult to acquire. Pluripotent stem cell (PSC)-derived cells represent a promising population for syngeneic transplantation, since they provide solutions to these hurdles. Using well-established protocols, mouse- or human-induced PSCs (iPSCs) can be generated from any individual without the invasive procedures needed to collect primary distal lung progenitors.[25–27] In the case of patients with genetic disorders, CRISPR gene editing can then be used to reverse disease-causing mutations, creating a gene-corrected, syngeneic, freezable, and expandable population of cells for generating differentiated donor cells of pulmonary lineages.[24]

Here, we present the derivation and *in vivo* engraftment of mouse PSC-derived alveolar epithelial progenitors that can durably reconstitute the injured distal lung epithelium of immunocompetent, syngeneic recipient mice. We first develop a protocol for the directed differentiation of PSCs into distal lung epithelial progenitors that are transcriptionally similar to transplantable cultured primary tip-like cells.[10] When transplanted into bleomycin-injured lungs, these PSC-derived cells integrate into the endogenous alveolar epithelium, reconstituting the desired facultative progenitor function to produce AT2-like and alveolar type 1 (AT1)-like cells. Importantly, these donor-derived cells can persist for at least 6 months in an immunocompetent host and feature functional AT2-specific organelles, such as lamellar bodies. These results demonstrate successful engraftment of PSC-derived cells into an immunocompetent host and provide an important guidepost for developing clinically relevant PSC-derived pulmonary cell therapy without the need for immunosuppression.

RESULTS

Lung epithelial specification

In order to generate PSC-derived tip-like cells for transplantation, we sought to emulate pulmonary development in which anterior foregut endoderm is specified into early NKX2-1+ primordial lung progenitors,[28] which later give rise to fetal distal lung bud tip cells, the developmental precursors of bronchial and alveolar epithelia.[29] Since NKX2-1 is expressed by all known lung epithelia,[28] we used a mouse embryonic stem cell (ESC) line with an mCherry reporter targeted to the 3′ untranslated region (UTR) of the endogenous *Nkx2-1* locus (hereafter NKX2-1^{mCherry}) to track, quantify, and purify putative ESC-derived lung epithelial cells.[30,31] To provide an initial basis for our lung specification protocol, we used our previously published approach[28] stimulating ESC-derived foregut cells with WNT3a and BMP4 to induce NKX2-1^{mCherry} expression (WB protocol, Figure 1A). Recent single-cell RNA sequencing (scRNA-seq) of murine lung specification *in vivo* has not only verified that WNT and BMP pathways are active throughout lung lineage specification but also revealed a switch from retinoic acid (RA) to Fgf signaling soon after *Nkx2-1* is first expressed.[32] To recapitulate this switch *in vitro*, we added supplemental RA to our specification media until day 8, when NKX2-1^{mCherry} is still barely detectable (Figure 1B) and then added rmFGF10 after day 8 (WBRF protocol, Figure 1A). This protocol resulted in a significant increase in the frequency of NKX2-1^{mCherry}+ epithelial cells by day 14 (53.95% ± 3.84%) compared with either the original WB protocol[28] or the addition of either supplement alone (Figures 1B–1D). WBRF increased both the percent of epithelial cells that were NKX2-1^{mCherry}+ and the overall yield of NKX2-1^{mCherry}+/EpCAM+ double-positive putative lung epithelial progenitors (Figure 1E).

Consistent with their putative primordial progenitor state,[28] sorted Nkx2-1^{mCherry}+/EpCAM+ cells from the WB and WBRF protocols expressed similar levels of early lung epithelial marker transcripts (*Nkx2-1* and *Foxp2*) and progenitor markers (*Sox9*, *Id2*, and *Sox2*) (reverse transcriptase quantitative PCR [RT-qPCR]; Figure S1A). Distal differentiation markers (*Sftpc* and *Etv5*) were expressed at similar but low levels in cells from each protocol, emphasizing their primordial state. Both protocols produced cells expressing only low levels of non-lung epithelial lineage markers, with WB inducing low but detectable expression of the thyroid marker *Pax8* (CT>30) and WBRF-inducing expression of liver markers (*Afp* and *Alb*). Finally, low-level expression of basal cell markers (*Trp63* and *Krt5*) was present in cells grown in WB, but not those grown in WBRF. This suggests that WBRF results in reduced proximal airway fate capacity, consistent with the known distalizing role of FGF10 during lung development.[33] To test the lung differentiation competence of each specified population, we plated day 13 NKX2-1^{mCherry}+/EpCAM+ double-positive cells from WB and WBRF protocols in our previously published culture conditions that promote expression of distal alveolar or proximal airway epithelial lineage markers.[34,35] Cells from either specification protocol maintained a high percentage of NKX2-1^{mCherry}+ cells in distalizing conditions and expressed similar levels of distal markers (*Sox9*, *Ager*, and *Sftpc*) (Figures S1B and S1C). Alternatively, proximalizing conditions induced expression of proximal markers (*Sox2*, *Trp63*, *Krt5*, and *Scgb3a2*) in both cell populations, but as expected, WBRF-specified cells gave rise to fewer NKX2-1^{mCherry}+ cells with lower expression of these markers. Thus, although cells specified in either protocol are competent to upregulate both airway and alveolar markers, cells specified in WBRF have a reduced efficiency for proximal airway differentiation.

To better understand the heterogeneity of cells generated through the WBRF protocol, we profiled all live EpCAM+ cells on day 13 by scRNA-seq. Uniform manifold approximation and projection (UMAP) analysis revealed that the vast majority of cells localized to two *Nkx2-1*+ clusters, which were predominantly distinguished by expression of cell cycle genes (Figures 1F, 1G, and S1F). Cells in these two clusters also expressed other primordial lung epithelial-associated transcripts (*Cpm*, *Foxa2*, *Foxp2*, *Irx2*, and *Sox9*) but featured minimal expression of the proximal marker *Sox2* or more differentiated lineage markers (*Sftpc* and *Trp63*) (Figure 1H). In line with these results, immunohistochemistry revealed that day 13 epithelial cells express nearly ubiquitous NKX2-1, low but detectable levels of the distal progenitor marker SOX9, and no detectable proSFTPC (Figure S1E). In addition to the two lung lineage clusters, a minor third cell cluster (4.07% of cells) expressed liver markers (*Ttr* and *Afp*) and minimal *Nkx2-1*. This cluster likely represents the small percentage of NKX2-1^{mCherry}−/EpCAM+ non-lung endodermal cells expected in culture (Figures 1D and 1E).[28,36] Altogether these data indicate that the novel WBRF protocol can efficiently generate a population of early NKX2-1+ lung epithelial progenitors competent to subsequently differentiate toward airway or alveolar fates.

CellPress
OPEN ACCESS

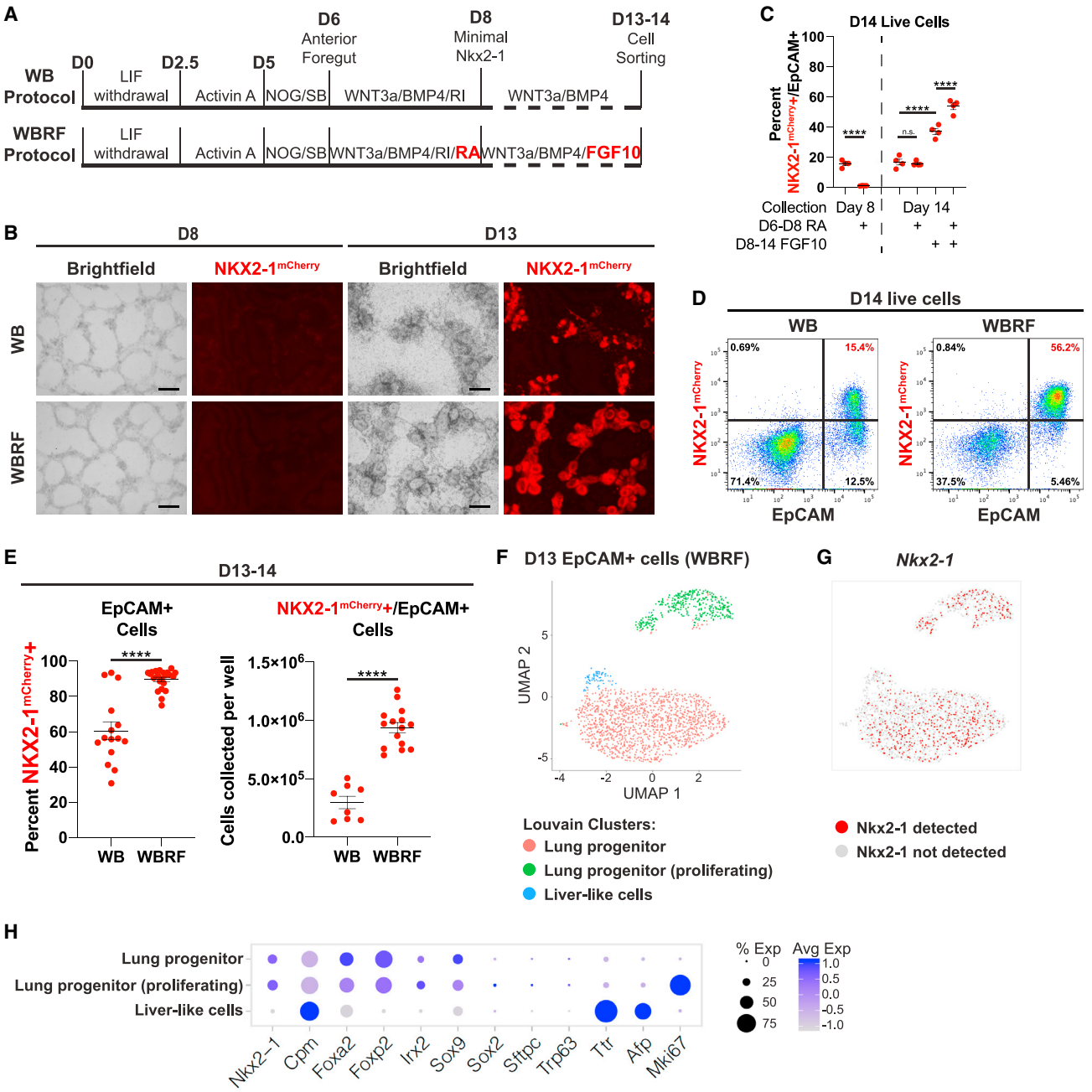


Figure 1. Modified lung specification protocol leads to increased yield of lung progenitors

(A) Schematic of WB and WBRF lung specification protocols for directed differentiation of mouse ESCs carrying an NKX2-1^{mCherry} reporter.

(B) Fluorescent microcopy images of days 8 and 13 of WB and WBRF lung specification protocols. Scale bars, 200 μm.

(C) Quantification of the percent of live cells that are NKX2-1^{mCherry}+/EpCAM+ double-positive lung epithelial progenitors on day 8 or day 14 depending on the inclusion of RA and/or FGF10. n.s., not significant, ****$p < 0.0001$ by one-way ANOVA. n = 4 biological replicates. Error bars = mean ± SEM.

(D) Fluorescence-activated cell sorting (FACS) plots on day 14 of WB and WBRF lung specification protocols.

(E) Quantification of the percent of epithelial cells that are NKX2-1^{mCherry}+ and the yield of NKX2-1^{mCherry}+/EpCAM+ cells per well of a 6-well plate on days 13 or 14 of lung specification protocols. ****$p < 0.0001$ by unpaired, two-tailed Student's t test. n = 15, 22, 8, 15 biological replicates. Error bars = mean ± SEM.

(F) UMAP plot for scRNA-seq of EpCAM+ sorted day 13 cells from WBRF lung specification protocol. Plot displays three clusters identified by Louvain clustering.

(G) UMAP plot displaying presence or absence of detectable *Nkx2-1* in these same day 13 cells.

(H) Expression of pulmonary and liver genes in annotated clusters of scRNA-seq dataset.

See also Figure S1 and Tables S1 and S2.

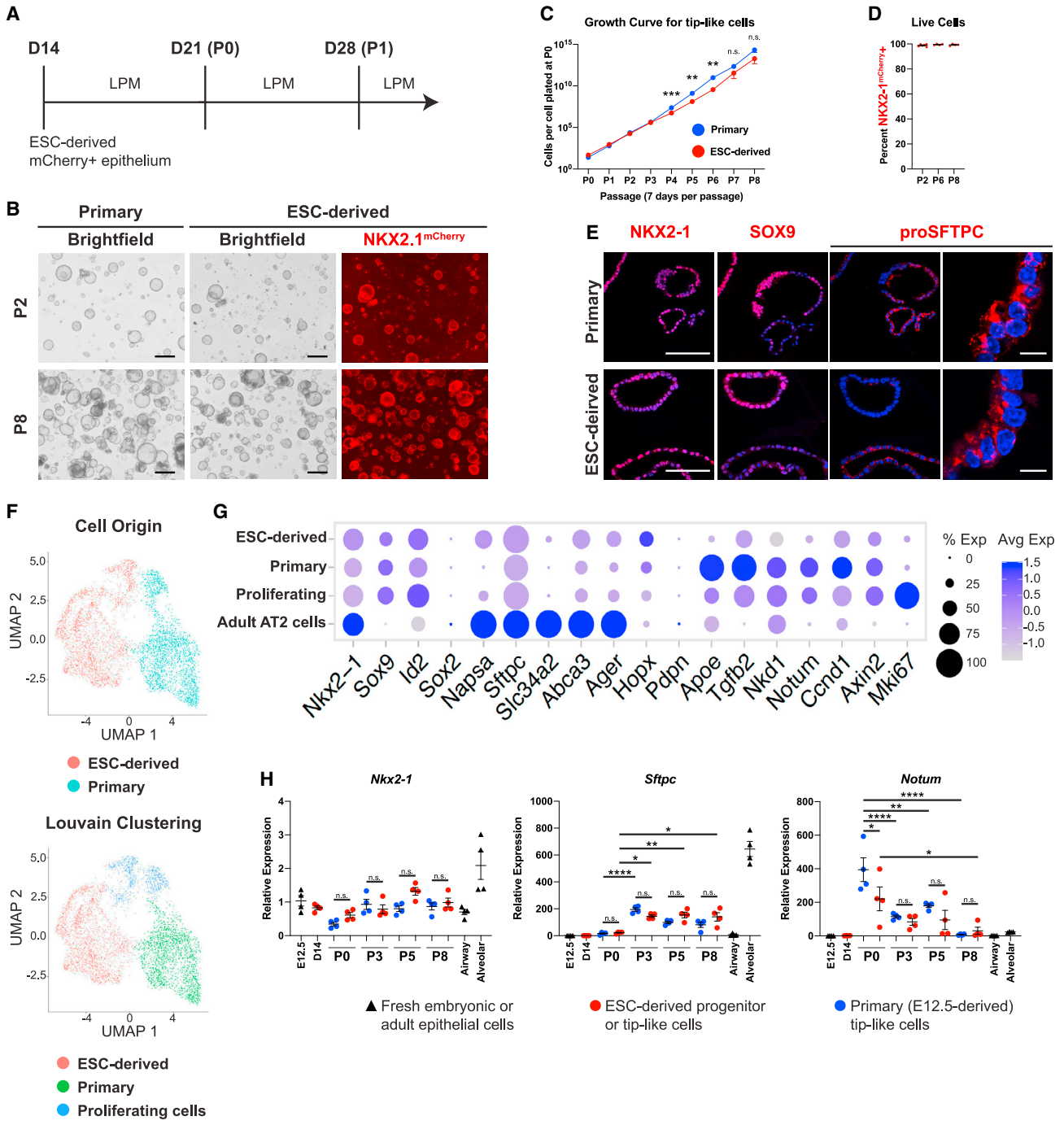


Figure 2. ESC-derived tip-like cells are morphologically and transcriptionally similar to cultured primary tip-like cells

(A) Schematic of differentiation and passaging of ESC-derived tip-like cells in lung progenitor media (LPM) following cell sorting of Nkx2-1^{mCherry}+ cells on day 14.

(B) Fluorescent microcopy images of primary and ESC-derived tip-like cells at P2 and P8. Scale bars, 500 μm.

(C) Quantification of cell proliferation for primary and ESC-derived tip-like cells across nine passages. n.s., not significant, **p < 0.01, ***p < 0.001 by unpaired, two-tailed Student's t test. n = 4 biological replicates. Error bars = mean ± SEM.

(D) Assessment of NKX2-1^{mCherry} expression throughout passaging of ESC-derived tip-like cells. n = 4 biological replicates. Error bars = mean ± SEM.

(E) Immunofluorescence microscopy for NKX2-1, SOX9, and proSFTPC in primary and ESC-derived tip-like cells. Nuclei stained with Hoechst, scale bars: 100 μm (first three columns) or 10 μm (rightmost column).

(F) UMAP plot for scRNA-seq of primary and ESC-derived tip-like cells. Top plot distinguishes cells by sample origin; bottom plot displays Louvain clusters.

(G) Expression of genes, including those identified as differently expressed between primary and ESC-derived tip-like cells. Cells from clusters annotated in (F) are compared against primary adult AT2 cells collected and sequenced at the same time.

(legend continued on next page)

Generation of ESC-derived tip-like progenitor cells

Having generated an early lung epithelial progenitor population, we next sought to differentiate these cells into tip-like cells for pulmonary cell transplantation. To do this, we plated day 14 NKX2-1^{mCherry}+/EpCAM+ double-positive cells in lung progenitor media (LPM) culture conditions, similar to those published for maintenance of the progenitor state of murine primary embryonic day 12.5 (E12.5) tip cells (Figure 2A).[10] In parallel, we generated primary control lines through culturing E12.5 lung epithelial cells from syngeneic mice (hereafter 129X1/S1) in identical LPM conditions. Both ESC-derived and primary cells grew out as hollow monolayered epithelial spheres that could be passaged multiple times with a stable karyotype and without losing their proliferative capacity (Figures 2B, 2C, and S1F). ESC-derived tip-like cells could be frozen down and thawed for later use, similar to primary tip-like cells.[10] Furthermore, passaged cells maintained their lung lineage identity, as indicated by retained expression of both the NKX2-1^{mCherry} reporter and NKX2-1 nuclear protein (Figures 2D and 2E). In both the primary and ESC-derived cells, we saw a mixture of SOX9^{High}/proSFTPC^{Low} and SOX9^{Low}/proSFTPC^{High} spheres, suggesting some heterogeneity in alveolar epithelial maturation in LPM conditions (Figure 2E).

We then performed scRNA-seq of primary and ESC-derived cells from parallel cultures in LPM conditions 7 days post-passaging to compare gene expression within these two populations. UMAP visualization of their global transcriptomes (Figure 2F) with Louvain clustering analysis indicated that the cells segregated partially based on their sample of origin, forming two major clusters, plus a third cluster of proliferating cells from both samples (Figure S1G). Although this third cluster accounted for only 16.9% of ESC-derived tip-like cells, EdU labeling across longer time periods indicated that the majority of ESC-derived tip-like cells are capable of proliferating (Figure S1H).

These datasets confirmed similar expression of tip cell markers (*Sox9* and *Id2*) in cells of all clusters, with relatively low expression of mature AT2 markers, suggesting that both samples contained primarily tip-like cells (Figure 2G). To further investigate the maturation state of these cells, we compared them with alveolar epithelial cells at different developmental time points.[37] By both spearman correlation analysis assessing the 1,000 most variable genes and hierarchical clustering based on markers of AT2 development, the cultured cells were most similar to E12.5 and E15.5 samples, suggesting an early embryonic tip-like identity (Figures S1I and S1J).[36] Analysis of differentially expressed genes (DEGs) between the primary and ESC-derived tip-like cells indicated that the ESC-derived cells had higher expression of genes associated with AT2 cells and surfactant metabolism (Figures S2A and S2B; Tables S1 and S2). However, maturation markers were still expressed at levels well below those of mature primary AT2 cells (Figure 2G). On the other hand, primary tip-like cells had higher expression of genes involved in WNT and transforming growth factor β (TGF-β) signaling (Figure S2).

To verify these results and screen for any potential drift in gene expression over serial passaging in LPM, we performed RT-qPCR on passage 0 (P0), P3, P5, and P8 of primary and ESC-derived tip-like cells and compared these samples against day 14 cells, freshly collected E12.5 tip cells, adult airway epithelial cells, and adult alveolar epithelial cells (Figures 2H, S2C, and S2D). Lung epithelial progenitor markers (*Nkx2-1*, *Sox9*, and *Sox2*) were expressed at similar levels in both primary and ESC-derived tip-like cells without alteration after passaging. There were no consistent significant gene expression differences between passage-matched primary and ESC-derived cells for any of the AT2 markers analyzed (*Sftpc*, *Abca3*, *Sftpb*, *Lamp3*, *Slc34a2*, and *Napsa*). Although *Sftpc* and *Abca3* expression did increase from P0 to P3, these increases did not differ based on cell of origin, and there were no further increases after subsequent passaging through P8. Similarly, we observed no consistent differential expression based on cell of origin for other selected genes by RT-qPCR (*Notum*, *Nkd1*, *Axin2*, *Tgfb2*, and *Apoe*), but these genes seemed to steadily decline over the first few passages. Altogether, this suggests that ESC-derived tip-like cells are transcriptionally similar to primary tip-like cells and maintain their progenitor profile even after expansion in cell culture over multiple passages.

Transplantation of primary tip-like cells into immunocompetent recipients

One of the ultimate goals of cell therapy is the transplantation of syngeneic cells, alleviating the need for immunosuppression. Although primary tip-like cells have been successfully transplanted into NOD-SCID *Il2rg*$^{-/-}$ (NSG) mice, it is still unclear whether these transplants can survive in an immunocompetent recipient.[10] To test this, we generated primary tip-like cells from UBC-green fluorescent protein (GFP) C57BL/6 mice with ubiquitous GFP expression, thus enabling tracking of transplanted cells. Syngeneic C57BL/6J recipient mice were given bleomycin intratracheally to injure endogenous alveolar epithelial cells (Figure S3A). 10 days later, 6e5 cultured primary tip-like cells were intratracheally instilled. GFP+ donor-derived cells were detected in recipient distal lungs at 9 weeks post-transplantation, indicating long-term survival of transplanted cells. These cells appeared in alveolar regions as cuboidal cells with punctate proSFTPC protein immunostaining, characteristic of AT2 cells, as well as thin cells expressing PDPN, characteristic of AT1 cells (Figure S3B white arrowheads and yellow arrows, respectively). Donor-derived cells were not found in the airway. These results suggest that primary tip-like cell transplants can survive and differentiate in syngeneic immunocompetent recipients, similar to published transplants into NSG recipients.[10]

ESC-derived tip-like cells give rise to persistent AT2- and AT1-like cells following transplantation

Given the transcriptional similarity between ESC-derived and primary tip-like cells, we next sought to determine whether

(H) Analysis of gene expression by RT-qPCR. Primary and ESC-derived tip-like cells from multiple passages are compared against lung epithelial progenitors from day 14 of the WBRF protocol (day 14) and freshly sorted lung epithelial cells from embryonic (E12.5) and adult (airway and alveolar) mouse lungs. Reference gating for primary controls can be found in Figures S2C and S2D. n.s., not significant, *p < 0.05, **p < 0.01, ****p < 0.0001 by one-way ANOVA. n = 4 biological replicates. Error bars = mean ± SEM.
See also Figures S1 and S2 and Tables S1 and S2.

