UC Office of the President

Title

Development of human alveolar epithelial cell models to study distal lung biology and disease

Permalink

https://escholarship.org/uc/item/7f63h111

Journal

iScience, 25(2)

ISSN

2589-0042

Authors

Tran, Evelyn
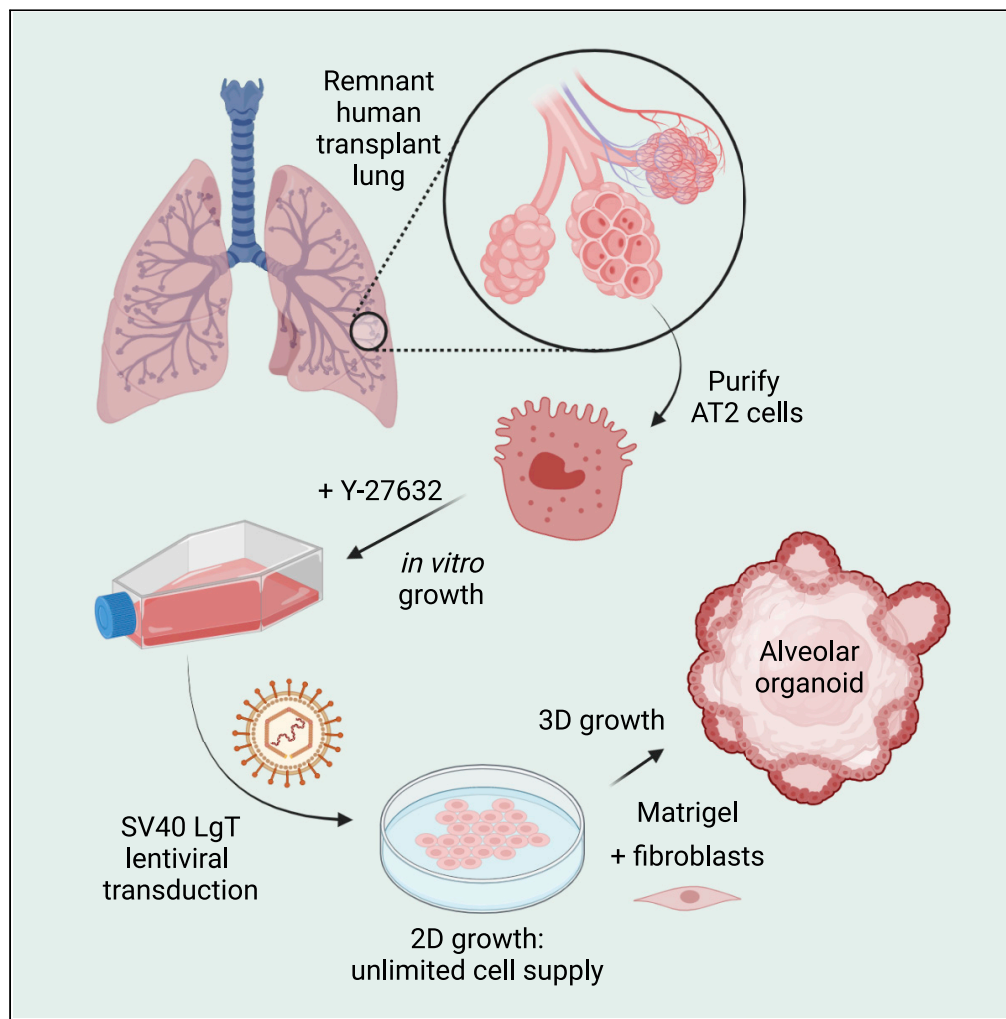Shi, Tuo
Li, Xiuwen
et al.

Publication Date

2022-02-01

DOI

10.1016/j.isci.2022.103780

Peer reviewed

Article

Development of human alveolar epithelial cell models to study distal lung biology and disease

Remnant human transplant lung

Purify AT2 cells

+ Y-27632

in vitro growth

SV40 LgT lentiviral transduction

2D growth: unlimited cell supply

3D growth

Matrigel + fibroblasts

Alveolar organoid

Evelyn Tran, Tuo Shi, Xiuwen Li, ..., Beiyun Zhou, Zea Borok, Ite A. Offringa

ilaird@usc.edu

Highlights

Human AT2 cells grown in Y-27632 medium are immortalized by SV40 Large T antigen

Immortalized AT2 cells are SOX9+/SOX2+ in the absence of mature alveolar markers

Immortalized AT2 cells form NKX2-1+/AQP5+/GPRC5A+ organoids in 3D culture

Specific 3D culture conditions induce AT2 marker HTII280 expression in organoids

Article

Development of human alveolar epithelial cell models to study distal lung biology and disease

Evelyn Tran,[1,2,3] Tuo Shi,[1,2,3] Xiuwen Li,[2,4] Adnan Y. Chowdhury,[5] Du Jiang,[5] Yixin Liu,[6] Hongjun Wang,[6] Chunli Yan,[1,2] William D. Wallace,[7] Rong Lu,[5] Amy L. Ryan,[5,6] Crystal N. Marconett,[1,2,3] Beiyun Zhou,[2,6] Zea Borok,[2,3,6,8] and Ite A. Offringa[1,2,3,9,*]

SUMMARY

Many acute and chronic diseases affect the distal lung alveoli. Alveolar epithelial cell (AEC) lines are needed to better model these diseases. We used de-identified human remnant transplant lungs to develop a method to establish AEC lines. The lines grow well in 2-dimensional (2D) culture as epithelial monolayers expressing lung progenitor markers. In 3-dimensional (3D) culture with fibroblasts, Matrigel, and specific media conditions, the cells form alveolar-like organoids expressing mature AEC markers including aquaporin 5 (AQP5), G-protein-coupled receptor class C group 5 member A (GPRC5A), and surface marker HTII280. Single-cell RNA sequencing of an AEC line in 2D versus 3D culture revealed increased cellular heterogeneity and induction of cytokine and lipoprotein signaling in 3D organoids. Our approach yields lung progenitor lines that retain the ability to differentiate along the alveolar cell lineage despite long-term expansion and provides a valuable system to model and study the distal lung *in vitro*.

INTRODUCTION

Diseases affecting the distal lung, including lung cancer, chronic obstructive pulmonary disease (COPD), idiopathic pulmonary fibrosis (IPF), and pulmonary viral infections are a significant health burden across communities worldwide. In the United States, lung cancer is still the leading cause of cancer-related deaths (Siegel et al., 2020), and chronic lower respiratory diseases account for approximately 5.7% of all deaths (Centers for Disease Control and Prevention, 2018). Because of the susceptibility of the distal lung to severe damage, and public health concerns regarding long-term effects of prior lung damage, there is an urgent need to generate *in vitro* models to study mechanisms underlying distal lung diseases. Such models would facilitate screening for novel therapeutics.

The distal lung encompasses saccular respiratory units called alveoli. Alveoli are composed of two epithelial cell types: cuboidal alveolar type 2 cells (AT2) responsible for producing and secreting surfactant, and large, delicate type 1 cells (AT1) enveloped by a network of capillaries that mediate gas exchange (Rackley and Stripp, 2012; Rock and Hogan, 2011; Tata and Rajagopal, 2017). AT2 cells have been extensively investigated due to their role as stem cells in the regeneration of damaged lung epithelium (Barkauskas et al., 2013; Khalil et al., 1994; Liu et al., 2011). AT2 cells can proliferate or differentiate into AT1 cells in response to lung injury (Adamson and Bowden, 1974; Beers and Morrisey, 2011; Evans and Bils, 1969). AT1 cells have been less extensively studied due to the difficulty in purifying and culturing these fragile cells. Initially, AT1 cells were thought to be terminally differentiated (Evans and Hackney, 1972; Hackney et al., 1975). However, it is now suggested that AT1 cells, or at least a subpopulation thereof, may have the ability to reenter the cell cycle and differentiate into AT2 cells (Borok et al., 1998; Jain et al., 2015; Little et al., 2019; Wang et al., 2018; Yang et al., 2016). Therefore, AT1 cells may also play a role in alveolar regeneration.

In the last 30 years, numerous human lung cell lines have been used to study lung disease and homeostasis *in vitro*. The most common cell models available are immortalized airway epithelial cells established using viral and non-viral immortalization methods to overcome cellular senescence and crisis (Hahn and Weinberg, 2002; Herranz and Gil, 2018; Reddel, 2000), such as BEAS-2B cells immortalized by simian virus 40 large T antigen (SV40 LgT) transduction (Reddel et al., 1988; Schiller et al., 1994) and human bronchial epithelial cell (HBEC) and small airway epithelial cell (SAEC) lines immortalized by overexpression of

[1]Department of Surgery, Keck School of Medicine, University of Southern California (USC), Los Angeles, CA 90033, USA

[2]USC Norris Comprehensive Cancer Center, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[3]Department of Biochemistry and Molecular Medicine, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[4]Department of Translational Genomics, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[5]Department of Stem Cell Biology and Regenerative Medicine, Eli and Edythe Broad CIRM Center for Regenerative Medicine and Stem Cell Research, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[6]Hastings Center for Pulmonary Research and Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[7]Department of Pathology, Keck School of Medicine, USC, Los Angeles, CA 90033, USA

[8]Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[9]Lead contact

*Correspondence: ilaird@usc.edu

https://doi.org/10.1016/j.isci.2022.103780

telomerase (hTERT) in combination with cyclin-dependent kinase 4 (CDK4) mutant CDK4^{R24C} or knockdown of cell-cycle proteins p16^{INK4A}/p19^{ARF} (Ramirez et al., 2004; Sasai et al., 2011; Smith et al., 2016).

In contrast to these cell lines derived from airway cells, few human alveolar epithelial cell (AEC) lines have become available despite advances in cell culturing and stem cell reprogramming techniques. Kemp et al. and Kuehn et al. generated AT1-like immortalized cell lines from spontaneously transdifferentiated cultured human AT2 cells using SV40 LgT plus hTERT (Kemp et al., 2008) or a cocktail of proprietary "immortalizing" genes (Kuehn et al., 2016). Bove et al. (2014) achieved short-term culturing of AT1-like cells from human AT2 cells grown in medium containing the small molecule Rho-associated kinase (ROCK) inhibitor, Y-27632. These approaches, however, did not address whether the resulting AECs were able to form organoids, a feature exhibited by primary AT2 cells. An alternative approach has been to derive AECs from induced pluripotent stem cells (iPSCs). These are more capable of forming 3D structures under organotypic culture but have shown marked variability in differentiation efficiencies and in AT2:AT1 compositions (Dye et al., 2015; Gotoh et al., 2014; Jacob et al., 2017; Kanagaki et al., 2020; McCauley et al., 2018; Yamamoto et al., 2017). The formidable time and experience required to properly differentiate iPSCs (Meng et al., 2012; Volpato and Webber, 2020) are a drawback in its application as a tool for the general lung community. A simple method for deriving long-term proliferating human AECs that are rapidly expandable under standard culturing conditions and retain the ability to form 3D alveolar structures would therefore be of great value to the scientific community.

Here, we report the establishment and characterization of a collection of immortalized cell lines derived from isolated adult human AT2 cells. Under 2D culture conditions, the cells grow as an epithelial monolayer and exhibit lung-progenitor-like expression patterns. In 3D organotypic culture, they form diverse organoid structures and express mature AEC markers, AQP5 and GPRC5A. Under specific 3D medium conditions, a subpopulation of cells can be induced to express the AT2 cell membrane marker, HTII280. Cell lines derived from the alveolar epithelium are urgently needed for the study of diseases arising from lung alveoli. Our novel alveolar epithelial-derived distal lung cell lines provide a new resource for studying genetic and environmental mechanisms underlying distal lung diseases as well as for investigating the regulation of alveolar epithelial cell homeostasis.

RESULTS

Expansion of isolated adult human AT2 cells in ROCK inhibitor medium followed by lentiviral transduction results in proliferative AEC lines

Primary human AECs do not proliferate in culture (Isakson et al., 2002; Mao et al., 2015), making them challenging to manipulate and limiting their suitability for *in vitro* modeling. We initially attempted direct transduction of both freshly isolated and previously cryopreserved human AT2 cells with lentiviruses carrying CDK4^{R24C} and hTERT based on reported success in human bronchial epithelial cells (Ramirez et al., 2004). Although this approach resulted in early proliferation of cells, it was followed by either growth arrest or adoption of a fibroblast-like morphology at later passages (Figures S1A–S1C). Repeat transduction of seemingly growth-arrested cells with SV40 LgT resulted in fibroblast-like cells (Figure S1A, last panel). Failure to derive proliferating AECs by direct transduction motivated us to refine our immortalization strategy. In culture, AT2 cells spontaneously transdifferentiate into AT1-like cells over the course of 6 days (Cheek et al., 1989; Dobbs, 1990; Foster et al., 2007; Fuchs et al., 2003; Marconett et al., 2013). These AT1-like cells have attenuated cell bodies and a high overall surface area, similar to AT1 cells *in vivo*. We considered that our failure to immortalize primary AECs by direct transduction with lentiviruses carrying immortalizing genes might be because cells were not proliferating and thus not molecularly responsive to mutant CDK4 and hTERT. In the literature, we found that successful immortalization appeared to be possible for purified primary human cells exhibiting at least some proliferative capacity (Herbert et al., 2002; O'Hare et al., 2001). Indeed, primary human bronchial epithelial cells were first subcultured before immortalization with retroviruses (Ramirez et al., 2004). We therefore hypothesized that successful immortalization of primary AECs would require first stimulating the cells to proliferate in culture, then transducing the dividing cells with lentiviruses carrying immortalizing genes (Figure S1D).

We performed a small-scale screen of media containing growth factors and small molecules using previously cryopreserved isolated human AT2 cells from one de-identified human lung from a deceased subject (Lung-FT, Tables 1 and S1). After several weeks, visual inspection indicated that medium containing 10 μM

Table 1. De-identified lung donor information and lung cell purity metrics

Lung Sample	Age (Years)	Sex	Race/ Ethnicity	Disease Status	Cause of Death	Tobacco Use	SFTPC[+a]	NKX2-1[+a]	CD45[+a]	VIM[+a]	TUBBA[+a]
Lung-FT	25	M	Caucasian	Normal	Head trauma	NS[b]	88%	96%	3%	4%	3%
Lung-ON	66	F	Caucasian	Normal	ICH[c]-stroke	NS	79%	–	15%	–	–
Lung-TN	62	M	Caucasian	Normal	Stroke	FS[d]	84%	86%	5%	12%	–

[a]Lung cell purity metrics: following anti-EpCAM bead purification, cell suspensions were assessed for AT2 cell enrichment by cytospin preparation (right five columns), then IF staining with AT2 cell markers (SFTPC, NKX2-1), hematopoietic marker (CD45), fibroblast marker (VIM), and lung airway cell marker (TUBB4A).
[b]Subject reported as nonsmoker (NS), but gross lung displayed black speckles/particulate matter, suggesting possible inhalant exposure.
[c]ICH, intracerebral hemorrhage.
[d]FS, former smoker. Subject smoked one pack per day for 18 years, then quit in 1991.

Y-27632 supported cell survival and proliferation. About 6 weeks postplating, we passaged cells for expansion. From this passage on, cells were maintained in medium containing Y-27632, grown as contact-inhibited monolayers, and referred to as "AEC-FT-ROCKinh" cells (Figure 1A, a).

Although Y-27632 greatly expanded the lifespan of our isolated human AT2 cells, as previously reported for other primary human cells (Chapman et al., 2014; Horani et al., 2013; Liu et al., 2012), this induced proliferative state was not sufficient for immortalization (Bove et al., 2014; Martinovich et al., 2017; Suprynowicz et al., 2012). Viral oncogenes, such as HPV 15/16, EBV, and SV40 LgT, have been shown previously to immortalize human cells even in the absence of hTERT (Band et al., 1991; Christian et al., 1987; Liu et al., 2006; Neufeld et al., 1987; Reddel et al., 1988). To maintain the proliferative potential of our AEC-FT-ROCKinh cells, we transduced these cells with lentiviruses carrying hTERT alone, CDK4[R24C] alone, CDK4[R24C] + hTERT, and SV40 LgT alone (Figure 1A). Each transduction yielded proliferative but contact-inhibited epithelial-like cells when maintained in Y-27632 medium (Figure 1A, a–e). Cells transduced with SV40 LgT (henceforth referred to as "AEC-FT" cells) were compact and more rapidly expandable in culture compared with cells transduced with hTERT alone, CDK4[R24C] alone, or CDK4[R24C] + hTERT. We therefore generated biological replicates of the AEC-FT cell line from two additional de-identified human donor lungs from deceased subjects (Table 1): "AEC-ON" (from Lung-ON) and "AEC-TN" (from Lung-TN). Like the AEC-FT cells, both lines exhibited epithelial-like morphologies and were contact-inhibited on confluence (Figures 1B and 1C).

AEC lines proliferate in culture but do not form tumors in nude mice

We next characterized the proliferative ability of our AEC lines. As a reference for a highly proliferative cell line presumed to be derived from alveolar epithelium, we used the human lung adenocarcinoma cell line, A549 (Giard et al., 1973). We found that AEC-FT cells proliferated almost as well as A549 cells (Figure 1D). All other FT-derived AEC lines exhibited an extended lag phase (Figure 1D). The growth potential of other FT-derived cells was not improved by seeding them at higher densities (Table S2). When seeded at low density, AEC-ON and AEC-TN cell lines grew slower than the AEC-FT line, but when plated at higher density, both cell lines proliferated readily (Figure 1E), reaching exponential growth 4 days after plating, with a population doubling time (PDT) of ~1 day for AEC-ON and ~2 days for AEC-TN cells (Table S2). These results indicate that the presence of SV40 LgT stimulates the growth potential of AECs in ways that CDK4[R24C] + hTERT cannot.

