Supplementary Materials and Methods

Mice (*Mus musculus*)

The day on which a vaginal plug was observed was designated as E1. The day on which mice were born was designated as P0. For the CreER driver screen, 2 mg tamoxifen (T5648, Sigma) dissolved in corn oil (C8267, Sigma) was injected intraperitoneally 2 days before harvest except that 60 ug tamoxifen was injected in P2 *Nkx2.1*^{CreER} and *Hopx*^{CreER} mice. For sparse cell labeling experiments, the tamoxifen dosage was determined empirically for each time point. For *Sox9*^{CreER}, 250 ug tamoxifen was injected intraperitoneally at E13 for E19 harvest; 0.4 mg (low dose) and 3.5 mg (high dose) tamoxifen was given via oral gavage at E19 for P20 harvest (Fig. S2B). For *Hopx*^{CreER}, 0.4 mg tamoxifen was given via oral gavage at E18 and E19 for E19 and postnatal harvest, respectively; postnatally, 7.5 ug, 10 ug, 25 ug, 100 ug and 400 ug tamoxifen injected intraperitoneally was determined to label isolated AT1 cells at P1, P3, P7, P14 and P24, respectively. Although at the same genomic locus, the *Rosa*^{Sox2} allele was activated much less efficiently than the *Rosa*^{mTmG} allele by *Hopx*^{CreER} and required daily gavage of 4 mg tamoxifen for three days after P35. For *Sox2*^{CreER}, 2 mg tamoxifen was injected intraperitoneally at P28 for P49 harvest. All animal experiments were approved by the institutional animal care and use committee at Texas A&M Health Science Center Institute of Biosciences and Technology (protocol number 2014-0001-IBT and 2014-0002-IBT).

Marker-based stereology

Inbred FVB/NJ (Stock 001800, Jackson Laboratory) mice of mixed gender were anaesthetized with Avertin (T48402, Sigma) and placed in the supine position. The trachea was cannulated with blunt needles of the appropriate size: 25 G (B25-50, SAI) for P0-P6, 23 G (B23-50, SAI) for P7-P14, 20 G (B20-50, SAI) for >P14. The descending aorta was cut and the rib cage partially removed. After perfusion through the right ventricle with phosphate-buffered saline (PBS, pH 7.4), the lung was inflated with 0.5% paraformaldehyde (PFA; P6148, Sigma) in PBS at 25 cm H_2O pressure. After 5 min, the trachea was tied to maintain the pressure and the lung was carefully dissected and submersion fixed in 0.5% PFA in PBS on a rocker at room temperature for 4~6 hr. The lung was subsequently washed in PBS at 4 °C overnight with the trachea cut to facilitate diffusion. This fixation protocol was compatible with fluorescence immunostaining for most antibodies and achieved fixation stability because the lung volume measured by the fluid displacement method (Scherle, 1970) was largely unchanged between immediately after fixation and after overnight PBS wash (Fig. S1B).

The left lobe, constituting ~1/3 of the total lung at all postnatal stages (Fig. S1D), was trimmed free of extra-pulmonary tissue and cryoprotected in 20% sucrose in PBS and 33% optimal cutting temperature compound (OCT; 4583, Tissue-Tek) at 4°C overnight and then embedded in OCT with its long axis perpendicular to the section plane. As illustrated in Fig. S1A, the left lobe was exhaustively sectioned at 60 um on a calibrated cryomicrotome (CM1850, Leica), with one in every 15 sections collected, resulting in a section interval of 900 um. To avoid dehydration and therefore vertical tissue shrinkage from section immunostaining, the sections were transferred to PBS to dissolve OCT and the tissue was processed for whole-mount immunostaining as described below.

After staining, the tissue was fixed with 2% PFA in PBS at room temperature for 2 hr and mounted under 22 mm square cover slips (48366067, VWR) that were spaced from the slide with tapes. Care was taken to avoid tissue folding during mounting. Confocal imaging verified that the mounted tissue maintained its original 60 um thickness. The tissue was first imaged at a low magnification with a stereomicroscope (M205C, Leica). The total area and airway/major vessel area were measured using STEPanizer (Tschanz et al., 2011) with at least 150 points counted for each lobe and the volume calculated using Cavalieri's principle. STEPanizer was used to systematic uniform randomly select 5~7 fields from all tissue sections of each lobe. The selected fields were located with a 10x objective and then imaged with a 40x oil objective with a field size of 318 um x 318 um x 30 um and a pixel dimension of 512 x 512 x 30 on a confocal microscope (A1plus, Nikon). Fields containing airways or major vessels were not imaged. To avoid surface artifacts from physical sectioning, images were collected several microns below the surface.

The resulting image stack was visualized using Imaris (Bitplane) and the number of AT1 and AT2 cells in the entire stack (N_v) was directly counted instead of using a dissector method (Williams and Rakic, 1988; Hsia et al., 2010). To account for edge effects, only cells with at least half of their nuclei visible were counted. Section view was used to resolve ambiguity in cell identification. Results from repeated counting vary by less than 5%. To measure alveolar surface area, the 10^{th} section in each confocal stack was quantified using the line probes in STEPanizer with the intercepts assigned to AT1 or AT2 cells based on HOPX and E-Cadherin staining.

Whole mount and section immunostaining

Immunostaining were performed following published protocols with minor modifications (Chang et al., 2013; Alanis et al., 2014). For section immunostaining, OCT sections at 10 um thickness were blocked in PBS with 0.3% Triton X-100 and 5% normal donkey serum (017-000-121, Jackson ImmunoResearch) and then incubated with primary antibodies diluted in PBS with 0.3% Triton X-100 in a humidified chamber at 4 °C overnight. The sections were washed with PBS in a coplin jar for 1.5 hr and incubated with secondary antibodies and 4',6-diamidino-2-phenylindole (DAPI) diluted in PBS with 0.3% Triton X-100 at room temperature for 2 hr. After another 1.5 hr wash with PBS, the sections were mounted with Aquamount (18606, Polysciences).

For whole mount immunostaining, 60 um frozen sections prepared as described above or 1~2 mm wide strips cut from around the edge of cranial and left lobes (Fig. 7A) were blocked in 2 mL Eppendorf tubes containing PBS with 0.3% Triton X-100 and 5% normal donkey serum and then incubated with primary antibodies diluted in PBS with 0.3% Triton X-100 on a rocker at 4 °C overnight. The following day, the tissues were washed in PBS with 1% Triton X-100 and 1% Tween-20 for 3~6 hr and incubated with secondary antibodies diluted in PBS with 0.3% Triton X-100 on a rocker at 4 °C overnight. The third day, the tissues were washed as before and fixed with 2% PFA in PBS for 2~3 hr. For tissues expressing green or red fluorescent protein, the native fluorescence but not that from secondary antibodies was quenched after immunostaining by incubation with methanol with 6% hydrogen peroxide (H1009, Sigma) at 4 °C overnight. Lobe strips were used to preserve the alveolar structure and whole AT1 cells and were thin enough for high resolution imaging with a 40x oil objective. Cell and tissue surface rendering was generated in Imaris with default settings: smoothing with surface area detail level at 0.621 um and thresholding by local contrast at 2.33 um. Vessel areas were measured using a threshold of 60 and a volume cut-off of 50 for both control and mutant lungs.