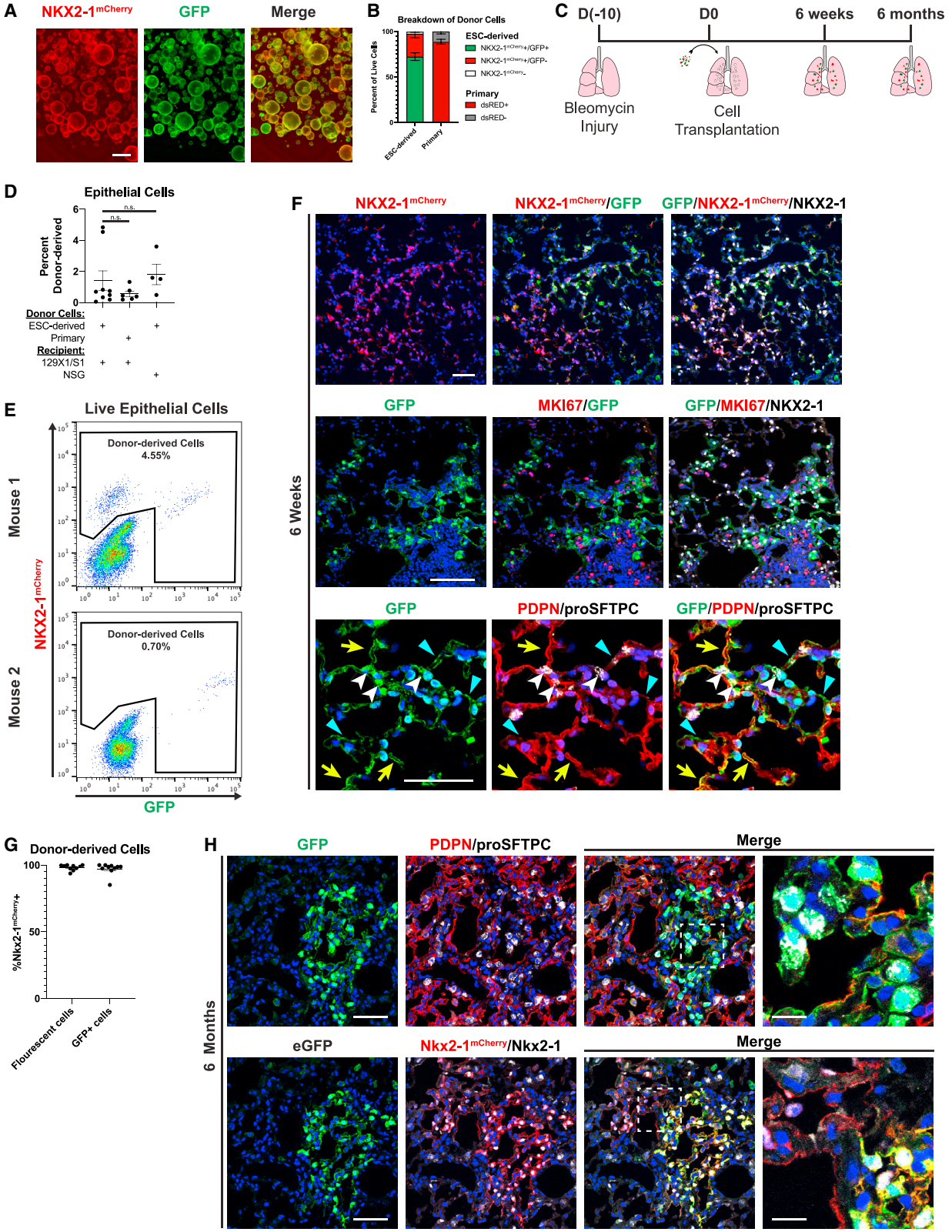


Figure 3. Transplanted ESC-derived tip-like cells give rise to AT2-like and AT1-like cells that persist for at least 6 months post-transplantation in immunocompetent mice

(A) Image of ESC-derived tip-like cells, carrying an NKX2-1^{mCherry} reporter (red), labeled with lentiviral GFP (green). Scale bars, 500 μm.

(legend continued on next page)

ESC-derived cells could also be transplanted into immunocompetent mouse lungs. ESC-derived tip-like cells were first labeled with lentiviral GFP for donor cell tracking (Figure 3A). Although these cells were sorted to enrich for GFP+ cells, not all cells maintained GFP expression, likely due to lentiviral silencing (Figure 3B). Syngeneic 129X1/S1 recipient mice were injured with bleomycin and received 5e5–7e5 donor cells, delivered intratracheally 10 days later (Figure 3C). At 3 days post-transplantation, we observed small, scattered clusters of cuboidal NKX2-1+ donor-derived cells (Figure S3C). At this time point, 52.6% ± 11.4% of NKX2-1+ donor-derived cells were MKI67+, relative to only 11.4% of neighboring endogenous NKX2-1+ cells (Figures S3D and S3E). Notably, these MKi67+ cells expressed SOX9 and proSFTPC, suggesting that they were still in a proliferative tip-like progenitor state (Figure S3F). By 2 weeks post-transplantation, these transplants gave rise to larger clusters of NKX2-1+ cells (Figure S4A). Although EdU labeling up to this time point confirmed proliferation of endogenous and donor-derived cells, few cells were SOX9+ or MKI67+ at 2 weeks (Figures S4B and S4C), suggesting the loss of tip-like progenitor identity and reduced proliferation. These donor-derived clusters instead contained both cuboidal cells with punctate proSFTPC and thin PDPN+ cells, suggesting differentiation into AT2-like and AT1-like cells, respectively (Figure S4D). Notably, a subset of donor-derived clusters featured thin AT1-like cells that were largely PDPN−, suggesting incomplete AT1 maturation (Figure S4D′). Similar to the primary cell transplants, donor-derived cells did not contribute to airway lineages at this or any other time point. Altogether, this suggests that within 2 weeks, transplanted cells quickly progress from proliferating tip-like progenitors to AT2-like and AT1-like cells.

To characterize the durability of donor-derived cells in immunocompetent recipients, we followed mice for longer periods post-transplantation of ESC-derived tip-like cells. By 6 weeks post-transplantation, donor-derived cells accounted for 1.4% of all live lung epithelial cells, similar to the results seen following transplantation of primary tip-like cells labeled with a lentiviral dsRed into syngeneic 129X1/S1 mice or transplantation of ESC-derived cells into immunocompromised NSG mice (Figures 3B–3E). Although transplantation efficiency was highly variable, increasing the number of donor cells significantly increased transplantation efficiency (Figure S4E). The vast majority of these donor-derived cells were NKX2-1^{mCherry}+/NKX2-1+/MKI67−, suggesting maintenance of a quiescent lung epithelial fate (Figure 3F). Flow cytometry confirmed that at best, a small fraction of donor-derived cells were GFP+/NKX2-1^{mCherry}−, and this population was not detectable in all transplant recipients, indicating that differentiation into non-lung lineages was rare (2.75% average and 1.2% median, Figures 3E, 3G, and S5A). Similar to 2 weeks post-transplantation, donor-derived cells included AT2-like cells, AT1-like cells, and thin PDPN− cells (Figure 3F white arrowheads, yellow arrows, and blue triangles, respectively). Finally, to assess the perdurance of transplanted cells in the presence of a functional immune system, we dissected mice at 6 months post-transplantation. Even at this late time point, we were able to find large clusters of donor-derived AT2-like and AT1-like cells (Figure 3H), suggesting long-term survival of donor-derived epithelial lineages. Together, these data suggest that ESC-derived cells transplanted into immunocompetent mice can differentiate into AT2-like and AT1-like cells and durably maintain these identities over time.

Profiling donor-derived and endogenous cells at single-cell resolution

To better characterize the fate of transplanted ESC-derived cells, we profiled recipient lungs by scRNA-seq 6 and 15 weeks post-transplantation. We first collected live epithelial cells (DRAQ7−/EpCAM+/CD45−/CD31−) from an uninjured control as well as the donor-derived (mCherry+ or GFP+) and endogenous (mCherry−/GFP−) epithelium from transplant recipients 6 and 15 weeks post-transplantation (Figures 4A and S5A–S5C). In order to determine whether donor-derived cells contributed to non-epithelial cell types, we collected all mCherry+ or GFP+ cells from our 15 weeks post-transplantation mouse. The resulting datasets were visualized with UMAP, and cells were clustered using the Louvain algorithm. Non-epithelial clusters were then identified based on expression of *Col1a2*, *Pecam1*, or *Ptprc* for mesenchymal, endothelial, or hematopoietic lineages, respectively. Importantly, out of the 2,092 non-epithelial cells characterized, only one cell expressed mCherry or GFP (Figure S5C). This cell expressed low levels of both AT2 and

(B) Percentage of NKX2-1^{mcherry}+/GFP+, NKX2-1^{mcherry}+/GFP−, and NKX2-1^{mcherry}− cells ESC-derived tip-like cells prior to transplantation (n = 5 distinct lines). Also shown is the average dsRed+ percentage of cultured primary cells similarly labeled with lentiviral dsRed (n = 3 technical replicates of the same line. Note: primary cells do not have an *Nkx2-1* reporter. Error bars = mean ± SEM.

(C) Schematic for transplantation of cells into bleomycin-injured immunocompetent lungs with subsequent histological or flow analysis.

(D) Flow cytometry quantitation of the percent of live epithelial (EpCAM+/CD45−/CD31−) cells that are donor-derived after transplantation of ESC-derived or primary tip-like cells based on flow analysis of whole recipient lungs. Recipient mice were either immunocompetent 129X1/S1 mice (6 weeks post-transplantation) or immunocompromised NSG mice (9 weeks post-transplantation). n.s., not significant by unpaired, two-tailed Student's t test. n = 9, 6, 4 biological replicates. Error bars = mean ± SEM.

(E) FACS plots identifying donor-derived cells within all live lung epithelial cells using mCherry and GFP expression. The two plots come from transplantations using different donor lines with different levels of GFP silencing. Few GFP+/mCherry− cells were detectable in these samples.

(F) Immunofluorescence confocal microscopy of lung tissue sections showing donor-derived cell clusters at 6 weeks post-transplantation, assessing markers of donor trackers, lung lineages, and proliferation. White arrowheads indicate cuboidal proSFTPC+/GFP+ cells, yellow arrows indicate thin PDPN+/GFP+ cells, and blue triangles indicate thin PDPN−/GFP+ cells. Nuclei stained with Hoechst, scale bars, 100 μm.

(G) The percent of all fluorescent (GFP+ or mCherry+) or GFP+ donor-derived cells that express NKX2-1^{mCherry} at 6 weeks post-transplantation as determined by flow cytometry (for lungs with donor-derived cells accounting for >0.5% of assessed epithelium). n = 10 biological replicates. Error bars = mean ± SEM.

(H) Histology of donor-derived cell clusters at 6 months post-transplantation. Nuclei stained with Hoechst, scale bars: 50 μm (leftmost panels) or 12.5 μm (rightmost panels). Lower panels indicate some mCherry+ cell clusters are GFP−, presumed due to lentiviral silencing before or after transplantation.

See also Figures S3 and S4.

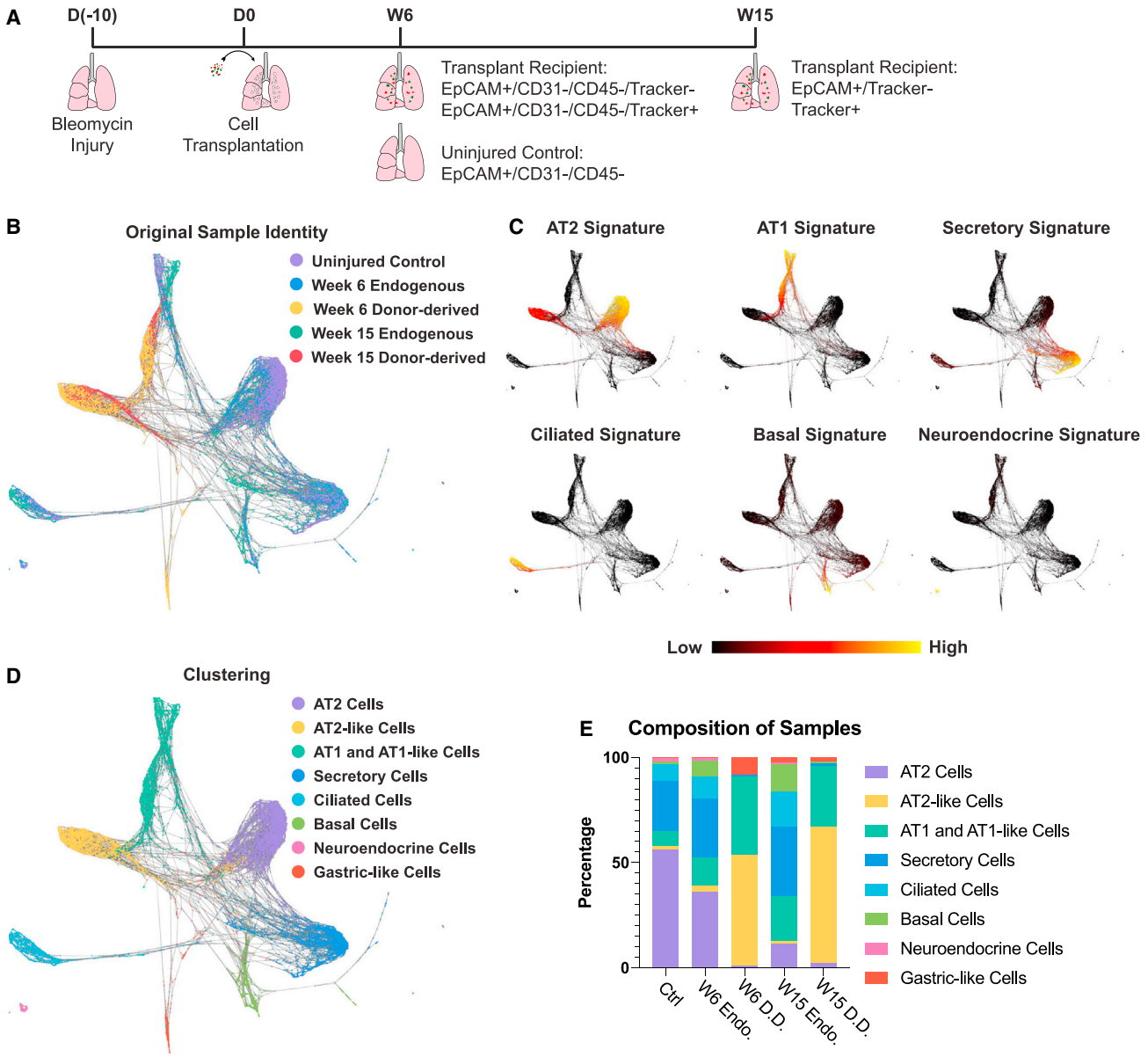


Figure 4. Single-cell transcriptomic profiling of donor-derived and endogenous lung epithelium

(A) Schematic for generation and collection of samples for scRNA-seq.

(B) SPRING plot of epithelial cells characterized by scRNA-seq labeled by sample origin.

(C) Expression of lung epithelial cell signatures. Gene sets comprising each signature can be found in Table S3.

(D) Cell-type annotation of clusters based on supervised Louvain clustering and expression of lung epithelial cell signatures.

(E) Composition of each sample based on clusters identified in (D).

See also Figure S5 and Table S3.

macrophage markers, indicating that it was likely a donor-derived AT2-like cell being phagocytosed by a macrophage. Altogether, this suggests that donor-derived cells primarily give rise to epithelial lineages.

In order to coalesce the datasets generated from our two transplant recipients, the epithelial cells from each dataset were combined using harmonization prior to plotting with UMAP (Figure S5D) or combined without harmonization and plotted on SPRING (Figure 4B).[38,39] The harmonized UMAP dataset was used to divide the cells based on Louvain clustering with overlapping or highly similar clusters being combined manually. The clusters were annotated based on expression of lung epithelial cell-type gene signatures and included all major lung epithelial cell types (Figures 4C, 4D, and S5E; Table S3). Donor-derived cells were predominantly found in three clusters (Figure 4E). Consistent with immunostaining results (Figure 3F), the donor-derived samples were mCherry+/GFP+ and did not express any proliferation markers, suggesting that they had

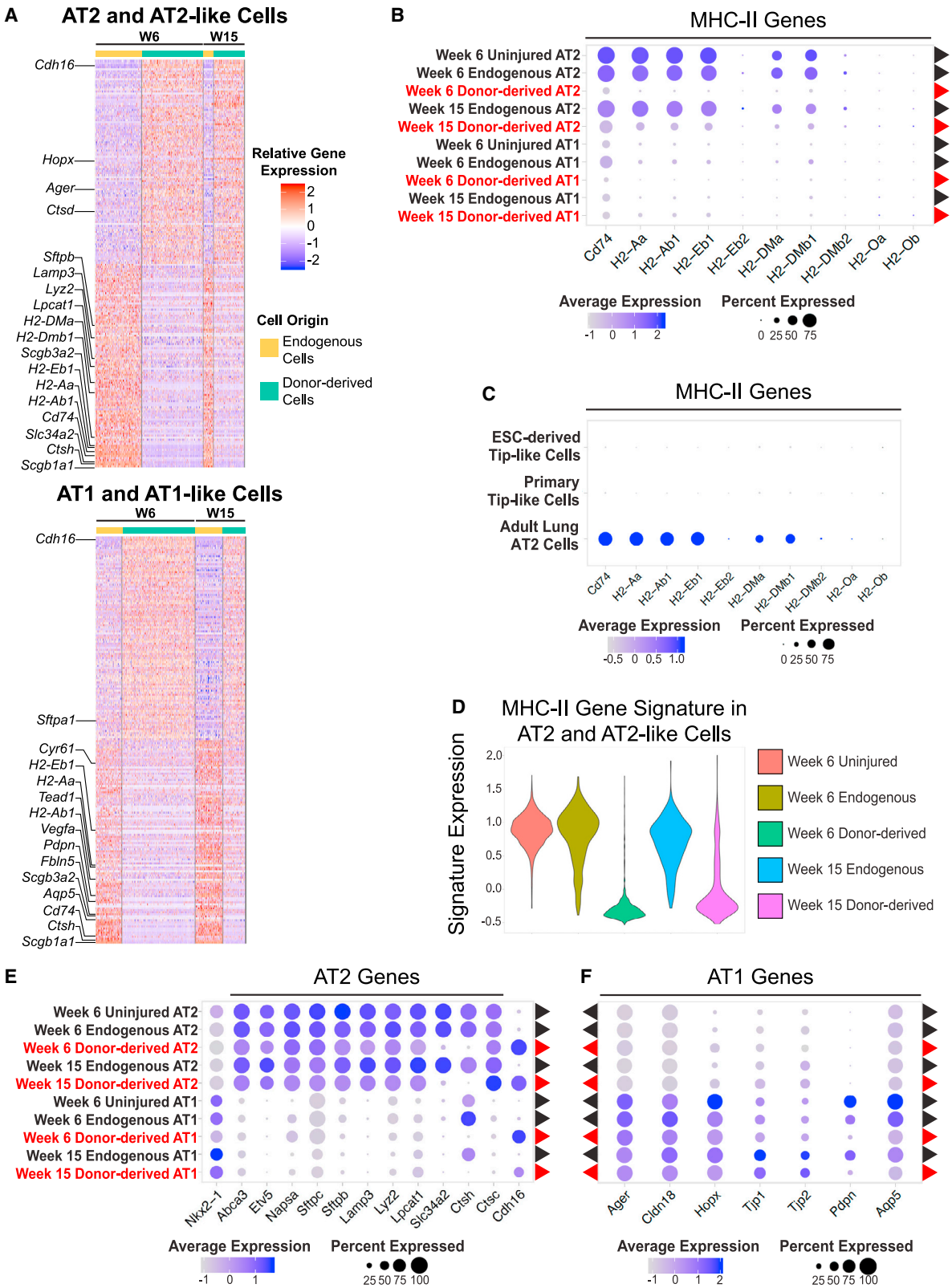


Figure 5. Donor-derived cells express lower levels of select MHC-II components and maturation markers

(A) Row-normalized heatmap of the 100 most upregulated and 100 most down-regulated genes (with adj. p value < 0.05, ordered by logFC) between donor-derived and endogenous cells for both AT2-like and AT1-like cells. Annotated genes are associated with lung epithelial lineages or MHC-II.

(legend continued on next page)

assumed a quiescent state similar to the endogenous epithelium at this stage (Figure S5F). The vast majority of donor-derived cells expressed high levels of either AT2 or AT1 gene signatures without expressing other cell-type signatures (hereafter AT2-like and AT1-like cells, respectively) (Figure 4C). These identities were verified by multimodal reference mapping that aligned the majority of donor-derived cells to endogenous AT2 and AT1 cells from the uninjured control (Figure S5G). Notably, a subpopulation of donor-derived AT1-like cells expressed a subset of AT2-to-AT1 transitional state cell markers, including Krt8 (Figure S5H).[40–42] RNA velocity suggested that these cells were in transition from AT2-like to AT1-like cells, potentially indicating the presence of donor-derived transitional state cells (Figure S5I). Finally, there was a small third donor-derived cluster from the 6-week post-transplantation sample that lacked *Nkx2-1* expression (Figure S5F). These cells (hereafter gastric-like cells) likely represent the rare GFP+/mCherry− cells seen in this sample and expressed gastric markers associated with loss of NKX2-1 in lung epithelium (Figure S5A).[43,44] Altogether, this indicates that donor-derived cells primarily give rise to cells transcriptionally similar to endogenous AT2 and AT1 cells.