BEAS-2B cells, derived from bronchial epithelial cells immortalized with SV40 LgT, exhibit anchorage independent growth (a common hallmark of cancer cells), but do not form tumors in immunocompromised (nude) mice. We therefore examined the AEC lines for these properties. Anchorage-independent growth is commonly tested by the ability of cells to form colonies in soft agar. We performed soft agar assays on all AEC lines, using A549 tumor cells as a positive control. Only AEC-FT cells consistently generated soft agar colonies (246 ± 55 colonies per well), comparable to A549 cells (284 ± 186 colonies, p value = 0.78, nonparametric Wilcoxon test) (Figure 1G). AEC-ON and AEC-TN cells formed rare small colonies visible under 10X brightfield magnification (Figure 1F, inset). Quantification of colonies yielded 2 ± 1 colonies per well for AEC-ON cells and 3 ± 2 colonies for AEC-TN, which was statistically significantly different
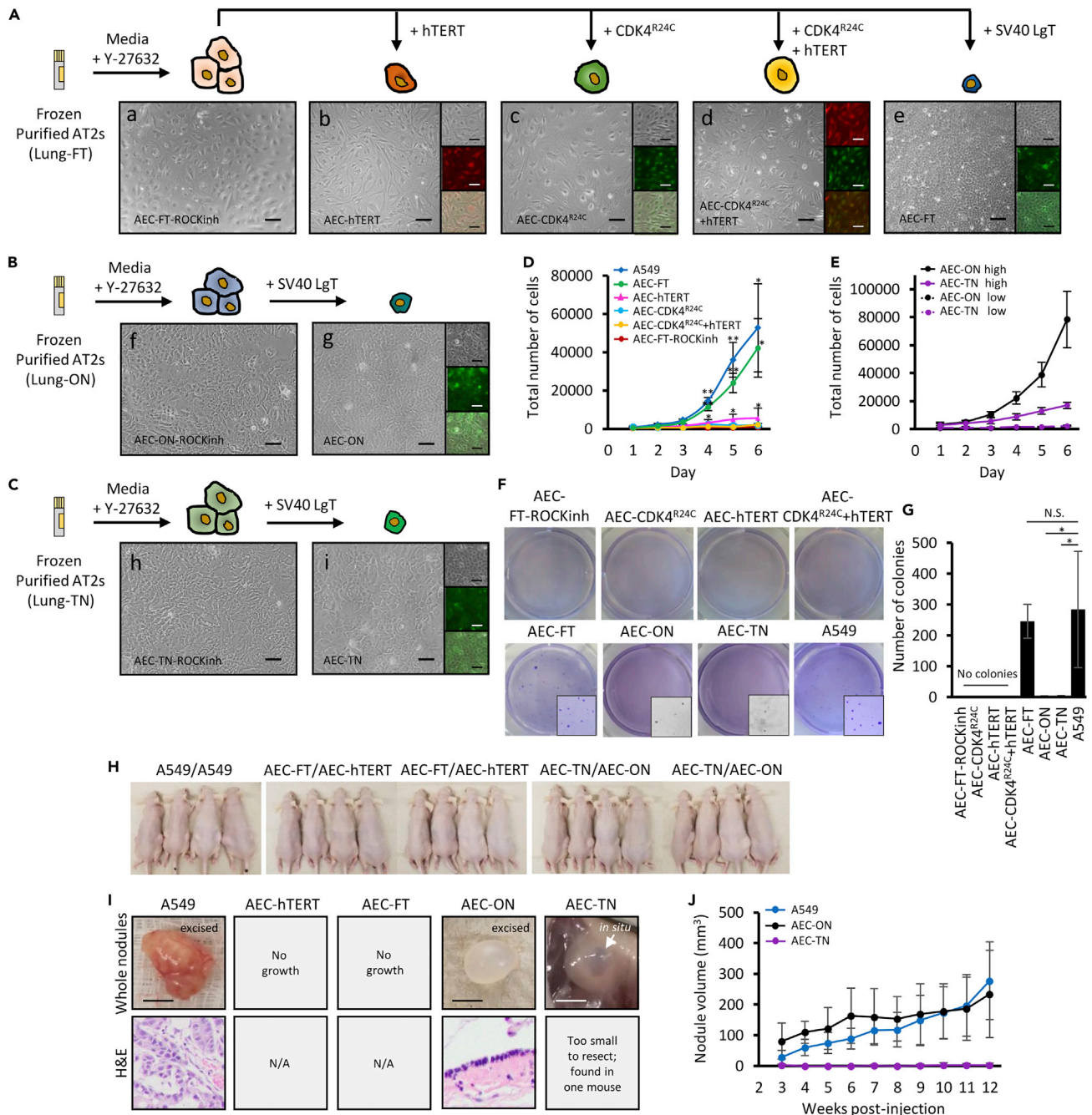
Figure 1. Derivation of AEC lines that proliferate in culture but do not form tumors in nude mice

(A) AEC line derivation scheme. Previously frozen purified human AT2 cells from Lung-FT were first cultured in Y-27632 medium. Proliferative cells (a) were then transduced with LeGO iT (tdTomato) and LeGO iG (eGFP) lentiviruses carrying hTERT (b), CDK4^{R24C} (c), CDK4^{R24C} + hTERT (d), and SV40 LgT (e). Scale bar, 100 μm.

(B, C) Additional SV40 LgT-transduced cell lines were established from AT2 cells purified from (B) Lung-ON and (C) Lung-TN in Y-27632 medium (f and h), then transduced with SV40 LgT (g and i). Scale bar, 100 μm.

(D) Cell proliferation assays in technical triplicates from three independent experiments (mean total number of cells ± std dev). *p<0.05, **p< 0.005 by independent two-sample t test relative to AEC-FT-ROCKinh.

(E) Cell proliferation assays for AEC-ON and AEC-TN, under standard (1,000 cells/well) and high-density (5,000 cells/well) conditions.

(F) Anchorage-independent growth assays. A549 cells, positive control. Colonies were stained with crystal violet and counted using ImageJ/Fiji after 1 month. Inset images, |✕2.5✕| magnified to show colonies.

Figure 1. *Continued*

(G) Colony quantification of six technical replicates from at least three independent experiments (mean ± std dev). N.S., not significant; *p< 0.05, nonparametric Wilcoxon test.

(H) Subcutaneous injection into NU/J mice to assess tumorigenicity of AEC-FT, AEC-ON, and AEC-TN cell lines over 3 months. A549 cells, positive control; AEC-hTERT line, negative control. Equal numbers of male and female mice per group. Sample labels indicate which cell line was injected into the left and right flanks (left/right).

(I) Photos of excised nodules (scale bar, 5 mm) and corresponding H&E stainings.

(J) Graph of nodule growth starting at 3 weeks postinjection. See also Figure S1, Tables S1 and S2.

from A549 cells (p = 1.4 × 10^{-7} and p = 1.5 × 10^{-7}, respectively, nonparametric Wilcoxon test) (Figure 1G). We next tested the ability of the AEC lines to form tumors in immunocompromised mice, again using A549 cells as a positive control. We subcutaneously injected the three AEC-LgT lines (AEC-FT, AEC-ON, AEC-TN) into the flanks of nude mice. After 3 months, mice transplanted with A549 cells and AEC-ON cells displayed prominent nodules (Figure 1H). Excised nodules were evaluated by an expert pathologist blind to sample identities. A549-derived nodules were solid tumors of adenocarcinoma histology, whereas AEC-ON-derived nodules were nonmalignant, fluid-filled cysts lined by a simple epithelium with areas of pseudostratified cells (Figure 1I). In one case (1 out of 8 flanks), AEC-TN cells formed a nodule less than 20 mm^3 that was too small to resect. This nodule did not significantly change in volume during the 3-month experimental period (Figure 1J). Neither AEC-FT nor negative control AEC-hTERT cells formed nodules. Taken together, although SV40 LgT transduction conferred cell immortality, it did not appear to induce malignant transformation.

AEC lines are transcriptomically distinct from primary human AECs and LUAD cell lines

To determine how similar the transcriptomic identities of our novel AEC lines were to their parental purified primary AECs, we performed comparative expression analyses using paired-end RNA-sequencing (RNA-seq). We also compared our AEC lines with lung fibroblasts to test for epithelial-to-mesenchymal transition gene signatures, as well as with LUAD cancer cell lines to determine if our novel cells bore transcriptional hallmarks of oncogenic transformation. We used RNA-seq data generated in-house as well as data from lung fibroblasts and LUAD cancer cell lines from the publicly available databases ENCODE (ENCODE: https://www.encodeproject.org, Davis et al., 2018) and DBTSS (DBTSS: https://dbtss.hgc.jp, Suzuki et al., 2018), analyzing a total of 42 samples (Table S3).

We performed Principal Component Analysis (PCA) and sample-sample distance matrix analyses using the top 500 most variable genes across the data set. Our analyses showed six clusters: lung fibroblasts, all AEC lines from Lung-FT, the biological replicate AEC-LgT lines from Lung-ON and -TN, fetal lung tissue, LUAD cancer lines, and primary AECs (Figure 2A). Human fetal lung tissue clustered away from all samples. Although the biological replicate lines AEC-ON and AEC-TN were generated by transduction of SV40 LgT in a similar fashion to AEC-FT cells, these two lines clustered separately from the collection of ROCKinh-derived cells from Lung-FT. Thirty-nine percent of the variation in data divided the sample set into clusters containing long-term cultured cells (LUAD, fibroblasts, AEC lines) and the other cells (primary AECs and fetal lung tissue), underlining the possible effects of growing cells in culture. Groups identified in the sample distance matrix were generally in agreement with the PCA analysis; however, lung fibroblasts were found to be grouped in a distinct cluster from AEC lines based on dissimilarity calculations represented by a dendrogram tree (Figures 2B and S2).

To determine which genes drive segregation of the different sample clusters, we performed unsupervised hierarchical clustering using the top 500 most variable genes across the data set. Primary AECs and fetal lung tissue were more related to each other than to the remaining groups of samples (Figure 2C). Nine distinct gene clusters were identified. Most of the clusters had at least one Gene Ontology (GO) term related to tissue development, morphogenesis, and cellular response to the environment (Table S4). Inspection of the differences in overall gene expression across the sample types (fetal lung tissue, primary AT2, *in-vitro*-derived AT1-like, AEC lines, lung fibroblasts, and LUAD) revealed a group of genes contained in gene clusters A, B, and C with higher expression in non-cancer samples compared with LUAD samples (Figure 2C, dashed box) and associated GO terms "surfactant homeostasis," "regulation of immune system process," "response to external biotic stimuli," and "tube development" (Figure 2D and Table S5).

We next performed differential gene expression analyses between the AEC lines and each of the other groups: AT2 cells, AT1-like cells, lung fibroblasts, and LUAD cancer lines (Figures 2E–2H). The largest
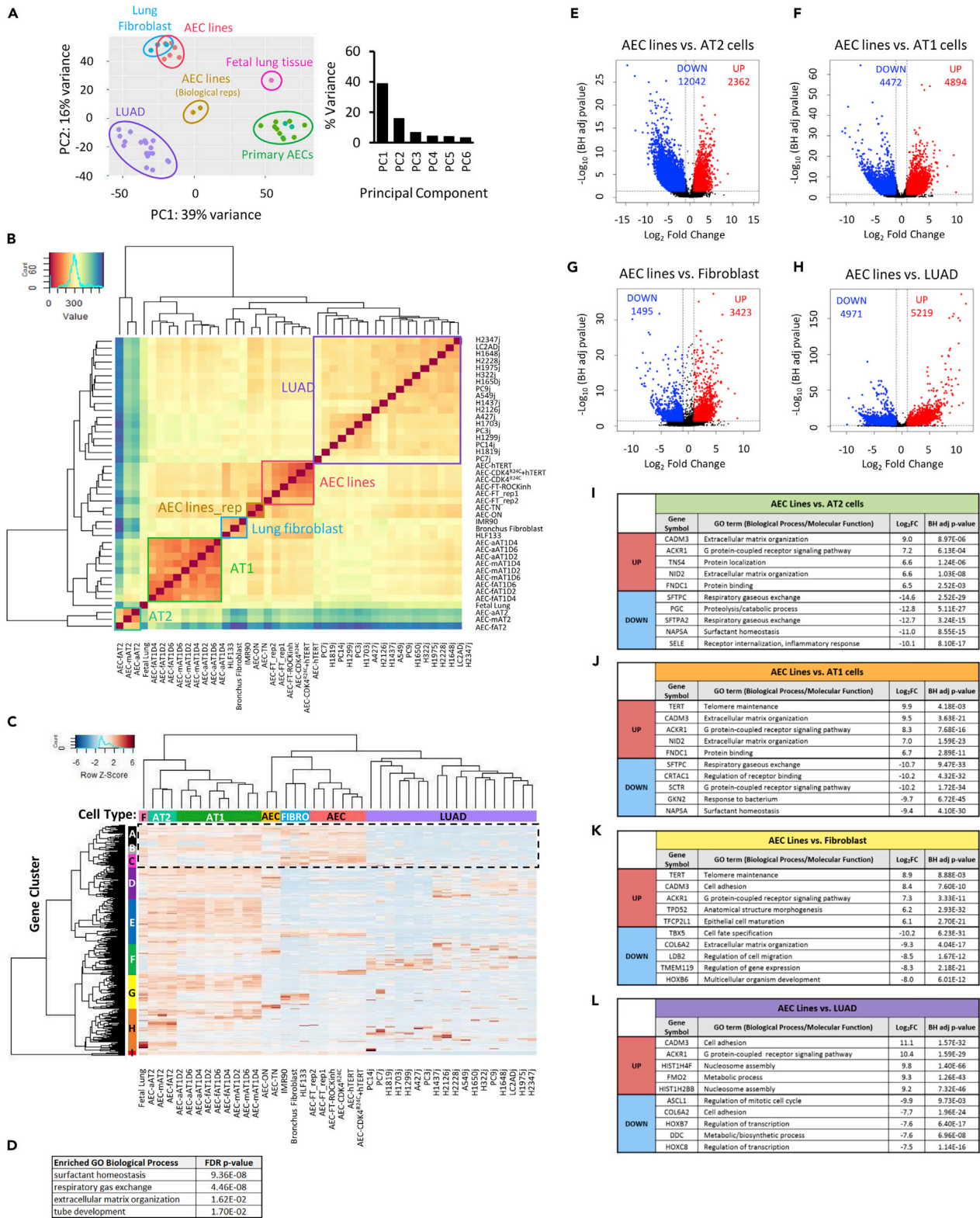
Figure 2. Transcriptomic profiling of AEC lines and comparison to lung fibroblasts, primary AECs, and LUAD cell lines

(A) Left: PCA plot of the top 500 most variable genes using DESeq2 pipeline. Each sample is represented by a dot colored according to sample type. Right: Bar chart of percent variance across PCs.

Figure 2. *Continued*

(B) Sample-sample distance matrix based on similarity calculations using Euclidean distance metric. Dark red, high similarity; blue, low similarity. Primary human AT2 cells are labeled with "f," "m," and "a" denoting three separate donor lungs. Human AT1-like cells *in vitro* differentiated from AT2 cells are labeled "—D2," "—D4," "—D6," representing days transdifferentiated in culture. LUAD cell line names were appended with "j" to indicate that RNA-seq data from the Japanese databased DBTSS was used.

(C) Unsupervised hierarchical clustering. Cell type categories: F, fetal lung tissue; AT2, primary human AT2 cells; AT1, human AT1-like cells; AEC (yellow), AEC-ON and AEC-TN replicate cell lines; Fibro, human lung fibroblasts; AEC (pink), AEC lines derived from Lung-FT; LUAD, human lung adenocarcinoma cancer cell lines.

(D) GO terms corresponding to gene clusters A, B, and C (dashed box) differentiating primary AECs, lung fibroblasts, and AEC lines from LUAD cancer lines; PANTHERv14.1 Overrepresentation Test (Released, 20200728) using Fisher's exact test from the Gene Ontology Consortium. Statistical significance, FDR-corrected p< 0.05.

(E–H) Differential gene expression analyses were performed between all derived AEC lines and (E) AT2, (F) AT1-like, (G) lung fibroblasts, and (H) LUAD cancer cell lines, shown as Volcano plots. Statistically significant differentially expressed genes relative to AEC lines shown in blue (downregulated in AEC lines) and red (upregulated in AEC lines). Significance cutoffs (dotted lines) were set at |fold change| $\geq$ 2 and Benjamini-Hochberg (BH) adjusted p< 0.05.

(I–L) Table of GO terms for the top 5 most significantly upregulated and downregulated protein-coding genes. Listed Log$_2$Fold Change and BH adjusted p values are from differential gene expression analyses by DESeq2. See also Figures S2, S3, and Tables S3–S5.

number of differentially expressed genes (DEGs) (14,404) resulted from comparing the AEC lines with primary AT2 cells (Figure 2E) and the fewest number of DEGs (4,918) resulted from comparing the AEC lines with lung fibroblasts (Figure 2G). When comparing AEC lines with LUAD cancer lines, we found the fold change values of resulting DEGs were orders of magnitude more significant than DEG fold changes from any other pairwise comparison (Figures 2E–2H). This corroborates our tumorigenicity data showing that the AEC lines, although genetically manipulated, remained distinct from transformed cancer cells. GO terms of the top upregulated and downregulated genes from each pairwise comparison revealed significant alterations in cell adhesion, matrix deposition, cell signaling, and metabolism, reinforcing our experimental findings that the AEC lines are morphologically different from and highly proliferative compared with primary AECs, while still distinguishable from cancer-derived cell lines (Figures 2I–2L).

Because the AEC lines appeared to be transcriptomically more similar to lung fibroblasts than to primary AECs, suggesting a possible loss in lung epithelial cell specificity and adoption of a more mesenchymal phenotype, we performed clustering on a subset of 75 lung-related genes manually curated from published RNA-seq data (Treutlein et al., 2014; Xu et al., 2016) (Figure S3A). As expected from our genome-wide transcriptomic analyses, expression of these lung-related genes was low in the AEC lines compared with isolated primary AECs (Figure S3A), but remarkably higher when compared with lung fibroblasts (Figure S3B), suggesting that the similarity between the AEC lines and fibroblasts calculated by PCA analysis could be driven by the fact that fibroblasts are the only other cell type that was non-cancer, of normal ploidy, and cultured on plastic.

AEC-LgT cells express features of lung progenitor cells

Y-27632 is a commonly used small molecule to promote stem cell survival and proliferation (Claassen et al., 2009; Vernardis et al., 2017). Combined with feeder cells, it has been shown to enhance culturing of primary epithelial cells from mammary, prostate, and upper airway lung tissues (Liu et al., 2012). However, in this process of facilitating cell survival, adult cells are reprogrammed to a stem-like state (Suprynowicz et al., 2012). In the mouse lung, SOX9 regulates distal lung cell fate, committing early cells to an alveolar epithelial cell lineage (Chang et al., 2013; Rockich et al., 2013), whereas SOX2 is an important regulator of proximal lung cell fate, committing early lung stem cells to the basal cell lineage (Daniely et al., 2004; Ochieng et al., 2014). In the human lung, progenitor cells located at budding distal epithelial tips during fetal lung development co-express SOX9 and SOX2 (Danopoulos et al., 2018; Nikolić et al., 2017). We therefore assessed SOX9 and SOX2 expression in the three AEC-LgT lines by immunofluorescence (IF) staining, finding all three lines positive for these progenitor markers (Figure 3A). Indeed, the AEC lines, as a group, expressed both *SOX9* and *SOX2* at the RNA level more highly than primary AECs (Figure S4).