Antibodies

The following antibodies were used: chicken anti-green fluorescent protein (GFP, 1:1,000, AB13970, Abcam), rabbit anti-HOPX (1:500, sc-30216, Santa Cruz Biotechnology), rat anti-E-Cadherin (ECAD, 1:2000, 131900, Invitrogen), rabbit anti-Ki67 (1:1000, RM9106S0, Thermofisher), rat anti-Ki67 (1:1000, 14-5698-80, eBioscience), rabbit anti-Krt5 (1:1000, RM2106S0, Thermofisher), rabbit anti-Krt14 (1:1000, RB9020P0, Thermofisher), goat anti-CCSP (1:2500, a gift from Dr. Barry Stripp), rabbit anti-CCND1 (1:1000, RM9104S0,

Thermofisher), rabbit anti-FOXJ1 (1:1000, HPA005714, Sigma), chicken anti-MUC5AC (1:1000, a gift from Dr. Samuel Ho), goat anti-SRY-box-containing gene 9 (SOX9, 1:1,000, AF3075, R&D Systems), rabbit anti-prosurfactant protein C (SFTPC, 1:1,000, AB3768, Millipore), rat anti-RAGE (1:1,000, MAB1179, R&D Systems), goat anti-Podocalyxin-like (PODXL, 1:250, AF1556, R&D Systems), rabbit anti-ZO1 (1:50, 187430, Invitrogen), rabbit anti-collagen type IV (COL4, 1:5000, LSL-LB-1403, CosmoBioUSA), Cy3-conjugated mouse anti-smooth muscle actin (SMA, 1:1000, C6198, Sigma), rabbit anti-Aquaporin 5 (AQP5, 1:2500, ab78486, Abcam; non-specific staining was observed in embryonic lungs), rat anti-Intercellular adhesion molecule 2 (ICAM2, 1:2500, 14-1021-80, eBioscience), rabbit anti-Avian erythroblastosis virus E-26 (v-ets) oncogene related (ERG, 1:5000, ab92513, Abcam), rat anti-Lysosomal-associated membrane protein 3 (LAMP3, 1:1000, DDX0192, Imgenex), goat anti-SRY-box-containing gene 2 (SOX2, 1:250, sc-17320, Santa Cruz Biotechnology), rabbit anti-deltaNp63 (DNp63, 1:1000, 619001, Biolegend) and rabbit anti-p63 (1:250, sc-8343, Santa Cruz Biotechnology). Fluorescent secondary antibodies were obtained from Invitrogen or Jackson ImmunoResearch and used at 1:1000. Apoptotic cells were detected with a TUNEL (terminal deoxynucleotidyl transferase-mediated dUTP nick end labeling) kit (11684795910, Roche) following the manufacturer's protocol.

Fluorescence and colorimetric in situ hybridization

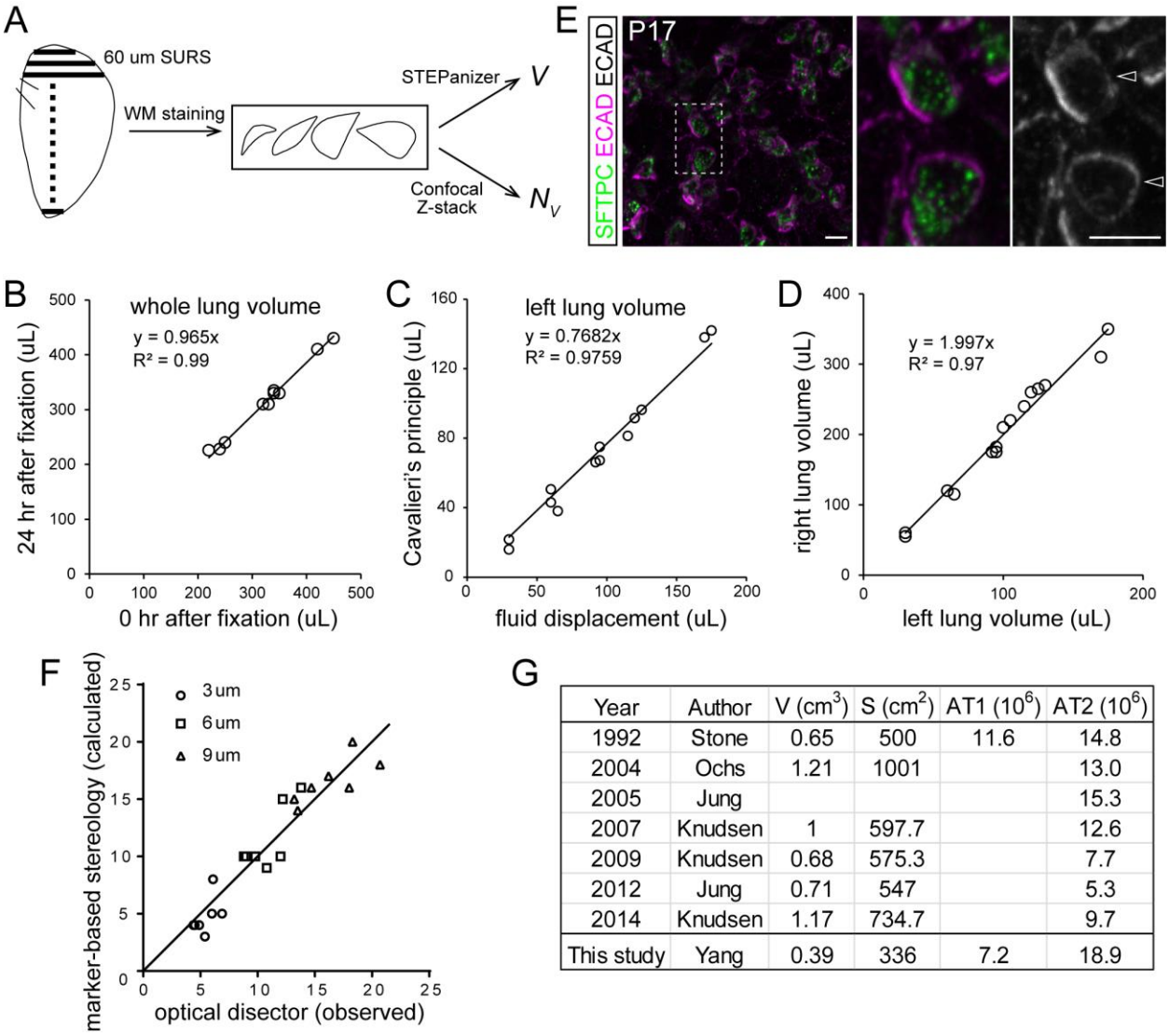
Postnatal lungs were harvested and processed as described for section immunostaining except 0.5% PFA was included for the sucrose/OCT overnight incubation to minimize RNA degradation. Colorimetric section in situ hybridization was carried out following published protocols (Chang et al., 2013; Alanis et al., 2014). For fluorescence in situ hybridization, sections were incubated with 1/20 of the amount of riboprobes (0.05 ug/mL). After washing with 0.2X SSC (sodium chloride and sodium citrate), sections were incubated with PBS with 3% hydrogen peroxide for 1 hr at room temperature to quench the endogenous peroxidase activity. An anti-digoxigenin POD Fab fragment (1:2000, Roche 11207733910) and a fluorescein tyramide signal amplification system (NEL741001KT, Perkin Elmer) were used to detect the hybridized riboprobes. The biotin tyramide signal amplification system (NEL749A001KT, Perkin Elmer) was not used due to the presence of endogenous biotin in AT2 and club cells (Kuhn, 1988). Subsequently, sections were immunostained for SFTPC, LAMP3, HOPX or NKX2.1 antibodies as described above. The *Vegfa* riboprobe spanned exon 2 to 6 and detected all *Vegfa* isoforms (Ng et al., 2001). The peri-nuclear localization of the *Vegfa* mRNA allowed unequivocal

identification of its cellular source. Control and mutant samples were processed on the same section to minimize technical variations.