Although the majority of donor-derived cells expressed alveolar epithelial lineage markers, these cells did not overlap perfectly with endogenous cells. In order to identify DEGs for both AT2-like and AT1-like cells, we compared donor-derived and endogenous cells within each of these cell types (Figure 5A; Tables S4 and S5). Both donor-derived AT2-like and AT1-like cells were deficient in the expression of major histocompatibility complex II (MHC-II) genes (Figure 5B; Tables S6 and S7), which corresponds to a nearly complete absence of MHC-II expression in sterile donor cells prior to transplantation (Figure 5C). In the adult lung epithelium, MHC-II is primarily expressed in AT2 cells, where it contributes to antigen presentation, which in turn regulates resident memory T cell function and barrier immunity during infection.[45,46] Although the majority of donor-derived AT2-like cells remained deficient for MHC-II genes even at 15 weeks post-transplantation, a portion of these did express endogenous levels of MHC-II components (Figure 5D). This suggests that with sufficient time and exposure to a non-sterile environment, donor-derived cells may upregulate MHC-II components.

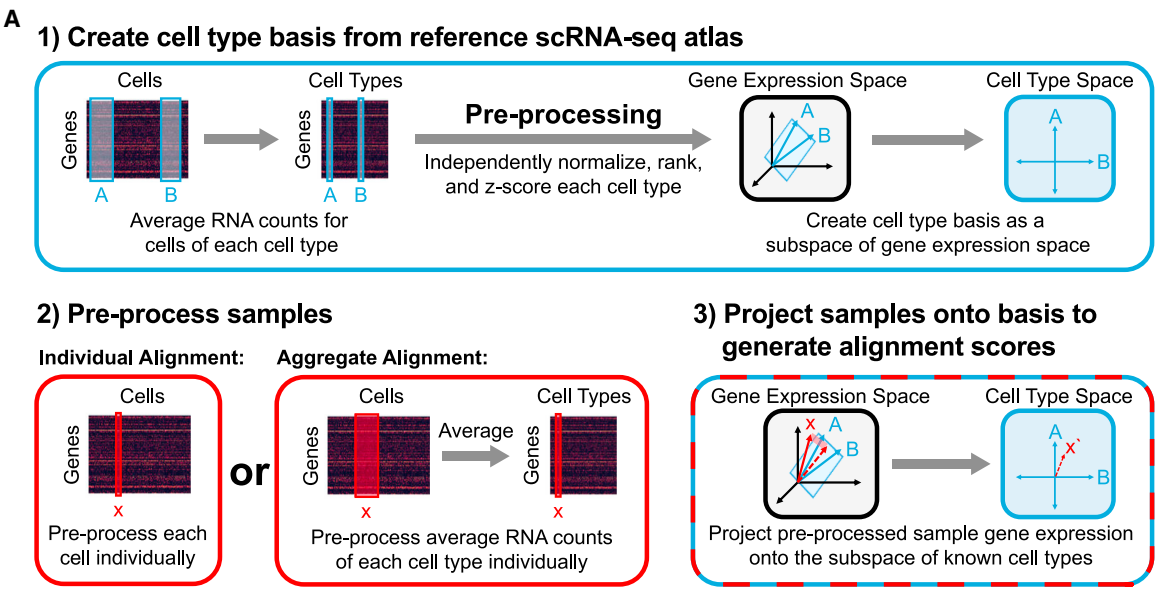
In addition to differences in MHC-II expression, both donor-derived AT2-like and AT1-like cells demonstrated subtle signs of incomplete maturation. Both cell types expressed many of the canonical lineage markers associated with the corresponding endogenous lineage, although some were expressed at lower levels (Figures 5A, 5E, and 5F), such as those associated with the late-stage maturation of AT2 (*Ctsh* and *Slc34a2*) or AT1 (*Pdpn* and *Aqp5*) cells. Although CTSH has been shown to play an important role in SFTPB processing, donor-derived cells expressed *Ctsc*, which can compensate for the absence of *Ctsh*.[47,48] Few noteworthy genes were significantly upregulated in donor-derived cells, but they did maintain high expression of an embryonic cadherin (*Cdh16*) (Figure 5E).[49] Altogether, this indicates that donor-derived cells express many of the genes necessary for alveolar epithelial cell function but may require further priming before or after transplantation in order to fully mature and respond to the non-sterile environment of the lung.

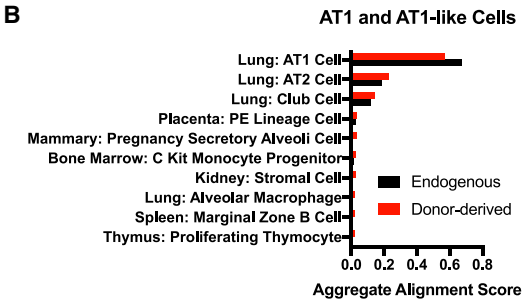
scTOP demonstrates global transcriptomic alignment between donor-derived and endogenous alveolar lineages

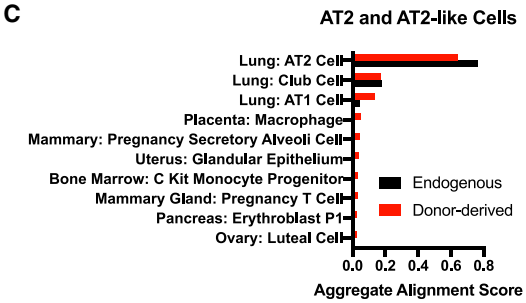
Differential gene expression analysis, SPRING, and UMAP are all designed to identify cell populations and highlight differences between them, but these methods do not readily provide means to quantify the overall similarity of non-identical cell populations. To provide unbiased quantitative assessments of how our donor-derived cells align with endogenous alveolar lineages on a global transcriptomic level, we developed a computational algorithm, single-cell type order parameters (scTOP) (Figure 6A; supplemental information).[51] Although other methods of dimensionality reduction rely on unsupervised machine learning to determine axes of relevance with no prior knowledge, scTOP uses established single-cell atlases as references to determine alignment with known cell types. This algorithm reduces the number of dimensions from the number of genes down to the number of known cell types, retaining more information than methods that reduce the data to 2 dimensions. These reference cell types are used to create vectors, which define the dimensions of a cell-type subspace. The population of interest is then projected in cell-type space as either individual cells or the average of the population, producing individual or aggregate alignment scores, respectively.

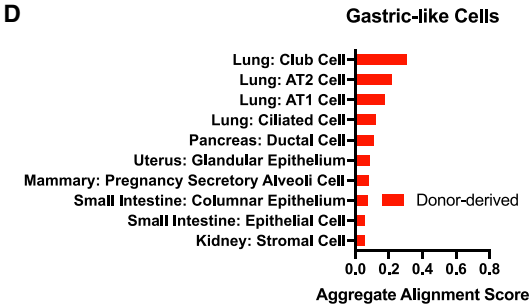
In order to determine the alignment scores of our donor-derived and endogenous epithelial cells in transplant recipients, we projected pre-processed cell transcriptomes onto reference datasets compiled from the Mouse Cell Atlas[50] and the previously described uninjured control mouse lung from Figure 4. Donor-derived and endogenous AT1 cells displayed similar aggregate alignment profiles, with both primarily aligning to the lung AT1 cell reference benchmark with high aggregate alignment scores (0.572 and 0.676, respectively; Figure 6B). For both populations, the next highest alignment was against lung AT2 cells (0.230 and 0.183), potentially reflecting the close lineage relationship between AT1 and AT2 cells. Donor-derived and endogenous AT2 cells also had similar aggregate alignment profiles, with both primarily aligning to lung AT2 cells by a considerable margin (alignment score 0.638 and 0.761, respectively; Figure 6C). Notably, donor-derived AT2-like cells had higher alignment to lung AT1-like cells compared with endogenously derived cells (0.138 vs. 0.045), which may reflect either incomplete maturation of donor-derived AT2-like cells or initiation of AT1 differentiation in a subset of this population. In contrast to these alveolar epithelial populations, the rare donor-derived gastric cells weakly aligned to multiple reference

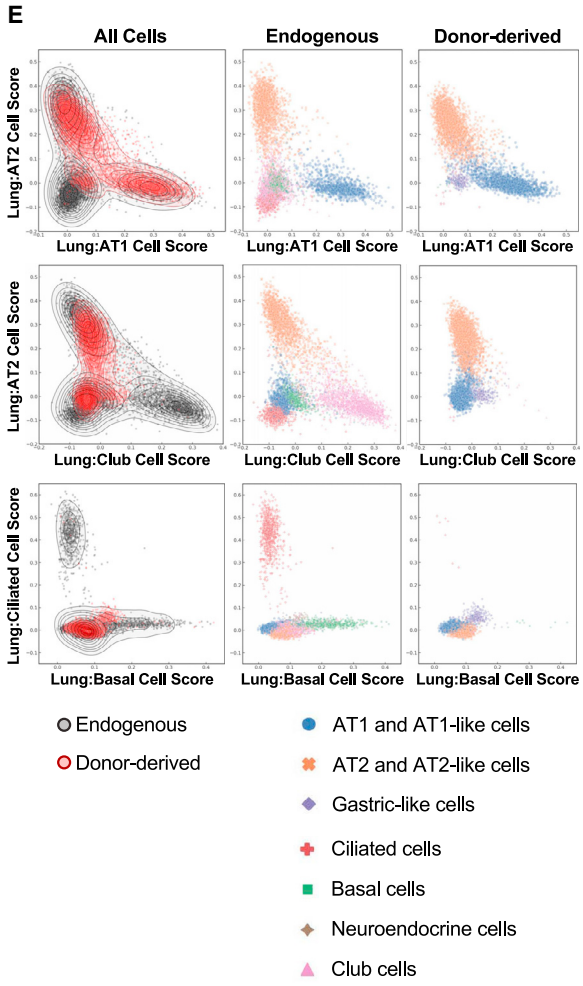
(B) Expression of MHC-II genes in donor-derived (red) and endogenous (black) cells.
(C) Expression of MHC-II genes in ESC-derived and primary tip-like cells compared with adult AT2 cells captured in the same experiment.
(D) Violin plots of an MHC-II gene signature composed of genes listed in (B).
(E) Expression of AT2 genes in donor-derived (red) and endogenous (black) cells.
(F) Expression of AT1 genes in donor-derived (red) and endogenous (black) cells.
See also Tables S4, S5, S6, and S7.











(legend on next page)

cell types (Figure 6D). Interestingly, although these cells do not express *Nkx2-1*, the top four alignment scores are all against lung references, suggesting a maintained lung transcriptional program even in the absence of this critical lung transcription factor.

To assess the individual alignment scores of transplanted cells, we then projected individual cells against the same reference basis (Figure 6E). In the resulting plots, cells were labeled as either donor-derived or endogenous and color-coded based on their previously assigned lineage identity from Figure 4. As expected, the majority of individual donor-derived cells aligned with either AT1 or AT2 cells, with all gastric-like cells aligning poorly to both cell types. Donor-derived and endogenous AT1 cells had nearly overlapping alignment distributions on all analyzed plots, reflecting substantial transcriptional similarity for these two populations, despite the DEGs identified above. Although endogenous and donor-derived AT2 cells did not overlap as precisely, there was a significant overlap between the two distributions, suggesting a relatively similar transcriptional profile for at least a subset of cells. As expected, based on our lineage marker analysis, few donor-derived cells aligned well with any airway lineage. Altogether, these results indicate that although select genes demarcate donor-derived and endogenous cell populations, the two populations are highly similar and exhibit alveolar cell states when scored on a global transcriptomic level.

Comparison of ESC-derived tip-like cell transplantation to primary cultured tip-like cell and primary adult cell transplantations

Although the above work identifies transcriptional differences between endogenous and donor-derived lineages, it is still unclear if these differences are a consequence of the engineered nature of ESC-derived donor cells, the maturation state of the donor cells, or an unavoidable consequence of alveolar cell transplantation. To begin to tease apart these possibilities, we first used scRNA-seq to compare donor-derived cells from parallel ESC-derived and primary tip-like cell transplants at 8 weeks post-transplantation (Figures S6A and S6B). As expected, both transplants primarily gave rise to AT2-like and AT1-like cells (Figure S6C). Primary transplant donor-derived cells had transcriptional deficiencies in maturation markers and MHC-II genes similar to those seen in ESC-derived transplants, with ESC-derived AT1-like cells having higher expression of several AT1 markers (Figures S6D–S6F; Tables S4 and S5). scTOP verified these findings with both ESC-derived and primary donor-derived AT2-like cells having lower AT2 alignment scores than endogenous AT2 cells (Figures S6G and S6H). Likewise, scTOP indicated that ESC-derived AT1-like cells have similar or higher AT1 alignment scores relative to primary donor-derived AT1-like cells. Together, this suggests that transcriptionally similar ESC-derived and primary donor cells give rise to comparable donor-derived cells following pulmonary cell transplantation.

To further benchmark tip-like cell transplantation, we compared our results with those of published adult lung cell transplantations.[18] In Louie et al., donor cells were generated by culturing adult murine lung cells enriched for SCA1− AT2 cells or SCA1+ bronchioalveolar stem cells (BASCs) with neonatal stromal cells for 3 weeks. These cells were then transplanted into immunocompromised mice at 1-day post-bleomycin injury and analyzed by scRNA-seq at 10–12 weeks post-transplantation. Reanalysis of these data confirmed that the majority of cells expressed markers of AT2 cells, ciliated cells, or the Krt17+/Krt8+ "transitional cell" clusters identified in the original paper (Figures S7A–S7E). Notably, donor-derived AT2 cells clustered with endogenous AT2 cells and did not exhibit the maturation or MHC-II expression deficiencies seen in tip-like cell transplants (Figure S7C). However, these adult primary cell transplants did not give rise to a significant AT1-like cell population, suggesting a block in AT2-to-AT1 differentiation.[18] Donor-derived ciliated cells aligned specifically to the lung ciliated cell reference (Figures S17D and S17E). Although a subset of transitional cells expressed markers of Krt8+ alveolar differentiation intermediates (ADIs), others expressed markers of club and basal cells[40] (Figure S17F). Thus, cultured adult lung cell transplantation generates donor-derived AT2 cells that are transcriptionally indistinguishable from endogenous AT2 cells, but these transplants did not give rise to AT1 cells and may generate other lineages with unknown impacts on alveolar function. Together with our previous results, this suggests that the transcriptional deficiencies seen in our ESC-derived tip-like cells are likely a consequence of the maturation state of the donor cells and are not fundamental to lung epithelial cell transplantation or the use of ESC-derived donor cells.

Donor-derived AT2-like cells are functionally similar to endogenous AT2 cells

In order to establish whether transplantation of ESC-derived tip-like cells models true cellular engraftment, it is necessary to determine whether donor-derived cells are functionally similar to endogenous alveolar epithelial cells despite any transcriptional differences.[24] AT2 cells utilize unique organelles, known as lamellar bodies, to secrete surfactant proteins and lipids that help protect the lung and maintain surface tension. The punctate localization of proSFTPC in donor-derived AT2-like cells suggested the presence of lamellar bodies (Figure 3F). To characterize the ultrastructural features of these cells, sort-purified GFP+ donor-derived and GFP− endogenous epithelial cells from the same mice were imaged using transition electron

Figure 6. Global transcriptomic comparison of endogenous and donor-derived lung epithelial cells using scTOP

(A) Schematic of scTOP (single-cell type order parameters).

(B) The top ten aggregate alignment scores for donor-derived AT1-like cells and the corresponding scores for endogenous AT1 cells. All reference cell types are from adult mice (Mouse Cell Atlas or control sample as delineated in Figure 4).[50]

(C) The top ten aggregate alignment scores for donor-derived AT2-like cells and the corresponding scores for endogenous AT2 cells. All reference cell types are from adult mice.

(D) The top ten aggregate alignment scores for donor-derived gastric-like cells. All reference cell types are from adult mice.

(E) Individual alignment scores for all donor-derived and endogenous epithelial cells against the indicated reference cells. Each cell is annotated (by color and shape shown in the key below the graphs) based on sample type or cell type as determined in Figure 4.

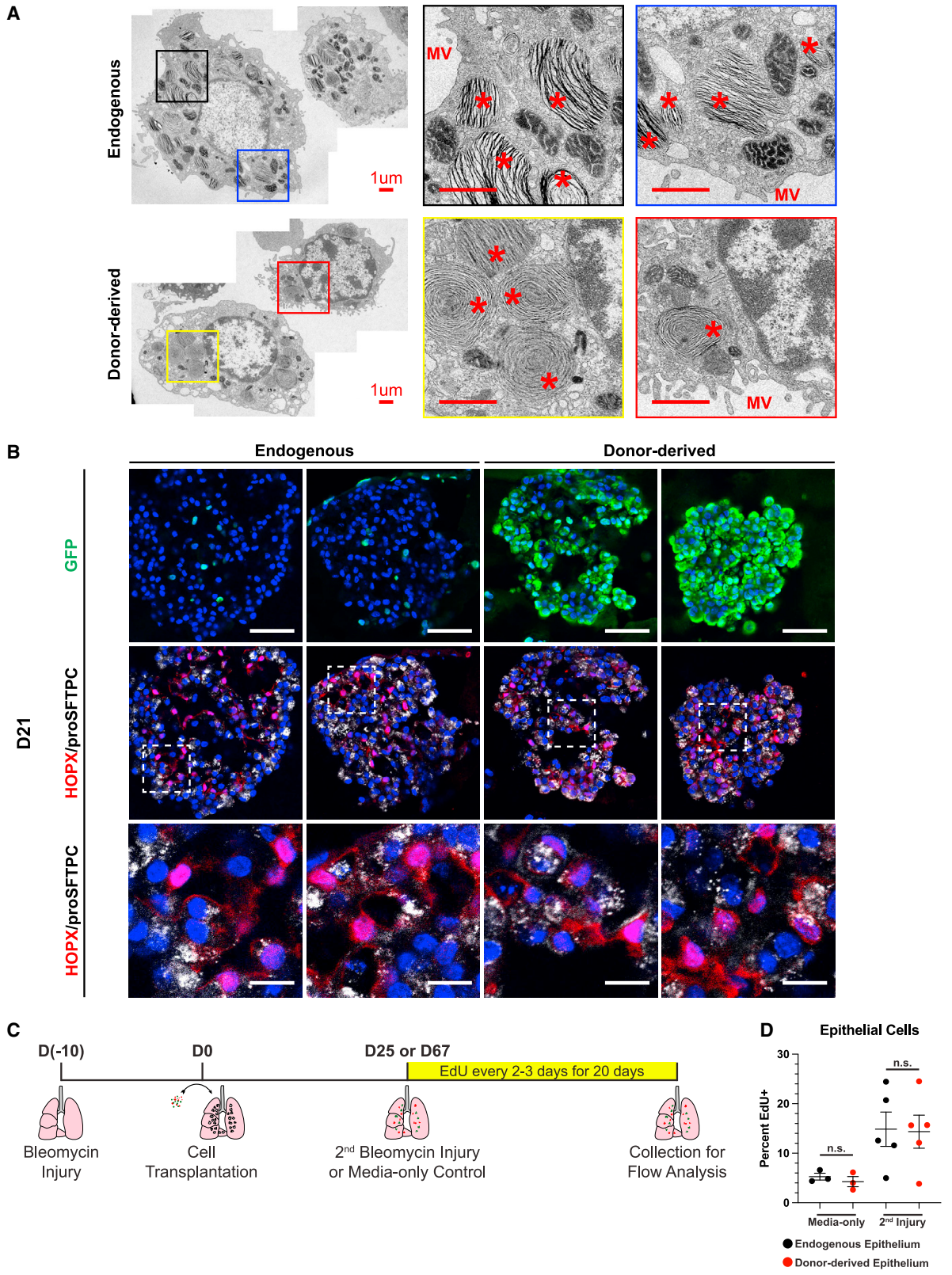


Figure 7. Functional assessment of donor-derived AT2-like cells

(A) Representative transmission electron micrographs of GFP− endogenous AT2 cells and GFP+ donor-derived AT2-like cells from the same mouse. Scale bars, 1.0 μm. *lamellar body; MV, microvilli.

(legend continued on next page)

microscopy (TEM; Figure 7A). Donor-derived AT2-like cells were ultrastructurally similar to endogenous AT2 cells and featured the microvilli and lamellar bodies characteristic of functional AT2 cells.

In addition to their secretory function, AT2 cells are the main facultative progenitors of the adult alveolar epithelium.[52–54] In response to injury or cell culture conditions, these normally quiescent cells can proliferate and differentiate into AT1 cells. To determine if donor-derived cells have a similar progenitor function, endogenous or donor-derived epithelial cells were sorted and cultured with PDGFRa^{nGFP+} lung mesenchymal cells.[54] After 21 days in culture, both endogenous and donor-derived cells plated as single-cell suspensions gave rise to large NKX2-1+ epithelial organoids, indicating that these previously quiescent cells (Figure 3F) can re-enter the cell cycle (Figure 7B). Both endogenous and donor-derived cell organoids contained cuboidal proSFTPC+/HOPX− cells as well as proSFTPC−/HOPX+ cells with thin protrusions, suggesting differentiation into AT2-like and AT1-like cells, respectively. In addition to these expected cell types, proSFTPC+/HOPX+ cells could be found in donor-derived organoids. This corresponds to increased *Hopx* expression (Figure 5A) and AT1 alignment (Figure 6C) in donor-derived AT2-like cells *in vivo* and potentially indicates instances of delayed or incomplete AT1 differentiation in this co-culture assay. To verify the facultative progenitor capacity of donor-derived cells *in vivo*, we reinjured transplant recipients with a second bleomycin injury and labeled proliferative cells with EdU for 20 days (Figure 7C). Flow cytometry indicated that endogenous and donor-derived had a similar proliferative response following a secondary injury, suggesting that these two populations have a similar progenitor capacity (Figure 7D). Altogether, this suggests that donor-derived cells can function as facultative progenitors and that the transplants described above model functional cellular engraftment capable of replacing endogenous epithelial cells.