We further examined the expression of mature lung markers by immunostaining, finding all three AEC-LgT lines negative for mature AT1 cell markers AQP5 and HOP homeobox (HOPX) and AT2 cell marker pro-surfactant protein C (pro-SFTPC). AEC-FT and AEC-TN cells were negative for lung epithelial marker NK2 Homeobox 1 (NKX2-1), whereas AEC-ON cells expressed NKX2-1, although at variable levels across the cell population (Figure 3B). Notably, although NKX2-1 regulates expression of the AT2 cell marker SFTPC (Kelly et al., 1996), we did not detect expression of SFTPC in AEC-ON cells, as indicated by its precursor
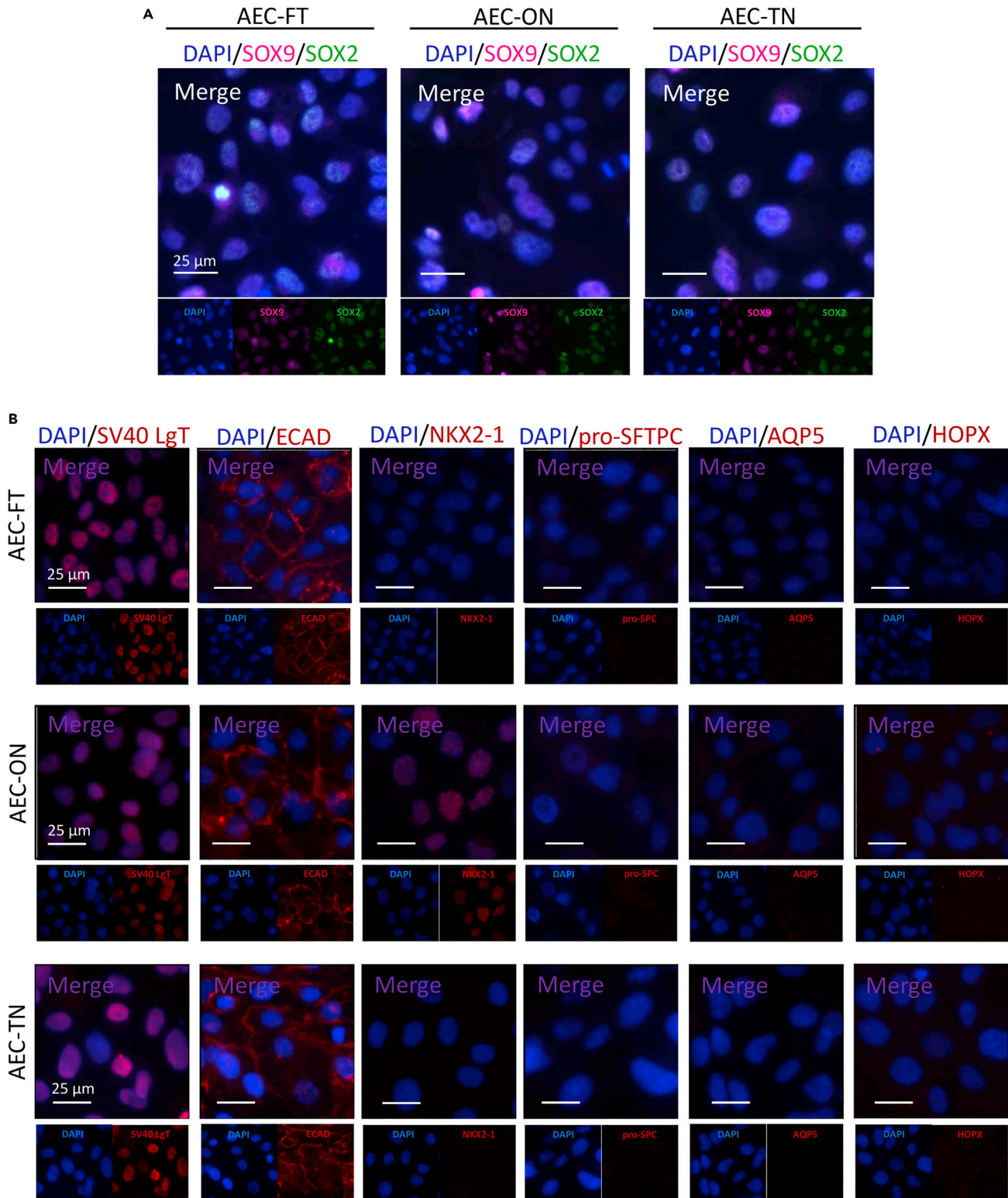
Figure 3. Alveolar epithelial cell lines prominently express lung progenitor markers over mature AEC markers

Representative IF staining of AEC-FT, AEC-ON, and AEC-TN cell lines grown on standard culture dishes.

(A) Co-staining of SOX9 and SOX2 lung progenitor markers.

(B) IF staining for SV40 LgT and epithelial marker ECAD and mature AEC markers NKX2-1, pro-SFTPC, APQ5, and HOPX.

Staining patterns were consistent across different passages. At least three independent experiments were performed in duplicate. Scale bar, 25 μm. See also Figure S4.

protein, pro-SFTPC. Taking our transcriptomic analyses and IF staining data together, we speculated that AEC-LgT cells in 2D culture favored a transcriptional program promoting cell proliferation and cell survival over one specifying alveolar epithelial cell lineage and perhaps represent a more immature precursor rather than fully differentiated AT2/AT1 cells. We therefore investigated whether AEC-LgT cells could be induced to differentiate to a phenotype resembling human adult AECs in 3D culture.

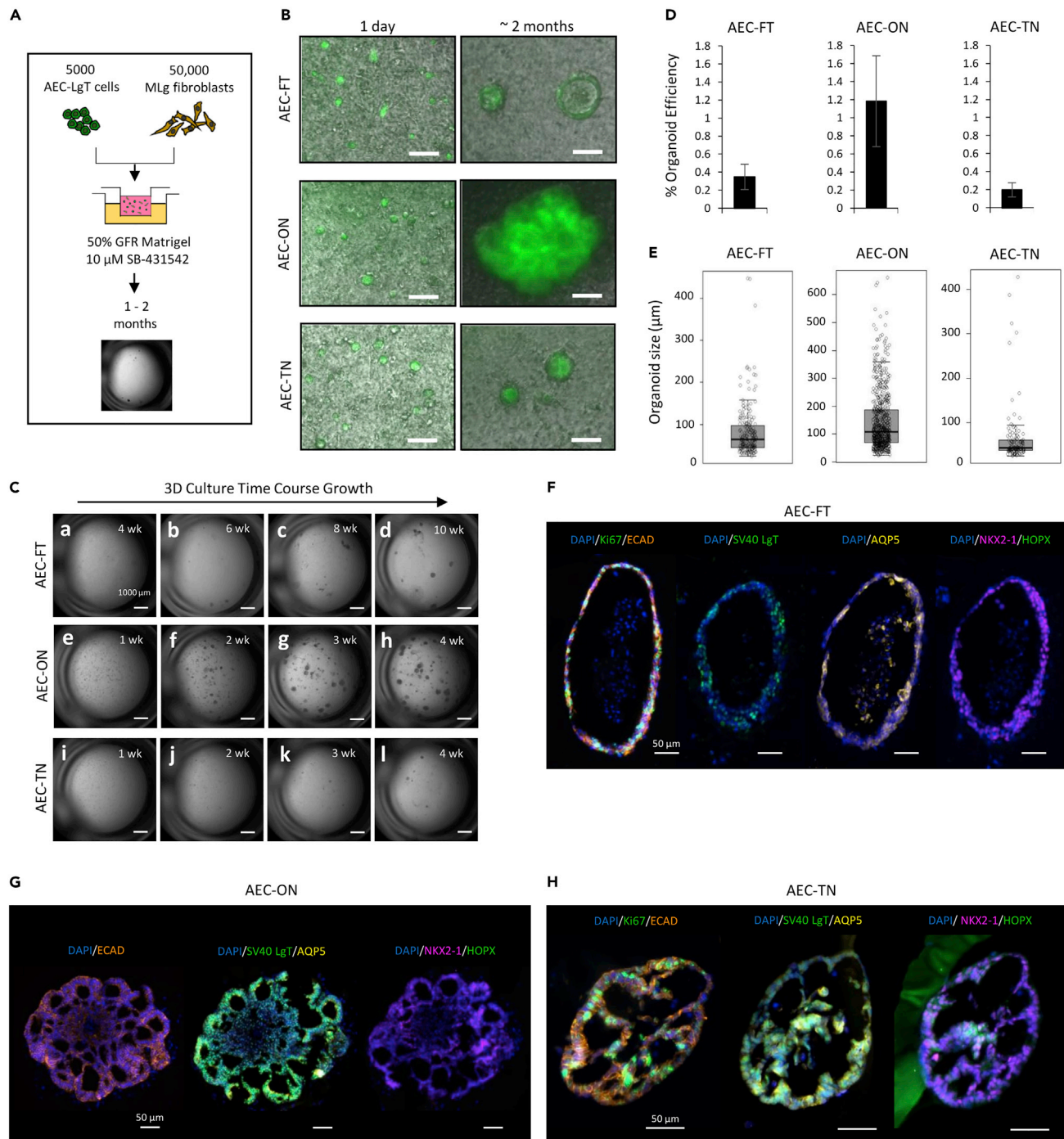## AEC-LgT cells form lung organoids in 3D co-culture

Purified AT2 cells from mouse and human lungs form organoids when co-cultured with stromal cells and suspended in Matrigel (Barkauskas et al., 2013; Jain et al., 2015; Zacharias et al., 2018; Zhou et al., 2018). To determine whether AEC-LgT cells possess the ability to form 3D structures, we mixed exponentially growing AEC-FT, AEC-ON, and AEC-TN cells with neonatal mouse lung fibroblasts (MLg) in Matrigel on Transwell inserts (Figure 4A). All three cell lines formed organoids from a single-cell suspension (Figure 4B), in contrast to the dense cell clusters formed by A549 cells under similar conditions (Figures S5A and S5B). In the absence of MLg fibroblasts, we did not observe organoid formation from any of the AEC-LgT cell lines after 1 month of culture (Figures S5C–S5E).

Mouse and human AT2 cells form organoids within several weeks of culturing (Barkauskas et al., 2013; Jain et al., 2015; Zacharias et al., 2018), whereas SOX9$^+$/SOX2$^+$ human distal tip progenitors form organoids in as little as 12 h (Nikolić et al., 2017). To examine the rate of organoid formation of all three AEC-LgT cell lines, we carried out a time course experiment. AEC-FT cells were noticeably slower in forming organoids than either AEC-TN or AEC-ON cells (Figure 4C), requiring at least 5 weeks to form detectable spheres (Figures 4C and 4B) compared with 3–4 weeks for AEC-TN (Figure 4C, k and l) and <2 weeks for AEC-ON cells (Figures 4C and 4F). We noted that the rates of organoid formation for the three AEC-LgT cell lines did not coincide with their rates of proliferation in 2D culture, suggesting this structural change was not merely determined by cell division. To quantitatively characterize organoid growth of the AEC-LgT lines, we calculated organoid formation efficiency and size after 2 months of culture. The mean organoid formation efficiency varied and was highest for AEC-ON cells (Figure 4D). The range of sphere sizes also varied, with AEC-ON cells showing the highest median sphere size (108 μm) (Figure 4E and Table S6).

Organoid shapes were variable across cultures for each AEC-LgT line. However, general growth patterns were observed. Under brightfield microscopy, AEC-FT cells formed organoids of predominantly round morphology with a single lumen (Figure 4F). Occasionally, across different 3D cultures of these cells, organoids containing multiple lumens were observed (Figure S6A). AEC-ON organoids were more heterogeneous in morphology. A subpopulation of them were large and floret-like, exhibiting an intricate lobulated structure (Figures 4G and S6B). AEC-TN cells formed rounded, single lumen organoids more commonly than multilumen organoids. However, when present, the multilumen organoids appeared more complex in structure than those from AEC-FT cells but less complex than those from the AEC-ON cells (Figure S6C).

AEC-LgT organoids robustly express alveolar epithelial markers

All three AEC-LgT cell lines formed organoids in 3D culture, an ability well documented in primary mouse and human AT2 cells grown under similar conditions. However, AT2-cell-derived organoids tend to be relatively dense with a small central lumen (Barkauskas et al., 2013; Zacharias et al., 2018; Zhou et al., 2018). AEC-LgT cell organoids, in contrast, were more reminiscent of organoids formed by multipotent distal tip progenitors (Nikolić et al., 2017), having marked spherical to lobulated morphologies and large lumens. To determine whether this morphological behavior was accompanied by changes in lung-specific marker expression, we performed IF staining for mature alveolar markers on organoid sections. Organoids were composed of SV40 LgT$^+$/ECAD$^+$/Ki67$^+$ proliferative epithelial cells (Figures 4F–4H). Upon probing for mature alveolar markers (NKX2-1, pro-SFTPC, AQP5, HOPX), we found that all organoids robustly expressed AQP5 and NKX2-1, whereas HOPX expression was variable (Figures 4F–4H), and pro-SFTPC was negative (Figure S7A). The commonly used AT2-cell-specific surface marker, HTII280, was also negative

Figure 4. AEC-LgT cell lines form organoids in 3D co-culture and express mature lung markers

(A) Schematic representation of 3D co-culture of AEC-LgT cells with MLg mouse neonatal fibroblasts in growth factor-reduced (GFR) Matrigel.
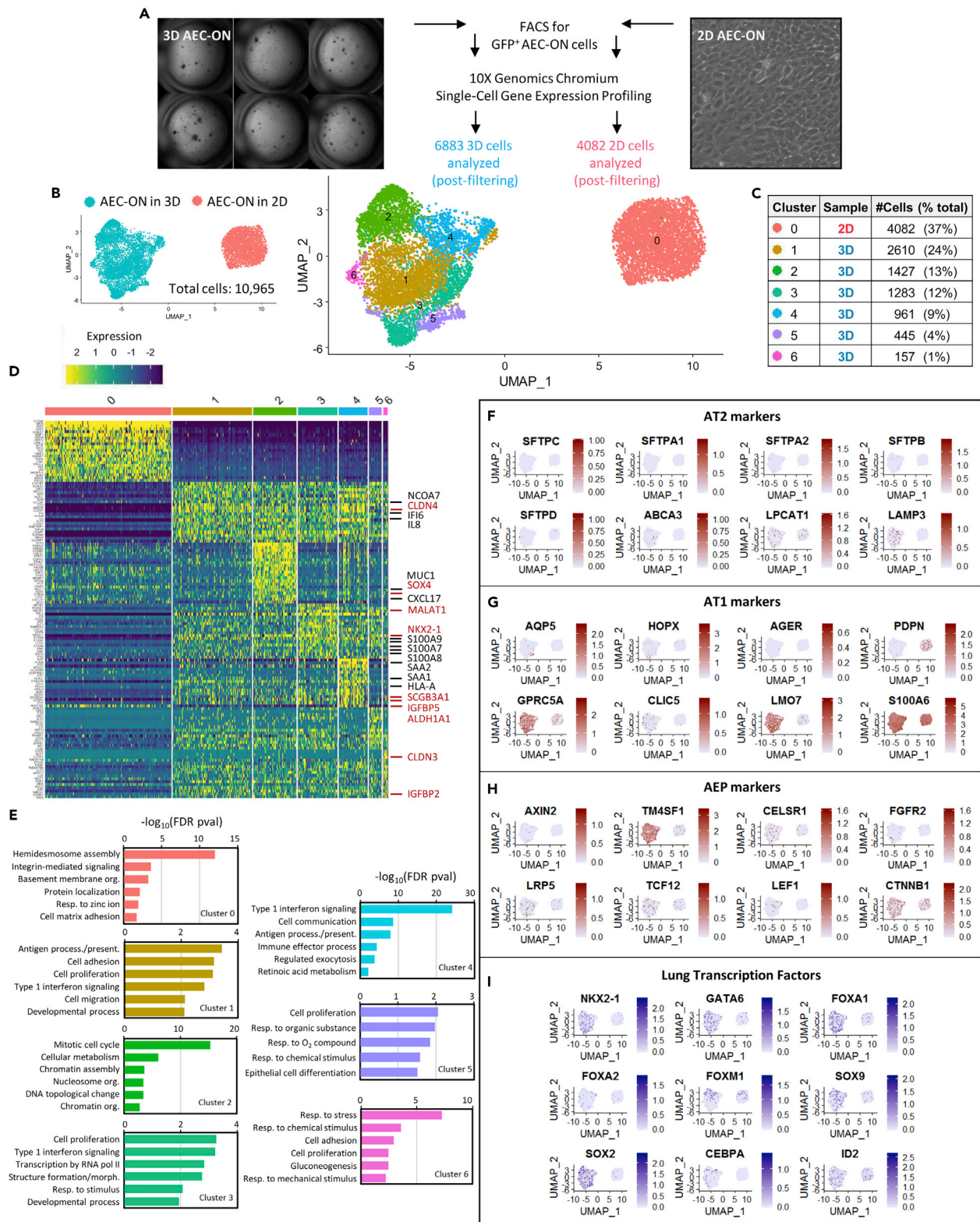
(B) Organoids arise from a single cell suspension of SV40 LgT$^+$/GFP$^+$ epithelial cells. Scale bar, 100 μm.

(C) Time course growth of AEC-FT, AEC-ON, and AEC-TN organoids. Imaging of slower-growing AEC-FT organoids began at 4 weeks (a–d). Imaging of AEC-ON (e-h) and AEC-TN (i-l) cultures began at 1 week. Scale bar, 1,000 μm.

(D and E) Organoid formation efficiencies and (E) size measurements for AEC-FT, AEC-ON, and AEC-TN cells at 2 months, three biological replicates, each with at least six inserts.

(F–H) IF staining of representative organoid sections from (F) AEC-FT, (G) AEC-ON, and (H) AEC-TN cell lines after 1 month of organoid growth. ECAD, epithelial marker; Ki67, proliferation marker; AQP5, HOPX, NKX2-1, mature AEC markers. Scale bar, 50 μm. See also Figures S5–S7 and Table S6.

Figure A: 3D AEC-ON / 2D AEC-ON. FACS for GFP+ AEC-ON cells. 10X Genomics Chromium Single-Cell Gene Expression Profiling. 6883 3D cells analyzed (post-filtering). 4082 2D cells analyzed (post-filtering).

Figure B: AEC-ON in 3D, AEC-ON in 2D. Total cells: 10,965.

Figure C: Cluster table

Cluster	Sample	#Cells (% total)
0	2D	4082 (37%)
1	3D	2610 (24%)
2	3D	1427 (13%)
3	3D	1283 (12%)
4	3D	961 (9%)
5	3D	445 (4%)
6	3D	157 (1%)

Figure D: Heatmap. Gene labels: NCOA7, CLDN4, IFI6, IL8, MUC1, SOX4, CXCL17, MALAT1, NKX2-1, S100A9, S100A7, S100A8, SAA2, SAA1, HLA-A, SCGB3A1, IGFBP5, ALDH1A1, CLDN3, IGFBP2.

Figure E: -log10(FDR pval) bar charts.
Cluster 0: Hemidesmosome assembly, Integrin-mediated signaling, Basement membrane org., Protein localization, Resp. to zinc ion, Cell matrix adhesion.
Cluster 1: Antigen process./present., Cell adhesion, Cell proliferation, Type 1 interferon signaling, Cell migration, Developmental process.
Cluster 2: Mitotic cell cycle, Cellular metabolism, Chromatin assembly, Nucleosome org., DNA topological change, Chromatin org.
Cluster 3: Cell proliferation, Type 1 interferon signaling, Transcription by RNA pol II, Structure formation/morph., Resp. to stimulus, Developmental process.
Cluster 4: Type 1 interferon signaling, Cell communication, Antigen process./present., Immune effector process, Regulated exocytosis, Retinoic acid metabolism.
Cluster 5: Cell proliferation, Resp. to organic substance, Resp. to O_2 compound, Resp. to chemical stimulus, Epithelial cell differentiation.
Cluster 6: Resp. to stress, Resp. to chemical stimulus, Cell adhesion, Cell proliferation, Gluconeogenesis, Resp. to mechanical stimulus.

Figure F: AT2 markers — SFTPC, SFTPA1, SFTPA2, SFTPB, SFTPD, ABCA3, LPCAT1, LAMP3.

Figure G: AT1 markers — AQP5, HOPX, AGER, PDPN, GPRC5A, CLIC5, LMO7, S100A6.

Figure H: AEP markers — AXIN2, TM4SF1, CELSR1, FGFR2, LRP5, TCF12, LEF1, CTNNB1.

Figure I: Lung Transcription Factors — NKX2-1, GATA6, FOXA1, FOXA2, FOXM1, SOX9, SOX2, CEBPA, ID2.

Figure 5. Single-cell transcriptomic analyses of AEC-ON cells reveal increased cellular heterogeneity in response to organotypic culture

(A) Brief workflow of scRNA-sequencing experiment using 10× Genomics Chromium Single Cell (v3) platform.

(B) Uniform Manifold Approximation and Projection (UMAP) dimensional reduction on 10,965 total cells showing grouping by sample type (left) and called clusters (right).

(C) Cell proportions per cluster.

(D) Heatmap showing the top 20 gene markers enriched in each cluster compared with all cells, ordered by average expression (Log$_2$FoldChange). Genes related to immune response and cytokine signaling are highlighted in black. Genes related to development/morphogenesis are highlighted in red.

(E) Statistically significant GO terms for the top 50 (top 75 genes, Cluster five only) most highly expressed genes in each cluster by PANTHERv14.1 Overrepresentation Test. Statistical significance cutoff of FDR-corrected p< 0.05. Abbreviations: org., organization; resp., response; process., processing; present., presentation; morph., morphogenesis.