Transcriptome profiling of FACS purified AT1 cells

Cell dissociation and purification was performed based on a previous protocol with modifications (Chang et al., 2013). Briefly, postnatal day 3 or 4 *Scnn1a-Cre; Rosa*$^{mTmG/+}$ and *Scnn1a-Cre; Rosa*$^{Sox2/+}$ littermate lungs were dissected in PBS. Around 3 mm wide strips were cut along the lobe edge, minced into pieces with forceps and digested in PBS with 2 mg/mL collagenase type I (Worthington, CLS-1) and 0.2 mg/mL DNase I (Worthington, D) for 40 min at 37 °C. An equal volume of 0.25% Trypsin/EDTA (Invitrogen, 25200-056) was added and incubated for another 10 min at 37 °C. Fetal bovine serum (FBS, Invitrogen, 10082-139) was added to a final concentration of 5% and the tissue was triturated for 40~50 times. Dissociated cells were spun down at 6000 rpm for 1 min, resuspended in red blood cell lysis buffer (15 mM NH_4Cl, 12 mM $NaHCO_3$, 0.1 mM EDTA, pH 8.0) for 3 min, washed with PBS with 1% FBS, SYTOX-red (Invitrogen, S34859) added to exclude dead cells and sorted on a BD Biosciences Influx sorter.

RNA was extracted from at least 10^5 purified cells using Trizol reagents (Invitrogen, 15596018) and an RNeasy Micro kit (Qiagen, 74004). 100~200 ng total RNA was used to prepare an RNAseq library using an mRNA isolation kit (New England BioLabs, E7490) and a NEBNext Ultra RNA library prep kit (New England BioLabs, E7530S) with a final double (0.65 x -1 x bead volume) size selection step using a SPRIselect reagent kit (Beckman Coulter, B23318). The libraries were bar-coded (New England BioLabs, E7335S) and sequenced on an Illumina HiSeq2000. 35~80 million 76 nucleotide pair-end reads were generated for each sample and aligned to the UCSC mm10 reference genome using tophat2 and bowtie2 in R (Langmead and Salzberg, 2012; Kim et al., 2013). Transcript abundance, differential expression, and isoform quantitation were calculated using the cufflinks suite in R from each pair of littermate samples or pooled control and mutant samples (Roberts et al., 2011a; Roberts et al., 2011b; Trapnell et al., 2012; Trapnell et al., 2013).



G

Year	Author	V (cm^3)	S (cm^2)	AT1 (10^6)	AT2 (10^6)
1992	Stone	0.65	500	11.6	14.8
2004	Ochs	1.21	1001		13.0
2005	Jung				15.3
2007	Knudsen	1	597.7		12.6
2009	Knudsen	0.68	575.3		7.7
2012	Jung	0.71	547		5.3
2014	Knudsen	1.17	734.7		9.7
This study	Yang	0.39	336	7.2	18.9

Fig. S1. Marker-based stereology. (**A**) Schematics of the marker-based stereology method. After volume measurement using the fluid displacement method, systematic uniform random sampling (SURS) is applied to the left lung by exhaustive sectioning at 60 um. One in every 15 sections is collected and whole mount (WM) immunostained to avoid dehydration. Tissue volume (*V*) is obtained using STEPanizer software and Cavalieri's principle. AT1 and AT2 cell densities (N_V) are obtained by counting HOPX and ECAD expressing cells in confocal Z-stacks using Imaris software. (**B**) Lung volumes measured by the fluid displacement method change minimally immediately or 24 hr after fixation indicating fixation stability. The intercept is set as 0 in the linear regression analysis and R denotes the correlation coefficient (same for (**C**) and (**D**)). (**C**) Left lung volumes measured by the fluid displacement method correlate linearly with those measured by Cavalieri's principle using STEPanizer. The volumes shrink by ~23% during embedding, sectioning and immunostaining. (**D**) Right lung volumes are twice left lung volumes as measured

by the fluid displacement method. (**E**) Confocal projection images of immunostained P17 lung strips showing that cuboidal ECAD-expressing cells (open arrowhead) express SFTPC. The boxed region is enlarged. Scale: 10 um. (**F**) Cell counting results are comparable using marker-based stereology or optical dissector method (Hsia et al., 2010). AT1 cells in six 30-um image stacks from P54 lungs are counted using the optical dissector method at an interval of 3, 6 or 9 um. The results are plotted against the numbers calculated as a fraction of the total cell number using marker-based stereology (Table S1). The diagonal line is drawn. (**G**) Comparison of results from our method and those in PubMed searched using “mouse lung stereology” as keywords (Stone et al., 1992; Ochs et al., 2004; Jung et al., 2005; Knudsen et al., 2007; Knudsen et al., 2009; Jung et al., 2012; Knudsen et al., 2014). Cell numbers were counted on single plane images and corrected by a shape factor by Stone et al. Measurements of P54 left lobes (Table S1) are multiplied by 3 (Fig. S1D) to represent those of the entire lung. V: lung parenchyma volume; S: total alveolar surface area.

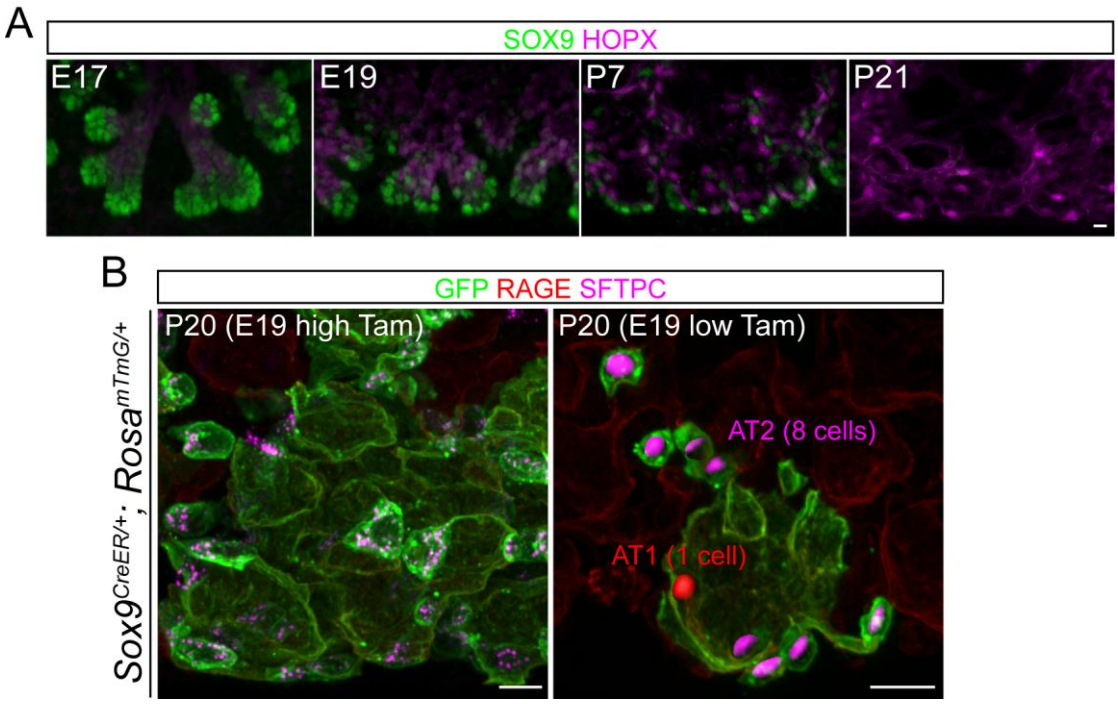
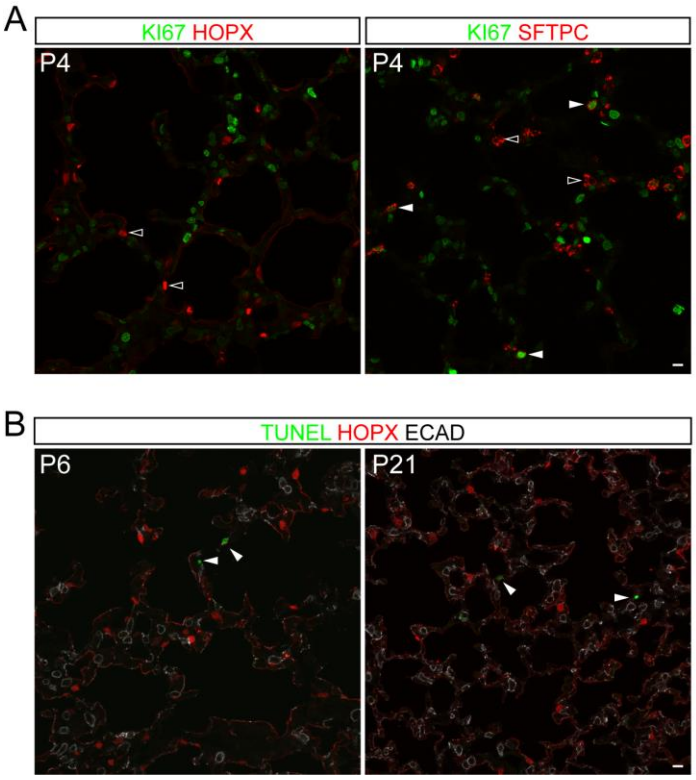