DISCUSSION

The ultimate goal of syngeneic pulmonary cell therapy is an effective treatment for pulmonary injury and disease that is not dependent on the availability of donor tissue or the use of detrimental immunosuppressants. Although previous studies have demonstrated cell transplantation in mouse lungs, these studies utilized difficult-to-collect primary cells or necessitated the use of immunocompromised recipients.[8–19] To address these issues, we developed an approach to engraft PSC-derived cells into the lungs of syngeneic and immunocompetent recipients. Following transplantation, these cells give rise to persistent AT2-like and AT1-like cells that are transcriptomically and functionally similar to endogenous alveolar epithelial cells. This model thus provides a valuable system to further characterize and optimize pulmonary cell therapy in an approachable and clinically relevant system.

To create an ESC-derived donor population, we focused on mimicking cultured primary tip-like cells, a cell population that can be transplanted in immunodeficient mice.[10] Utilizing this guidepost, along with published differentiation protocols and *in vivo* development signaling networks,[28,31,32] we developed a protocol for the directed differentiation of mouse ESCs into tip-like epithelial cells. The resulting progenitor population could be frozen down or expanded without losing its cell identity and was transcriptomically similar to cultured primary tip cells. To assess the viability of syngeneic transplantation, we then transplanted our ESC-derived cells into immunocompetent recipients. Initially, these cells maintained a tip-like identity *in vivo*. However, as early as 2 weeks post-transplantation, the donor-derived cells differentiated into AT2-like and AT1-like cells. Following our ESC-derived transplants, we saw differentiated donor-derived cells surviving *in vivo* for at least 6 months while maintaining a largely quiescent alveolar epithelial identity. This indicates progenitor cells engineered outside of the lung can differentiate following transplantation and survive for extended periods in the presence of a functional immune system without developing into tumorigenic cells.

PSC-derived cells differentiated *in vitro* are often substantially different from their mature primary counterparts on a transcriptomic level despite expression of key lineage markers.[23,55–59] Therefore, it was critical to determine whether our engineered cells could differentiate *in vivo* into cells transcriptomically similar to the endogenous alveolar epithelium. scRNA-seq analysis indicated that donor-derived cells were highly similar to the endogenous epithelial lineages with significant expression of canonical markers associated with these cell types. Through the use of scTOP, we observed that our donor-derived lineages had alignment profiles similar to those of paired endogenous cells. Despite these overall similarities, donor-derived AT2-like and AT1-like cells had notable deficiencies in expression of select maturation markers and components of the MHC-II complex. Importantly, these differences were common to both ESC-derived and primary cell transplants, indicating that this did not necessarily reflect a shortcoming of engineered donor cells. Further studies will be needed to understand whether these differences impact cellular function.

Although multiple studies have reported transplantation of cells into the lung, few have gone on to prove engraftment through assessment of the intrinsic functionality of donor-derived cells.[8–20,23] In this study, we demonstrated that donor-derived AT2-like cells produce lamellar bodies, the specialized organelles necessary for AT2 secretory function. Furthermore, quiescent donor-derived cells were capable of re-entering the cell cycle in culture, producing both AT2-like and AT1-like cells, as well as following a secondary injury

(B) Representative immunofluorescence confocal microscopy of GFP/HOPX/proSFTPC expression in cultured mouse lung alveolospheres, comparing endogenous or donor-derived epithelial cells co-cultured with PDGFRa^{nGFP}+ primary lung fibroblasts. PDGFRa-GFP is nuclear, while donor-derived cells have a cytoplasmic GFP. Nuclei stained with Hoechst, scale bars: 50 μm (top row) or 12.5 μm (bottom row).
(C) Schematic for EdU labeling of transplant recipient mice following a second bleomycin injury or media-only control.
(D) Percent EdU labeling of endogenous and donor-derived cells in ESC-derived tip-like cell recipients for the first 20 days following media-only delivery or a secondary bleomycin injury. Lobes were pruned down to regions containing GFP+ cells, or similar regions in no transplant controls, prior to digestion into single-cell suspension. n.s., not significant by unpaired, two-tailed Student's t test. n = 3, 5 biological replicates. Error bars = mean ± SEM.

in vivo, indicating the existence of facultative progenitors similar to endogenous AT2 cells.[52–54]

In summary, the current study establishes a model of PSC-derived pulmonary cell transplantation that results in durable engraftment of donor-derived cells in an immunocompetent recipient. This model thus provides an important foundation for further characterization and optimization of pulmonary cell engraftment, with the potential to yield important insights into pulmonary regeneration and the development of clinical cell therapies.

Limitations of the study

Although this work provides important insights into the feasibility of PSC-based cell therapy, future studies will be needed to further characterize the functionality of engrafted cells and transition toward clinical cell therapy. Future work will be needed to functionally test whether engraftment of engineered cells is sufficient to prevent mortality associated with severe acute injury or progressive pulmonary diseases. Although we used bleomycin, a known fibrotic agent, to clear out endogenous epithelium based on previous publications,[8,10,18] the dose used was not lethal and did not severally impact mouse behavior, preventing us from assessing the impact of transplantation on mouse viability. Furthermore, transplantation efficiency was highly variable and frequently below 1%. This is likely in part due to the variability of bleomycin-induced injury. To overcome these limitations, new methods will be needed to further develop the mouse system, improve engraftment efficiency, and apply lessons learned to the development of clinical pulmonary cell therapies. Alternative methods of cellular clearance, such as targeted decellularization,[60] may provide more precise and consistent removal of epithelial lineages. Finally, although ESC-derived cells were used in this study, iPSC-derived cells will be required for performing similar isogenic cell transplants in humans, and there will be a need to identify the ideal human iPSC-derived donor lines with current protocols to generate iPSC-derived tip-like or AT2-like cells providing promising candidates.[12,61–64]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2023.07.016.

ACKNOWLEDGMENTS

We thank Brian R. Tilton and the BUSM Flow Cytometry Core for their technical assistance and guidance. We would also like to thank Anne Hinds and Esther Bullitt for their help with processing and imaging samples for TEM. We would to like to thank the entire Kotton Lab and Center for Regenerative Medicine of Boston University and Boston Medical Center for their support and suggestions throughout the course of this research. We would like to thank Carla Kim for her insightful manuscript input as well as Jeff Whitsett and Andy Vaughan for their guidance and support. This work was supported by NIH grants T32HL007035 and F32HL149263 and an NHLBI Progenitor Cell Translational Consortium (PCTC) Jump Start Award to M.J.H.; Boston University Kilachand Multicellular Design Program Accelerator Grants to M.J.H., M.Y., P.M., and D.N.K.; and NIH grants (U01HL134745, U01HL134766, U01HL148692, and R01HL095993) and an Allen Distinguished Investigator grant from the Paul G. Allen Family Foundation to D.N.K. The graphical abstract was made with BioRender.com.

AUTHOR CONTRIBUTIONS

M.J.H. and D.N.K. designed the project and wrote the paper. M.J.H., B.R.T., J.H., C.-L.N., L.M., and M.M.M. performed experiments. J.L.-V., F.W., P.B., and C.V.-M. performed bioinformatic processing and analysis of scRNA-seq datasets. M.Y. and P.M. designed and performed scTOP analysis. D.N.K. and P.M. supervised research.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: July 13, 2022
Revised: June 9, 2023
Accepted: July 31, 2023
Published: August 24, 2023

REFERENCES

1. Thomas, E.D., Lochte, H.L., Lu, W.C., and Ferrebee, J.W. (1957). Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. N. Engl. J. Med. *257*, 491–496. https://doi.org/10.1056/NEJM195709122571102.
2. O'Connor, N., Mulliken, J., Banks-Schlegel, S., Kehinde, O., and Green, H. (1981). Grafting of burns with cultured epithelium prepared from autologous epidermal cells. Lancet *317*, 75–78. https://doi.org/10.1016/S0140-6736(81)90006-4.

3. Rama, P., Matuska, S., Paganoni, G., Spinelli, A., De Luca, M.D., and Pellegrini, G. (2010). Limbal stem-cell therapy and long-term corneal regeneration. N. Engl. J. Med. *363*, 147–155. https://doi.org/10.1056/NEJMoa0905955.

4. Schwartz, S.D., Regillo, C.D., Lam, B.L., Eliott, D., Rosenfeld, P.J., Gregori, N.Z., Hubschman, J.-P., Davis, J.L., Heilwell, G., Spirn, M., et al. (2015). Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. Lancet *385*, 509–516. https://doi.org/10.1016/S0140-6736(14)61376-3.

5. Xuan, K., Li, B., Guo, H., Sun, W., Kou, X., He, X., Zhang, Y., Sun, J., Liu, A., Liao, L., et al. (2018). Deciduous autologous tooth stem cells regenerate dental pulp after implantation into injured teeth. Sci. Transl. Med. *10*, eaaf3227. https://doi.org/10.1126/scitranslmed.aaf3227.

6. Kinoshita, S., Koizumi, N., Ueno, M., Okumura, N., Imai, K., Tanaka, H., Yamamoto, Y., Nakamura, T., Inatomi, T., Bush, J., et al. (2018). Injection of cultured cells with a ROCK inhibitor for bullous keratopathy. N. Engl. J. Med. *378*, 995–1003. https://doi.org/10.1056/NEJMoa1712770.

7. Schweitzer, J.S., Song, B., Herrington, T.M., Park, T.-Y., Lee, N., Ko, S., Jeon, J., Cha, Y., Kim, K., Li, Q., et al. (2020). Personalized iPSC-derived dopamine progenitor cells for Parkinson's disease. N. Engl. J. Med. *382*, 1926–1932. https://doi.org/10.1056/NEJMoa1915872.

8. Vaughan, A.E., Brumwell, A.N., Xi, Y., Gotts, J.E., Brownfield, D.G., Treutlein, B., Tan, K., Tan, V., Liu, F.C., Looney, M.R., et al. (2015). Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature *517*, 621–625. https://doi.org/10.1038/nature14112.

9. Rosen, C., Shezen, E., Aronovich, A., Klionsky, Y.Z., Yaakov, Y., Assayag, M., Biton, I.E., Tal, O., Shakhar, G., Ben-Hur, H., et al. (2015). Preconditioning allows engraftment of mouse and human embryonic lung cells, enabling lung repair in mice. Nat. Med. *21*, 869–879. https://doi.org/10.1038/nm.3889.

10. Nichane, M., Javed, A., Sivakamasundari, V., Ganesan, M., Ang, L.T., Kraus, P., Lufkin, T., Loh, K.M., and Lim, B. (2017). Isolation and 3D expansion of multipotent Sox9+ mouse lung progenitors. Nat. Methods *14*, 1205–1212. https://doi.org/10.1038/nmeth.4498.

11. Xi, Y., Kim, T., Brumwell, A.N., Driver, I.H., Wei, Y., Tan, V., Jackson, J.R., Xu, J., Lee, D.-K., Gotts, J.E., et al. (2017). Local lung hypoxia determines epithelial fate decisions during alveolar regeneration. Nat. Cell Biol. *19*, 904–914. https://doi.org/10.1038/ncb3580.

12. Miller, A.J., Hill, D.R., Nagy, M.S., Aoki, Y., Dye, B.R., Chin, A.M., Huang, S., Zhu, F., White, E.S., Lama, V., et al. (2018). In vitro induction and in vivo engraftment of lung bud tip progenitor cells derived from human pluripotent stem cells. Stem Cell Rep. *10*, 101–119. https://doi.org/10.1016/j.stemcr.2017.11.012.

13. Weiner, A.I., Jackson, S.R., Zhao, G., Quansah, K.K., Farshchian, J.N., Neupauer, K.M., Littauer, E.Q., Paris, A.J., Liberti, D.C., Worthen, G.S., et al. (2019). Mesenchyme-free expansion and transplantation of adult alveolar progenitor cells: steps toward cell-based regenerative therapies. npj Regen. Med. *4*, 17. https://doi.org/10.1038/s41536-019-0080-9.

14. Kathiriya, J.J., Brumwell, A.N., Jackson, J.R., Tang, X., and Chapman, H.A. (2020). Distinct airway epithelial stem cells hide among club cells but mobilize to promote alveolar regeneration. Cell Stem Cell *26*, 346–358.e4. https://doi.org/10.1016/j.stem.2019.12.014.

15. Liao, C.-C., Chiu, C.-J., Yang, Y.-H., and Chiang, B.-L. (2022). Neonatal lung-derived SSEA-1+ cells exhibited distinct stem/progenitor characteristics and organoid developmental potential. Iscience *25*, 104262. https://doi.org/10.1016/j.isci.2022.104262.

16. Kathiriya, J.J., Wang, C., Zhou, M., Brumwell, A., Cassandras, M., Le Saux, C.J.L., Cohen, M., Alysandratos, K.-D., Wang, B., Wolters, P., et al. (2022). Human alveolar type 2 epithelium transdifferentiates into metaplastic KRT5+ basal cells. Nat. Cell Biol. *24*, 10–23. https://doi.org/10.1038/s41556-021-00809-4.

17. Miyata, R., Hasegawa, K., Menju, T., Yoshizawa, A., Watanabe, A., Hirai, T., Date, H., and Sato, A. (2022). Lung fibrogenic microenvironment in mouse reconstitutes human alveolar structure and lung tumor. Iscience *25*, 104912. https://doi.org/10.1016/j.isci.2022.104912.

18. Louie, S.M., Moye, A.L., Wong, I.G., Lu, E., Shehaj, A., Garcia-de-Alba, C., Ararat, E., Raby, B.A., Lu, B., Paschini, M., et al. (2022). Progenitor potential of lung epithelial organoid cells in a transplantation model. Cell Rep. *39*, 110662. https://doi.org/10.1016/j.celrep.2022.110662.

19. Jaeger, B., Schupp, J.C., Plappert, L., Terwolbeck, O., Artysh, N., Kayser, G., Engelhard, P., Adams, T.S., Zweigerdt, R., Kempf, H., et al. (2022). Airway basal cells show a dedifferentiated KRT17highPhenotype and promote fibrosis in idiopathic pulmonary fibrosis. Nat. Commun. *13*, 5637. https://doi.org/10.1038/s41467-022-33193-0.

20. Ma, L., Thapa, B.R., Le-Suer, J., Tilson-Lünel, A., Herriges, M.J., Berical, A., Beerman, M.L., Wang, F., Bawa, P.S., Randell, S.H., et al. (2023). Airway stem cell reconstitution by transplantation of cultured primary or PSC-derived basal cells. Cell Stem Cell *30*, 1199–1216. https://doi.org/10.1016/j.stem.2023.07.01.

21. Boieri, M., Shah, P., Dressel, R., and Inngjerdingen, M. (2016). The role of animal models in the study of hematopoietic stem cell transplantation and GvHD: A historical overview. Front. Immunol. *7*, 333. https://doi.org/10.3389/fimmu.2016.00333.

22. Morgan, R.A., Gray, D., Lomova, A., and Kohn, D.B. (2017). Hematopoietic stem cell gene therapy: progress and lessons learned. Cell Stem Cell *21*, 574–590. https://doi.org/10.1016/j.stem.2017.10.010.

23. Alysandratos, K.-D., Rivas, C.G.de A., Yao, C., Pessina, P., Villacorta-Martin, C., Huang, J., Hix, O.T., Minakin, K., Konda, B., Stripp, B.R., et al. (2022). Impact of cell culture on the transcriptomic programs of primary and iPSC-derived human alveolar type 2 cells. https://doi.org/10.1101/2022.02.08.479591.

24. Alysandratos, K.-D., Herriges, M.J., and Kotton, D.N. (2021). Epithelial stem and progenitor cells in lung repair and regeneration. Annu. Rev. Physiol. *83*, 529–550. https://doi.org/10.1146/annurev-physiol-041520-092904.

25. Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676. https://doi.org/10.1016/j.cell.2006.07.024.

26. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872. https://doi.org/10.1016/j.cell.2007.11.019.

27. Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean, J.M., Smith, B.W., et al. (2010). Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells *28*, 1728–1740. https://doi.org/10.1002/stem.495.

28. Ikonomou, L., Herriges, M.J., Lewandowski, S.L., Marsland, R., Villacorta-Martin, C., Caballero, I.S., Frank, D.B., Sanghrajka, R.M., Dame, K., Kańduła, M.M., et al. (2020). The in vivo genetic program of murine primordial lung epithelial progenitors. Nat. Commun. *11*, 635. https://doi.org/10.1038/s41467-020-14348-3.

29. Rawlins, E.L., Clark, C.P., Xue, Y., and Hogan, B.L.M. (2009). The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Development *136*, 3741–3745. https://doi.org/10.1242/dev.037317.

30. Bilodeau, M., Shojaie, S., Ackerley, C., Post, M., and Rossant, J. (2014). Identification of a proximal progenitor population from murine fetal lungs with clonogenic and multilineage differentiation potential. Stem Cell Rep. *3*, 634–649. https://doi.org/10.1016/j.stemcr.2014.07.010.

31. Serra, M., Alysandratos, K.D., Hawkins, F., McCauley, K.B., Jacob, A., Choi, J., Caballero, I.S., Vedaie, M., Kurmann, A.A., Ikonomou, L., et al. (2017). Pluripotent stem cell differentiation reveals distinct developmental pathways regulating lung versus thyroid lineage specification. Development *144*, 3879–3893. https://doi.org/10.1242/dev.150193.

32. Han, L., Chaturvedi, P., Kishimoto, K., Koike, H., Nasr, T., Iwasawa, K., Giesbrecht, K., Witcher, P.C., Eicher, A., Haines, L., et al. (2020). Single cell transcriptomics identifies a signaling network coordinating endoderm

and mesoderm diversification during foregut organogenesis. Nat. Commun. *11*, 4158. https://doi.org/10.1038/s41467-020-17968-x.

33. Cardoso, W.V., and Lü, J. (2006). Regulation of early lung morphogenesis: questions, facts and controversies. Development *133*, 1611–1624. https://doi.org/10.1242/dev.02310.

34. McCauley, K.B., Hawkins, F., Serra, M., Thomas, D.C., Jacob, A., and Kotton, D.N. (2017). Efficient derivation of functional human airway epithelium from pluripotent stem cells via temporal regulation of Wnt signaling. Cell Stem Cell *20*, 844–857.e6. https://doi.org/10.1016/j.stem.2017.03.001.

35. Longmire, T.A., Ikonomou, L., Hawkins, F., Christodoulou, C., Cao, Y., Jean, J.C., Kwok, L.W., Mou, H., Rajagopal, J., Shen, S.S., et al. (2012). Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. Cell Stem Cell *10*, 398–411. https://doi.org/10.1016/j.stem.2012.01.019.

36. Hurley, K., Ding, J., Villacorta-Martin, C., Herriges, M.J., Jacob, A., Vedaie, M., Alysandratos, K.D., Sun, Y.L., Lin, C., Werder, R.B., et al. (2020). Reconstructed single-cell fate trajectories define lineage plasticity windows during differentiation of human PSC-derived distal lung progenitors. Cell Stem Cell *26*, 593–608.e8. https://doi.org/10.1016/j.stem.2019.12.009.

37. Zepp, J.A., Morley, M.P., Loebel, C., Kremp, M.M., Chaudhry, F.N., Basil, M.C., Leach, J.P., Liberti, D.C., Niethamer, T.K., Ying, Y., et al. (2021). Genomic, epigenomic, and biophysical cues controlling the emergence of the lung alveolus. Science *371*, eabc3172. https://doi.org/10.1126/science.abc3172.

38. Weinreb, C., Wolock, S., and Klein, A.M. (2018). Spring: a kinetic interface for visualizing high dimensional single-cell expression data. Bioinformatics *34*, 1246–1248. https://doi.org/10.1093/bioinformatics/btx792.

39. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

40. Strunz, M., Simon, L.M., Ansari, M., Kathiriya, J.J., Angelidis, I., Mayr, C.H., Tsidiridis, G., Lange, M., Mattner, L.F., Yee, M., et al. (2020). Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. Nat. Commun. *11*, 3559. https://doi.org/10.1038/s41467-020-17358-3.

41. Jiang, P., Rubio, R.G. de, Hrycaj, S.M., Gurczynski, S.J., Riemondy, K.A., Moore, B.B., Omary, M.B., Ridge, K.M., and Zemans, R.L. (2019). Ineffectual AEC2-to-AEC1 differentiation in IPF: persistence of KRT8hi transitional state. Am. J. Resp. Crit. Care *0*, 1443–1447. https://doi.org/10.1164/rccm.201909-1726le.

42. Kobayashi, Y., Tata, A., Konkimalla, A., Katsura, H., Lee, R.F., Ou, J., Banovich, N.E., Kropski, J.A., and Tata, P.R. (2020). Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis. Nat. Cell Biol. *22*, 934–946. https://doi.org/10.1038/s41556-020-0542-8.

43. Snyder, E.L., Watanabe, H., Magendantz, M., Hoersch, S., Chen, T.A., Wang, D.G., Crowley, D., Whittaker, C.A., Meyerson, M., Kimura, S., et al. (2013). Nkx2-1 represses a latent gastric differentiation program in lung adenocarcinoma. Mol. Cell *50*, 185–199. https://doi.org/10.1016/j.molcel.2013.02.018.

44. Herriges, M.J., Tischfield, D.J., Cui, Z., Morley, M.P., Han, Y., Babu, A., Li, S., Lu, M., Cendan, I., Garcia, B.A., et al. (2017). The NANCI–Nkx2.1 gene duplex buffers Nkx2.1 expression to maintain lung development and homeostasis. Genes Dev. *31*, 889–903. https://doi.org/10.1101/gad.298018.117.

45. Shenoy, A.T., Lyon De Ana, C.L.D., Arafa, E.I., Salwig, I., Barker, K.A., Korkmaz, F.T., Ramanujan, A., Etesami, N.S., Soucy, A.M., Martin, I.M.C., et al. (2021). Antigen presentation by lung epithelial cells directs CD4+ TRM cell function and regulates barrier immunity. Nat. Commun. *12*, 5834. https://doi.org/10.1038/s41467-021-26045-w.

46. Toulmin, S.A., Bhadiadra, C., Paris, A.J., Lin, J.H., Katzen, J., Basil, M.C., Morrisey, E.E., Worthen, G.S., and Eisenlohr, L.C. (2021). Type II alveolar cell MHCII improves respiratory viral disease outcomes while exhibiting limited antigen presentation. Nat. Commun. *12*, 3993. https://doi.org/10.1038/s41467-021-23619-6.