(F–I) Feature plots of selected AT2 and AT1 cell markers, AEP genes, and lung-related transcription factors. See also Figures S8, S9, and Table S8.

across all lines (Figures S7B and S7C). We also probed for the newly identified AT1 marker, GPRC5A (Horie et al., 2020), finding that all three AEC-LgT organoids possessed GPRC5A$^+$ cells specifically along the apical lining of the lumens (Figure S7C), in contrast to the ubiquitous GPRC5A expression observed in all three monolayer cultures (Figure S7B). In AEC-FT (Figure 4F) and AEC-TN organoids (Figure 4H), 3D co-culture conditions appeared to reactivate AQP5 and NKX2-1 expression from their silenced state in 2D culture, whereas AEC-ON cells maintained NKX2-1 expression in both 2D and 3D cultures and expressed AQP5 only in 3D culture (Figure 4G). Notably, the 3D culture conditions did not seem to induce such marked changes when applied to AEC-LgT cells grown as monolayers (Table S7), suggesting that the suspension of cells in a matrix aids in induction of a more lung-like phenotype and expression profile.

Single-cell transcriptomic analyses of AEC-ON cells reveal increased cellular heterogeneity in response to organotypic culture

As shown in Figure 4G, AEC-ON organoids exhibited the most dramatic morphologies and marked expression of the AT1-enriched gene *AQP5*, whereas the AT2 cell marker *SFTPC* was not expressed, despite the presence of its upstream regulator NKX2-1. Currently, purified primary human AT2 cells have been shown to form organoids; as of yet, the same ability has not been reported for human AT1 cells. Because AEC-ON cells may possess expression profiles of AT2 and AT1 cells that are too nuanced to detect by IF staining, we investigated the transcriptomes of these cells grown in 2D and 3D by single-cell RNA-sequencing (scRNA-seq). In parallel, AEC-ON cells grown in 2D and 3D were FACS sorted for GFP positivity to capture only SV40 LgT$^+$ cells and then processed using the 10× Genomics Chromium platform (Figure 5A). AEC-ON cells grown under standard 2D conditions clustered separately from cells comprising lung organoids (Figure 5B, left). Cluster analyses revealed 7 distinct groups: one cluster (Cluster 0) encompassed all 2D cells and the remaining clusters (Cluster 1–6) were identified in the 3D sample, indicating increased cellular heterogeneity among cells comprising AEC-ON organoids versus cells in the monolayer population (Figure 5B, right and 5C).

To determine the set of marker genes distinguishing each called cluster, we performed differential gene expression analysis. Figure 5D shows a heatmap indicating the top 20 genes in each cluster ranked by average Log$_2$Fold Change over all cells. Associated GO terms for the top 50 genes ordered by p value are shown in Figure 5E. For Cluster 5, GO analysis was performed on the top 75 genes, because the top 50 genes yielded no statistically significant results. Cluster 0, containing all cells from the 2D sample, showed marked enrichment for genes associated with cell adhesion processes, consistent with cell growth as a monolayer (Figure 5E). Cluster 2 was enriched with cell cycling genes, representing actively proliferating cells within AEC-ON organoids. The remaining clusters (1, 3–6) were enriched in genes associated with response to external stimuli such as immune response, inflammation, and type 1 interferon signaling, suggesting that cells interacted with the surrounding microenvironment by secreting cytokines and lipoproteins. Clusters 1 and 3–5 were also significantly enriched for genes involved in development, morphogenesis, and epithelial cell differentiation, consistent with structural changes required for organoid formation (Figure 5E).

Feature maps of marker genes representative of AT2 cells, AT1 cells, and alveolar epithelial progenitor (AEP) cells (Zacharias et al., 2018) revealed a higher proportion of positive cells in 3D organoids than in 2D cultured cells (Figures 5F–5H). Examining the AT2 cell panels, we observed numerous cells in the 3D sample expressing *ABCA3* and *LAMP3*, encoding markers of specialized organelles in AT2 cells called lamellar bodies, where surfactant is produced and stored. Consistent with expression of these genes, we detected several cells expressing surfactant proteins A2 (*SFTPA2*), B (*SFTPB*), and D (*SFTPD*) (Figure 5F). Based on scRNA-seq data from purified human lung cells, mRNA expression of *LAMP3* and *SFTPA2* is

highly AT2-cell-specific (Adams et al., 2020). We did not observe a significant number of cells expressing *SFTPC*, which is exclusively produced by AT2 cells. Examining the plots of AT1-enriched markers, we found that cells comprising the organoids highly expressed transcripts of the actin-binding gene *LMO7* and, to a lesser extent, the chloride channel gene *CLIC5* but expressed *AQP5* at a lower level than expected from our IF staining (Figure 5G). At the protein level, AQP5 is highly AT1-cell-specific, although in human lung scRNA-seq data, *AQP5* RNA is not highly expressed in AT1 cells (Adams et al., 2020; Habermann et al., 2020; Morse et al., 2018; Reyfman et al., 2019). In AT2 to AT1 *in vitro* transdifferentiation experiments using purified human AT2 cells, derived AT1-like cells clearly expressed AQP5 protein, although RNA levels were quite low (Marconett et al., 2013). Our inability to detect *AQP5* at the mRNA level may thus be due to insufficient transcript capture rate. The challenge in detecting the full transcript complement of each cell in scRNA-seq data may lead to an underestimate of the actual cell distribution. Overall, the mixed population of organoid cells expressing AT2- and AT1-cell-associated genes suggests a very limited presence of AT2 cells and the possible presence of immature AT2-like cells or an AT2-AT1 intermediate cell type (Liebler et al., 2016).

AEPs were previously found to highly express the surface marker TM4SF1 and the cytoplasmic protein AXIN2 and to be WNT- and FGF-responsive (Zacharias et al., 2018). Compared with cells grown in 2D, we found a number of AEC-ON cells in 3D expressing *AXIN2* and a much greater proportion expressing *TM4SF1*. In addition, AEC-ON cells expressed genes of the WNT (*TCF12*, *LEF1*, and *CTNNB1*) and FGF pathways (*FGFR2*) (Figure 5H). We examined lung-related transcription factor expression in AEC-ON cells in 2D versus 3D, finding a higher number of cells expressing *NKX2-1*, *GATA6*, *FOXA1*, *FOXA2*, and *CEBPA* in 3D than in 2D. *SOX2* and *SOX9* were also highly expressed in 3D samples (Figure 5I).

Since AEC-ON organoids contained AEP-like cells expressing genes related to WNT and FGF pathways, we investigated whether activation of these pathways can modulate organoid growth characteristics. We treated AEC-LgT 3D cultures with the glycogen synthase kinase 3 (GSK3) inhibitor, CHIR99021 (CHIR), or a mixture of fibroblast growth factor 7 (FGF7) and FGF10 protein ligands (FGF7+10). GSK3 negatively regulates the WNT/β-catenin pathway by maintaining the phosphorylation state of β-catenin (Wu and Pan, 2010; Little et al., 2021); inhibition of GSK3 activity therefore results in activation of WNT signaling. FGF7 and 10 are growth factors that promote differentiation and maintenance of alveolar epithelium (Zhang et al., 2004). After 2 months, treatment of AEC-ON cultures with either CHIR or FGF7+10 resulted in about a ~2-fold increase in sphere size (Figures S8D and S8E), whereas sphere formation efficiencies did not change (Figure S8F). In contrast, sphere size did not change for either AEC-FT nor AEC-TN cells upon treatment (Figures S8A, S8B, S8G, and S8H), whereas FGF7+10 treatment of AEC-FT cells increased sphere formation efficiency by ~2-fold (p value = 0.03) (Figure S8C). Table S8 summarizes the growth metrics for the treatment study. We then ascertained whether modulation of WNT signaling in AEC-ON cells resulted in concomitant changes in AEC marker expression by IF staining. No change in expression of AT1 cell markers AQP5 and GPRC5A nor induction of AT2 cells markers pro-SFTPC, mature SFTPC, and HTII280 was detected (data not shown). Moreover, increasing the dose of CHIR treatment of AEC-ON cells did not result in changes in AEC marker expression (data not shown), despite significant increases in overall sphere size (Figure S9). In sum, we found that changes in sphere size did not correlate with changes in sphere formation efficiency and that, although AEC-ON organoids were WNT- and FGF-responsive, activation of these signaling pathways was not sufficient to induce robust expression of AT2 markers. We therefore speculated that AEC-ON cells may be more akin to a reprogrammed alveolar epithelial cell that is primed to mature under proper microenvironmental conditions, rather than a defined alveolar progenitor cell type, such as AEPs.

Activation of WNT, FGF, and glucocorticoid signaling induces HTII280 expression

Isolated AT2 cells are known to lose expression of cell-type-specific markers such as SFTPC and HTII280 when cultured *in vitro* (Dobbs et al., 1985; Gonzalez et al., 2010). Identification of crucial alveolar epithelial cell lineage pathways during lung development have resulted in improved methods for maintaining or inducing an AT2 cell phenotype *in vitro* (Gonzales et al., 2001; Nikolić et al., 2017; Jacob et al., 2017; Katsura et al., 2020). In particular, stimulating glucocorticoid and cyclic adenosine monophosphate (cAMP) pathways by the addition of dexamethasone, cAMP, and the phosphodiesterase inhibitor 3-isobutyl-1-methyl-xanthine (IBMX), together called "DCI," has been shown to induce and maintain surfactant production, a hallmark of AT2 cells (Ballard and Ballard, 1974; Gonzales et al., 2002). To further investigate whether SFTPC or HTII280 can be induced in AEC-ON cells under specific conditions, we tested the effect of three published AEC culture media conditions on these cells in 2D as well as in 3D: human alveolar differentiation
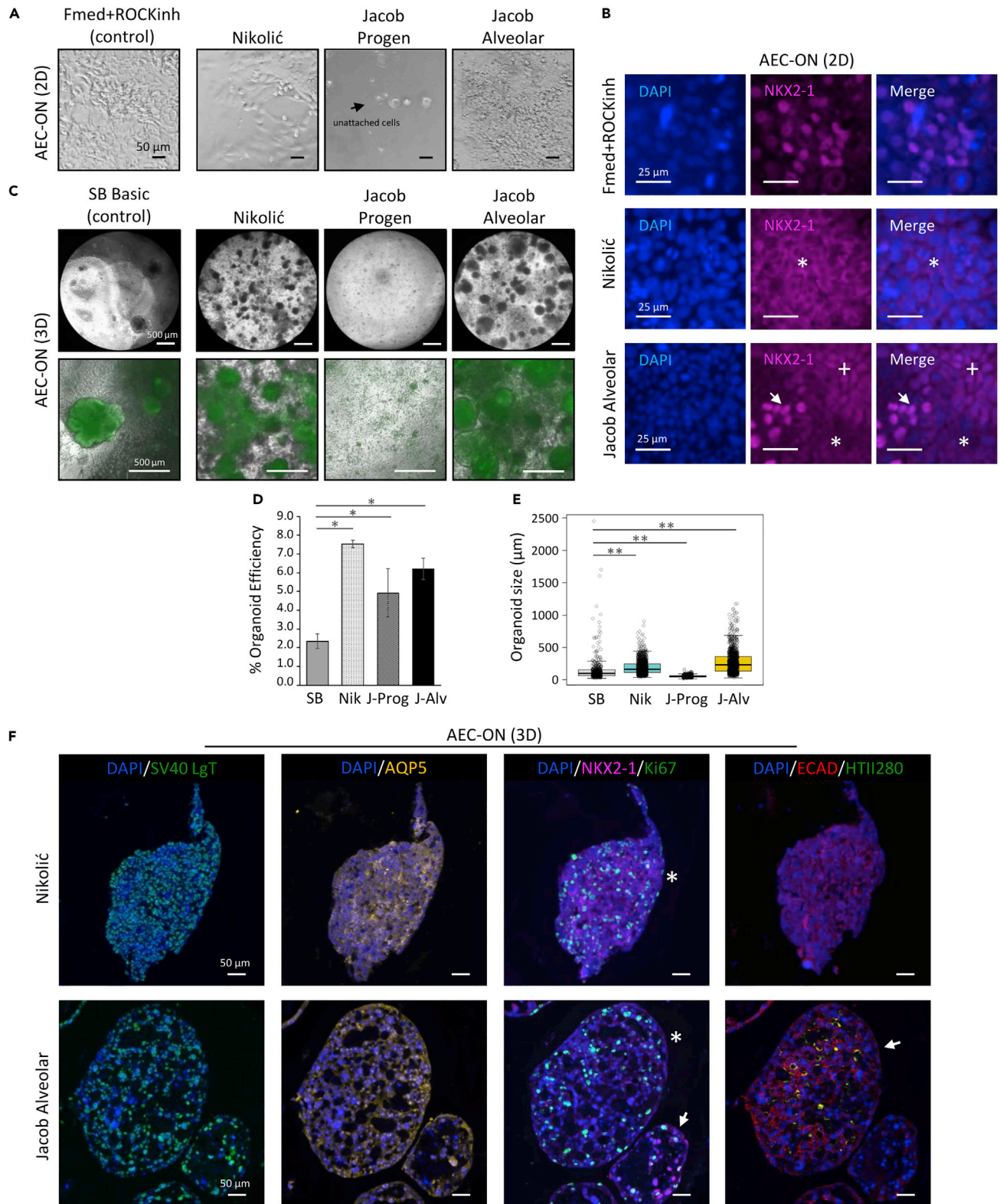
Figure 6. Activation of WNT, FGF, and glucocorticoid signaling results in HTII280 AT2 marker expression in AEC-ON cells grown in 3D

(A) AEC-ON cells were plated in 2D in either Fmed + ROCKinh medium (control) or Nikolić, Jacob Progen, or Jacob Alveolar media for 5 days with medium refreshment every other day. Cells grown in Jacob Progen medium did not survive after 2 days in culture and detached into the medium (arrow). Scale bar, 50 μm.

(B) IF staining of AEC-ON cells in 2D for NKX2-1 under control, Nikolić, and Jacob Alveolar media conditions showing cytoplasmic (asterisk) and nuclear localization (arrow). Scale bar, 25 μm.

(C) Top: representative brightfield images taken at 4X magnification showing organoid growth after media treatment for 5 weeks. Scale bar, 500 μm. Bottom: representative fluorescence images taken at 10X magnification of organoids. Scale bar, 500 μm.

(D) Organoid size and (E) formation efficiencies after 5 weeks (mean ± std dev, n = 4). *p< 0.05, **p< 0.005 by nonparametric Wilcoxon test.

(F) IF staining of AEC-ON organoids grown in either Nikolić or Jacob Alveolar medium for alveolar epithelial markers showing cytoplasmic (asterisk) and nuclear (arrow) expression of NKX2-1 (third panel) and induction of AT2 marker HTII280 (arrow, last panel). Scale bar, 50 μm. See also Figures S10 and S11, Table S9.

medium reported by Nikolić et al. (2017) ("Nikolić" medium), progenitor specification medium from Jacob et al. (2017) ("Jacob Progen" medium), and alveolar maturation medium described by Jacob et al. (2017) ("Jacob Alveolar" medium) (Figure S10A).

In 2D, compared with standard 2D culture medium ("Fmed + ROCKinh"), Nikolić and Jacob Alveolar media supported AEC-ON cell survival and proliferation (Figure 6A). In contrast, Jacob Progen medium resulted in widespread cell death (Figure 6A). In the two media conditions supporting 2D cell growth, cells did not express HTII280, pro-SFTPC, nor AQP5, but did maintain GPRC5A expression (Figure S10B). AEC-ON cells expressed NKX2-1 in standard Fmed + ROCKinh medium; however, when cultured in Nikolić and Jacob Alveolar media, NKX2-1 localization was greatly altered. We detected cytoplasmic expression of NKX2-1 in Nikolić medium (Figure 6B, asterisk) and a mix of localized nuclear expression and general cytoplasmic expression of NKX2-1 across cells grown in Jacob Alveolar medium (Figure 6B, arrow and asterisk).

Under 3D co-culture conditions, AEC-ON cells grown with Nikolić, Jacob Progen, and Jacob Alveolar media survived (Figure 6C). Nikolić and Jacob Alveolar media conditions increased organoid formation efficiency by at least 2-fold (p = 0.028) and statistically significantly increased median size of AEC-ON organoids from 103 μm to 167 and 234 μm, respectively (p < 2.2 × 10^{-16}) (Figures 6D, 6E, and Table S9). In addition to their larger sizes, the organoids formed in Nikolić and Jacob Alveolar media were generally more spherical in structure compared with AEC-ON organoids formed in control SB Basic medium, which were characteristically lobulated and floret-like in shape (Figure 6C). Although Jacob Progen medium also supported greater organoid formation efficiency (from 2.3%–4.9%), organoids that formed were dramatically smaller (median size 55 μm, range 23–164 μm) and appeared to resemble cell clusters more than lumen-containing structures (Figures 6C–6E). Because Nikolić and Jacob Alveolar media significantly improved AEC-ON organoid growth, we next determined whether each resulted in changes to AT1 or AT2 cell marker expression by IF staining. As shown in Figure 6F, AQP5 expression was maintained under both Nikolić and Jacob Alveolar conditions but appeared to be expressed in the cytoplasm rather than at the membrane. In particular, although we were able to detect enrichment of AQP5 expression on the luminal side of the membrane of AEC-ON organoids grown under control 3D medium, we did not observe this localization pattern in either Nikolić- nor Jacob Alveolar-grown organoids (Figures S10C and S11). Assessing NKX2-1 expression, we found a similarly altered localization in AEC-ON cells grown in 3D under Nikolić and Jacob Alveolar media as we did in 2D: NKX2-1 expression was cytoplasmic in Nikolić medium (Figures 6F, asterisk and S10C), whereas NKX2-1 was both nuclear and cytoplasmic in Jacob Alveolar medium (Figures 6F, asterisk, arrow and S10C). Probing for AT2 markers pro-SFTPC and HTII280, we discovered strong induction of HTII280 in a subpopulation of AEC-ON cells comprising organoids grown in Jacob Alveolar medium (Figure 6F, last panel, arrow). Cells that were positive for HTII280 expressed cytoplasmic NKX2-1, but not all cytoplasmic NKX2-1 expressing cells were positive for HTII280, indicating that these expression changes were independent of each other. Interestingly, neither pro-SFTPC nor mature SFTPC was detected. From these observations, we conclude that under certain growth conditions, it is possible to induce a more AT2-like cell fate in a proportion of AEC-ON cells. Furthermore, the added NOTCH inhibition and thyroid and cytokine signaling activation provided by the 10-reagent Nikolić medium may counteract the induction of HTII280 expression in AEC-ON organoids.

DISCUSSION

The lung is highly susceptible to environmental damage but is exquisitely engineered to deal with these insults. Upon lung injury, specific cells are mobilized to aid in repair. However, when repair and

regeneration are deficient, depending on the region injured, a variety of acute and chronic pulmonary diseases arise. Having a full spectrum of cell models available to investigate the complete pathophysiology of these diseases would greatly benefit the field. Here, we established a practical method for immortalizing alveolar epithelial cells from human adult lung tissue by first expanding isolated AT2 cells in ROCK inhibitor medium followed by SV40 LgT lentiviral transduction. We show that the method works robustly, using cells from three different donor lungs. Gene expression analyses of the cell lines show they are transcriptomically distinct from primary alveolar epithelial cells, lung fibroblasts, and LUAD cancer cell lines. Our novel AEC-LgT lines proliferate well in 2D culture and can form lung organoids expressing markers of alveolar epithelial cells in 3D co-culture with fibroblasts.