Fig. S2. AT1 cells are descendants of SOX9 progenitors. (**A**) Confocal projection images of immunostained lung strips at indicated embryonic and postnatal stages showing that SOX9-expressing progenitors are distal to HOPX-expressing AT1 cells and are present at P7. Scale: 10 um. (**B**) Confocal images of immunostained P20 lung strips that have been lineage-labeled at E19 (Tam, tamoxifen). Most alveolar cells near the lobe edge are labeled at a high dose (left), whereas a cluster of 1 AT1 cell and 8 AT2 cells are labeled at a low dose (right, nuclei rendered with Imaris). Scale: 10 um.



	HOPX⁺KI67⁺/HOPX⁺	SFTPC⁺KI67⁺/SFTPC⁺
P4	0/70 (0%)	35/99 (35%)
P8	0/32 (0%)	37/149 (25%)
P14	0/46 (0%)	17/165 (10%)
P21	0/21 (0%)	0/120 (0%)

	HOPX⁺TUNEL⁺/HOPX⁺	ECAD⁺TUNEL⁺/ECAD⁺
P4	0/49 (0%)	0/107 (0%)
P8	0/42 (0%)	0/133 (0%)
P14	0/35 (0%)	1/118 (0.8%)
P21	0/42 (0%)	0/175 (0%)

Fig. S3. Epithelial cell proliferation and apoptosis. (**A**) Confocal images of lung sections showing KI67 positive AT2 cells (filled arrowhead) and KI67 negative AT1 and AT2 cells (open arrowhead). AT2 cells are labeled with the rabbit anti-SFTPC antibody instead of the rat anti-ECAD antibody so that the same rat anti-KI67 antibody can be used for both AT1 and AT2 cells to allow direct comparison. Scale: 10 um. The number and percentage of KI67 positive AT1 and AT2 cells at indicated postnatal days are tabulated. (**B**) Confocal images of lung sections containing occasional TUNEL positive cells (arrowhead), which do not co-localize with HOPX or ECAD. Scale: 10 um. The number of percentage of TUNEL positive AT1 and AT2 cells at indicated postnatal days are tabulated.

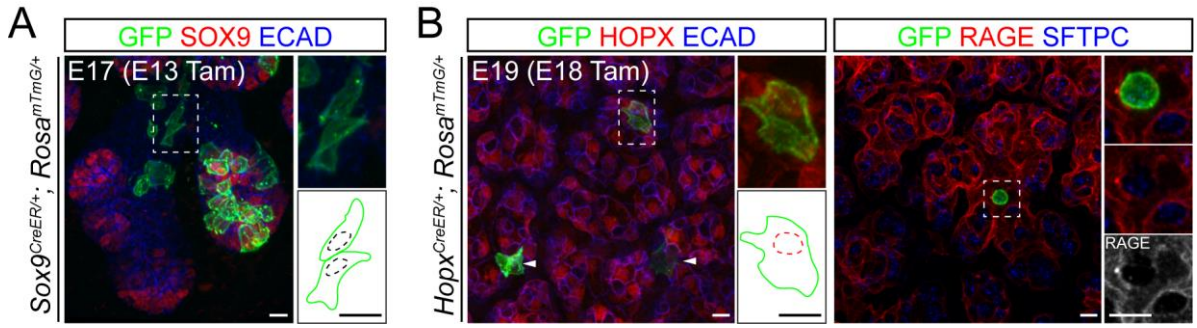


Fig. S4. AT1 cell morphology during embryonic development. (**A**) Confocal projection images of immunostained strips from E17 *Sox9*$^{CreER/+}$; *Rosa*$^{mTmG/+}$ lungs with recombination induced at E13 (Tam, tamoxifen). The boxed region is enlarged on the right. Occasional elongated presumable AT1 cells are found in branch stalks and outlined with green lines and their nuclei marked with black dashed lines. Scale: 10 um. (**B**) Confocal projection images of immunostained strips from E19 *Hopx*$^{CreER/+}$; *Rosa*$^{mTmG/+}$ lungs with recombination induced at E18 (Tam, tamoxifen). Boxed regions are enlarged on the right. HOPX-expressing elongated AT1 cells are indicated with arrowheads or outlined with green lines and the nucleus marked with red dashed lines. They have a similar morphology to those labeled with *Sox9*^{CreER} (Fig. 2A). Infrequent cuboidal SFTPC-expressing AT2 cells are also labeled (right panels) possibly due to the low diffuse expression of HOPX in the progenitors. Scale: 10 um.