47. Ueno, T., Linder, S., Na, C.-L., Rice, W.R., Johansson, J., and Weaver, T.E. (2004). Processing of pulmonary surfactant protein B by napsin and cathepsin H. J. Biol. Chem. *279*, 16178–16184. https://doi.org/10.1074/jbc.M312029200.

48. Bühling, F., Kouadio, M., Chwieralski, C.E., Kern, U., Hohlfeld, J.M., Klemm, N., Friedrichs, N., Roth, W., Deussing, J.M., Peters, C., et al. (2011). Gene targeting of the cysteine peptidase cathepsin H impairs lung surfactant in mice. PLoS One *6*, e26247. https://doi.org/10.1371/journal.pone.0026247.

49. Wertz, K., and Herrmann, B.G. (1999). Kidney-specific cadherin (cdh16) is expressed in embryonic kidney, lung, and sex ducts. Mech. Dev. *84*, 185–188. https://doi.org/10.1016/s0925-4773(99)00074-x.

50. Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. (2018). Mapping the mouse cell atlas by microwell-seq. Cell *172*, 1091–1107.e17. https://doi.org/10.1016/j.cell.2018.02.001.

51. Yampolskaya, M., Herriges, M., Ikonomou, L., Kotton, D., and Mehta, P. (2023). scTOP: physics-inspired order parameters for cellular identification and visualization. https://doi.org/10.1101/2023.01.25.525581.

52. Adamson, I.Y., and Bowden, D.H. (1974). The type 2 cell as progenitor of alveolar epithelial regeneration. A cytodynamic study in mice after exposure to oxygen. Lab. Invest. *30*, 35–42.

53. Evans, M.J., Cabral, L.J., Stephens, R.J., and Freeman, G. (1975). Transformation of alveolar Type 2 cells to Type 1 cells following exposure to NO2. Exp. Mol. Pathol. *22*, 142–150. https://doi.org/10.1016/0014-4800(75)90059-3.

54. Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B.L.M. (2013). Type 2 alveolar cells are stem cells in adult lung. J. Clin. Invest. *123*, 3025–3036. https://doi.org/10.1172/JCI68782.

55. Pavlovic, B.J., Blake, L.E., Roux, J., Chavarria, C., and Gilad, Y. (2018). A comparative assessment of human and chimpanzee iPSC-derived cardiomyocytes with primary heart tissues. Sci. Rep. *8*, 15312. https://doi.org/10.1038/s41598-018-33478-9.

56. Abo, K.M., Sainz de Aja, J.S. de, Lindstrom-Vautrin, J., Alysandratos, K.D., Richards, A., Garcia-de-Alba, C., Huang, J., Hix, O.T., Werder, R.B., Bullitt, E., et al. (2022). Air-liquid interface culture promotes maturation and allows environmental exposure of pluripotent stem cell–derived alveolar epithelium. JCI Insight *7*, e155589. https://doi.org/10.1172/jci.insight.155589.

57. Xia, N., Zhang, P., Fang, F., Wang, Z., Rothstein, M., Angulo, B., Chiang, R., Taylor, J., and Reijo Pera, R.A.R. (2016). Transcriptional comparison of human induced and primary midbrain dopaminergic neurons. Sci. Rep. *6*, 20270. https://doi.org/10.1038/srep20270.

58. Fidanza, A., Stumpf, P.S., Ramachandran, P., Tamagno, S., Babtie, A., Lopez-Yrigoyen, M., Taylor, A.H., Easterbrook, J., Henderson, B.E.P., Axton, R., et al. (2020). Single-cell analyses and machine learning define hematopoietic progenitor and HSC-like cells derived from human PSCs. Blood *136*, 2893–2904. https://doi.org/10.1182/blood.2020006229.

59. Schwartz, R.E., Fleming, H.E., Khetani, S.R., and Bhatia, S.N. (2014). Pluripotent stem cell-derived hepatocyte-like cells. Biotechnol. Adv. *32*, 504–513. https://doi.org/10.1016/j.biotechadv.2014.01.003.

60. Dorrello, N.V., Guenthart, B.A., O'Neill, J.D., Kim, J., Cunningham, K., Chen, Y.-W., Biscotti, M., Swayne, T., Wobma, H.M., Huang, S.X.L., et al. (2017). Functional vascularized lung grafts for lung bioengineering. Sci. Adv. *3*, e1700521. https://doi.org/10.1126/sciadv.1700521.

61. Huang, S.X.L., Islam, M.N., O'Neill, J., Hu, Z., Yang, Y.-G., Chen, Y.-W., Mumau, M., Green, M.D., Vunjak-Novakovic, G., Bhattacharya, J., et al. (2014). Efficient generation of lung and airway epithelial cells from human

pluripotent stem cells. Nat. Biotechnol. *32*, 84–91. https://doi.org/10.1038/nbt.2754.

62. Jacob, A., Morley, M., Hawkins, F., McCauley, K.B., Jean, J.C., Heins, H., Na, C.L., Weaver, T.E., Vedaie, M., Hurley, K., et al. (2017). Differentiation of human pluripotent stem cells into functional lung alveolar epithelial cells. Cell Stem Cell *21*, 472–488.e10. https://doi.org/10.1016/j.stem.2017.08.014.

63. Yamamoto, Y., Gotoh, S., Korogi, Y., Seki, M., Konishi, S., Ikeo, S., Sone, N., Nagasaki, T., Matsumoto, H., Muro, S., et al. (2017). Long-term expansion of alveolar stem cells derived from human iPS cells in organoids. Nat. Methods *14*, 1097–1106. https://doi.org/10.1038/nmeth.4448.

64. Chen, Y.-W., Huang, S.X., Carvalho, A.L.R.T. de, Ho, S.-H., Islam, M.N., Volpi, S., Notarangelo, L.D., Ciancanelli, M., Casanova, J.-L., Bhattacharya, J., et al. (2017). A three-dimensional model of human lung development and disease from pluripotent stem cells. Nat. Cell Biol. *19*, 542–549. https://doi.org/10.1038/ncb3510.

65. Tirosh, I., Izar, B., Prakadan, S.M., II, Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196. https://doi.org/10.1126/science.aad0501.

66. Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210. https://doi.org/10.1093/nar/30.1.207.

67. Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res. *29*, e45. https://doi.org/10.1093/nar/29.9.e45.

68. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

69. Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. *16*, 278. https://doi.org/10.1186/s13059-015-0844-5.

70. Hamilton, T.G., Klinghoffer, R.A., Corrin, P.D., and Soriano, P. (2003). Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. Mol. Cell. Biol. *23*, 4013–4025. https://doi.org/10.1128/MCB.23.11.4013-4025.2003.

71. Rock, J.R., Onaitis, M.W., Rawlins, E.L., Lu, Y., Clark, C.P., Xue, Y., Randell, S.H., and Hogan, B.L.M. (2009). Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc. Natl. Acad. Sci. USA *106*, 12771–12775. https://doi.org/10.1073/pnas.0906850106.

72. Alysandratos, K.-D., Russo, S.J., Petcherski, A., Taddeo, E.P., Acín-Pérez, R., Villacorta-Martin, C., Jean, J.C., Mulugeta, S., Rodriguez, L.R., Blum, B.C., et al. (2021). Patient-specific iPSCs carrying an SFTPC mutation reveal the intrinsic alveolar epithelial dysfunction at the inception of interstitial lung disease. Cell Rep. *36*, 109636. https://doi.org/10.1016/j.celrep.2021.109636.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Anti-NKX2-1 Antibody	Abcam	ab76013; RRID: AB_1310784
Rabbit Anti-Sox9 Antibody	Abcam	ab185966; RRID: AB_2728660
APC anti-mouse CD45 Antibody	BioLegend	Cat#103112; RRID: AB_312977
APC anti-mouse CD31 Antibody	BioLegend	Cat#102410; RRID: AB_312905
BV421 Rat Anti-Mouse CD326	BD Biosciences	Cat#563214; RRID: AB_2738073
Alexa Fluor 488 anti-GFP Antibody	BioLegend	Cat#338007; RRID: AB_2563287
Goat Anti-GFP Antibody	US Biological	G8965-01E
Chicken Anti-GFP Antibody	Aves Labs	GFP-1020; RRID: AB_10000240
Hamster Anti-PDPN Antibody	Thermo Fisher	Cat#14-5381-82; RRID: AB_1210505
Rabbit Anti-Pro-Sftpc Antibody	Abcam	ab211326; RRID: AB_2927746
Goat Anti-RFP Antibody	My Biosource	Cat#MBS448122
Rat Anti-Mki67 Antibody	Thermo Fisher	Cat#14-5698-82; RRID: AB_10854564
PE anti-mouse ITGB4 Antibody	BioLegend	Cat#123610; RRID: AB_2563544
Mouse Anti-Hopx Antibody	Santa Cruz	sc-398703; RRID: AB_2687966
Alexa Fluor 546 goat anti-hamster IgG (H+L)	Thermo Fisher	Cat#PA1-32045; RRID: AB_10985178
Alexa Fluor 647 donkey anti-rabbit IgG (H+L)	Thermo Fisher	Cat#A32795; RRID: AB_2762835
Alexa Fluor 546 donkey anti-goat IgG (H+L)	Thermo Fisher	Cat#A-11056; RRID: AB_2534103
Alexa Fluor 488 donkey anti-goat IgG (H+L)	Thermo Fisher	Cat#A-11055; RRID: AB_2534102
Cy3 AffiniPure F(ab')2 fragment donkey anti-rat IgG (H+L)	Jackson Immuno Research	712-166-153; RRID: AB_2340669
Alexa Fluor 488 donkey anti-chicken IgY (IgG) (H+L)	Jackson Immuno Research	703-545-155; RRID:AB_2340375
Bacterial and virus strains		
pHAGE-EF1α_L-GFP-W lentivirus	Darrell Kotton Lab	N/A
pHAGE-EF1α_L-dsRed-W lentivirus	Darrell Kotton Lab	N/A
Chemicals, peptides, and recombinant proteins		
IMDM	Gibco	12440053
Ham's F12 Media	Corning	10-080-CV
B27 supplement with RA	Gibco	17504044
N2 Supplement	Gibco	17502048
Glutamax	Gibco	35050061
Bovine Albumin Fraction V	Gibco	15260037
Ascorbic Acid	Sigma	A4544-25G
Thioglycerol	Sigma	M6145-25ML
Primocin	InvivoGen	ant-pm-2
rhFGF2	R&D Systems	233-FB
rhFGF10	R&D Systems	345-FG
Heparin	Sigma-Aldrich	H3149
Y-27632	Tocris	1254
rmWnt3a	R&D Systems	1324-WN
Advanced DMEM/F12	Gibco	12634010

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
rmFgf9	R&D Systems	7399-F9
rmFgf10	R&D Systems	6224-FG
CHIR99021	Tocris	4423
rmEGF	R&D Systems	2028-EG
A 83-01	Tocris	2939
BIRB796	Tocris	5989
Insulin	Roche	11376497001
Transferrin	Roche	10652202001
DMEM	Gibco	Cat#2414671
2-Mercaptoethanol	Gibco	Cat#21985023
rhBmp4	R&D Systems	Cat#314-BP
rmNoggin	R&D Systems	Cat#1967-NG
SB431542	Sigma	Cat#S4317
Retinoic acid	Sigma	Cat#R2625
fetal bovine serum	Gibco	Cat#16141079
DMEM	Gibco	Cat#2414671
Dispase	Gibco	Cat#17105-041
Collagenase Type IV	ThermoFisher	Cat#17104019
Papain	Worthington	LS003119
TrypLE express	Thermo Fisher	12604013
0.05% Trypsin-EDTA	Gibco	Cat#25300062
Hoecsht 33342	Thermo Fisher	Cat # H3570
Normal Donkey Serum	Jackson Immuno Research	017-000-121
Antigen Unmasking Solution, Citric Acid Based	Vector Laboratories	H-3300-250
Bleomycin sulfate from *Stretomyces verticillus*	Millipore Sigma	B8416-15UN
Matrigel	Corning	356231
ProLong Diamond Antifade Mountant	Invitrogen	Cat#P36965
HistoGel	Thermo Fisher	22-110-678
Paraformaldehyde	Ted Pella	18505
Critical commercial assays		
RNeasy Mini Kit	QIAGEN	Cat#74104
QIAzol Lysis Reagent QIAGEN	QIAGEN	Cat#79306
RLT Plus lysis buffer	QIAGEN	Cat#1053393
TaqMan Fast Universal PCR Master Mix (2X)	Thermo Fisher	Cat#4364103
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	Cat#4368814
Click-iT EdU Cell Proliferation Kit	ThermoFisher	Cat#C10340
Deposited data		
Sequence Data	This Paper	GEO Super Series GEO: GSE200886
Experimental models: Cell lines		
Nkx2-1^{mCherry} mouse ES Cell Line	Rossant Lab[29]	N/A
Experimental models: Organisms/strains		
UBC-GFP mice	Jackson Labs	JAX004353
C57BL/6J mice	Jackson Labs	JAX000664
129X1/SvJ mice	Jackson Labs	JAX000691
129S1/SvImJ mice	Jackson Labs	JAX002448
NSG mice	Jackson Labs	JAX005557
PDGFRa-EGFP	Jackson Labs	JAX007669

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Sox9-IRES-EGFP	Jackson Labs	JAX030137
Oligonucleotides		
Abca3	Thermo Fisher	Mm00550501_m1
Afp	Thermo Fisher	Mm00431715_m1
Ager	Thermo Fisher	Mm01161340_g1
Alb	Thermo Fisher	Mm00802090_m1
Apoe	Thermo Fisher	Mm01307192_m1
Axin2	Thermo Fisher	Mm00443610_m1
Etv5	Thermo Fisher	Mm00465816_m1
Foxj1	Thermo Fisher	Mm01267279_m1
Foxp2	Thermo Fisher	Mm00475030_m1
Id2	Thermo Fisher	Mm00711781_m1
Krt5	Thermo Fisher	Mm01305291_g1
Lamp3	Thermo Fisher	Mm00616604_m1
Mki67	Thermo Fisher	Mm01278617_m1
Napsa	Thermo Fisher	Mm00492829_m1
Nkd1	Thermo Fisher	Mm00471902_m1
Nkx2-1	Thermo Fisher	Mm00447558_m1
Notum	Thermo Fisher	Mm01253273_m1
Pax8	Thermo Fisher	Mm00440623_m1
Scgb3a2	Thermo Fisher	Mm00504412_m1
Sftpb	Thermo Fisher	Mm00455678_m1
Sftpc	Thermo Fisher	Mm00488144_m1
Slc34a2	Thermo Fisher	Mm01215846_m1
Sox2	Thermo Fisher	Mm03053810_s1
Sox9	Thermo Fisher	Mm00448840_m1
Tgfb2	Thermo Fisher	Mm00436955_m1
Trp63	Thermo Fisher	Mm00495793_m1
Software and algorithms		
FlowJo Software v.10.8.1	Becton Dickinson & Company	https://www.flowjo.com/solutions/flowjo
Seurat v.3	Sajita Lab[65]	https://github.com/satijalab/seurat
SPRING	Klein Lab[37]	https://github.com/AllonKleinLab/SPRING_dev
Graphpad Prism v.9.5.1	GraphPad Software	https://www.graphpad.com/features
ImageJ v.2.1.0/1.53i	NIH	https://imagej.net/ij/index.html
scTOP	Pankaj Mehta Lab	Figshare:
scTOP (analysis code to make figures)	Pankaj Mehta Lab	Figshare: https://doi.org/10.6084/m9.figshare.23796063

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Darrell Kotton (dkotton@bu.edu).

Materials availability

Research reagents generated in this study will be distributed upon request to other investigators.

Data and code availability

- The scRNA-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus[66] and are accessible through GEO Series accession numbers GEO: GSE200886, GEO: GSE200883, GEO: GSE200884, and GEO: GSE200885 and will also be available on the Kotton Lab's Bioinformatics Portal at http://www.kottonlab.com.

- All original code has been deposited at Figshare and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

C57BL/6J (JAX Strain #000664), 129X1/SvJ (JAX Strain #000691), 129S1/SvImJ (JAX Strain #002448), UBC-GFP (JAX Strain #004353), NSG (JAX Strain #005557), Sox9-IRES-EGFP (JAX Strain #030137) and PDGFRa-EGFP (JAX Strain #007669) mice were obtained from Jackson Laboratory. 129X1/S1 transplant recipient mice were generated by crossing 129X1/SvJ females with 129S1/SvImJ males. For all transplant experiments healthy 8-15-week-old male and female mice were used. All mouse studies were approved by the Institutional Animal Care and Use Committee of Boston University School of Medicine. All mice were maintained in facilities overseen by the Animal Science Center at Boston University.

Mouse cell lines

Primary tip-like cells were generated from individual male and female E12.5 mouse lungs. These lungs were digested in TrypLE Express Enzyme for 15 minutes and broken up through repeated pipetting. EpCAM+/CD45-/CD31- live cells were then sort purified and cultured at 40-200 cells/ul in LPM conditions. For P0 RT-qPCR multiple lungs were pooled to generate sufficient sample, but lines used for continued culturing were all generated from individual lungs. The sex of primary lines was determined by SRY PCR and male lines were used for all transplants.

ESC cells were generated by and obtained from the Rossant Lab.

METHOD DETAILS

Specification and purification of ESC-derived lung epithelial progenitors

As previously described,[28] NKX2-1^{mCherry} mouse ESCs were differentiated into definitive endoderm by culturing in cSFDM for 2.5 days, trypsinizing cells to a single-cell suspension, and culturing in cSFDM supplemented with Activin A (50 ng/ml) for another 2.5 days. The resulting embryoid bodies were then grown in suspension in cSFDM supplemented with SB431542 (10 uM) and rmNoggin (100ng/ml). After one day in culture, these embryoid bodies were trypsinized and plated on six well plates coated in 100ul of Matrigel at 2e6 cells/well in cSFDM supplemented with rhBMP4 (10ng/ml), Wnt3a (100ng/ml), and Y-27632 ROCK inhibitor (10uM). Cells were fed the same media the next day and then fed daily with cSFDM supplemented with just rhBMP4 and Wnt3a. Where indicated in the text results, the media was supplemented with RA (100nM) from day 6 to day 8 or rmFgf10 (50ng/ml) from day 8 to day 14. On day 13 or day 14 the cells were incubated at 37C for 1 hour in 1mg/ml each of Collagenase IV and Dispase to digest the Matrigel bed. In cases where the percent of EpCAM+ cells was not being measured two slow spins (100xg) and washes were used to enrich for the undigested epithelial spheres. Epithelial spheres were then trypsinized to generate a single-cell suspension, and NKX2-1^{mCherry}+/EpCAM+ live cells were assessed by flow cytometry or sort purified for further cell culture.

Differentiation of ESC-derived lung epithelial progenitors

In order to further differentiate day 13 or day 14 NKX2-1^{mCherry}+/EpCAM+ live cells into lung bud tip-like cells, sorted cells were resuspended in Matrigel droplets at either 200 cells/ul (LPM) or 500 cells/ul (Proximal and Distal Media) and fed every two days until collection. In order to passage cells grown in LPM, the Matrigel droplets were incubated at 37C for 1 hour in 1mg/ml each of Collagenase IV, Dispase, and Papain with pipetting every 30 minutes. The resulting single-cell suspension was resuspended in LPM and counted on a hemocytometer. These cells were then resuspended in Matrigel droplets as above. Cells are passaged once every 7 days. These cells were frozen in fetal bovine serum with 10% DMSO.

EdU labeling of ESC-derived tip-like cells in vitro

EdU labeling was performed according to manufacturer's protocol (ThermoFisher, Cat #C10634). In brief, ESC-derived tip-like cells were seeded in 3D Matrigel in 6-well plates in LPM media. Cells were treated with EdU (10μM) or DMSO vehicle control for the last 6, 24, 30, or 48 hours prior to collection of day 7 post-passaging. Subsequently, cells were dissociated using papain, dispase, and collagenase-containing solution for 1h prior to fixation with 4% paraformaldehyde for 15 min at room temperature. In order to label mCherry+ cells, fixed cells were treated with a goat anti-RFP antibody (1:250) for 1h at room temperature in 1X saponin-based permeabilization buffer and then with Alexa Fluor 546 donkey anti-goat (1:500) 30 mins at room temperature. Cells were then incubated in the Click-iT Plus detection cocktail for 30 min, washed, and analyzed by flow cytometry.

Reverse transcriptase quantitative PCR (RT-qPCR)

RNA was isolated according to manufacturer's instructions using the QIAGEN miRNeasy mini kit (QIAGEN). cDNA was generated by reverse transcription of up to 100ng RNA from each sample using the Applied Biosystems High-Capacity cDNA Reverse Transcription Kit. For qPCR, technical duplicates of each of at least three biological replicates were run for 40 cycles as 10ul reactions. All

primers were TaqMan probes Table S8 from Applied Biosystems and the qPCR reactions were performed on an Applied Biosysytems Quantstudio 6 Flex. Relative gene expression was normalized to an 18S control and reported as a fold change relative to a control sample in the experiment (i.e. fold change calculated as 2^{-DDCT} based on the method of Pfaffl[67]). Samples that were undetectable were assigned a CT value of 40 to allow for fold change calculations.

Immunohistochemistry

Mice were euthanized and lungs were inflation-fixed with 4% paraformaldehyde (PFA) prior to overnight fixation in 4% PFA. Cells grown in vitro were embedded in HistoGel prior to overnight fixation in 4% PFA. All samples were then dehydrated and embedded in paraffin for sectioning with a microtome (8um sections). The resulting slides were deparaffinized, blocked using normal donkey serum, and then stained with up to three primary antibodies overnight at 4C (see key resources table for antibodies used). The next day slides were stained with Hoecsht and up to three secondary antibodies (see key resources table for antibodies used) for 1 hour at room temperature and mounted with ProLonged Diamond Antifade Mountant. Stained slides were imaged using a Leica SP5 Confocal Microscope. For MKI67 quantification at 3 days post transplantation, 2 separate slides were analyzed for each mouse and pictures were taken for all visible GFP+/NKX2-1+ donor-derived cells.