Our original goal was to derive immortalized cell lines from purified human AT2 cells. In the process of accomplishing this, we incidentally found that the cells exhibited gene signatures suggestive of either an "intermediate" alveolar epithelial cell state or a distal lung progenitor cell state. In the lung, the alveolar epithelium is maintained by proliferation of AT2 cells and transdifferentiation of a subpopulation of AT2 daughter cells into AT1 cells (Adamson and Bowden, 1974; Barkauskas et al., 2013; Nabhan et al., 2018). Liebler et al. (2016) found remarkable cellular heterogeneity even within the adult lung under homeostatic conditions, identifying populations of "intermediate" or "transitory" AECs expressing combinations of AT2 and AT1 cell markers. Although early *in vitro* transdifferentiation studies did not suggest functions for these intermediate cells, recent studies using scRNA-seq reveal potentially new roles for these and transitory lung cells in injury resolution, which when disrupted may facilitate disease progression, particularly in IPF (Adams et al., 2020; Strunz et al., 2020). Our findings suggest that cells comprising the AEC-LgT organoids generally represent an intermediate AEC state, exhibiting AT1-like cell expression patterns (AQP5+/GPRC5A+; SFTPC−/HTII280−) and possessing the AT2-like ability to form structures similar to "alveolospheres" recently reported by Katsura et al. (2020).

Distal lung progenitors in the mouse adult lung have been identified as rare subpopulations within the greater "bulk" AT2 population and are not all equally fated. AXIN2$^+$/TM4SF1$^+$ resident AEPs, comprising ~20% of bulk AT2 cells, give rise to lineage-labeled AT2 and AT1 cells at sites of lung injury after H1N1 influenza infection (Zacharias et al., 2018). AT2 "ancillary" progenitors express AXIN2 when activated upon lung injury, induced by WNT signaling from surrounding stromal cells (Nabhan et al., 2018). Although intermediate alveolar cells and alveolar progenitors exhibit different gene expression patterns, the presence of these cell types indicates that the alveolar epithelium has considerable cellular plasticity, which primes the lung to respond to injury expeditiously despite the normally slow turnover of lung cells (Bowden, 1983). By scRNA-seq, we discovered that AEC-ON organoids contained AXIN2$^+$/TM4SF1$^+$ cells; however, although AEC-ON organoids were WNT and FGF responsive, reminiscent of AEPs, activation of these pathways alone was not sufficient to induce AT2 cell differentiation or enrich for AT2-like cells. Robust expression of AT2-cell-specific marker HTII280 was induced with Jacob Alveolar medium, which contains not only CHIR99021 and FGF7 but also the glucocorticoid dexamethasone, cAMP, and IBMX (which raises intracellular cAMP and cGMP levels). Interestingly, the Jacob Alveolar medium also resulted in heterogeneous expression of NKX2-1 in the nucleus and cytoplasm. This change in localization may reflect changes in NKX2-1 regulation and interactions with co-factors, which may then affect AEC differentiation state. Yang et al. (2016) found that Sox2-expressing mutant AT1 cells marked by *Aqp5* exhibited diffuse Nkx2-1 expression, downregulated AT1 marker expression, and re-entered the cell cycle as marked by Ki67$^+$ proliferation. Although these mutant cells tended to cluster together and morphologically resembled airway cells, the authors concluded that these changes did not indicate a physiologically functional airway cell. Our observations may therefore suggest that activation of WNT, FGF7, and glucocorticoid plus cAMP signaling induces a switch in transcriptional program that initiates a period of cell proliferation before differentiation. Indeed, NKX2-1 has recently been found to regulate AT2 versus AT1 cell fates by epigenetically altering chromatin states (Little et al., 2021).

Although further exploration of growth conditions will be important, our immortalized AEC-derived distal lung cell lines provide an important unmet resource to the research community. Our derivation methodology can be readily applied to lung tissues of diverse racial/ethnic individuals with different genetic and epigenetic landscapes, generating appropriate models for health disparities research. The cells, which proliferate well in 2D, can be subjected to genetic and epigenetic manipulation prior to organoid formation, allowing their use to study differentiation of the distal lung compartment and diseases affecting the distal lung, such as cancer, emphysema, and pulmonary viral infections. We also envision that our 3D

AEC-LgT cell culture system can be expanded to incorporate additional elements of alveolar physiology, such as the vasculature, or be integrated with current lung-on-a-chip applications (Hassell et al., 2017).

Limitations of the study

In our characterization of the three AEC-LgT cell lines, we found that although these lines were more similar to each other than to primary human AECs or lung fibroblasts, there were differences among the three lines in 2D growth kinetics and in the frequency and complexity of organoids formed. These differences could arise from many sources, including genetic differences between subjects; epigenetic differences related to numerous factors including age, gender, environmental exposures, manner of death, and/or ventilation time; and technical differences between experiments related to AEC preparation or cellular response to culture conditions. These aspects will be important to study, as additional cell lines are made from a wider range of individuals.

Each of the AEC-LgT cell lines we derived will likely have specific applications of interest for further study; however, some limitations must be considered. First, although the reprogramming effect of ROCK inhibitor Y-27632 allowed for expansion of isolated adult AT2 cells as distal progenitor cells, reversion of the transcriptional and epigenetic changes must be further explored in order to fully re-differentiate the cell lines into mature alveolar cells. Recently, Katsumiti et al. (2020) reported the establishment of clonal lines from one human adult lung donor by transduction of hTERT and SV40 LgT. The cell lines expressed low levels of SFTPC at the RNA and protein levels, but the absence of *in situ* staining showing proper localization of surfactant emphasizes the challenge of maintaining proliferating AT2-like cells *in vitro*. A second limitation is that clear mRNA expression profiles of AT1 cells are still lacking and this hampers optimal characterization of our cells. AT1 cells are morphologically large and delicate and have been difficult to collect in large numbers, especially for bulk RNA sequencing. Although whole lung scRNA-seq studies have been performed (ie, IPF Cell Atlas, Neumark et al., 2020; Adams et al., 2020; Habermann et al., 2020; Reyfman et al., 2019; Morse et al., 2018), identified AT1 cells only comprise a small proportion of the total epithelial cell population. As AT1 cells are inherently fragile and the single cell sequencing platforms are sensitive to cell-to-cell noise, the AT1 gene signatures that arise from these analyses may not completely reflect homeostatic AT1 cells. Recent nuclear RNA scRNA-seq experiments have aided in capturing more AT1 cell expression profiles (Koenitzer et al., 2020) and will be important to expand on, while taking into account the possible limitations of examining nuclear RNA, which may be incompletely processed. Lastly, although we were able to induce HTII280 expression in a subset of AEC-ON cells under 3D co-culture conditions, SFTPC expression was still not detected, suggesting the presence of a potentially primed AT2 cell population in need of further external maturation cues, such as those provided by the abutting capillary network from endothelial cells as reported by Reyfman et al. (2019). Indeed, plans to implant the AEC-LgT cell lines into decellularized mouse lungs, as was done for airway cells (Gilpin et al., 2016), are in progress to test whether physiological cues can further induce alveolar epithelial cell differentiation and maturation.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mouse studies
  - Cell lines and primary cultures
- METHOD DETAILS
  - Mycoplasma and rodent pathogens testing
  - Derivation of human alveolar epithelial cell lines
  - Construction of lentiviral plasmids and production of viral particles
  - Lentiviral transduction of human AECs
  - Proliferation assay
  - Anchorage-independent growth assay
  - Three-dimensional (3D) co-culture
  - Treatment of 3D co-cultures

- ○ Histological processing of organoids
- ○ Immunofluorescence staining
- ○ Bulk RNA-sequencing analyses
- ○ Single cell RNA-sequencing analyses
- ● QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ Statistical analyses

AUTHOR CONTRIBUTIONS

ET primarily designed and performed the experiments. TS and XL assisted in preparing 3D organotypic cultures and immunostainings. YL and HW provided staining support. CY assisted with *in vivo* subcutaneous injection. IAO supervised the entire study. ET and IAO wrote the paper. AYC assisted with sample preparation for scRNA-seq. DJ performed scRNA-seq data processing and quality control. RL provided access to 10× Genomics Chromium device for scRNA-seq and consultation on data analysis approaches. WDW evaluated tissue H&E slides. ALR, CNM, BZ, and ZB provided resources including primary AEC. ET, BZ, ZB, and IAO interpreted results. All authors edited and approved the paper. Correspondence should be addressed to IAO.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure sex balance in the selection of non-human subjects. One or more of the authors of this paper received support from a program designed to increase minority representation in science.

REFERENCES

Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., Deluliis, G., Januszyk, M., et al. (2020). Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. Sci. Adv. 6, eaba1983. https://doi.org/10.1126/sciadv.aba1983.

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.

Ballard, P.L., and Ballard, R.A. (1974). Cytoplasmic receptor for glucocorticoids in lung of the human fetus and neonate. J. Clin. Invest. 53, 477–486. https://doi.org/10.1172/JCI107581.

Ballard, P.L., Lee, J.W., Fang, X., Chapin, C., Allen, L., Segal, M.R., Fischer, H., Illek, B., Gonzales, L.W., Kolla, V., et al. (2010). Regulated gene expression in cultured type II cells of adult human lung. Am. J. Physiol. Lung Cell. Mol. Physiol. 299, L36–L50. https://doi.org/10.1152/ajplung.00427.2009.

Band, V., De Caprio, J.A., Delmolino, L., Kulesa, V., and Sager, R. (1991). Loss of p53 protein in human papillomavirus type 16 E6-immortalized human mammary epithelial cells. J. Virol. 65, 6671–6676. https://doi.org/10.1128/JVI.65.12.6671-6676.1991.

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. (2013). Type 2 alveolar cells are stem cells in adult lung. J. Clin. Invest. 123, 3025–3036. https://doi.org/10.1172/JCI68782.

Beers, M.F., and Morrisey, E.E. (2011). The three R's of lung health and disease: repair, remodeling, and regeneration. J. Clin. Invest. 121, 2065–2073. https://doi.org/10.1172/JCI45961.

Borok, Z., Lubman, R.L., Danto, S.I., Zhang, X.L., Zabski, S.M., King, L.S., Lee, D.M., Agre, P., and Crandall, E.D. (1998). Keratinocyte growth factor modulates alveolar epithelial cell phenotype in vitro: expression of aquaporin 5. Am. J. Respir. Cell Mol. Biol. 18, 554–561. https://doi.org/10.1165/ajrcmb.18.4.2838.

Borowicz, S., Van Scoyk, M., Avasarala, S., Karuppusamy Rathinam, M.K., Tauler, J., Bikkavilli, R.K., and Winn, R.A. (2014). The soft agar colony formation assay. J. Vis. Exp. 92, e51998. https://doi.org/10.3791/51998.

Bove, P.F., Dang, H., Cheluvaraju, C., Jones, L.C., Liu, X., O'Neal, W.K., Randell, S.H., Schlegel, R., and Boucher, R.C. (2014). Breaking the in vitro alveolar type II cell proliferation barrier while retaining ion transport properties. Am. J. Respir. Cell Mol. Biol. 50, 767–776. https://doi.org/10.1165/rcmb.2013-0071OC.

Bowden, D.H. (1983). Cell turnover in the lung. Am. Rev. Respir. Dis. 128, S46–S48. https://doi.org/10.1164/arrd.1983.128.2P2.S46.

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.

Centers for Disease Control and Prevention (2018). Table 6. Leading causes of death and numbers of deaths, by sex, race, and hispanic origin: United States, 1980 and 2017 [fact sheet]. https://www.cdc.gov/nchs/hus/contents2018.htm#Table_006.

Chang, D.R., Martinez Alanis, D., Miller, R.K., Ji, H., Akiyama, H., McCrea, P.D., and Chen, J. (2013). Lung epithelial branching program antagonizes alveolar differentiation. Proc. Natl. Acad. Sci. U S A 110, 18042–18051. https://doi.org/10.1073/pnas.1311760110.

Chapman, S., McDermott, D.H., Shen, K., Jang, M.K., and McBride, A.A. (2014). The effect of Rho kinase inhibition on long-term keratinocyte proliferation is rapid and conditional. Stem Cell Res. Ther. 5, 60. https://doi.org/10.1186/scrt449.

Cheek, J.M., Evans, M.J., and Crandall, E.D. (1989). Type I cell-like morphology in tight alveolar epithelial monolayers. Exp. Cell Res. 184, 375–387. https://doi.org/10.1016/0014-4827(89)90337-6.

Christian, B.J., Loretz, L.J., Oberley, T.D., and Reznikoff, C.A. (1987). Characterization of human uroepithelial cells immortalized in vitro by simian virus 40. Cancer Res. 47, 6066–6073.

Claassen, D.A., Desler, M.M., and Rizzino, A. (2009). ROCK inhibition enhances the recovery and growth of cryopreserved human embryonic stem cells and human induced pluripotent stem cells. Mol. Reprod. Dev. 76, 722–732. https://doi.org/10.1002/mrd.21021.

Counter, C.M., Hahn, W.C., Wei, W., Caddle, S.D., Beijersbergen, R.L., Lansdorp, P.M., Sedivy, J.M., and Weinberg, R.A. (1998). Dissociation among in vitro telomerase activity, telomere maintenance, and cellular immortalization. Proc. Natl. Acad. Sci. U S A 95, 14723–14728. https://doi.org/10.1073/pnas.95.25.14723.

Daniely, Y., Liao, G., Dixon, D., Linnoila, R.I., Lori, A., Randell, S.H., Oren, M., and Jetten, A.M. (2004). Critical role of p63 in the development of a normal esophageal and tracheobronchial epithelium. Am. J. Physiol. Cell Physiol. 287, C171–C181. https://doi.org/10.1152/ajpcell.00226.2003.

Danopoulos, S., Alonso, I., Thornton, M.E., Grubbs, B.H., Bellusci, S., Warburton, D., and Al Alam, D. (2018). Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9. Am. J. Physiol. Lung Cell. Mol. Physiol. 314, L144–L149. https://doi.org/10.1152/ajplung.00379.2017.

Danto, S.I., Shannon, J.M., Borok, Z., Zabski, S.M., and Crandall, E.D. (1995). Reversible transdifferentiation of alveolar epithelial cells. Am. J. Respir. Cell Mol. Biol. 12, 497–502. https://doi.org/10.1165/ajrcmb.12.5.7742013.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. 46, D794–D801. https://doi.org/10.1093/nar/gkx1081.

Dobbs, L.G. (1990). Isolation and culture of alveolar type II cells. Am. J. Physiol. 258, L134–L147. https://doi.org/10.1152/ajplung.1990.258.4.L134.

Dobbs, L.G., Williams, M.C., and Brandt, A.E. (1985). Changes in biochemical characteristics and pattern of lectin binding of alveolar type II cells with time in culture. Biochim. Biophys. Acta 846, 155–166. https://doi.org/10.1016/0167-4889(85)90121-1.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. J. Virol. 72, 8463–8471. https://doi.org/10.1128/JVI.72.11.8463-8471.1998.

Dye, B.R., Hill, D.R., Ferguson, M.A., Tsai, Y.H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. Elife 4, e05098. https://doi.org/10.7554/eLife.05098.

Evans, M.J., and Bils, R.F. (1969). Identification of cells labeled with tritiated thymidine in the pulmonary alveolar walls of the mouse. Am. Rev. Respir. Dis. 100, 372–378. https://doi.org/10.1164/arrd.1969.100.3.372.

Evans, M.J., and Hackney, J.D. (1972). Cell proliferation in lungs of mice exposed to elevated concentrations of oxygen. Aerosp. Med. 43, 620–622.

Foster, C.D., Varghese, L.S., Skalina, R.B., Gonzales, L.W., and Guttentag, S.H. (2007). In vitro transdifferentiation of human fetal type II cells toward a type I-like cell. Pediatr. Res. 61, 404–409. https://doi.org/10.1203/pdr.0b013e3180332c6d.

Fuchs, S., Hollins, A.J., Laue, M., Schaefer, U.F., Roemer, K., Gumbleton, M., and Lehr, C.M. (2003). Differentiation of human alveolar epithelial cells in primary culture: morphological characterization and synthesis of caveolin-1 and surfactant protein-C. Cell Tissue Res. 311, 31–45. https://doi.org/10.1007/s00441-002-0653-5.

Giard, D.J., Aaronson, S.A., Todaro, G.J., Arnstein, P., Kersey, J.H., Dosik, H., and Parks, W.P. (1973). In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J. Natl. Cancer Inst. 51, 1417–1423. https://doi.org/10.1093/jnci/51.5.1417.

Gilpin, S.E., Charest, J.M., Ren, X., Tapias, L.F., Wu, T., Evangelista-Leite, D., Mathisen, D.J., and Ott, H.C. (2016). Regenerative potential of human airway stem cells in lung epithelial engineering. Biomaterials 108, 111–119. https://doi.org/10.1016/j.biomaterials.2016.08.055.

Gonzales, L.W., Angampalli, S., Guttentag, S.H., Beers, M.F., Feinstein, S.I., Matlapudi, A., and Ballard, P.L. (2001). Maintenance of differentiated function of the surfactant system in human fetal lung type II epithelial cells cultured on plastic.

Pediatr. Pathol. Mol. Med. *20*, 387–412. https://doi.org/10.1080/15513810109168622.

Gonzales, L.W., Guttentag, S.H., Wade, K.C., Postle, A.D., and Ballard, P.L. (2002). Differentiation of human pulmonary type II cells in vitro by glucocorticoid plus cAMP. Am. J. Physiol. Lung Cell. Mol. Physiol. *283*, L940–L951. https://doi.org/10.1152/ajplung.00127.2002.

Gonzalez, R.F., Allen, L., Gonzales, L., Ballard, P.L., and Dobbs, L.G. (2010). HTII-280, a biomarker specific to the apical plasma membrane of human lung alveolar type II cells. J. Histochem. Cytochem. *58*, 891–901. https://doi.org/10.1369/jhc.2010.956433.

Gotoh, S., Ito, I., Nagasaki, T., Yamamoto, Y., Konishi, S., Korogi, Y., Matsumoto, H., Muro, S., Hirai, T., Funato, M., et al. (2014). Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. Stem Cell Rep. *3*, 394–403. https://doi.org/10.1016/j.stemcr.2014.07.005.

Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L., Chung, M.I., Taylor, C.J., Jetter, C., et al. (2020). Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Sci. Adv. *6*, eaba1972. https://doi.org/10.1126/sciadv.aba1972.

Hackney, J.D., Evans, M.J., and Christie, B.R. (1975). Effects of 60 and 80% oxygen on cell division in lung alveoli of squirrel monkeys. Aviat. Space Environ. Med. *46*, 791–794.

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296. https://doi.org/10.1186/s13059-019-1874-1.

Hahn, W.C., and Weinberg, R.A. (2002). Rules for making human tumor cells. N. Engl. J. Med. *347*, 1593–1603. https://doi.org/10.1056/NEJMra021902.

Hahn, W.C., Dessain, S.K., Brooks, M.W., King, J.E., Elenbaas, B., Sabatini, D.M., DeCaprio, J.A., and Weinberg, R.A. (2002). Enumeration of the simian virus 40 early region elements necessary for human cell transformation. Mol. Cell. Biol. *22*, 2111–2123. https://doi.org/10.1128/MCB.22.7.2111-2123.2002.

Hassell, B.A., Goyal, G., Lee, E., Sontheimer-Phelps, A., Levy, O., Chen, C.S., and Ingber, D.E. (2017). Human organ chip models recapitulate orthotopic lung cancer growth, therapeutic responses, and tumor dormancy in vitro. Cell Rep. *21*, 508–516. https://doi.org/10.1016/j.celrep.2017.09.043.

Herbert, B.S., Wright, W.E., and Shay, J.W. (2002). p16(INK4a) inactivation is not required to immortalize human mammary epithelial cells. Oncogene *21*, 7897–7900. https://doi.org/10.1038/sj.onc.1205902.