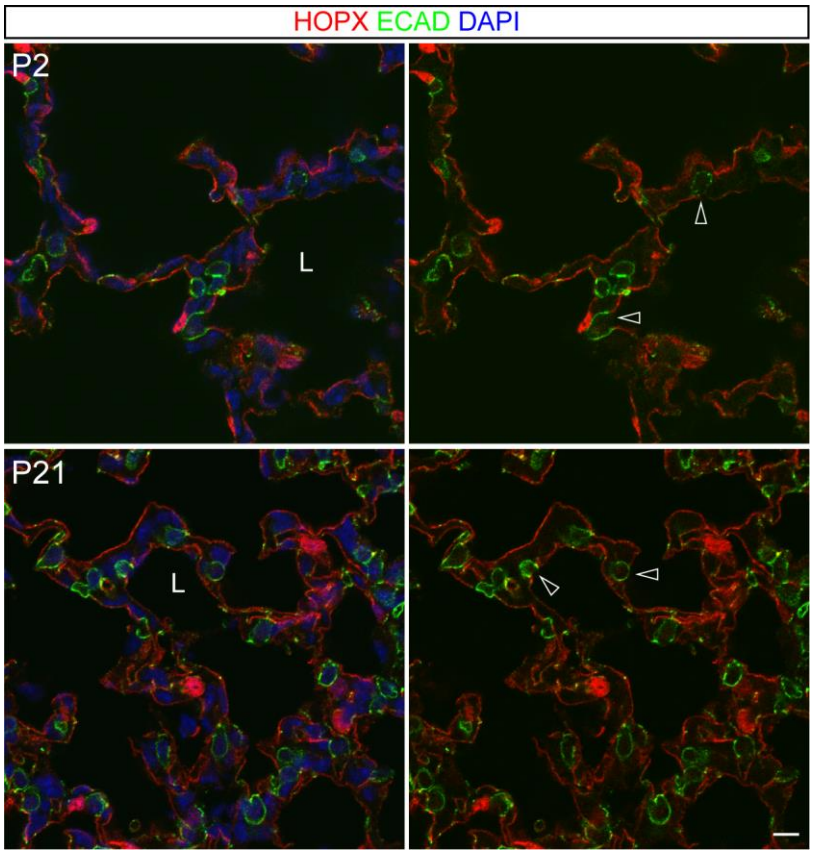


Fig. S5. Interstitial positioning of AT2 cells. Confocal images of immunostained lung sections at indicated postnatal days. AT2 cells (open arrowhead, cuboidal ECAD staining) are in the interstitial regions instead of protruding into the air lumen (L). Scale: 10 um.

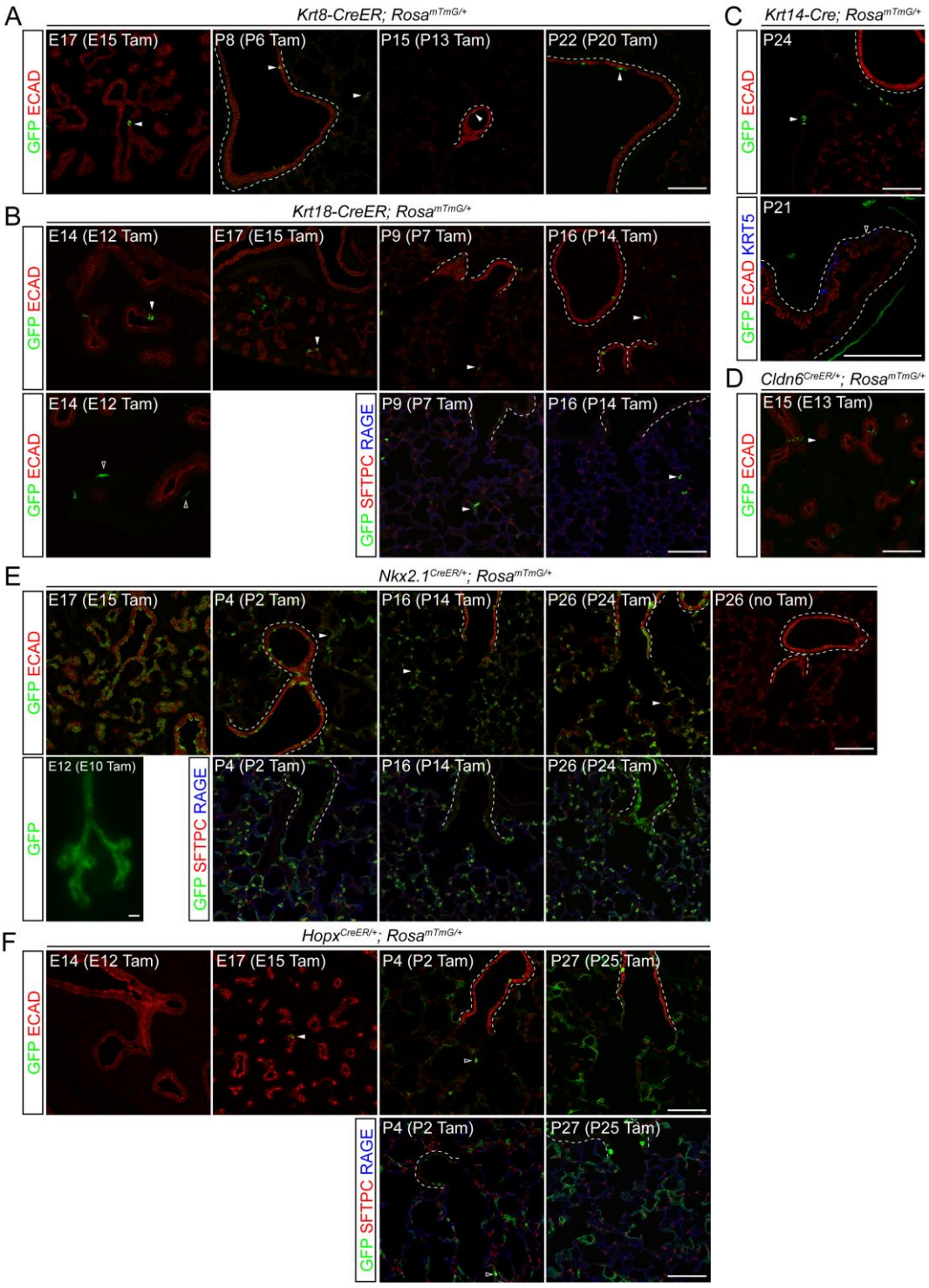


Fig. S6. Cre driver screen. Confocal images of immunostained sections from *Krt8-CreER* (**A**), *Krt18-CreER* (**B**), *Krt14-Cre* (**C**), *Cldn6*^{CreER} (**D**), *Nkx2.1*^{CreER} (**E**) and *Hopx*^{CreER} (**F**) lungs with recombination of the *Rosa*^{mTmG} reporter induced at indicated stages (Tam, tamoxifen). Occasional epithelial cells are labeled with *Krt8-CreER*, *Krt18-CreER*

and *Cldn6*^{CreER} (filled arrowhead) at all examined stages. No *Cldn6*^{CreER} activity is observed postnatally. *Krt18-CreER* also labels mesenchymal cells (open arrowhead). *Krt14-Cre* does not label KRT5-expressing basal cells (open arrowhead), but instead occasional vasculature-associated cells (filled arrowhead). *Nkx2.1*^{CreER} labels most epithelial cells (arrowhead, AT1 cells) at all stages examined. *Hopx*^{CreER} recombination is rare before E15 (filled arrowhead) and is observed in occasional AT2 cells at P2 (open arrowhead). Dashed lines denote conducting airways. Scale: 100 um.