Transplant protocol including recipient generation

Transplant recipients were syngeneic to the cells they received, unless otherwise indicated in the the text. For UBC-GFP primary tip-like cell transplants all recipients were C57BL/6J male mice. For all other transplants, recipient mice (129X1/S1) were generated by crossing 129X1/SvJ females with 129S1/SvImJ males. Recipient mice were at least 8 weeks old at the time of injury and we used both male and female 129X1/S1 mice with no clear difference in transplantation success. Ten days prior to cell transplantation mice were given 1.5U/kg bleomycin by oral pharyngeal delivery. Mice were anesthetized with isoflurane until they displayed agonal breathing, at which point mice were hung vertically by their teeth, their tongue was pulled out using blunt forceps, and a p200 was used to administer the liquid orally. On the day of transplantation donor cells were digested down to a single-cell suspension, as described above for passaging. These cells were suspended in LPM and left in a 37C incubator for 2-3 hours, with flicking every 30 minutes, to recover from digestion. At the end of this period cells were counted on a hemocytometer and resuspended in LPM (no more than 50ul/mouse with cell numbers indicated in the text results). These cell suspensions were delivered intratracheally, similar to bleomycin. Transplantation into NSG mice was performed by identical methods, except bleomycin was given 3 days prior to transplantation.

Adult lung digestion to a single-cell suspension

Mice were euthanized and perfused through injection of PBS into the right ventricle. The lungs were washed three times with 1ml of PBS administered through the trachea. Lungs were then inflated with 1.5 ml of digestion buffer (9.5U/mL Elastase, 20U/mL Collagenase, 5U/mL Dispase) followed immediately by up to 0.5ml of 1% low melt agarose and tied off with suture. These lungs were incubated in PBS on ice for 5 minutes before dissecting off lobes and placing them into 3.5 ml of digestion buffer. Lungs were incubated at 37C on a rocker for 40 minutes before being dissociated with frequent pipetting using a 10ml pipette. Cells were passed through 70um and then 40um cell strainers. If necessary, red cell lysis buffer was used to remove red blood cells. Cells were then resuspended in FACS buffer (2% FBS in PBS) and stained as described below.

Flow cytometry analysis and fluorescence-activated cell sorting (FACS)

Single-cell suspensions were prepared as described for passaging (in vitro samples) or lung digestion. When necessary, cells were stained with conjugated antibodies for 30 minutes in FACS buffer (2% FBS in PBS) at 4C and then resuspended in FACS buffer with 1:100 DRAQ7 (live/dead stain). FACS was performed on either a Beckman Coulter MoFLo Astrios or BD FACSARIA II SORP. Flow analysis was performed on a Beckman Coulter MoFLo Astrios, BD LSR II SORP, or Stratedigm S1000EXi. Resulting plots were further analyzed using FlowJo v10.7.1. Lung epithelial cells were isolated by selecting for EpCAM+/CD31-/CD45-/DRAQ7- cells, with donor-derived lineages being identified by the presence of GFP, DsRed, or mCherry depending on the experiment. Flow analysis of primary and donor-derived epithelial cells contained both airway and alveolar epithelial cells, except in Figures 2H and S2D, where ITGB4 was used to separate the two populations.

Single-cell RNA-sequencing

Single-cell suspensions were prepared and FACS purified on a Beckman Coulter MoFlo Astrios cell sorter as described above to collect live cell populations described in the results section. Single-cell RNA-sequencing was performed using the Chromium Single Cell 3′ system (10X Genomics) at the Single Cell Sequencing Core at Boston University Medical Center according to the manufacturer's instructions (10X Genomics). The resulting samples were demultiplexed using Cell Ranger and mapped using STARsolo to the GRCm38 mouse genome reference extended with GFP and mCherry transcripts. Downstream analysis and quality controls were performed on Seurat v3.2.3.[68] We excluded from analysis cell doublets, cells containing more than 15% of mitochondrial RNA reads and cells with less than 800 genes detected (indicative of dying cells). We used SCTransform for normalization, regressing out the effect of unwanted sources of variation like that of the mitochondrial reads percentage. Cell cycle regression was likewise used to remove the differences between G2M and S phase cells. The Nearest Neighbors graph and Louvain clustering was based on the top 20 Principal Components. The data was plotted first using UMAP. In the case of in vivo samples, UMAP and Louvain clustering was used to identify and remove non-epithelial cells, as detailed in Figure S5. Samples from distinct runs were then combined using

harmonization,[39] and again clustered using Louvain clustering. Clusters that contained fewer than 10 cells and/or overlapped other clusters were combined with overlapping clusters to avoid the multiplicity of clusters that aren't biologically relevant. In order to facilitate cluster annotation, cell cycle and other molecular signature enrichment were scored using the method described in Tirosh et al.[65] Multimodal reference mapping was performed using Seurat v4 reference mapping. Differential expression tests were run using MAST,[69] with prior gene filters to reduce the burden of multiple test corrections (min.pct = 0.25, logfc.threshold = 0.25). DEG heatmaps included a designated number of both up- and down-regulated genes with the greatest log fold change that have an adjusted p value less than 0.05. FGSEA was performed using REACTOME, KEGG, and C2 reference gene sets or just REACTOME and KEGG. When FGSEA was graphically visualized (Figure S3B), only the 15 REACTOME or KEGG gene sets with the lowest p value were visualized to reduce visualization of redundant gene sets.

For comparison to data from Zepp et al.[37] the indicated timepoints were combined into a single UMAP and analyzed using Louvain clustering. Cell types of interest were identified based on known lineage markers (*EpCAM*, *Nkx2-1*, *Sox9*, *Sftpc*). The selected cell populations were then integrated with our scRNA-seq data on primary and ESC-derived tip-like cells for analysis.

Cell type signatures

Cell type signatures were used to identify lung epithelial cell types in scRNA-seq. Cell types were originally identified in two separate uninjured wild-type lung data sets (Control [GSM606035] from GSE200884 and Huang Protocol Epithelial Cells [GSM6046033] from GSE200883) based on Louvain clustering and known cell type markers. We then identified genes that were upregulated in a specific cell relative to all other lung epithelial cells. For cell types identified in both data sets (AT1, AT2, secretory, and ciliated cells) we took all genes that were in the 60 most enriched genes (by z score) for both data sets. For cell types found only in one data set (basal and neuroendocrine cells) we took the 20 most enriched genes (by z score) for that data set (Table S3).

scTOP Methods

The Python package Single-Cell Type Order Parameters (scTOP) was used to calculate alignment scores for endogenous and donor-derived cell populations. This algorithm pre-processes scRNA-seq data then finds the projection of a sample onto the space of known cell types and is described in detail in Yampolskaya et al.[51] In brief, data was first pre-processed and normalized to reduce batch effects. To pre-process individual cells, the vector of raw RNA counts for each cell was normalized independently, 1 was added to each entry of the vector, the logarithm was taken, and the resulting data was fit onto a log-normal distribution. Next, a z-score was assigned to each gene. To do so, the vector components were first assigned a rank from least to greatest. Each rank was then divided by the total number of genes, which gave the probability that the value of a variable drawn from a normal distribution is equal to or less than that data point. Finally, the resulting percentile function was applied to a normal distribution with mean 0 and standard deviation 1. To pre-process aggregates of cells and thus find the pre-processed gene expression profile of a particular cell population, the same process is used as for individual cells, except in the very first step the average raw RNA counts of the population is used rather than the individual counts.

To find the cell type alignments for a sample, each sample's gene expression profile was projected onto the subspace of cell types. In the following equation, the reference basis of cell types is denoted by ξ, which is a p (number of cell types in the reference basis) by n (number of genes) matrix. The sample is represented by a vector in gene expression space S. This is a vector of length n. Each sample in gene expression space is projected onto the hyperplane of cell type space. Thus, the sample vector in gene expression space is broken into a component that lies on the hyperplane and a component perpendicular to the hyperplane ($S_\perp$).

$$S = \xi^T a + S_\perp$$

The component that lies on the hyperplane is a linear combination of all the cell type vectors. a is a p-length vector of the cell type components of the projected sample. The equation to find these components is:

$$a = \left(\xi\xi^T\right)^{-1}\xi S$$

The alignment scores given in the paper represent the alignment of the sample with the indicated cell type. Individual or aggregate alignment scores were found by pre-processing and projecting the gene expression of a single cell or the average gene expression profile of a cell population, respectively.

To create the reference basis ξ for adult mouse cell types, we pre-processed data from the Mouse Cell Atlas.[50] Since the number of lung cells sampled by the Mouse Cell Atlas were relatively low for the cell types of interest, the AT1, AT2, ciliated, club, and basal cell expression profiles were taken from an uninjured control lung sample (Figure 4), as indicated in the text. To create the reference basis from the raw count data of the atlases, the scRNA-seq counts were averaged across all cells of each cell type, then the aggregate gene expression profile was pre-processed as previously described.

Primary cell coculture assay

Coculture protocol was used as previously described.[54] Pdgfra^{nGFP} animals were used to sort lung fibroblasts for cocultures.[70] GFP+ donor-derived cells and GFP- endogenous cells were sorted from a transplant recipient at least 6 weeks after transplantation. In both cases lungs were digested as described above, but with a different digestion buffer (4U/mL Elastase, 400U/mL Collagenase, 5U/mL Dispase) and no agarose. Donor-derived or endogenous epithelia (5,000 cells) were then cultured with Pdgfra^{nGFP}+ fibroblasts at a

1:20 ratio in 1:1 growth factor reduced 3D Matrigel with MTEC-plus medium.[71] 90ul of matrigel and cell suspension was added to a 24-well 0.4-μm transwell insert (Falcon). Cells were cultured in MTEC-plus medium for 21 days. The resulting organoids were fixed with 4% neutral PFA and embedded in paraffin. Sections (7um thickness) were stained and imaged using standard immunofluorescence confocal microscopy protocols, as detailed above.

TEM imaging of sorted cells

TEM imaging was performed as previously described.[72] In short, GFP+ and GFP- lung epithelial cells were collected as described above and combined with mouse embryonic fibroblasts to generate sufficient cell pellets. These cells were fixed 3 hours total in 2.0% glutaraldehyde (Ladd Research) + 1% paraformaldehyde in 0.1% cacodylate buffer (pH 7.4) at room temperature as follows: an equal volume of 4% glutaraldehyde + 2% paraformaldehyde/ 0.1 M cacodylate was added to Eppendorf tubes containing the cells in known volume of FACS buffer (2% FBS in PBS), fixed for 1.5 hours, and centrifuged gently (300 g for 1 minute) followed by the addition of fresh 2% glutaraldehyde + 1% paraformaldehyde/ 0.1 M cacodylate for an additional 1.5 hours at room temperature. The samples were then washed 3 times in 0.1 M cacodylate and post-fixed overnight in 1.5% osmium tetroxide (Polysciences) in 0.1 M cacodylate buffer in dark at 4°C. After washing, the pellet was embedded in 2% low gelling temperature agarose in Cacodylate buffer heated to 60C, quickly spun at 300rcf and placed on ice. The pellet was removed from the Eppendorf tube, placed in a glass vial, washed 3-4 times in 0.05 M Na Maleate buffer (pH 5.2) and block stained in 1.5% Uranyl acetate (Electron Microscopy Sciences, EMS) in 0.025 M Na Maleate buffer (pH 6.0). Next, the samples were dehydrated quickly through acetone on ice, from 70% to 80% to 90%. The samples were then incubated 2 times in 100% acetone at room temperature for 10 minutes each, and in propylene oxide at room temperature for 15 minutes each. Finally, the samples were changed into EMbed 812 (EMS), left for 2 hours at room temperature, changed into fresh Embed 812 and left overnight at room temperature, after which they were embedded in fresh EMbed 812 and polymerized overnight at 60°C. Plastic embedded samples were thin sectioned at 70 nm and grids were stained in 4% aqueous Uranyl Acetate for 5 minutes at 60°C followed by Lead Citrate for 10 minutes at room temperature. Electron microscopy was performed on a Philips CM12 EM operated at 100kV, and images were recorded on a TVIPS F216 CMOS camera with a pixel size of 4.18 nm per pixel.

EdU labeling of transplant recipient mice

EdU labeling was performed according to manufacturer's protocol (ThermoFisher, Cat #C10634). In brief, EdU was dissolved in PBS to a concentration of 5mg/ml. This solution was given through intraperitoneal injections (10ul/g mouse weight) every 2-3 days until mouse collection at either 14 days post transplant or 20 days post secondary injury. Mouse lungs were pruned and dissected down to a single cell suspension as described above. Cells were then fixed using 4% paraformaldehyde for 30 minutes at room temperature. Fixed cells were labeled with the Click-iT Plus detection cocktail for 30 min, washed with Brilliant Violet 421™ rat anti-mouse CD326 (1:200), PE rat anti-mouse CD31 (1:500), PE rat anti-mouse CD45 (1:500), and Alexa 488 rat anti-GFP (1:200) for 1h at room temperature in 1X saponin-based permeabilization buffer, washed, and analyzed by flow cytometry.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details relevant to RT-qPCR or Flow assessment are outlined in the figure legends. Unpaired, two-tailed Student's t tests were used for comparisons involving only two groups, while ANOVA was used when considering multiple groups. Significance was defined as $p < 0.05$.

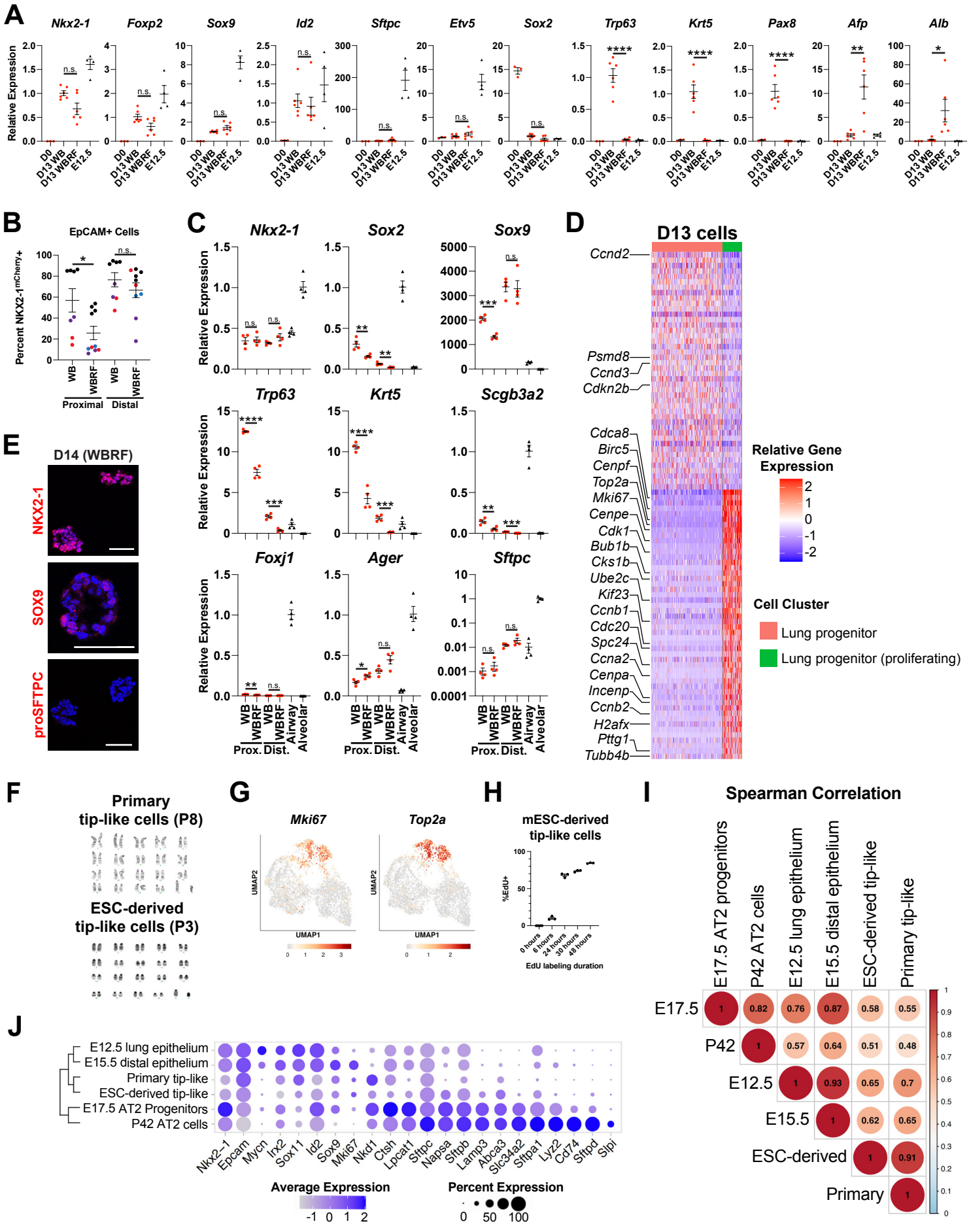
Cell Stem Cell, Volume *30*

Supplemental Information

Durable alveolar engraftment of PSC-derived lung epithelial cells into immunocompetent mice

Michael J. Herriges, Maria Yampolskaya, Bibek R. Thapa, Jonathan Lindstrom-Vautrin, Feiya Wang, Jessie Huang, Cheng-Lun Na, Liang Ma, McKenna M. Montminy, Pushpinder Bawa, Carlos Villacorta-Martin, Pankaj Mehta, and Darrell N. Kotton

Fig. S1


A
Nkx2-1
Foxp2
Sox9
Id2
Sftpc
Etv5
Sox2
Trp63
Krt5
Pax8
Afp
Alb
Relative Expression
D0
D13 WB
D13 WBRF
E12.5
B
EpCAM+ Cells
Percent NKX2-1mCherry+
WB
WBRF
Proximal
Distal
C
Nkx2-1
Sox2
Sox9
Trp63
Krt5
Scgb3a2
Foxj1
Ager
Sftpc
Relative Expression
Prox.
Dist.
Airway
Alveolar
D
D13 cells
Ccnd2
Psmd8
Ccnd3
Cdkn2b
Cdca8
Birc5
Cenpf
Top2a
Mki67
Cenpe
Cdk1
Bub1b
Cks1b
Ube2c
Kif23
Ccnb1
Cdc20
Spc24
Ccna2
Cenpa
Incenp
Ccnb2
H2afx
Pttg1
Tubb4b
Relative Gene Expression
Cell Cluster
Lung progenitor
Lung progenitor (proliferating)
E
D14 (WBRF)
NKX2-1
SOX9
proSFTPC
F
Primary tip-like cells (P8)
ESC-derived tip-like cells (P3)
G
Mki67
Top2a
UMAP1
UMAP2
H
mESC-derived tip-like cells
%EdU+
0 hours
6 hours
24 hours
30 hours
48 hours
EdU labeling duration
I
Spearman Correlation
E17.5 AT2 progenitors
P42 AT2 cells
E12.5 lung epithelium
E15.5 distal epithelium
ESC-derived tip-like
Primary tip-like
E17.5
P42
E12.5
E15.5
ESC-derived
Primary
1
0.82
0.76
0.87
0.58
0.55
0.57
0.64
0.51
0.48
0.93
0.65
0.7
0.62
0.65
0.91
J
E12.5 lung epithelium
E15.5 distal epithelium
Primary tip-like
ESC-derived tip-like
E17.5 AT2 Progenitors
P42 AT2 cells
Nkx2-1
Epcam
Mycn
Irx2
Sox11
Id2
Sox9
Mki67
Nkd1
Ctsh
Lpcat1
Sftpc
Napsa
Sftpb
Lamp3
Abca3
Slc34a2
Sftpa1
Lyz2
Cd74
Sftpd
Slpi
Average Expression
Percent Expression

Supplemental Figure 1: Characterization of ESC-derived Lung Epithelial Progenitors and ESC-derived Tip-like Cells, Related to Figures 1, 2

(A) Analysis of gene expression by RT-qPCR at day 13 of WB and WBRF lung specification protocols compared to ESCs (D0) and freshly collected E12.5 SOX9+ epithelium. n.s. not significant, * $p<0.05$, ** $p<0.01$, **** $p<0.0001$ by unpaired, two-tailed Student's t-test. n= 3, 6, 6, 4 biological replicates. Error bars = mean +/- SEM.

(B) The percent of cells that were NKX2-1^{mCherry}+ based on specification (WB vs WBRF) and differentiation (Proximal vs Distal) protocols. Dot color indicates biological replicates performed in the same batch. n.s. not significant, * $p<0.05$ by one-way ANOVA. n= 8, 10, 8, 10 biological replicates. Error bars = mean +/- SEM.

(C) RT-qPCR analysis of gene expression in day 26 NKX2-1^{mCherry}+ cells of proximal and distal differentiation protocols. Cultured cells were compared to freshly collected epithelial cells from adult mouse lungs (Airway and Alveolar). n.s. not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$ by unpaired, two-tailed Student's t-test. n= 4 biological replicates. Error bars = mean +/- SEM.

(D) Row-normalized heatmap of the top 50 most up-regulated and top 50 most down-regulated genes (with adj. p-value <0.05, ordered by logFC) between the clusters Lung progenitor and Lung progenitor (proliferating) in day 13 of the WBRF specification protocol. Annotated genes are associated with proliferation.

(E) Representative immunofluorescence confocal microsocopy of paraffin tissue sections prepared from day 14 cells cultured in the WBRF specification protocol. Staining performed with antibodies against NKX2-1, SOX9, or proSFTPC. Nuclei stained with Hoechst, scale bars are 200um.

(F) Representative G-banding indicates karyotypically normal primary and ESC-derived tip-like cells.