Herranz, N., and Gil, J. (2018). Mechanisms and functions of cellular senescence. J. Clin. Invest. *128*, 1238–1246. https://doi.org/10.1172/JCI95148.

Horani, A., Nath, A., Wasserman, M.G., Huang, T., and Brody, S.L. (2013). Rho-associated protein kinase inhibition enhances airway epithelial basal-cell proliferation and lentivirus transduction. Am. J. Respir. Cell Mol. Biol. *49*, 341–347. https://doi.org/10.1165/rcmb.2013-0046TE.

Horie, M., Castaldi, A., Sunohara, M., Wang, H., Ji, Y., Liu, Y., Li, F., Wilkinson, T.A., Hung, L., Shen, H., et al. (2020). Integrated single-cell RNA-sequencing analysis of aquaporin 5-expressing mouse lung epithelial cells identifies GPRC5A as a novel validated type I cell surface marker. Cells *9*, 2460. https://doi.org/10.3390/cells9112460.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., et al. (2021). Ensembl 2021. Nucleic Acids Res. *49*, D884–D891. https://doi.org/10.1093/nar/gkaa942.

Isakson, B.E., Seedorf, G.J., Lubman, R.L., and Boitano, S. (2002). Heterocellular cultures of pulmonary alveolar epithelial cells grown on laminin-5 supplemented matrix. In Vitro Cell. Dev. Biol. Anim. *38*, 443–449. https://doi.org/10.1290/1071-2690(2002)038<0443:HCOPAE>2.0.CO;2.

Ishikawa, Y., Kozakai, T., Morita, H., Saida, K., Oka, S., and Masuo, Y. (2006). Rapid detection of mycoplasma contamination in cell cultures using SYBR Green-based real-time polymerase chain reaction. In Vitro Cell. Dev. Biol. Anim. *42*, 63–69. https://doi.org/10.1290/0505035.1.

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.

Jain, R., Barkauskas, C.E., Takeda, N., Bowie, E.J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L.J., Gupta, M., Li, D., et al. (2015). Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. Nat. Commun. *6*, 6727. https://doi.org/10.1038/ncomms7727.

Jiang, Y., Cui, L., Yie, T.A., Rom, W.N., Cheng, H., and Tchou-Wong, K.M. (2001). Inhibition of anchorage-independent growth and lung metastasis of A549 lung carcinoma cells by IkappaBbeta. Oncogene *20*, 2254–2263. https://doi.org/10.1038/sj.onc.1204293.

Kanagaki, S., Ikeo, S., Suezawa, T., Yamamoto, Y., Seki, M., Hirai, T., Hagiwara, M., Suzuki, Y., and Gotoh, S. (2020). Directed induction of alveolar type I cells derived from pluripotent stem cells via Wnt signaling inhibition. Stem Cells *39*, 156–169. https://doi.org/10.1002/stem.3302.

Kang, Y., Omura, M., Suzuki, A., Oka, T., Nakagami, Y., Cheng, C., Nagashima, Y., and Inoue, T. (2006). Development of an orthotopic transplantation model in nude mice that simulates the clinical features of human lung cancer. Cancer Sci. *97*, 996–1001. https://doi.org/10.1111/j.1349-7006.2006.00276.x.

Katsumiti, A., Ruenraroengsak, P., Cajaraville, M.P., Thorley, A.J., and Tetley, T.D. (2020). Immortalisation of primary human alveolar epithelial lung cells using a non-viral vector to study respiratory bioreactivity in vitro. Sci. Rep. *10*, 20486. https://doi.org/10.1038/s41598-020-77191-y.

Katsura, H., Sontake, V., Tata, A., Kobayashi, Y., Edwards, C.E., Heaton, B.E., Konkimalla, A.,
Asakura, T., Mikami, Y., Fritch, E.J., et al. (2020). Human lung stem cell-based alveolospheres provide insights into SARS-CoV-2-mediated interferon responses and pneumocyte dysfunction. Cell Stem Cell *27*, 890–904.e8. https://doi.org/10.1016/j.stem.2020.10.005.

Kelly, S.E., Bachurski, C.J., Burhans, M.S., and Glasser, S.W. (1996). Transcription of the lung-specific surfactant protein C gene is mediated by thyroid transcription factor 1. J. Biol. Chem. *271*, 6881–6888. https://doi.org/10.1074/jbc.271.12.6881.

Kemp, S.J., Thorley, A.J., Gorelik, J., Seckl, M.J., O'Hare, M.J., Arcaro, A., Korchev, Y., Goldstraw, P., and Tetley, T.D. (2008). Immortalization of human alveolar epithelial cells to investigate nanoparticle uptake. Am. J. Respir. Cell Mol. Biol. *39*, 591–597. https://doi.org/10.1165/rcmb.2007-0334OC.

Khalil, N., O'Connor, R.N., Flanders, K.C., Shing, W., and Whitman, C.I. (1994). Regulation of type II alveolar epithelial cell proliferation by TGF-beta during bleomycin-induced lung injury in rats. Am. J. Physiol. *267*, L498–L507. https://doi.org/10.1152/ajplung.1994.267.5.L498.

Koenitzer, J.R., Wu, H., Atkinson, J.J., Brody, S.L., and Humphreys, B.D. (2020). Single-nucleus RNA-sequencing profiling of mouse lung. Reduced dissociation bias and improved rare cell-type detection compared with single-cell RNA sequencing. Am. J. Respir. Cell Mol. Biol. *63*, 739–747. https://doi.org/10.1165/rcmb.2020-0095MA.

Kuehn, A., Kletting, S., de Souza Carvalho-Wodarz, C., Repnik, U., Griffiths, G., Fischer, U., Meese, E., Huwer, H., Wirth, D., May, T., et al. (2016). Human alveolar epithelial cells expressing tight junctions to model the air-blood barrier. ALTEX *33*, 251–260. https://doi.org/10.14573/altex.1511131.

Liebler, J.M., Marconett, C.N., Juul, N., Wang, H., Liu, Y., Flodby, P., Laird-Offringa, I.A., Minoo, P., and Zhou, B. (2016). Combinations of differentiation markers distinguish subpopulations of alveolar epithelial cells in adult lung. Am. J. Physiol. Lung Cell. Mol. Physiol. *310*, L114–L120. https://doi.org/10.1152/ajplung.00337.2015.

Little, D.R., Gerner-Mauro, K.N., Flodby, P., Crandall, E.D., Borok, Z., Akiyama, H., Kimura, S., Ostrin, E.J., and Chen, J. (2019). Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1. Proc. Natl. Acad. Sci. U S A *116*, 20545–20555. https://doi.org/10.1073/pnas.1906663116.

Little, D.R., Lynch, A.M., Yan, Y., Akiyama, H., Kimura, S., and Chen, J. (2021). Differential chromatin binding of the lung lineage transcription factor NKX2-1 resolves opposing murine alveolar cell fates in vivo. Nat. Commun. *12*, 2509. https://doi.org/10.1038/s41467-021-22817-6.

Liu, J.P., Cassar, L., Pinto, A., and Li, H. (2006). Mechanisms of cell immortalization mediated by EB viral activation of telomerase in nasopharyngeal carcinoma. Cell Res. *16*, 809–817. https://doi.org/10.1038/sj.cr.7310098.

Liu, Y., Sadikot, R.T., Adami, G.R., Kalinichenko, V.V., Pendyala, S., Natarajan, V., Zhao, Y.Y., and

Malik, A.B. (2011). FoxM1 mediates the progenitor function of type II epithelial cells in repairing alveolar injury induced by Pseudomonas aeruginosa. J. Exp. Med. *208*, 1473–1484. https://doi.org/10.1084/jem.20102041.

Liu, X., Ory, V., Chapman, S., Yuan, H., Albanese, C., Kallakury, B., Timofeeva, O.A., Nealon, C., Dakic, A., Simic, V., et al. (2012). ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. Am. J. Pathol. *180*, 599–607. https://doi.org/10.1016/j.ajpath.2011.10.036.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

Luo, J., Chimge, N.O., Zhou, B., Flodby, P., Castaldi, A., Firth, A.L., Liu, Y., Wang, H., Yang, C., Marconett, C.N., et al. (2018). CLDN18.1 attenuates malignancy and related signaling pathways of lung adenocarcinoma in vivo and in vitro. Int. J. Cancer *143*, 3169–3180. https://doi.org/10.1002/ijc.31734.

Mao, P., Wu, S., Li, J., Fu, W., He, W., Liu, X., Slutsky, A.S., Zhang, H., and Li, Y. (2015). Human alveolar epithelial type II cells in primary culture. Physiol. Rep. *3*, e12288. https://doi.org/10.14814/phy2.12288.

Marconett, C.N., Zhou, B., Rieger, M.E., Selamat, S.A., Dubourd, M., Fang, X., Lynch, S.K., Stueve, T.R., Siegmund, K.D., Berman, B.P., et al. (2013). Integrated transcriptomic and epigenomic analysis of primary human lung epithelial cell differentiation. PLoS Genet. *9*, e1003513. https://doi.org/10.1371/journal.pgen.1003513.

Martinovich, K.M., Iosifidis, T., Buckley, A.G., Looi, K., Ling, K.M., Sutanto, E.N., Kicic-Starcevich, E., Garratt, L.W., Shaw, N.C., Montgomery, S., et al. (2017). Conditionally reprogrammed primary airway epithelial cells maintain morphology, lineage and disease specific functional characteristics. Sci. Rep. *7*, 17971. https://doi.org/10.1038/s41598-017-17952-4.

McCauley, K.B., Alysandratos, K.D., Jacob, A., Hawkins, F., Caballero, I.S., Vedaie, M., Yang, W., Slovik, K.J., Morley, M., Carraro, G., et al. (2018). Single-cell transcriptomic profiling of pluripotent stem cell-derived SCGB3A2+ airway epithelium. Stem Cell Rep. *10*, 1579–1595. https://doi.org/10.1016/j.stemcr.2018.03.013.

Meng, G., Liu, S., and Rancourt, D.E. (2012). Synergistic effect of medium, matrix, and exogenous factors on the adhesion and growth of human pluripotent stem cells under defined, xeno-free conditions. Stem Cells Dev. *21*, 2036–2048. https://doi.org/10.1089/scd.2011.0489.

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.P., Mushayamaha, T., and Thomas, P.D. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. Nucleic Acids Res. *49*, D394–D403. https://doi.org/10.1093/nar/gkaa1106.

Morse, C., Tabib, T., Sembrat, J., Buschur, K.L., Bittar, H.T., Valenzi, E., Jiang, Y., Kass, D.J., Gibson, K., Chen, W., et al. (2018). Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. Eur. Respir. J. *54*, 1802441. https://doi.org/10.1183/13993003.02441-2018.

Nabhan, A.N., Brownfield, D.G., Harbury, P.B., Krasnow, M.A., and Desai, T.J. (2018). Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. Science *359*, 1118–1123. https://doi.org/10.1126/science.aam6603.

Neufeld, D.S., Ripley, S., Henderson, A., and Ozer, H.L. (1987). Immortalization of human fibroblasts transformed by origin-defective simian virus 40. Mol. Cell. Biol. *7*, 2794–2802. https://doi.org/10.1128/mcb.7.8.2794.

Neumark, N., Cosme, C., Jr., Rose, K.A., and Kaminski, N. (2020). The idiopathic pulmonary fibrosis cell Atlas. Am. J. Physiol. Lung Cell. Mol. Physiol. *319*, L887–L893. https://doi.org/10.1152/ajplung.00451.2020.

Nikolić, M.Z., Caritg, O., Jeng, Q., Johnson, J.A., Sun, D., Howell, K.J., Brady, J.L., Laresgoiti, U., Allen, G., Butler, R., et al. (2017). Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. Elife *6*, e26575. https://doi.org/10.7554/eLife.26575.

Ochieng, J.K., Schilders, K., Kool, H., Boerema-De Munck, A., Buscop-Van Kempen, M., Gontan, C., Smits, R., Grosveld, F.G., Wijnen, R.M., Tibboel, D., et al. (2014). Sox2 regulates the emergence of lung basal cells by directly activating the transcription of Trp63. Am. J. Respir. Cell Mol. Biol. *51*, 311–322. https://doi.org/10.1165/rcmb.2013-0419OC.

O'Hare, M.J., Bond, J., Clarke, C., Takeuchi, Y., Atherton, A.J., Berry, C., Moody, J., Silver, A.R., Davies, D.C., Alsop, A.E., et al. (2001). Conditional immortalization of freshly isolated human mammary fibroblasts and endothelial cells. Proc. Natl. Acad. Sci. U S A *98*, 646–651. https://doi.org/10.1073/pnas.98.2.646.

Rackley, C.R., and Stripp, B.R. (2012). Building and maintaining the epithelium of the lung. J. Clin. Invest. *122*, 2724–2730. https://doi.org/10.1172/JCI60519.

Ramirez, R.D., Sheridan, S., Girard, L., Sato, M., Kim, Y., Pollack, J., Peyton, M., Zou, Y., Kurie, J.M., Dimaio, J.M., et al. (2004). Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res. *64*, 9027–9034. https://doi.org/10.1158/0008-5472.CAN-04-3703.

Reddel, R.R. (2000). The role of senescence and immortalization in carcinogenesis. Carcinogenesis *21*, 477–484. https://doi.org/10.1093/carcin/21.3.477.

Reddel, R.R., Ke, Y., Gerwin, B.I., McMenamin, M.G., Lechner, J.F., Su, R.T., Brash, D.E., Park, J.B., Rhim, J.S., and Harris, C.C. (1988). Transformation of human bronchial epithelial cells by infection with SV40 or adenovirus-12 SV40 hybrid virus, or transfection via strontium phosphate coprecipitation with a plasmid containing SV40 early region genes. Cancer Res. *48*, 1904–1909.

Reyfman, P.A., Walter, J.M., Joshi, N., Anekalla, K.R., McQuattie-Pimentel, A.C., Chiu, S., Fernandez, R., Akbarpour, M., Chen, C.I., Ren, Z., et al. (2019). Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. Am. J. Respir. Crit. Care Med. *199*, 1517–1536. https://doi.org/10.1164/rccm.201712-2410OC.

Rock, J.R., and Hogan, B.L. (2011). Epithelial progenitor cells in lung development, maintenance, repair, and disease. Annu. Rev. Cell Dev. Biol. *27*, 493–512. https://doi.org/10.1146/annurev-cellbio-100109-104040.

Rockich, B.E., Hrycaj, S.M., Shih, H.P., Nagy, M.S., Ferguson, M.A., Kopp, J.L., Sander, M., Wellik, D.M., and Spence, J.R. (2013). Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proc. Natl. Acad. Sci. U S A *110*, E4456–E4464. https://doi.org/10.1073/pnas.1311847110.

Sasai, K., Sukezane, T., Yanagita, E., Nakagawa, H., Hotta, A., Itoh, T., and Akagi, T. (2011). Oncogene-mediated human lung epithelial cell transformation produces adenocarcinoma phenotypes in vivo. Cancer Res. *71*, 2541–2549. https://doi.org/10.1158/0008-5472.CAN-10-2221.

Schiller, J., Sabatini, L., Bittner, G., Pinkerman, C., Mayotte, J., Levitt, M., and Meisner, L. (1994). Phenotypic, molecular and genetic-characterization of transformed human bronchial epithelial-cell strains. Int. J. Oncol. *4*, 461–470. https://doi.org/10.3892/ijo.4.2.461.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

Siegel, R.L., Miller, K.D., and Jemal, A. (2020). Cancer statistics, 2020. CA Cancer J. Clin. *70*, 7–30. https://doi.org/10.3322/caac.21590.

Smith, J.L., Lee, L.C., Read, A., Li, Q., Yu, B., Lee, C.S., and Luo, J. (2016). One-step immortalization of primary human airway epithelial cells capable of oncogenic transformation. Cell Biosci. *6*, 57. https://doi.org/10.1186/s13578-016-0122-6.

Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA *9*, 493–501. https://doi.org/10.1261/rna.2192803.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

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.

Suprynowicz, F.A., Upadhyay, G., Krawczyk, E., Kramer, S.C., Hebert, J.D., Liu, X., Yuan, H., Cheluvaraju, C., Clapp, P.W., Boucher, R.C., Jr., et al. (2012). Conditionally reprogrammed cells represent a stem-like state of adult epithelial cells. Proc. Natl. Acad. Sci. U S A *109*, 20035–

20040. https://doi.org/10.1073/pnas.1213241109.

Suzuki, A., Kawano, S., Mitsuyama, T., Suyama, M., Kanai, Y., Shirahige, K., Sasaki, H., Tokunaga, K., Tsuchihara, K., Sugano, S., et al. (2018). DBTSS/DBKERO for integrated analysis of transcriptional regulation. Nucleic Acids Res. 46, D229–D238. https://doi.org/10.1093/nar/gkx1001.

Tata, P.R., and Rajagopal, J. (2017). Plasticity in the lung: making and breaking cell identity. Development 144, 755–766. https://doi.org/10.1242/dev.143784.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375. https://doi.org/10.1038/nature13173.

Vernardis, S.I., Terzoudis, K., Panoskaltsis, N., and Mantalaris, A. (2017). Human embryonic and induced pluripotent stem cells maintain phenotype but alter their metabolism after exposure to ROCK inhibitor. Sci. Rep. 7, 42138. https://doi.org/10.1038/srep42138.

Volpato, V., and Webber, C. (2020). Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. Dis. Model. Mech. 13, dmm042317. https://doi.org/10.1242/dmm.042317.

Wang, Y., Tang, Z., Huang, H., Li, J., Wang, Z., Yu, Y., Zhang, C., Li, J., Dai, H., Wang, F., et al. (2018).

Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. Proc. Natl. Acad. Sci. U S A 115, 2407–2412. https://doi.org/10.1073/pnas.1719474115.

Weber, K., Bartsch, U., Stocking, C., and Fehse, B. (2008). A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. Mol. Ther. 16, 698–706. https://doi.org/10.1038/mt.2008.6.

Wu, D., and Pan, W. (2010). GSK3: a multifaceted kinase in Wnt signaling. Trends Biochem. Sci. 35, 161–168. https://doi.org/10.1016/j.tibs.2009.10.002.

Xu, Y., Mizuno, T., Sridharan, A., Du, Y., Guo, M., Tang, J., Wikenheiser-Brokamp, K.A., Perl, A.T., Funari, V.A., Gokey, J.J., et al. (2016). Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight 1, e90558. https://doi.org/10.1172/jci.insight.90558.

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.

Yang, J., Hernandez, B.J., Martinez Alanis, D., Narvaez del Pilar, O., Vila-Ellis, L., Akiyama, H., Evans, S.E., Ostrin, E.J., and Chen, J. (2016). The development and plasticity of alveolar type 1 cells. Development 143, 54–65. https://doi.org/10.1242/dev.130005.

Yang, C., Stueve, T.R., Yan, C., Rhie, S.K., Mullen, D.J., Luo, J., Zhou, B., Borok, Z., Marconett, C.N., and Offringa, I.A. (2018). Positional integration of lung adenocarcinoma susceptibility loci with primary human alveolar epithelial cell epigenomes. Epigenomics 10, 1167–1187. https://doi.org/10.2217/epi-2018-0003.

Zacharias, W.J., Frank, D.B., Zepp, J.A., Morley, M.P., Alkhaleel, F.A., Kong, J., Zhou, S., Cantu, E., and Morrisey, E.E. (2018). Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. Nature 555, 251–255. https://doi.org/10.1038/nature25786.

Zhang, F., Nielsen, L.D., Lucas, J.J., and Mason, R.J. (2004). Transforming growth factor-beta antagonizes alveolar type II cell proliferation induced by keratinocyte growth factor. Am. J. Respir. Cell Mol. Biol. 31, 679–686. https://doi.org/10.1165/rcmb.2004-0182OC.