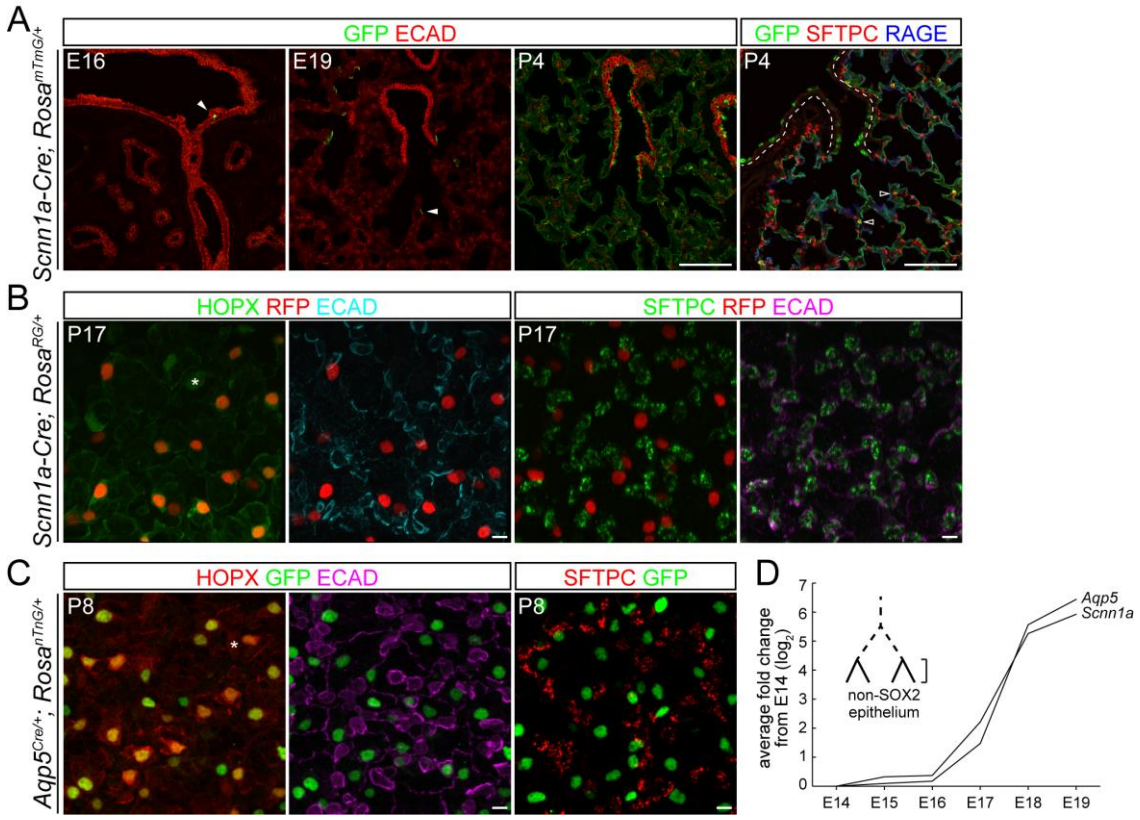


Fig. S7. *Scnn1a-Cre* and *Aqp5^{Cre}* activity in the lung. Confocal images of immunostained lung sections (**A**) and strips (**B**, **C**; projection view). *Scnn1a-Cre* recombination is rare before E19 (filled arrowhead) and is observed in occasional AT2 cells at P4 (open arrowhead). The nuclear reporter *Rosa^{RG}* and *Rosa^{nTnG}* allows unambiguous identification of recombined AT2 cells that are surrounded by AT1 cells. Note that HOPX is expressed at a variable level in AT1 cells and therefore gives different shades of yellow when overlapping with GFP. Asterisks denote unrecombined AT1 cells. Dashed lines denote conducting airways. Scale: 100 um (**A**) and 10 um (**B**, **C**). (**D**) Average fold changes of gene expression showing upregulation of *Scnn1a* and *Aqp5* in FACS-isolated distal epithelial cells at indicated embryonic days. The original data and experimental details are previously reported (Chang et al., 2013).

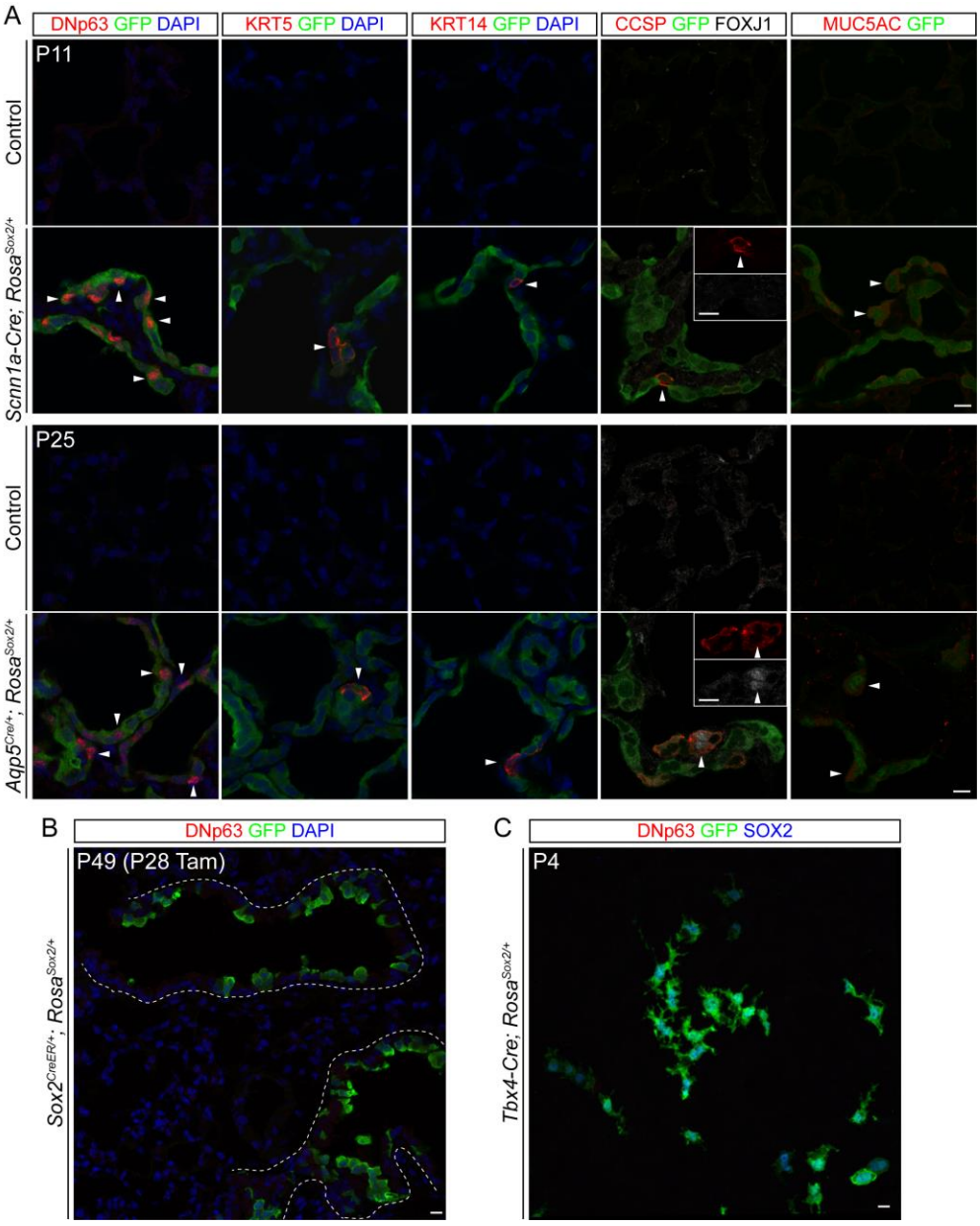


Fig. S8. Marker analyses of *Scnn1a-Cre*, *Aqp5*^{Cre}, *Sox2*^{CreER} and *Tbx4-Cre* targeted cells. (**A**, **B**) Confocal images of immunostained sections from littermate control and mutant lungs. Mutant AT1 cells (GFP, arrowhead) activate an isoform of P63 (DNp63) that is specific for basal cells. Occasional mutant AT1 cells express additional basal cell markers (KRT5, KRT14) and a club cell marker (CCSP). A ciliated cell marker (FOXJ1) is observed only in *Aqp5*^{Cre} targeted AT1 cells but aberrantly co-expressed with CCSP (inset showing single channel images). A low but detectable level of a goblet cell marker (MUC5AC) is present in mutant AT1 cells, consistent with the RNA-seq data (Fig. 4G and Table S2). Scale: 10 um. (**C**) A confocal image of immunostained section from the *Sox2*$^{CreER/+}$; *Rosa*$^{Sox2/+}$ lung induced with 2 mg tamoxifen (Tam) at P28. No DNp63 cells are observed. Airways are outlined with dashes. Scale: 10 um. (**D**) Projection view of a confocal stack of an immunostained *Tbx4-Cre; Rosa*$^{Sox2/+}$ lung strip. No DNp63 cells are observed. Scale: 10 um.