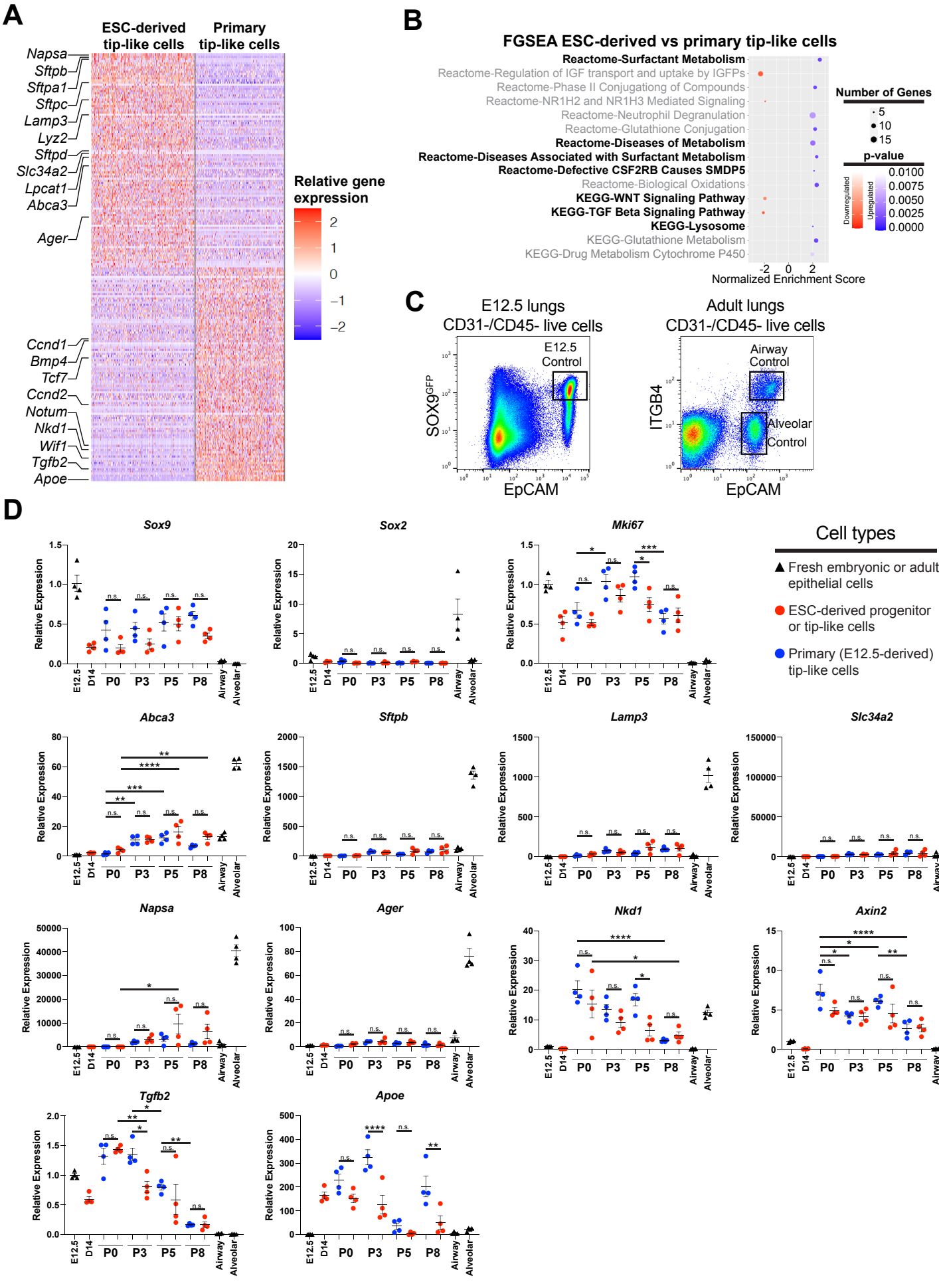
(G) UMAP plots displaying expression of *Mki67* and *Top2a* in primary and ESC-derived tip-like cells.

(H) Percent EdU labeling of ESC-derived tip-like cells with different durations of EdU labeling prior to collection.

(I) Spearman Correlation using the top 1000 most variable genes when comparing ESC-derived tip-like cells, primary tip-like cells, and primary distal lineages[36].

(J) Hierarchical clustering of cell populations from figure S1I based on expression of lung epithelial markers and AT2 differentiation markers.

Fig. S2


A
ESC-derived tip-like cells
Primary tip-like cells
Napsa
Sftpb
Sftpa1
Sftpc
Lamp3
Lyz2
Sftpd
Slc34a2
Lpcat1
Abca3
Ager
Ccnd1
Bmp4
Tcf7
Ccnd2
Notum
Nkd1
Wif1
Tgfb2
Apoe
Relative gene expression
2
1
0
−1
−2
B
FGSEA ESC-derived vs primary tip-like cells
Reactome-Surfactant Metabolism
Reactome-Regulation of IGF transport and uptake by IGFPs
Reactome-Phase II Conjugationg of Compounds
Reactome-NR1H2 and NR1H3 Mediated Signaling
Reactome-Neutrophil Degranulation
Reactome-Glutathione Conjugation
Reactome-Diseases of Metabolism
Reactome-Diseases Associated with Surfactant Metabolism
Reactome-Defective CSF2RB Causes SMDP5
Reactome-Biological Oxidations
KEGG-WNT Signaling Pathway
KEGG-TGF Beta Signaling Pathway
KEGG-Lysosome
KEGG-Glutathione Metabolism
KEGG-Drug Metabolism Cytochrome P450
-2 0 2
Normalized Enrichment Score
Number of Genes
5
10
15
p-value
0.0100
0.0075
0.0050
0.0025
0.0000
Downregulated
Upregulated
C
E12.5 lungs
CD31-/CD45- live cells
E12.5 Control
SOX9GFP
EpCAM
Adult lungs
CD31-/CD45- live cells
Airway Control
Alveolar Control
ITGB4
EpCAM
D
Sox9
Sox2
Mki67
Abca3
Sftpb
Lamp3
Slc34a2
Napsa
Ager
Nkd1
Axin2
Tgfb2
Apoe
Relative Expression
E12.5
D14
P0
P3
P5
P8
Airway
Alveolar
Cell types
Fresh embryonic or adult epithelial cells
ESC-derived progenitor or tip-like cells
Primary (E12.5-derived) tip-like cells

Supplemental Figure 2: Transcriptomic Analysis of ESC-derived and Primary Tip-like Cells, Related to Figure 2

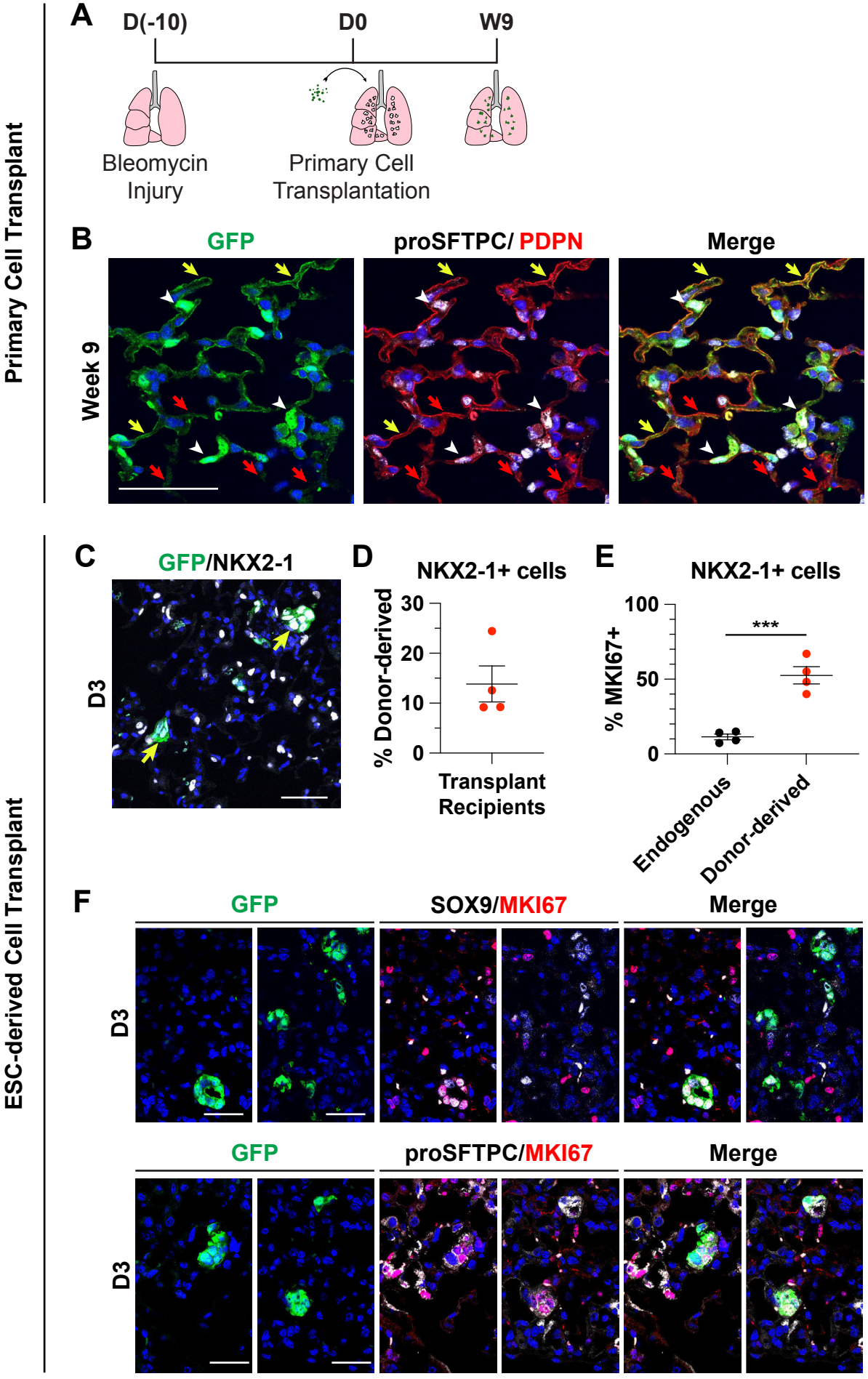
(A) Row-normalized heatmap of the top 100 most up-regulated and top 100 most down-regulated genes (with adj. p-value <0.05, ordered by logFC) between ESC-derived and Primary Tip-like Cells. Annotated genes are associated with surfactant metabolism or signaling pathways identified in supplemental figure 3B.

(B) FGSEA-identified Reactome and KEGG categories that are differentially regulated between ESC-derived and primary tip-like cells. The highlighted categories contain genes associated with AT2 function or prominent signaling pathways in lung development. Shown here are the 15 categories with the lowest p-value, a full list can be found in Table S2.

(C) Representative gating used to collect E12.5 SOX9+ epithelial cells and adult epithelial cells as primary controls for RT-qPCR.

(D) Analysis of gene expression by RT-qPCR. Primary and ESC-derived tip-like cells from multiple passages are compared against lung epithelial progenitors from day 14 of the WBRF protocol (D14) and freshly sorted lung epithelial cells from embryonic (E12.5) and adult (Airway and Alveolar) mouse lungs. n.s. not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$ by one-way ANOVA. n= 4 biological replicates. Error bars = mean +/- SEM.

Fig. S3


Primary Cell Transplant
A
D(-10)
D0
W9
Bleomycin Injury
Primary Cell Transplantation
B
GFP
proSFTPC/ PDPN
Merge
Week 9
ESC-derived Cell Transplant
C
GFP/NKX2-1
D3
D
NKX2-1+ cells
% Donor-derived
30
20
10
0
Transplant Recipients
E
NKX2-1+ cells
% MKI67+
100
50
0
***
Endogenous
Donor-derived
F
GFP
SOX9/MKI67
Merge
D3
GFP
proSFTPC/MKI67
Merge
D3

Supplemental Figure 3: Transplantation of Primary Tip-like cells into an Immunocompetent Recipient and Progenitor State of ESC-derived Tip-like Cells at 3 Days Post Transplantation, Related to Figure 3

(A) Schematic for transplantation of GFP+ primary cells into bleomycin injured lungs with later histological assessment of recipient lungs.

(B) Representative immunofluorescence confocal microscopy of donor-derived cells at 9 weeks post-transplantation with antibodies detecting GFP, proSFTPC, and PDPN. White arrowheads indicate cuboidal proSFTPC+/GFP+ cells, yellow arrows indicate thin PDPN+/GFP+ cells, and red arrows indicate PDPN+/GFP- endogenous AT1 cells. Nuclei stained with Hoechst, scale bar is 50um.

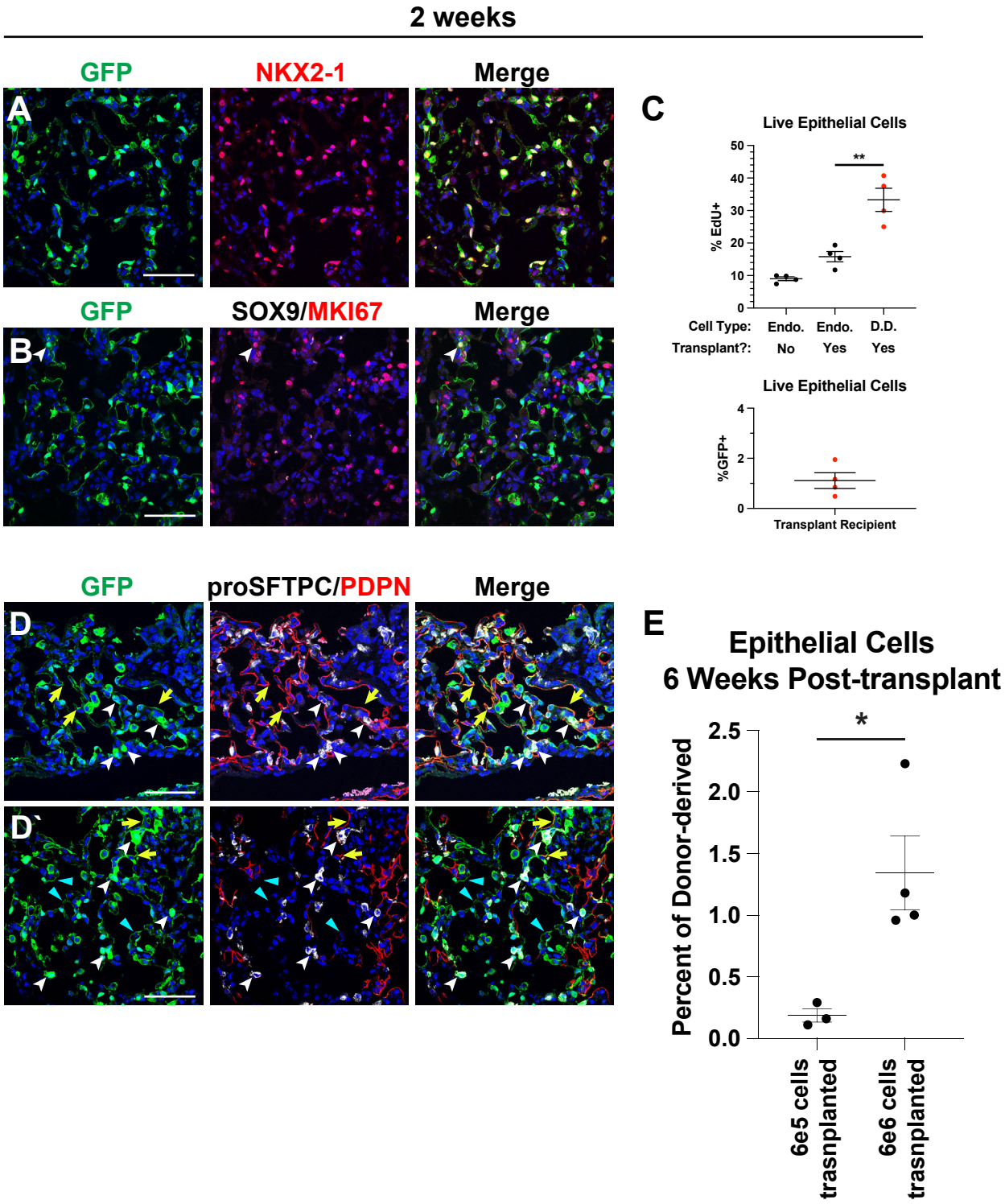
(C) Representative immunofluorescence confocal microscopy of lung tissue sections containing donor-derived cells at 3 days post-transplantation. Sections stained with antibodies detecting GFP and NKX2-1. Yellow arrows indicate representative GFP+/NKX2-1+ cells. Nuclei stained with Hoechst, scale bar is 50um.

(D) Percent of NKX2-1+ cells that were GFP+ donor-derived cells in analyzed confocal images from 3 days post-transplantation.

(E) Percent of endogenous and donor-derived NKX2-1+ cells that expressed MKI67.

(F) Representative immunofluorescence confocal microscopy of lung tissue sections containing donor-derived cells at 3 days post-transplantation. Sections stained with antibodies detecting GFP, proSFTPC, SOX9, and MKI67. Nuclei stained with Hoechst, scale bars are 25um.

Fig. S4


2 weeks
GFP
NKX2-1
Merge
A
GFP
SOX9/MKI67
Merge
B
C
Live Epithelial Cells
% EdU+
50
40
30
20
10
0
**
Cell Type:
Endo.
Endo.
D.D.
Transplant?:
No
Yes
Yes
Live Epithelial Cells
%GFP+
4
2
0
Transplant Recipient
GFP
proSFTPC/PDPN
Merge
D
D`
E
Epithelial Cells
6 Weeks Post-transplant
*
Percent of Donor-derived
2.5
2.0
1.5
1.0
0.5
0.0
6e5 cells trasnplanted
6e6 cells trasnplanted

Supplemental Figure 4: Donor-derived Cells Have Differentiated and Become Less Proliferative by 2 Weeks Post-transplantation, Related to Figure 3

(A) Representative immunofluorescence confocal microscopy of lung tissue sections indicating donor-derived cells (GFP+) at 2 weeks post-transplantation of ESC-derived tip-like cells. The majority of GFP+ donor-derived cells are NKX2-1+. Nuclei stained with Hoechst, scale bars are 50um.

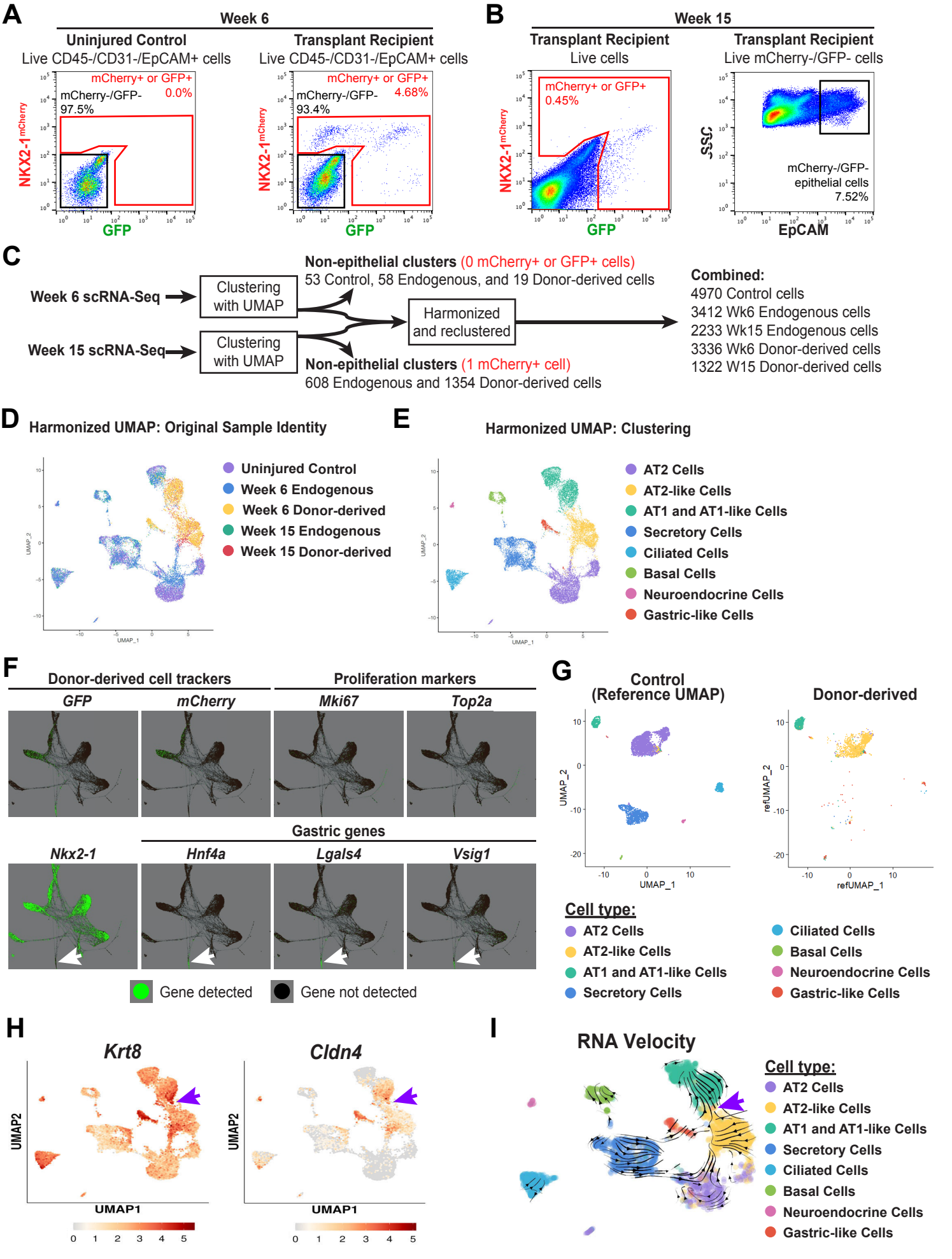
(B) At this time point donor-derived cells are SOX9- and only a small fraction are MKI67+ (white arrowhead). Nuclei stained with Hoechst, scale bars are 50um.

(C) Percent EdU labeling of endogenous and donor-derived cells for the first 2 weeks following transplantation of ESC-derived tip-like cells or media only and the percent of EdU+ cells that were GFP+. Lobes were pruned down to regions containing GFP+ cells, or similar regions in no transplant controls, prior to digestion into single cell suspension. Endo. = endogenous, D.D. = donor-derived. ** = p. value < 0.01 by unpaired, two-tailed Student's t-test. n= 4 biological replicates. Error bars = mean +/- SEM.

(D) At this time point donor-derived cells include both cuboidal proSFTPC+ (white arrowheads) and thin PDPN+ (yellow arrows) donor-derived cells. Some donor-derived clusters (D`) are primarily composed of cuboidal proSFTPC+ (white arrowheads) and thin PDPN- cells (blue triangles). Nuclei stained with Hoechst, scale bars are 50um.

(E) Flow cytometry quantitation of the percent of live epithelial (EpCAM+/CD45-/CD31-) cells that are donor-derived in whole lungs 6 weeks after transplantation of 6e5 or 6e6 ESC-derived tip-like cells. * = p. value < 0.05 by unpaired, two-tailed Student's t-test. n= 3,4 biological replicates. Error bars = mean +/- SEM.