Zhao, J.J., Gjoerup, O.V., Subramanian, R.R., Cheng, Y., Chen, W., Roberts, T.M., and Hahn, W.C. (2003). Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase. Cancer Cell 3, 483–495. https://doi.org/10.1016/s1535-6108(03)00088-6.

Zhou, B., Flodby, P., Luo, J., Castillo, D.R., Liu, Y., Yu, F.X., McConnell, A., Varghese, B., Li, G., Chimge, N.O., et al. (2018). Claudin-18-mediated YAP activity regulates lung stem and progenitor cell homeostasis and tumorigenesis. J. Clin. Invest. 128, 970–984. https://doi.org/10.1172/JCI90429.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Monoclonal rabbit anti-AQP5	Abcam	ab92320; RRID:AB_2049171
Monoclonal mouse anti-E Cadherin	BD Transduction Laboratories	610181; RRID:AB_397581
Polyclonal rabbit anti-GPRC5A	Abbexa	abx005719; N/A
Monoclonal mouse anti-HTII280 (IgM)	Terrace Biotech	TB-27AHT2-280; RRID:AB_2665381
Polyclonal rabbit anti-HOPX	SCBT	sc-30216; RRID:AB_2120833
Monoclonal rabbit anti-Ki67	Abcam	ab16667; RRID:AB_302459
Monoclonal mouse anti-NKX2-1 (NCL-L-TTF1)	Leica	TTF-1-L-CE; RRID:AB_442138
Polyclonal rabbit anti-proSFTPC	Seven Hills Bioreagent	WRAB-9337; RRID:AB_2335890
Monoclonal mouse anti-SOX2	SCBT	sc-365823; RRID:AB_10842165
Polyclonal rabbit anti-SOX9	SCBT	sc-20095; RRID:AB_661282
Monoclonal mouse anti-SV40 Large T antigen	SCBT	sc-147; RRID:AB_628305
Mouse IgG control	Vector Laboratories	I-2000; RRID:AB_2336354
Rabbit IgG control	Vector Laboratories	I-1000; RRID:AB_2336355
Mouse IgM control	Sigma-Aldrich	M5909; RRID:AB_1163655
Biotinylated horse anti-mouse IgG	Vector Laboratories	BA-2000; RRID:AB_2313581
Biotinylated goat anti-rabbit IgG	Vector Laboratories	BA-1000; RRID:AB_2313606
Biotinylated goat anti-mouse IgM	Vector Laboratories	BA-2020; RRID:AB_2336183
Streptavidin-Alexa Fluor 647 conjugate	ThermoFisher	S21374; RRID:AB_2336066
Streptavidin-Alexa Fluor 488 conjugate	ThermoFisher	S11223; N/A
Streptavidin-FITC conjugate	ThermoFisher	SA-10002; N/A
Donkey anti-mouse Alexa Fluor 488 secondary	ThermoFisher	A21202; RRID:AB_141607
Donkey anti-rabbit Alexa Fluor 488 secondary	ThermoFisher	A21206; RRID:AB_2535792
Bacterial and virus strains		
Stbl3 chemically competent *E.coli*	ThermoFisher	C737303
Biological samples		
Human Lung donor tissue samples	This paper	
Chemicals, peptides, and recombinant proteins		
Y-27632	Enzo Life Sciences	270-333
SB-431542	BioVision	1674
CHIR99021	Sigma-Aldrich	SML1046
FGF7	Peprotech	100-19
FGF10	Peprotech	100-26
Insulin	Sigma-Aldrich	I0516
hEGF	ThermoFisher	PHG0311
Cholera toxin	Sigma-Aldrich	C8052
Hydrocortisone	Sigma-Aldrich	H0888
Insulin-Transferrin-Selenium (ITS) mix	Gibco	41400-045
Dexamethasone	Sigma-Aldrich	D4902
8-Bromo-cyclic-AMP	Cayman Chemical	14431
IBMX	Enzo Life Sciences	BML-PD140

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Matrigel Matrix Growth Factor Reduced	Corning	354230
Matrigel Membrane Matrix	Corning	354234
Histogel	ThermoFisher	HG-4000
DAPI mounting solution	Vector Laboratories	H-1200
Deposited data		
Raw data SuperSeries	This paper	GEO: GSE164515
Raw and analyzed bulk RNA-seq data SubSeries	This paper	GEO: GSE164513
Raw and analyzed scRNA-seq data SubSeries	This paper	GEO: GSE164514
Experimental models: Cell lines		
HEK 293T cells	ATCC	CRL-3216; RRID:CVCL_0063
Human A549 lung adenocarcinoma cells	ATCC	CRM-CCL-185; RRID:CVCL_0023
Mouse MLg fibroblasts	ATCC	CCL-206; RRID:CVCL_0437
AEC-FT-ROCKinh cells	This paper	
AEC-ON-ROCKinh cells	This paper	
AEC-TN-ROCKinh cells	This paper	
AEC-FT cells	This paper	
AEC-ON cells	This paper	
AEC-TN cells	This paper	
Experimental models: Organisms/strains		
NU/J mice, 6 weeks, homozygous $Foxn1^{nu}$	Jackson Laboratories	002019; RRID:IMSR_JAX:002019
Recombinant DNA		
Plasmid: pBABE-hygro-CDK4 R24C	Hahn et al., Mol Cell Biol 2002 Apr;22(7):2111-23.	RRID:Addgene_11254
Plasmid: pBABE-puro-hTERT	Counter et al. Proc Natl Acad Sci U S A 1998 Dec 8;95(25):14723-8.	RRID:Addgene_1771
Plasmid: pBABE-puro SV40 LgT	Zhao et al. Cancer Cell. 2003 May; 3(5):483-95.	RRID:Addgene_13970
Plasmid: pMDLg/pRRE	Dull et al. J Virol. 1998 Nov; 72(11):8463-71.	RRID:Addgene_12251
Plasmid: pRSV-Rev	Dull et al. J Virol. 1998 Nov; 72(11):8463-71.	RRID:Addgene_12253
Plasmid: pCMV-VSVG	Stewart et al., RNA 2003 Apr;9(4):493–501	RRID:Addgene_8454
Plasmid: LeGO iG	Weber et al. Mol Ther. 2008 Apr; 16(4):698–706	RRID:Addgene_27358
Plasmid: LeGO iT	Weber et al. Mol Ther. 2008 Apr; 16(4):698–706	RRID:Addgene_27361
Plasmid: LeGO iG-CDK4^{R24C}	This paper	
Plasmid: LeGO iG-SV40 LgT	This paper	
Plasmid: LeGO iT-hTERT	This paper	
Software and algorithms		
ImageJ	Schneider et al. Nat Methods. 2012, 9:671–675	https://imagej.nih.gov/ij/index.html
Partek Flow	Partek Inc. (2019). Partek®Flow®(Version 10.0) [Computer software]	https://www.partek.com/partek-flow/
STAR aligner (v2.6.1d)	Dobin et al. Bioinformatics. 2013 Jan; 29(1):15–21	(Implemented within Partek Flow software)
R (v3.6.0)	R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2019	https://www.r-project.org

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DESeq2	Love et al. Genome Biol. 2014; 15(12): 550	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
PANTHER14.1	Mi et al. Nucl Acids Res. 2021 Dec; 49(D1):D394-D403	http://www.pantherdb.org/news/news20190327.jsp
Ensemble gene database	Howe et al. Nucl Acids Res. 2021, 49(1):884–891	https://uswest.ensembl.org/index.html
Cell Ranger (v3.1.0)	10× Genomics Cell Ranger 3.1.0 [Computer Software]	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
Seurat (v2)	Butler et al., Nat Biotechnol 2018 Apr; 36:411–420	https://satijalab.org/seurat/articles/install.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ite A. Offringa (ilaird@usc.edu).

Materials availability

Immortalized alveolar epithelial cell lines generated in this study are available upon request with the submission of the signed material transfer agreement supplied by USC.

Data and code availability

- Bulk RNA-seq and single-cell RNA seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.

- This paper also analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.

- Microscopy data reported in this paper will be shared by the lead contact upon request.

- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse studies

Mouse experiments were performed under the guidance of the University of Southern California Institutional Animal Care and Use Committee (IACUC protocol ID 21116). Tumorigenicity of the three AEC-LgT cell lines was assessed by subcutaneous injection of 1×10^6 cells into each flank of 6-week-old male and female homozygous Foxn1^{nu} NU/J mice (Jackson Laboratories, strain 002019). NU/J mice were kept in sterile housing and given irradiated rodent feed *ad libitum*. Ten male and ten female NU/J mice were used in the study. Equal numbers of male and female mice were included in experimental and control groups. On the day of injection, cells expanded on standard tissue culture dishes were detached with Accutase, collected in Fmed + ROCKinh growth medium, then centrifuged. Supernatant was removed and then cells were resuspended in 1 mL growth medium and manually counted using a hemocytometer. For each flank, a master mix of cells resuspended in sterile phosphate-buffered saline (PBS) and 50% final concentration of Matrigel Membrane Matrix solution (Corning 354234) was prepared such that 1×10^6 cells were delivered in 150 μL of solution. Mixes were kept on ice as injections were performed to prevent premature solidification. Cell mixtures were injected using a tuberculin syringe with attached 27G needle (BD 305620). Mice were anesthetized by inhaled isoflurane according to IACUC-approved procedures. AEC-hTERT cells were used as negative controls as they were moderately proliferative out of the four slow-growing cell lines but did not form colonies in soft agar (see Figure 1D). A549 cells (RRID:CVCL_0023) were used as positive controls (Jiang et al., 2001; Kang et al., 2006). Four mice (8 flanks total) for A549 cells and eight mice (8 flanks total) for each AEC line were used. Injections were performed as a single-blind

study, where sample identities were unknown to the experimenter performing injections. Mice were monitored every 3 days, weighed every week, and nodule length (L) and width (W) were measured weekly with a digital caliper (iKKEGOL model 714838802360, Shenzen, China) starting 3 weeks post-injection. Nodule volume (V) was calculated as V = L X (W/2)2 (Luo et al., 2018). Excised nodules were fixed in 4% paraformaldehyde (PFA) overnight at 4°C, processed for paraffin embedding and sectioning by the Norris Cancer Center Pathology Core at USC. An expert lung pathologist blind to sample identities was consulted to evaluate H&E nodule sections.

Cell lines and primary cultures

A549 (RRID:CVCL_0023; a human lung carcinoma cell line originally derived from a 58-year old white male) and HEK293T (RRID:CVCL_0063; a human embryonic kidney cell line originally derived from a fetus, gender unknown) cell lines were authenticated by DNA fingerprinting through the University of Arizona Genetics Core and were routinely tested for mycoplasma. A549 cells were grown in RPMI 1640 medium (Gibco 11875119) with 10% FBS (X&Y Cell Culture LLC FBS-500) and 1% Pen/Strep (ThermoFisher Scientific 15140122) at 37°C. HEK293T cells were grown in DMEM medium (Corning 10-013-CV) with 10% FBS and 1% Pen/Strep at 37°C. De-identified remnant human transplant lungs from deceased subjects were obtained in compliance with the University of Southern California Institutional Review Board-approved protocols for the use of human source material in research (HS-07-0060). Lungs were processed within 3 days of death. Lung tissue was collected from cancer-free donors: Lung-FT, 25 years old (yr), Caucasian, male; Lung-ON, 66 yr, Caucasian, female; Lung-TN, 62 yr, Caucasian, male (Table 1). Human AT2 cells were isolated and purified as previously described (Ballard et al., 2010; Marconett et al., 2013), modified by using anti-EpCAM beads and purity was assessed by staining of cytospins (Table 1). AT2 cells were resuspended in 50:50 growth medium [50% DMEM-F12 (Sigma-Aldrich D64421), 50% DMEM High glucose (Gibco 21063) supplemented with 10% FBS (Omega Scientific FB-11), Pen/Strep, Gentamycin, and Amphotericin B]. The ability to differentiate into AT1-like cells was assessed by *SFTPC* and *AQP5* expression by qPCR and Western blot analyses as described by Marconett et al. (2013). Remaining purified AT2 cells were tested for mycoplasma then resuspended in 90% growth medium-10% DMSO, frozen in cryovials at 1-2x10^6 cells/ml, and stored in liquid nitrogen.

METHOD DETAILS

Mycoplasma and rodent pathogens testing

Cells used in the study were negative for mycoplasma and rodent pathogens. Cells were routinely tested using an in-house qPCR-based method adapted from Ishikawa et al. (2006). Briefly, cells to be tested were passaged two times in antibiotic- and antimycotic-free media, then collected for genomic DNA (gDNA) extraction using Qiagen DNeasy Blood and Tissue kit (Qiagen 69504) following the manufacturer's instructions with the exception of the last step in which gDNA was eluted in DNase-RNase-free water. Genomic DNA was diluted to a concentration of 10 ng/μL in water, then 50 ng gDNA was used per qPCR reaction using iQ SYBR Green Supermix (Bio-Rad 1708880). *Mycoplasma*-specific primer sequences used: forward primer for *Acholeplasma laidlawii* species, 5'-GGAATCCCGTTTGAAGATAGGA, for *Mycoplasma pirum* species, 5'-GGAAAATGTTATTTTGACGGAACCT, for six other *Mycoplasma* species, 5'-TCTGAAT(C/T)TGCCGGGACCACC; reverse primer for all 8 species of *Mycoplasma* listed above, 5'-CTTTCC(A/C)TCAC(G/T)GTACT(A/G)GTTCACT. Each assayed primer set was tested in technical triplicates. Per 50 ng gDNA (5 μL), 0.375 μL 3 μM forward primer, 0.375 μL 3 μM reverse primer, and 6.25 μL SYBR Supermix was combined, mixed gently, then run on an MJ DNA Engine Opticon 2 Research thermocycler. Cycling conditions: Initial 94°C, 3 min, followed by 40 cycles of: (i) 94°C for 15 s; (ii) 65°C for 30 s, (iii) 72°C for 30 s, followed by final extension at 72°C for 10 min. A melting curve (55°C–95°C) was performed at the end of the PCR to confirm the identity of each product and verify controls. Results were considered negative if Ct values were above 30 cycles. Negative controls typically have Ct between 35 and 40, positive controls around Ct 15–20. In cases where Ct values were not clear, samples were sent to the Norris Cancer Center Bioreagent and Cell Culture Core for additional testing. For cell line injections into mice, rodent pathogen testing was managed through the USC Department of Animal Resources, which sent cell samples and Matrigel to Charles River Laboratories for testing.

Derivation of human alveolar epithelial cell lines

Previously frozen isolated AT2 cells were quick-thawed in a 37°C water bath, spun down to remove freezing medium, and resuspended in Fmed + ROCKinh medium, modified from Liu et al. (2012) [3:1 (v/v)

DMEM/F12 (Corning 10-090-CV) to DMEM (Gibco 21063-029), 5% FBS, 0.4 μg/mL hydrocortisone (Sigma-Aldrich H0888), 5 μg/mL insulin (Sigma-Aldrich I0516), 8.4 ng/mL cholera toxin (Sigma-Aldrich C8052), 10 ng/mL human recombinant EGF (ThermoFisher PHG0311), Antibiotic-Antimycotic (Gibco 15240-062), and 10 μM Y-27632 (Enzo Life Sciences 270–333)]. Resuspended AT2 cells were plated in 96-well Primaria culture plates (BD Falcon 3872), ∼6250 cells per well and allowed to attach for 2 days. On the second day, Fmed + ROCKinh medium was completely replaced with fresh medium. Wells were monitored every day for surviving cells and proliferation. Media were changed every 2–3 days. Once cells reached 90–100% confluence, cells were detached with Accutase (Innovative Cell Technologies AT-104) and re-plated onto 48-well culture plates ("Passage 1"). This procedure was performed again, re-plating cells onto 24-well culture plates ("Passage 2"), then 12-well culture plates ("Passage 3"), then 6-well culture plates ("Passage 4"). Stocks of cells were frozen down starting at passages 3 and 4 in medium containing 10% DMSO and 90% 0.22 μm filtered FBS.

Construction of lentiviral plasmids and production of viral particles

LeGO iG and LeGO iT plasmids (Weber et al., 2008; http://www.lentigo-vectors.de/) were kind gifts from Dr. Kate Lawrenson (Cedars Sinai Medical Center, Los Angeles, CA). CDK4^{R24C} coding sequence (CDS) was subcloned from pBABE-hygro-CDK4^{R24C} plasmid (Hahn et al., 2002; Addgene 11254) by PCR amplification into LeGO iG vector between *Bam*HI and *Sbf*I sites, upstream of enhanced green fluorescent protein (eGFP). hTERT CDS was subcloned from pBABE-puro-hTERT plasmid (Counter et al., 1998; Addgene 1771) into LeGO iT vector between *Bgl*II and *Eco*RI sites, upstream of tdTomato. SV40 Large T antigen CDS was subcloned from pBABE-puro-SV40 LgT plasmid (Zhao et al., 2003; Addgene 13970) into LeGO iG vector between *Bam*HI and *Eco*RI sites, upstream of eGFP. Plasmids were propagated in Stbl3 chemically competent *E.coli* (ThermoFisher C737303). Plasmid sequences were verified by Sanger sequencing (GENEWIZ Inc). Third generation lentiviral particles were produced in low passage HEK293T cells (RRID:CVCL_0063) by transfection. Per 10 cm dish of 4-5x10^6 HEK293T cells, we used: 15 μL BioT (Bioland Scientific LLC B01-00), 2 μg pCMV-VSVG (Stewart et al., 2003; Addgene 8454), 2 μg pMDLg/pRRE (Dull et al., 1998; Addgene 12,251), 2 μg pRSV-Rev (Dull et al., 1998; Addgene 12253), and 2 μg lentiviral plasmid carrying transgene (LeGO iG-CDK4^{R24C}, LeGO iT-hTERT, LeGO iG-SV40 LgT). Viral supernatant was collected at 48- and 72-h post-transfection, pooled, spun down at 300 g to remove cell debris, filtered through 0.45 μm PES filters, and concentrated using Lenti-X Concentrator (Takara Bio 631231). Lentiviral pellets were resuspended in DMEM, aliquoted, and stored at −80°C. Transduction efficiency was tested empirically on HEK293T cells.

Lentiviral transduction of human AECs

AEC-ROCKinh cells (passage 4) were plated onto 96-well culture plate in Fmed + ROCKinh medium. The following day when cells were at 40–50% confluence, different volumes of the lentiviruses, either singly or in combination, were mixed with 8 μg/mL polybrene in Fmed + ROCKinh medium and added to each well. Of 300 μL resuspended viral supernatant, 1, 2, and 3 μL of LeGO iG-CDK4^{R24C}, LeGO iT-hTERT, and LeGO iG-SV40 LgT viruses were used to transduce cells in a total media volume of 50 μL. The following day, 100 μL of Fmed + ROCKinh media were added to each well to dilute out the viral supernatant. Two days post-transduction when cells were 90–100% confluent, cells were detached using Accutase and re-plated onto 48-well culture plates. At four days post-transduction, expression of CDK4^{R24C}, hTERT, and SV40 LgT was checked by fluorescence microscopy (Nikon Eclipse Ti-U inverted fluorescence microscope). Cells transduced with LeGO iT-hTERT only, LeGO iT-hTERT + LeGO iG-CDK4^{R24C}, or LeGO iT-hTERT + LeGO iG-SV40 LgT were sorted by fluorescence-activated cell sorting (FACS) on eGFP, tdTomato, or dual fluorescence at the USC Flow Cytometry Core Facility (FACS Aria II, BD Biosciences). Negative fluorescence was set by AEC-ROCKinh cells. All cell lines were maintained in Fmed + ROCKinh medium.