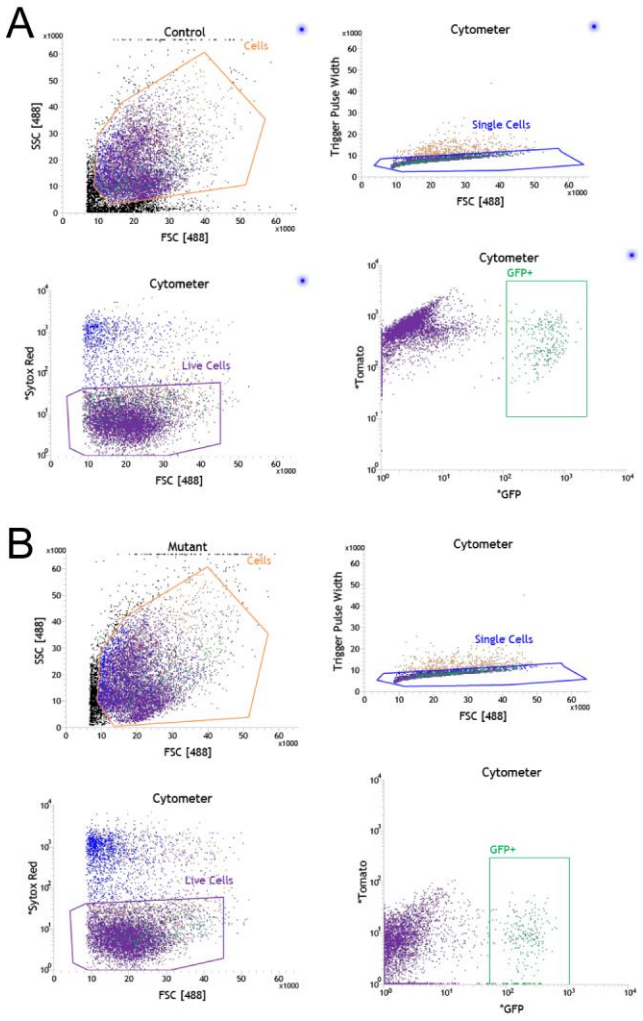


Fig. S9. FACS purification of AT1 cells. Representative FACS plots of dissociated live AT1 cells from P3 littermate *Scnn1a-Cre; Rosa*$^{mTmG/+}$ (**A**, control) and *Scnn1a-Cre; Rosa*$^{Sox2/+}$ (**B**, mutant) lungs. GFP positive cells represent 2~3% of total cells. Note that GFP expression from *Rosa*$^{Sox2/+}$ is lower than that from *Rosa*$^{mTmG/+}$ and Tomato expression perdures after recombination.

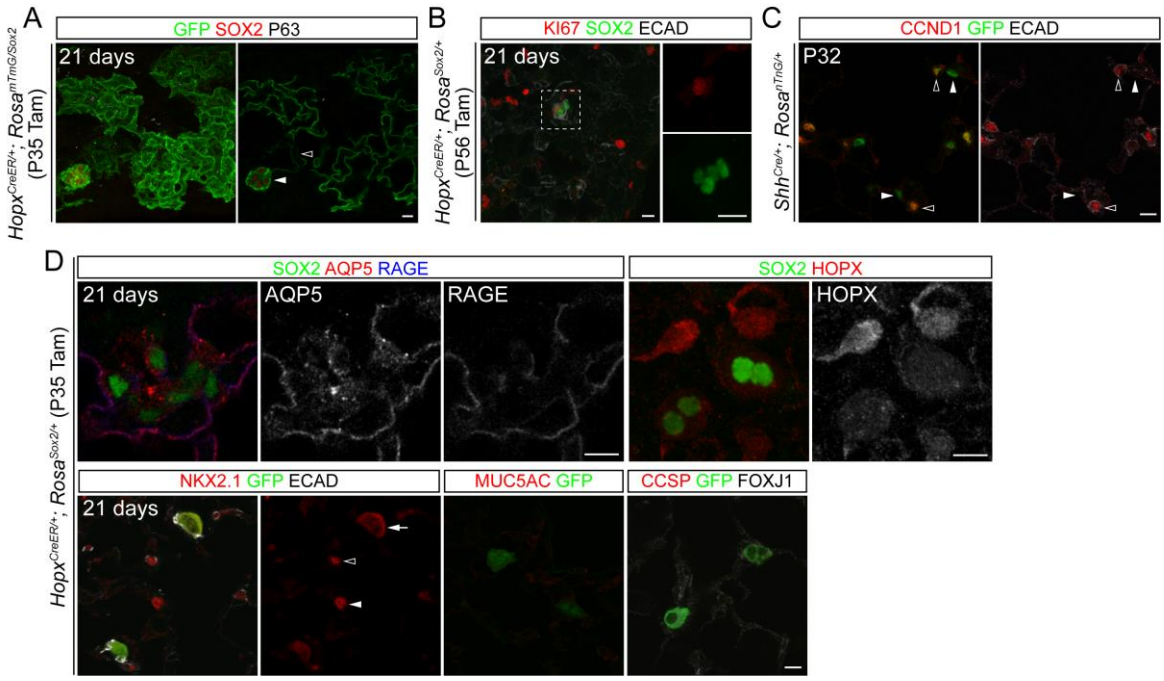


Fig. S10. Marker analyses of targeted mature AT1 cells. (**A**) Projection (left) and sectional (right) views of confocal images of immunostained strips from *Hopx*$^{CreER/+}$; *Rosa*$^{mTmG/Sox2}$ lungs at 21 days after recombination induced at P35 (Tam, tamoxifen). *Rosa*^{mTmG} is easier to recombine than *Rosa*^{Sox2}; therefore GFP marks both control (open arrowhead) and SOX2-expressing mutant (filled arrowhead) AT1 cells. No P63 staining is observed. Scale: 10 um. (**B**) Confocal projection images of immunostained strips at 21 days after recombination induced at P56. The boxed region is shown as single channel images. While absent in control AT1 cells (Fig. S3A), KI67 is activated in a subset of mutant AT1 cells (11%, n=72 cells). (**C**) Confocal images of immunostained sections showing that in a control lung CCND1 is expressed in AT2 cells (open arrowhead, cuboidal ECAD staining), but not AT1 cells (arrowhead). All epithelial cell nuclei are genetically marked by GFP to allow co-staining with the rabbit anti-CCND1 antibody. (**D**) Confocal images of immunostained strips (top row) and sections (bottom row) at 21 days after recombination induced at P35 (Tam, tamoxifen). AQP5, RAGE and HOPX are mostly unaffected. NKX2.1 is diffuse in mutant AT1 cells (arrow, marked by GFP) as compared to non-targeted AT1 cells (arrowhead) and AT2 cells (open arrowhead). No MUC5AC, CCSP and FOXJ1 are detected in mutant AT1 cells. Scale: 10 um.