Fig. S5


A
Week 6
Uninjured Control
Live CD45-/CD31-/EpCAM+ cells
mCherry+ or GFP+ 0.0%
mCherry-/GFP- 97.5%
NKX2-1 mCherry
GFP
Transplant Recipient
Live CD45-/CD31-/EpCAM+ cells
mCherry+ or GFP+ 4.68%
mCherry-/GFP- 93.4%
NKX2-1 mCherry
GFP
B
Week 15
Transplant Recipient
Live cells
mCherry+ or GFP+ 0.45%
NKX2-1 mCherry
GFP
Transplant Recipient
Live mCherry-/GFP- cells
SSC
mCherry-/GFP- epithelial cells 7.52%
EpCAM
C
Week 6 scRNA-Seq
Week 15 scRNA-Seq
Clustering with UMAP
Clustering with UMAP
Non-epithelial clusters (0 mCherry+ or GFP+ cells)
53 Control, 58 Endogenous, and 19 Donor-derived cells
Harmonized and reclustered
Non-epithelial clusters (1 mCherry+ cell)
608 Endogenous and 1354 Donor-derived cells
Combined:
4970 Control cells
3412 Wk6 Endogenous cells
2233 Wk15 Endogenous cells
3336 Wk6 Donor-derived cells
1322 W15 Donor-derived cells
D
Harmonized UMAP: Original Sample Identity
UMAP_2
UMAP_1
Uninjured Control
Week 6 Endogenous
Week 6 Donor-derived
Week 15 Endogenous
Week 15 Donor-derived
E
Harmonized UMAP: Clustering
UMAP_2
UMAP_1
AT2 Cells
AT2-like Cells
AT1 and AT1-like Cells
Secretory Cells
Ciliated Cells
Basal Cells
Neuroendocrine Cells
Gastric-like Cells
F
Donor-derived cell trackers
Proliferation markers
GFP
mCherry
Mki67
Top2a
Gastric genes
Nkx2-1
Hnf4a
Lgals4
Vsig1
Gene detected
Gene not detected
G
Control (Reference UMAP)
Donor-derived
UMAP_2
UMAP_1
refUMAP_2
refUMAP_1
Cell type:
AT2 Cells
AT2-like Cells
AT1 and AT1-like Cells
Secretory Cells
Ciliated Cells
Basal Cells
Neuroendocrine Cells
Gastric-like Cells
H
Krt8
Cldn4
UMAP2
UMAP1
0 1 2 3 4 5
I
RNA Velocity
Cell type:
AT2 Cells
AT2-like Cells
AT1 and AT1-like Cells
Secretory Cells
Ciliated Cells
Basal Cells
Neuroendocrine Cells
Gastric-like Cells

Supplemental Figure 5: Sort Purification and Identification of Endogenous and Donor-derived Cell Types, Related to Figure 4

(A) Gating for FACS-based collection of endogenous (mCherry-/GFP-) and donor-derived (mCherry+ or GFP+) lung epithelial cells from an uninjured control and a transplant recipient at 6 weeks post-transplantation of ESC-derived tip-like cells.

(B) Gating for FACS-based collection of endogenous epitheilal (mCherry-/GFP-/EpCAM+) and donor-derived (mCherry+ or GFP+) cells at 15 weeks post-transplantation of ESC-derived tip-like cells.

(C) Pipeline for analysis of scRNA-seq data including exclusion of non-epithelial lineages based on minimal presence of true donor-derived cells in these populations.

(D) Epithelial cells from both timepoints were combined using harmonization to generate a single UMAP plot.

(E) Cells were clustered using the Louvain algorithm followed by combining overlapping clusters. Clusters were then identified based on cell type signatures outlined in Table S3.

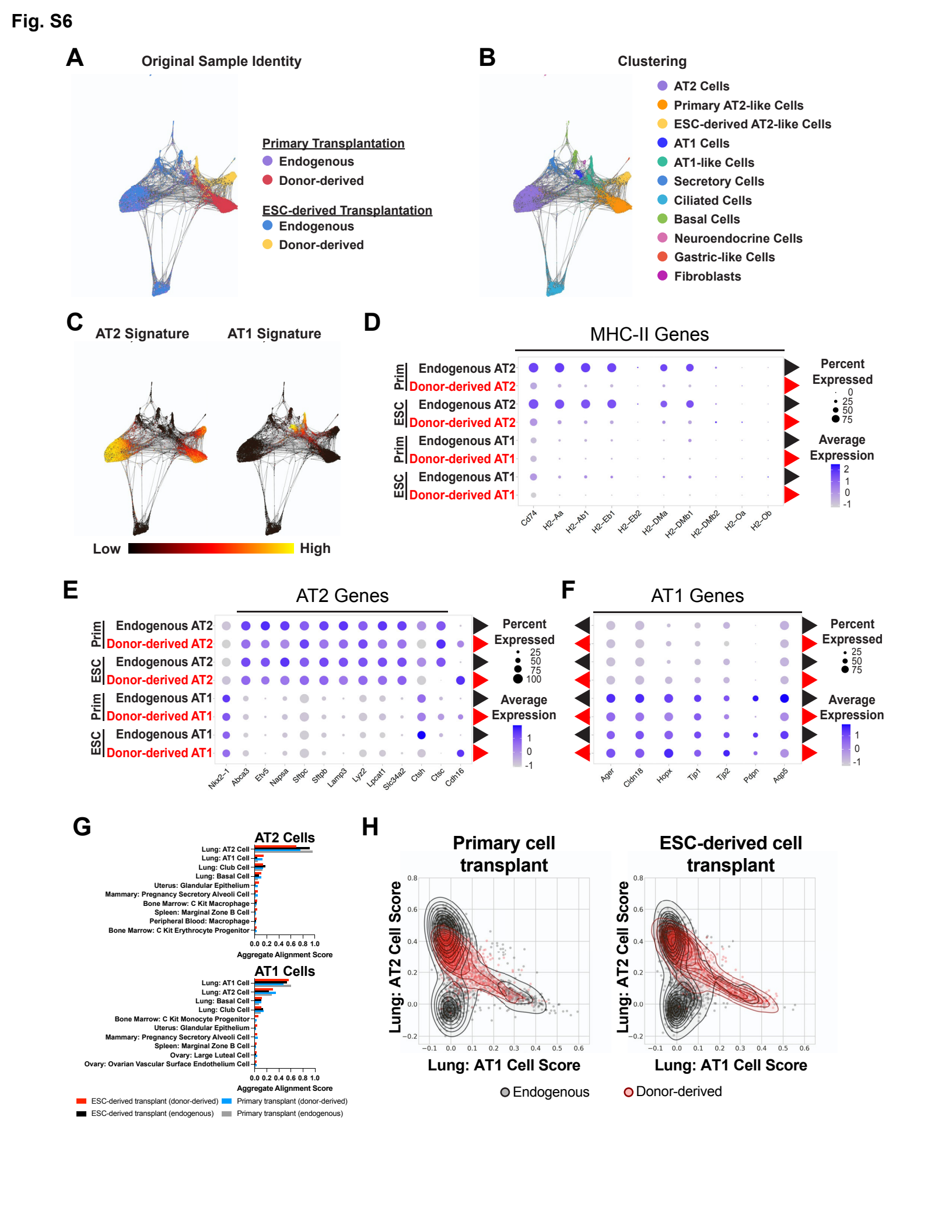
(F) SPRING plots indicating cells with detectable expression of *GFP*, *mCherry*, proliferation markers, *Nkx2-1*, and gastric genes in endogenous and donor-derived cells at 6 and 15 weeks post-transplantation. See figure 4 for annotation of SPRING plot by sample origin or cell type. White arrows indicate the gastric-like cells.

(G) Multimodal reference mapping using samples from figure S5D, E and figure 4. Donor-derived cells are mapped onto the reference created from the uninjured control.

(H) UMAP of combined cell transplantation samples from S5D, E and figure 4 displaying expression of transitional cell markers. The purple arrow indicates a subpopulation high for transitional cell markers.

(I) RNA velocity analysis of endogenous and donor-derived cells from figure S5E.

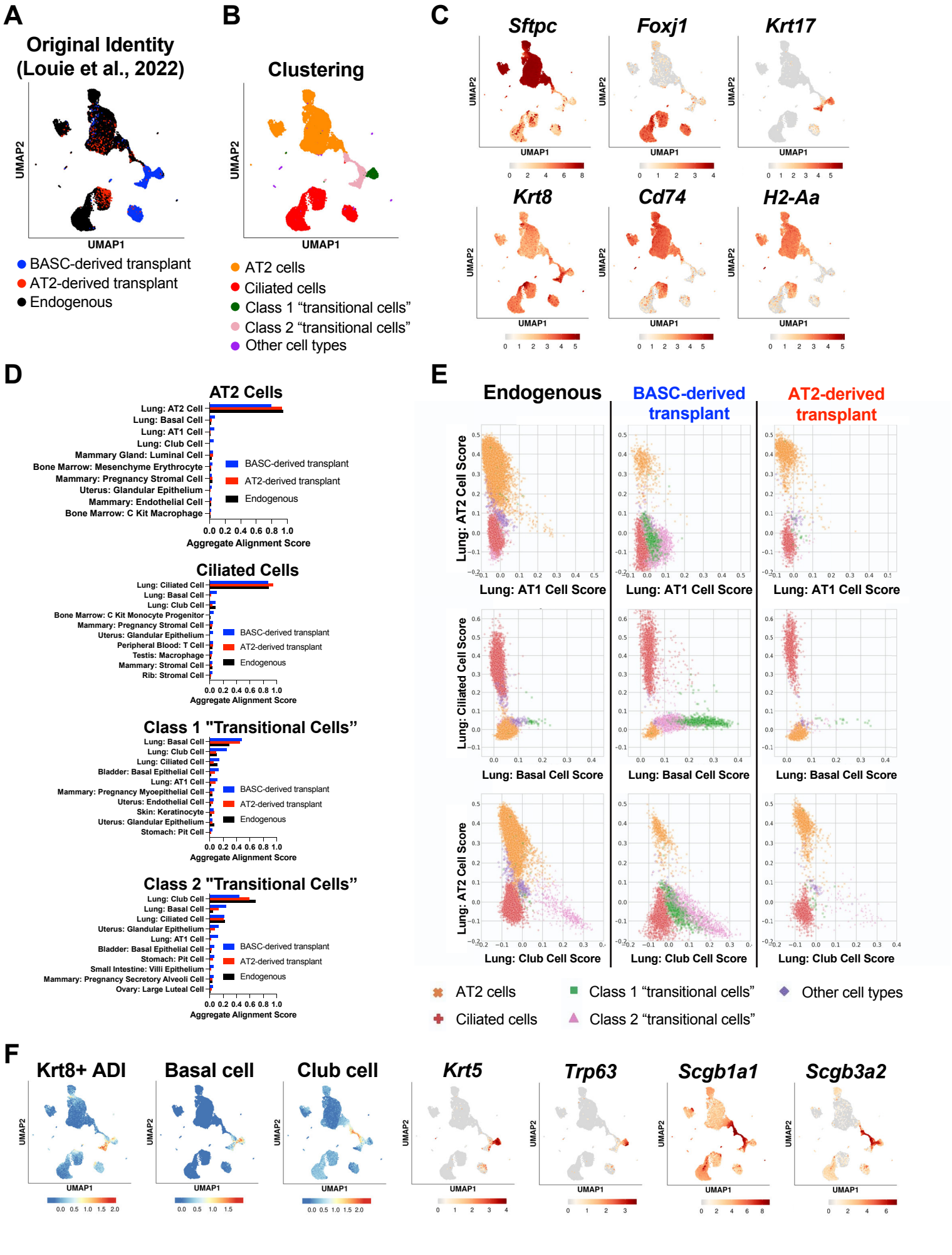
Fig. S6


A
Original Sample Identity
Primary Transplantation
Endogenous
Donor-derived
ESC-derived Transplantation
Endogenous
Donor-derived
B
Clustering
AT2 Cells
Primary AT2-like Cells
ESC-derived AT2-like Cells
AT1 Cells
AT1-like Cells
Secretory Cells
Ciliated Cells
Basal Cells
Neuroendocrine Cells
Gastric-like Cells
Fibroblasts
C
AT2 Signature
AT1 Signature
Low
High
D
MHC-II Genes
Prim
Endogenous AT2
Donor-derived AT2
ESC
Endogenous AT2
Donor-derived AT2
Prim
Endogenous AT1
Donor-derived AT1
ESC
Endogenous AT1
Donor-derived AT1
Cd74
H2-Aa
H2-Ab1
H2-Eb1
H2-Eb2
H2-DMa
H2-DMb1
H2-DMb2
H2-Oa
H2-Ob
Percent Expressed
0
25
50
75
Average Expression
2
1
0
-1
E
AT2 Genes
Prim
Endogenous AT2
Donor-derived AT2
ESC
Endogenous AT2
Donor-derived AT2
Prim
Endogenous AT1
Donor-derived AT1
ESC
Endogenous AT1
Donor-derived AT1
Nkx2-1
Abca3
Etv5
Napsa
Sftpc
Sftpb
Lamp3
Lyz2
Lpcat1
Slc34a2
Ctsh
Ctsc
Cdh16
Percent Expressed
25
50
75
100
Average Expression
1
0
-1
F
AT1 Genes
Ager
Cldn18
Hopx
Tjp1
Tjp2
Pdpn
Aqp5
Percent Expressed
25
50
75
Average Expression
1
0
-1
G
AT2 Cells
Lung: AT2 Cell
Lung: AT1 Cell
Lung: Club Cell
Lung: Basal Cell
Uterus: Glandular Epithelium
Mammary: Pregnancy Secretory Alveoli Cell
Bone Marrow: C Kit Macrophage
Spleen: Marginal Zone B Cell
Peripheral Blood: Macrophage
Bone Marrow: C Kit Erythrocyte Progenitor
0.0 0.2 0.4 0.6 0.8 1.0
Aggregate Alignment Score
AT1 Cells
Lung: AT1 Cell
Lung: AT2 Cell
Lung: Basal Cell
Lung: Club Cell
Bone Marrow: C Kit Monocyte Progenitor
Uterus: Glandular Epithelium
Mammary: Pregnancy Secretory Alveoli Cell
Spleen: Marginal Zone B Cell
Ovary: Large Luteal Cell
Ovary: Ovarian Vascular Surface Endothelium Cell
0.0 0.2 0.4 0.6 0.8 1.0
Aggregate Alignment Score
ESC-derived transplant (donor-derived)
Primary transplant (donor-derived)
ESC-derived transplant (endogenous)
Primary transplant (endogenous)
H
Primary cell transplant
ESC-derived cell transplant
Lung: AT2 Cell Score
Lung: AT1 Cell Score
Endogenous
Donor-derived

Supplemental Figure 6: Primary and ESC-derived Tip-like Cell Transplants Give Rise to Transcriptionally Similar Donor-derived Cells, Related to Figure 5, 6

(A) SPRING plot of data generated from scRNA-seq of parallel ESC-derived and primary tip-like cell transplant recipients at 8 weeks post-transplant.
(B) Cell-type annotation of clusters based on supervised Louvain clustering and expression of lung epithelial cell signatures.
(C) Expression of AT2 and AT1 cell signatures. Gene sets comprising each signature can be found in Supplementary Table 3.
(D) Expression of MHC-II genes in donor-derived (red) and endogenous (black) cells.
(E) Expression of AT2 genes in donor-derived (red) and endogenous (black) cells.
(F) Expression of AT1 genes in donor-derived (red) and endogenous (black) cells.
(G) The top ten aggregate alignment scores for donor-derived AT2-like cells and the corresponding scores for endogenous AT2 cells. The top ten aggregate alignment scores for donor-derived AT1-like cells and the corresponding scores for endogenous AT1 cells. All reference cell types are from adult mice (Mouse Cell Atlas or Control sample as delineated in Figure 4) [50].
(H) Individual alignment scores for all donor-derived and endogenous epithelial cells against reference adult lung AT1 and AT2 cells. Each cell is annotated based on sample type. See also Tables S8-S11.

Fig. S7


A
Original Identity
(Louie et al., 2022)
UMAP2
UMAP1
BASC-derived transplant
AT2-derived transplant
Endogenous
B
Clustering
UMAP2
UMAP1
AT2 cells
Ciliated cells
Class 1 “transitional cells”
Class 2 “transitional cells”
Other cell types
C
Sftpc
Foxj1
Krt17
Krt8
Cd74
H2-Aa
UMAP2
UMAP1
D
AT2 Cells
Lung: AT2 Cell
Lung: Basal Cell
Lung: AT1 Cell
Lung: Club Cell
Mammary Gland: Luminal Cell
Bone Marrow: Mesenchyme Erythrocyte
Mammary: Pregnancy Stromal Cell
Uterus: Glandular Epithelium
Mammary: Endothelial Cell
Bone Marrow: C Kit Macrophage
BASC-derived transplant
AT2-derived transplant
Endogenous
0.0 0.2 0.4 0.6 0.8 1.0
Aggregate Alignment Score
Ciliated Cells
Lung: Ciliated Cell
Lung: Basal Cell
Lung: Club Cell
Bone Marrow: C Kit Monocyte Progenitor
Mammary: Pregnancy Stromal Cell
Uterus: Glandular Epithelium
Peripheral Blood: T Cell
Testis: Macrophage
Mammary: Stromal Cell
Rib: Stromal Cell
Class 1 "Transitional Cells”
Lung: Basal Cell
Lung: Club Cell
Lung: Ciliated Cell
Bladder: Basal Epithelial Cell
Lung: AT1 Cell
Mammary: Pregnancy Myoepithelial Cell
Uterus: Endothelial Cell
Skin: Keratinocyte
Uterus: Glandular Epithelium
Stomach: Pit Cell
Class 2 "Transitional Cells”
Lung: Club Cell
Lung: Basal Cell
Lung: Ciliated Cell
Uterus: Glandular Epithelium
Lung: AT1 Cell
Bladder: Basal Epithelial Cell
Stomach: Pit Cell
Small Intestine: Villi Epithelium
Mammary: Pregnancy Secretory Alveoli Cell
Ovary: Large Luteal Cell
E
Endogenous
BASC-derived transplant
AT2-derived transplant
Lung: AT2 Cell Score
Lung: Ciliated Cell Score
Lung: Basal Cell Score
Lung: AT1 Cell Score
Lung: Club Cell Score
AT2 cells
Ciliated cells
Class 1 “transitional cells”
Class 2 “transitional cells”
Other cell types
F
Krt8+ ADI
Basal cell
Club cell
Krt5
Trp63
Scgb1a1
Scgb3a2
UMAP2
UMAP1

Supplemental Figure 7: Transplantation of cultured adult lung epithelial cells gives rise to mature AT2 cells and non-alveolar epithelial lineages, Related to Figure 5, 6

(A) UMAP of combined cell transplantation samples from (Louie et al., 2022)[15].
(B) UMAP with annotation of cell types based on supervised Louvain clustering and expression of marker genes.
(C) UMAP plots displaying expression of the indicated genes.
(D) The top ten aggregate alignment scores for donor-derived AT2 cells following a BASC-derived transplant as well as the corresponding scores for donor-derived AT2 cells following a AT2-derived transplant and endogenous AT2 cells. Similar graphs are shown for ciliated cells and “transitional cells”. All reference cell types are from adult mice (Mouse Cell Atlas or Control sample as delineated in Figure 4) [50].
(E) Individual alignment scores for all donor-derived and endogenous epithelial cells against the indicated reference cells. Each cell is annotated (by color and shape shown in the key below the graphs) based on cell type as determined in figure S7B.
(F) UMAP plots displaying expression of cell type signature (top 30 DEGs for each cell type)[39] or the indicated gene.

Supplemental Table 3: Lung Epithelial Cell Type Gene Signatures, Related to Figure 4

AT1 Cell Gene Signature	AT2 Cell Gene Signature	Ciliated Cell Gene Signature	Secretory Cell Gene Signature	Basal Cell Gene Signature	Neuroendocrine Gene Signature
Ager	Abca3	1110017D15Rik	5330417C22Rik	Apoe	Nov
Akap5	Acot7	1700001C02Rik	Acsm1	Dcn	Resp18
Aqp5	Acoxl	1700007K13Rik	Aldh1a7	Dapl1	Ascl1
Cldn18	Acsl4	1700016K19Rik	Cckar	Aqp3	Scg5
Clic3	Ank3	Ak7	Cldn10	Krt5	Chgb
Clic5	Atp8a1	BC051019	Cyp2f2	Krt15	Pcsk1
Col4a3	Cd74	Ccdc113	Fmo3	Fxyd3	Calca
Col4a4	Cebpa	Ccdc153	Gabrp	Ltbp4	Meis2
Cryab	Cpm	Cfap126	Gpx2	Dlk2	Nnat
Cyp2b10	Ctsh	Cfap45	Gsta2	Trp63	Cplx2
Emp2	Cxcl15	Cfap53	Gsta4	Wnt4	Cd9
Fam189a2	Dram1	Cyp2s1	Hp	Krt17	Piezo2
Gprc5a	Egfl6	Dnali1	Iyd	Tpt1	Col8a1
Hopx	Elovl1	Drc1	Kcnk2	Tmem176b	Pnmal2
Hs2st1	Etv5	Fam183b	Ldhb	Dst	Cdc14b
Igfbp2	Fabp5	Foxj1	Lrrc26	Col17a1	Pkib
Krt7	Fasn	Gm867	Lypd2	Aqp4	Tmem158
Lmo7	Hc	Lrriq1	Mgat3	Hcar2	Ptn
Mal2	Lamp3	Mlf1	Mgst1	Defb1	Hist3h2ba
Pdpn	Lgi3	Mns1	Pigr	Spon2	Pcsk1n
Prdx6	Lpcat1	Nme5	Pon1		
Pxdc1	Lyz2	Nme9	Por		
Rtkn2	Mlc1	Pifo	Rassf9		
Scnn1g	Muc1	Riiad1	Reep6		
Spock2	Napsa	Rsph1	Retnla		
Tspan8	Npc2	Smim5	Scgb1a1		
Vegfa	Ppp1r14c	Sntn	Scgb3a2		
	S100g	Spaca9	Selenbp1		
	Scd1	Tekt1	Slc16a11		
	Sfta2		Wfdc2		
	Sftpa1				
	Sftpb				
	Sftpc				
	Sftpd				
	Slc34a2				
	Zdhhc3				

This table contains lists of genes that are selectively upregulated in the predominant lung epithelial cell types and were used to generate cell type gene expression signatures used to identify cell types in scRNA-seq datasets. See the methods section for cutoffs used to generate these lists.