Proliferation assay

One thousand cells were plated on 24-well culture plates in quadruplicate in Fmed + ROCKinh medium and monitored for seven days. Twenty-four hours post seeding (day 1), cells were detached with Trypsin-EDTA (0.05% Trypsin, 0.02% EDTA) and resuspended in Fmed + ROCKinh medium. Cells were counted manually using a hemocytometer. Cell counts were reported as mean total cell number ±SD from at least three biological replicates. Population doubling time (PDT) was calculated based on the linear part of the growth curve using the equation [(t$_2$-t$_1$)/3.32] X (log n$_2$ – log n$_1$), where n$_2$ was the number of cells on day 6 and n$_1$ was the number of cells on day 4. High density proliferation assays were performed as described above with an initial cell seed count of 5,000 cells per well.

Anchorage-independent growth assay

Per well of a 6-well culture plate, 1.5 mL of 0.6% (w/v) Difco Noble Agar (BD Biosciences 214220) in Fmed + ROCKinh medium was added to form the bottom layer of the soft agar assay. For the top layer, in 1.5 mL, 5,000 cells were mixed with 0.3% final concentration Noble Agar in Fmed + ROCKinh medium. The top layer was allowed to solidify at room temperature (RT) before 1 mL Fmed + ROCKinh medium was carefully added. For A549 positive control cells (ATCC CCL-185), RPMI 1640 medium with 10% FBS was used to set up soft agar layers. Media for all cell lines were changed every 3 days. Colony growth was monitored for 1 month. Colonies were visualized by staining soft agar samples with crystal violet solution (crystal violet dissolved in 10% ethanol) according to Borowicz et al. (2014) and counted using ImageJ software (Schneider et al., 2012). At least three biological replicates were performed each with six technical replicates. Data were reported as mean number of colonies ±SD.

Three-dimensional (3D) co-culture

Actively dividing AEC-LgT cells between passages 6 and 19 were used to set up 3D co-culture with neonatal mouse lung fibroblasts, MLg (ATCC CCL-206). MLgs were cultured in DMEM/F12 medium (Corning 10-090-CV) with 10% FBS and maintained at sub-confluence. Noticeably lower organoid formation efficiency was observed when MLgs that had been grown beyond 70% confluence were used in the 3D co-culture. Five thousand AEC-LgT cells were mixed with 50,000 MLgs in Basic medium [phenol red free DMEM/F12 (Gibco 11039021), 1X ITS (Gibco 41400-045), 10% FBS, 1X Antibiotic-Antimycotic] with 50% Growth Factor Reduced Matrigel (Corning 354230) and plated on Clear Transwell inserts (Corning 3470), 100 μL per insert. Basic medium supplemented with 10 μM SB-431542 (BioVision 1674), a transforming growth factor beta (TGFβ) inhibitor, was added to the outer chamber and replaced every 2 days. Organoid formation was monitored under brightfield microscopy for 1–2 months. Whole-well images were captured using a Leica MZ16 F fluorescence stereomicroscope and Spot Advanced software (v4.5.8) through the USC Hastings Center for Pulmonary Research Core. Brightfield and fluorescence images at 4X magnification were captured using ECHO Revolve R4 fluorescence microscope (San Diego, CA). Organoid size (diameter) was measured as the largest distance between two points on the organoid membrane using the "Annotation length" feature on the ECHO Revolve R4 microscope. Organoid formation efficiency was defined as the total number of organoids divided by the initial cell seeding number (5,000 epithelial cells). Organoid number was determined using the "Annotation count" feature on ECHO Revolve R4 microscope. Only organoids of diameter >20 μm were counted. Values were reported as the mean percentage ±standard deviation.

Treatment of 3D co-cultures

Three-dimensional co-cultures were set up as described above. Treatment with 1 μM WNT agonist CHIR99021 (Sigma-Aldrich SML1046) dissolved in DMSO or 10 ng/mL FGF7 (Peprotech 100-19) and 10 ng/mL FGF10 (Peprotech 100-26) dissolved in sterile PBS began two days following culture set up, where SB Basic medium was replaced with fresh SB media containing either CHIR99021 or FGF7+FGF10, or DMSO (vehicle control). Media were changed every two days. Cultures were maintained for two months before sphere size and number were assessed. Treatment of 3D cultures with increasing concentrations of CHIR (1, 3, 5 μM) were performed as detailed above; spheres were assessed after five weeks. Size and formation efficiency values were reported as mean ± SD.

AT2 cell induction experiments using media published by Nikolić et al. (2017) and Jacob et al. (2017) were performed as described above. Two days following 3D culture set up, SB Basic medium was changed to either Nikolić "human alveolar differentiation" medium [(Advanced DMEM/F12 (ThermoFisher 12634010), 1X N-2 supplement (ThermoFisher 17502048), 1X B-27 supplement (ThermoFisher 12587010), 1.25 mM N-acetylcysteine (Sigma-Aldrich A9165), 1X Glutamax (ThermoFisher 35050061), 3 μM CHIR99021, 100 ng/mL FGF7, 100 ng/mL FGF10, 50 nM Dexamethasone (Sigma-Aldrich D4902), 0.1 mM 8-bromo-cyclic-AMP (Cayman Chemical 14431), 0.1 mM 3-isobutyl-1-methylxanthine (IBMX) (Enzo Life Sciences BML-PD140), 50 μM DAPT (Enzo Life Sciences ALX-270-416), 6.7 ng/mL T3 (Sigma-Aldrich T6397), 1 μg/mL human IGF-1 (ThermoFisher PHG0078), 1 μg/mL human IL-6 (ThermoFisher PHC0064)], Jacob "progenitor specification" medium [Advanced DMEM/F12, 1X Glutamax, 3 μM CHIR99021, 10 ng/mL human BMP4 (ThermoFisher PHC9534), 100 nM all-*trans* retinoic acid (Sigma-Aldrich R2625)], or Jacob "alveolar differentiation" medium [Advanced DMEM/F12, 1X Glutamax, 3 μM CHIR99021, 10 ng/mL FGF7, 50 nM Dexamethasone, 0.1 mM 8-bromo-cyclic-AMP, 0.1 mM IBMX]. Media were changed every two

days for ~1 month (5 weeks) at which point sphere size and number were assessed and organoids were fixed for IF staining. Size and formation efficiency values were reported as mean $\pm$ SD.

Histological processing of organoids

Once organoids formed, inserts were removed and fixed with 4% PFA for 30 min. PFA solution was removed by inverting inserts. Inserts were then submerged in 1X PBS for 15 min with two changes and dehydrated in 70% ethanol for 30 min with three changes. Matrigel samples were removed from insert housing by cutting the filter out from the bottom face with a feather razor, then embedded in Histogel (ThermoFisher HG-4000) and equilibrated in 70% ethanol for 1 h at RT. Histogel "buttons" were further dehydrated using standard methods, then embedded in paraffin wax. Paraffin blocks were sliced to 5 μm sections in-house using Microm HM 314 microtome through the USC Hastings Center for Pulmonary Research Core.

Immunofluorescence staining

For 2D cultures. Cells were plated on a standard multiwell culture plate to reach confluence the following day. Cells were rinsed with filtered 1X PBS, fixed with ice-cold methanol for 10 min, washed three times with PBS, blocked with 5% filtered bovine serum albumin (BSA) in PBS, then probed overnight at 4°C with respective primary antibodies in 5% BSA-PBS solution. Horse serum (RMBIO DES-BBT) diluted to 30% in PBS was used as blocking solution for pro-SFTPC antibody. The following day, cells were washed with 1X TBST (20 mM Tris, 150 mM NaCl, 0.01% Tween 20, pH 7.5), probed in PBS with biotinylated secondary antibodies for 1 h, washed, then probed with Streptavidin-Alexa Fluor 647 conjugate (ThermoFisher S21374). For double staining, cells were blocked again, then probed with either donkey anti-rabbit Alexa Fluor 488 secondary antibody (ThermoFisher A21206) or donkey anti-mouse Alexa Fluor 488 secondary antibody (ThermoFisher A21202). Mounting solution with 4′,6-diamidino-2-phenylindole (DAPI) was used as nuclear counterstain (Vector Laboratories H-1200). Stained cells were viewed using Nikon Eclipse Ti-U inverted fluorescence microscope and imaging software, NIS-Elements Br (v4.00.12, build 802, 64-bit).

For 3D cultures. Organoid paraffin sections were baked in a 60°C oven for 12 hours. Excess paraffin was wiped off with Kimwipes. Slides were then submerged in xylene, two changes, 5 min each, rehydrated through a series of ethanol baths, each with two changes (100% ethanol, 95%, 85%, 75%, 50%), and rinsed with distilled water. Samples were boiled for 6 min on high power, then for 5 min at 10% power in Tris-based antigen unmasking solution (Vector Laboratories H-3301) in a standard microwave oven, cooled to RT, then permeabilized with 2% Triton X-100 in PBS for 15 min, washed with PBS, then blocked with 5% BSA-PBS or 30% horse serum in PBS (for pro-SFTPC antibody) for 1 h RT. Citrate-based unmasking solution was used for pro-SFTPC probed samples (Vector Laboratories H-3300). Primary antibodies were diluted in 5% BSA-PBS or 30% horse serum (pro-SFTPC antibody) and probed overnight at 4°C. Subsequently, all washes were with 1X TBST. For single and double stainings, biotinylated secondary antibodies, fluorochrome-conjugated secondaries, Streptavidin-Alexa Fluor 647, Streptavidin-Alexa Fluor 488 (ThermoFisher S11223), and Streptavidin-FITC antibodies (ThermoFisher SA-10002) in PBS were incubated 1 h at RT for each step. Sections were mounted with Prolong Gold antifade reagent with DAPI (Invitrogen P36931), sealed, cured overnight at RT, then stored at 4°C. Slides were visualized the following day using the ECHO Revolve R4 fluorescence microscope.

Antibodies. Primary antibodies were: HOPX (SCBT sc-30216), pro-SFTPC (Seven Hills Bioreagent WRAB-9337), SV40 LgT (SCBT sc-147), AQP5 (Abcam ab92320), ECAD (BD Transduction Laboratories 610181), NKX2-1 (Leica TTF-1-L-CE), Ki67 (Abcam ab16667), SOX2 (SCBT sc-365823), SOX9 (SCBT sc-20095), GPRC5A (Abbexa abx005719), HTII280 (Terrace Biotech TB-27AHT2-280), mouse control IgG (Vector Laboratories I-2000), rabbit control IgG (Vector Laboratories I-1000), mouse control IgM (Sigma-Aldrich M5909). Secondary antibodies were: biotinylated horse anti-mouse IgG (Vector Laboratories BA-2000), biotinylated goat anti-rabbit IgG (Vector Laboratories BA-1000), biotinylated goat anti-mouse IgM (Vector Laboratories BA-2020). See key resources table for additional reagent details.

Bulk RNA-sequencing analyses

Data are available through the NCBI Gene Expression Omnibus (GEO) repository under the SuperSeries GSE164515 which contains bulk RNA-seq (17 samples, SubSeries GSE164513). RNA-seq analyses comparing derived AEC lines with normal lung cells and LUAD cancer cell lines were performed using

publicly available data from ENCODE (https://www.encodeproject.org/, Davis et al., 2018) and the Database of Transcriptional Start Sites (DBTSS) (https://dbtss.hgc.jp/, Suzuki et al., 2018) databases (Table S3). For the AEC lines (AEC-FT-ROCKinh, AEC-CDK4^{R24C}, AEC-hTERT, AEC-CDK4^{R24C}+hTERT, and AEC-FT) and the adult human lung fibroblast cell line, HLF-133, total RNA was isolated from subconfluent, exponentially dividing cells using Illustra TriplePrep kit (GE Healthcare 28-9425-44) following manufacturer's instructions. For the primary AECs, purified alveolar epithelial cells from three de-identified human donor lungs were used. Cells from each lung were distinguished from each other by the sample name extension "-m," "-f," or "-a." From each lung, purified primary AT2 cells and AT2 cells transdifferentiated on filters for 6 days into AT1-like cells (Danto et al., 1995) were isolated and total RNA extracted. Total RNA from purified AT2 cells and AT1-like cells from one of the donor lungs (labeled "AEC-m" in Table S3) had been previously characterized (Yang et al., 2018; GEO record GSE84273). Total RNA from the remaining two donor lungs ("AEC-f" and "AEC-a" in Table S3) was also previously isolated in our laboratory and sequenced, however without subsequent publication. For the AEC lines, HLF-133 cell line, and primary AECs, 2 μg of RNA were submitted to the USC Molecular Genomics Core facility for sequencing. RNA quality was assessed on the Bioanalyzer (Agilent) then rRNA-depleted using Ribo-Zero rRNA Removal kit for human samples (Illumina MRZH11124) before proceeding with library preparation (TruSeq mRNA Stranded Library preparation kit, Illumina, 20020594). These samples were sequenced paired-end 75 bp (PE75) at a depth of ∼20–30 million reads per sample, on HiSeq2000/2500 (Illumina). For AEC-ON and AEC-TN cell lines, total RNA was isolated using the Illustra TriplePrep kit as detailed above and 1 μg of RNA was submitted to the UCLA Technology Center for Genomics and Bioinformatics for sequencing. Samples were rRNA depleted and libraries were prepared at the UCLA facility, PE75, sequenced at a depth of ∼30 million reads per sample on NextSeq500 Mid Output (Illumina). Raw fastq files were retrieved and processed as follows:

Raw fastq files generated from our samples and taken from ENCODE and DBTSS databases were uploaded to Partek Flow through the USC Norris Medical Library Bioinformatics Core and the USC High-Performance Computing nodes. Files were quality controlled using Partek's QC tool and trimmed at both ends using Partek default parameters, then aligned using STAR RNA-sequence aligner (v 2.6.1d) (Dobin et al., 2013) to the human genome assembly, hg38 GENCODE Genes, release 29. Raw read counts were generated by quantification to the transcriptome using Partek E/M algorithm under default parameters, using hg38 GENCODE, release 29. Raw counts were rounded to the nearest integer then analyzed and processed in R (v3.6.0) using the DESeq2 package (Love et al., 2014). Genes with zero reads across samples were first filtered out, then counts were transformed using the regularized log (rlog) transformation method. Transformed counts were used to generate a sample-sample (Euclidean) distance matrix and PCA plot and to perform unsupervised hierarchical clustering. Differential gene expression analysis was performed under default DESeq2 parameters using the *results()* function and setting the appropriate pairwise comparison *contrast* parameter; the reference group comprised all established AEC lines (8 samples). Log$_2$ fold change shrinkage *lfcShrink()* was performed using the "normal" shrinkage estimator for visualization and ranking of genes. The most statistically significant differentially expressed genes were identified as those having log2 fold change > |2| and BH-adjusted p-value < 0.05.

Due to data acquisition and availability constraints, batch effect was accounted for in the DESeq2 *design* formula since the LUAD data used in this study were downloaded entirely from DBTSS. In lieu of performing standard batch removal (ie, *removeBatchEffect()* from the limma package), clustering was performed on the original study dataset plus additional LUAD RNA-seq data available from ENCODE to ensure RNA-sequencing data sources were not strongly driving sample segregation (see Figures S2A–S2D).

Gene clusters from unsupervised clustering. GO terms were analyzed using PANTHER14.1 (2018_04 release) PANTHER Overrepresentation Test (Released 20200728) with the default reference gene list of all *Homo sapiens* genes in the GO Ontology database (released 2019-07-03) (www.geneontology.org) (Mi et al., 2021). Statistically significant GO terms were calculated using Fisher's Exact Test with FDR corrected p-value cutoff of < 0.05.

GO terms of differentially expressed genes. Associated GO terms for each gene were taken from Ensembl gene database of annotated genes (release 98, Human genome assembly GRCh38.p13, www.ensembl.org) (Howe et al., 2021).

Single cell RNA-sequencing analyses

Data are available through the NCBI Gene Expression Omnibus (GEO) repository under the SuperSeries GSE164515 which contains scRNA-seq (2 samples, SubSeries GSE164514). AEC-ON cells (passage 9) grown in 2D and as organoids in 3D co-culture were harvested in parallel by sequential Dispase protease treatment (StemCell Technologies 07923) and gentle centrifugation. Dissociation of organoids into a single-cell suspension was assessed by brightfield microscopy using a hemocytometer. Surrounding single cells in the 3D culture that did not form organoids were excluded from collection by first incubating samples with Dispase briefly for ~15 min in a 37°C water bath, followed by gentle centrifugation. The subsequent iterations of Dispase treatment were for 30 min in a 37°C water bath, followed by centrifugation. Cells were immediately FACS sorted for GFP⁺ epithelial cells using MLg fibroblasts as a negative control for GFP fluorescence gating. Retrieved cells for each sample were then washed in 0.04% BSA-PBS solution. Barcoding and library preparation for scRNA-seq (n = 1) was performed following the manufacturer's protocol for the 10x Genomics Chromium Single Cell 3′ GEM, Library & Gel Bead Kit v3 (PN-1000092). cDNA libraries were assessed for quality and quantification according to the Chromium Single Cell 3′ Reagent kit v3 user guide using the High Sensitivity D5000 ScreenTape (Agilent 5067-5592) and 2200 TapeStation Controller software (Agilent, Santa Clara, CA). Sequencing was performed using Illumina HiSeq 3000/4000 kit at a coverage of 50,000 raw reads per cell (Paired-end; Read 1: 28 cycles, i7 Index: 8 cycles, i5 Index: 0 cycles, Read 2: 91 cycles). Raw data were processed using the CellRanger function *cellranger count* (10x Genomics, v3.1.0, default settings) to align to the human reference genome (hg19, 10x Genomics, v1.2.0) and identify 4320 2D cells and 7144 3D cells (11,464 total cells), which were then aggregated using the CellRanger pipeline (10x Genomics, v3.1.0, *cellranger aggr* function on default settings). Pre-processed outputs were then analyzed in R using the Seurat package for additional quality control assessment and downstream analyses (Butler et al., 2018; Stuart et al., 2019). Cells with less than 500 transcripts profiled and more than 18% of their transcriptome of mitochondrial origin were removed, leaving a total of 10,965 cells (4082 2D and 6883 3D cells) used for clustering and visualization. Read counts were normalized using the SCTransform method (Hafemeister and Satija, 2019). Dimensionality reduction and clustering analyses were performed as outlined in the Seurat vignette (https://satijalab.org/seurat/v3.2/pbmc3k_tutorial.html) with modifications: a Shared Nearest Neighbor (SNN) graph was constructed using the *FindNeighbors()* function on 20 dimensions of reduction and clusters were determined using the *FindClusters()* function with a reduction of 0.3.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses

For cell proliferation assays, cell counts were reported as the mean total cell number ±SD from at least three independent biological replicates each with technical triplicates. A test for normal distribution using the Shapiro-Wilk's test was first performed in R (v3.6.0, "Planting of a Tree") using the function *shapiro.test()* and then the student's t-test was performed using the function *t.test()*. Statistical details can be found in Figure 1D figure legend. For anchorage-independent growth assays, colony counts were reported as the mean ± standard deviation of at least three independent biological replicates each with six technical replicates. Since values were not normally distributed as evaluated by the Shapiro-Wilk's test, statistical significance was determined by the Wilcoxon nonparametric rank sum test in R using the function *wilcox.test()*. Statistical details can be found in Figure 1G figure legend. For bulk RNA-seq differential gene expression analyses, the DESeq2 package (Bioconductor release version 3.14) in R was used and the most statistically significant differentially expressed genes were identified as those having log2 fold change > |2| and BH-adjusted p-value < 0.05 (see additional details in method details section). For 3D organoid size and percent organoid formation efficiency experiments, values were reported as the mean size and mean percent ± respectively, from three biological replicates (different cell passage numbers) each with at least six technical replicates (number of inserts). Statistical details can be found in Figures 4D–4E figure legends.