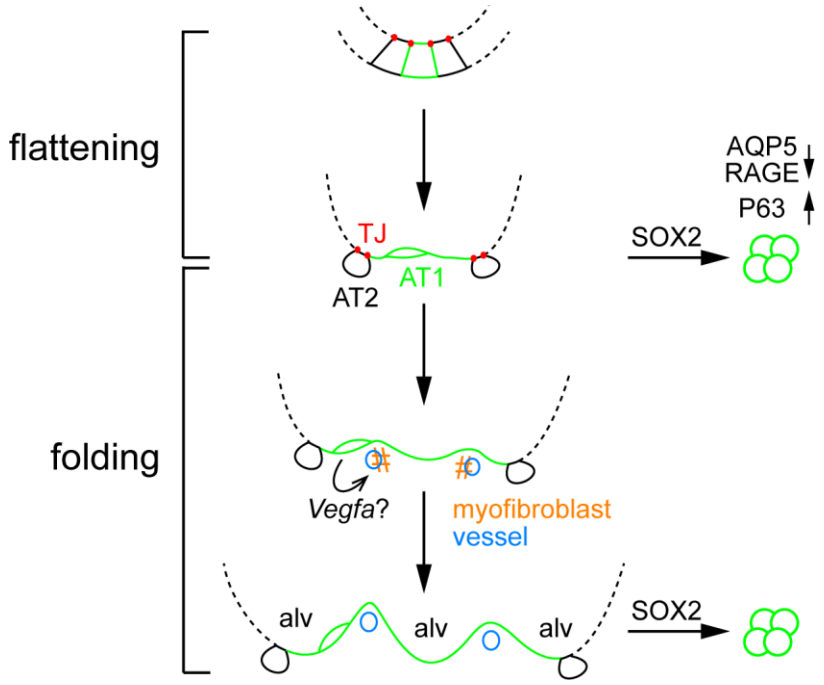


Fig. S11. A model summarizing AT1 cell development and plasticity combining Fig. 2C, 3D and the ectopic SOX2 expression experiments. AT1 cells have a potential signaling role by expressing *Vegfa*.

Table S1: Quantification of AT1 and AT2 cells at postnatal stages.

The lung volume and total AT1 and AT2 alveolar surface are obtained using STEPanizer. Major vessels are defined as those that parallel airways; therefore the parenchyma volumes include small vessels in the gas exchange region. AT1 and AT2 cell numbers are averaged over image fields and corresponding standard errors calculated. The ratio between AT1 and AT2 cell numbers is calculated for each image field and then averaged. The total cell number in the left lung is calculated by multiplying average cell number per image field (0.003037537 uL, 318.2 um x 318.2 um x 30 um) by the parenchyma volume. The average AT1 apical surface area is calculated by dividing the total AT1 apical surface area by the total AT1 cell number.

Click here to Download Table S1

Table S2: Transcriptome comparison of FACS purified littermate control and SOX2-expressing mutant AT1 cells from two biological replicates. Analyses are performed for individual pairs and pooled samples. Significance is defined as false discovery rate by Benjamini-Hochberg procedure (q) < 0.05.

Click here to Download Table S2

Supplementary References

Alanis, D. M., Chang, D. R., Akiyama, H., Krasnow, M. A. and Chen, J. (2014). Two nested developmental waves demarcate a compartment boundary in the mouse lung. *Nature communications* **5**, 3923.
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.
Hsia, C. C., Hyde, D. M., Ochs, M., Weibel, E. R. and Structure, A. E. J. T. F. o. Q. A. o. L. (2010). An official research policy statement of the American Thoracic Society/European Respiratory Society: standards for quantitative assessment of lung structure. *Am J Respir Crit Care Med* **181**, 394-418.
Jung, A., Allen, L., Nyengaard, J. R., Gundersen, H. J., Richter, J., Hawgood, S. and Ochs, M. (2005). Design-based stereological analysis of the lung parenchymal architecture and alveolar type II cells in surfactant protein A and D double deficient mice. *The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology* **286**, 885-890.
Jung, K., Schlenz, H., Krasteva, G. and Muhlfeld, C. (2012). Alveolar epithelial type II cells and their microenvironment in the caveolin-1-deficient mouse. *Anatomical record* **295**, 196-200.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36.
Knudsen, L., Atochina-Vasserman, E. N., Guo, C. J., Scott, P. A., Haenni, B., Beers, M. F., Ochs, M. and Gow, A. J. (2014). NOS2 is critical to the development of emphysema in Sftpd deficient mice but does not affect surfactant homeostasis. *PLoS One* **9**, e85722.
Knudsen, L., Wucherpfennig, K., Mackay, R. M., Townsend, P., Muhlfeld, C., Richter, J., Hawgood, S., Reid, K., Clark, H. and Ochs, M. (2009). A recombinant fragment of human surfactant protein D lacking the short collagen-like stalk fails to correct morphological alterations in lungs of SP-D deficient mice. *Anatomical record* **292**, 183-189.
Knudsen, L., Ochs, M., Mackay, R., Townsend, P., Deb, R., Muhlfeld, C., Richter, J., Gilbert, F., Hawgood, S., Reid, K. et al. (2007). Truncated recombinant human SP-D attenuates emphysema and type II cell changes in SP-D deficient mice. *Respiratory research* **8**, 70.
Kuhn, C. (1988). Biotin stores in rodent lungs: localization to Clara and type II alveolar cells. *Exp Lung Res* **14**, 527-536.
Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-359.
Ng, Y. S., Rohan, R., Sunday, M. E., Demello, D. E. and D'Amore, P. A. (2001). Differential expression of VEGF isoforms in mouse during development and in the adult. *Dev Dyn* **220**, 112-121.
Ochs, M., Knudsen, L., Allen, L., Stumbaugh, A., Levitt, S., Nyengaard, J. R. and Hawgood, S. (2004). GM-CSF mediates alveolar epithelial type II cell changes, but not emphysema-like pathology, in SP-D-deficient mice. *Am J Physiol Lung Cell Mol Physiol* **287**, L1333-1341.
Roberts, A., Pimentel, H., Trapnell, C. and Pachter, L. (2011a). Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **27**, 2325-2329.
Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L. and Pachter, L. (2011b). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome biology* **12**, R22.
Scherle, W. (1970). A simple method for volumetry of organs in quantitative stereology. *Mikroskopie* **26**, 57-60.
Stone, K. C., Mercer, R. R., Gehr, P., Stockstill, B. and Crapo, J. D. (1992). Allometric relationships of cell numbers and size in the mammalian lung. *Am J Respir Cell Mol Biol* **6**, 235-243.

Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L. and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology* **31**, 46-53.
Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L. and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578.
Tschanz, S. A., Burri, P. H. and Weibel, E. R. (2011). A simple tool for stereological assessment of digital images: the STEPanizer. *Journal of microscopy* **243**, 47-59.
Williams, R. W. and Rakic, P. (1988). Three-dimensional counting: an accurate and direct method to estimate numbers of cells in sectioned material. *J Comp Neurol* **278**, 344-352.