SUPPLEMENTAL MATERIALS AND METHODS

Animal Studies: All mouse studies were approved by an Institutional Animal Care and Use Committee. Mice were maintained on a 12-hour light/dark cycle, and were given food and water *ad libitum*. Animals were housed in individually ventilated units in the specific pathogens-free facility at the University of Iowa. Female and male mice of the following strains were used and maintained on a C57BL/6 background: B6.129-*Gt(ROSA)26Sor*$^{tm1(cre/ERT2)Tyj}$/J (ROSA-Cre^{ERT2}; The Jackson Laboratory, stock number 008463), B6.129(Cg)-Gt(ROSA)26Sor$^{tm4(ACTB-tdTomato,-EGFP)Luo}$/J x (ROSA-TG) Cre-reporter mice (The Jackson Laboratory, stock number 007676), B6.Cg-Lef1^{tm1Hhx}/J (Lef-1cKO; The Jackson Laboratory, stock number 030908), and ROSA26-CAG-^{LoxP}EGFP^{StopLoxP}(Lef-1 mice[1]). All transgenic mouse lines used can be found in Table S1, along with their corresponding abbreviated names that were used throughout this manuscript. Animals were genotyped at weaning from tail clippings. For lineage tracing experiments, 6-7 week old animals were induced by i.p. injection of 75 µg tamoxifen dissolved in corn oil per gram bodyweight every 24 hours for five days. For injury experiments, mice were rested for 7-10 days after tamoxifen induction, then given one i.p. injection of 200 µg naphthalene dissolved in corn oil per 1 gram of bodyweight (200 mg/kg); one i.p. injection of corn oil was used for uninjured control animals.

Table S1: Genetic mouse lines and abbreviated names.

Mouse Lines	Abbreviated Name	Function
ROSA-CAG-^{fl}tdTomato$^{stop\text{-}fl}$EGFP	ROSA-TG	Cre Reporter
ROSA-CAG-Cre^{ERT2}	ROSA-Cre^{ERT2}	Cre Induction
Lef-1$^{fl\text{-}Exon7/8\text{-}fl}$	Lef-1$^{fl/fl}$	Conditional Lef-1 Knockout
ROSA-^{fl}EGFP$^{stop\text{-}fl}$Lef-1	Lef-1KI	Conditional Lef-1 Knock-in
ROSA-Cre^{ERT2}:ROSA-TG:Lef-1$^{fl/fl}$	Lef-1cKO:ROSA-TG	-
ROSA-Cre^{ERT2}:ROSA-TG	Lef-1WT:ROSA-TG	-
ROSA-Cre^{ERT2}:Lef-1$^{fl/fl}$	Lef-1cKO	-
ROSACre^{ERT2}:ROSA^{fl}EGFP$^{stop\text{-}fl}$Lef-1:Lef-1$^{fl/fl}$	Lef-1cKO:Lef-1KI	-

Primary Cell Isolation and Culture:

Primary Cell Isolation: Surface epithelial cells were isolated from resected murine tracheae using enzymatic digestion. Tracheae were opened longitudinally to expose the airway epithelium, then digested in 1.5 mg/mL Pronase (Roche) in DMEM w/ 1% Penicillin/Streptomycin and 0.1% amphotericin B for 60 min at 37ºC with gentle nutation; tracheae from 3-6 animals were digested together in batches. Tissues were agitated to remove surface airway epithelium (SAE), then thoroughly rinsed in Epithelial Wash Media (DMEM with 1% Penicillin/Streptomycin, 0.1% amphotericin B, and 10% FBS). The resulting solution was passed through a 100 µm cell strainer to remove tissue clumps. Cells were pelleted by centrifuging at 1000 RPM for 5 min and resuspended in SAGM prior to plating for culture.

Primary Cell Culture: Culture conditions and media used to propagate BCs were previously described[1]. Primary murine BCs were cultured in Small Airway Growth Medium (SAGM) (Lonza, Cat#CC-3118) modified with the addition of 10 µM Y-27632 (Tocris, Cat#1254), 1 µM DMH-1 (Tocris, Cat#4126), 1 µM A83-01 (Tocris, Cat#2939), and 1 µM CHIR 99021 (Tocris, Cat#4423) and 1% Penicillin/Streptomycin. Monolayer cultures were grown on tissue culture plastic pre-treated with filter-sterilized laminin-enriched 804G-conditioned media as previously described[2]. For air-liquid interface (ALI) cultures, BCs were initially grown in SAGM then seeded onto 0.33 cm^2 polyester transwell membranes (Corning). Membranes were pre-treated with 804G-conditioned media. After 2-3 days of expansion in SAGM, cells were “lifted” to air-liquid interface, then maintained in PneumaCult ALI media (StemCell Technologies, Inc.). For all experiments, Accutase (StemCell Technologies) was used to dissociate cells from plastic plates. Centrifugation of cells was done at 1000 RPM for 5 min. Cells used in immunofluorescent analyses were grown on Tek Chamber Permanox Slides (ThermoFisher Scientific).

Cell Proliferation and Competition Assays:

Cell proliferation assays were performed on cultured primary BCs isolated from ROSA-Cre^{ERT2}:$Lef\text{-}1^{fl/fl}$:ROSA-TG mice. Cells were first passaged three times after isolation, then treated with either 4-hydroxytamoxifen (OH-Tam) (Sigma-Aldrich) or 100% ethanol (vehicle control). OH-Tam was first

dissolved in 100% ethanol, then added to SAGM at a final concentration of 1 μM. Cells were grown in treatment media for 3 days, with the media replaced every day. Cells were then expanded until ~85-90% confluency, then passaged onto 6 well dishes. Cell number was determined each day 24 hours after passaging using a Countess Automated Cell Counter (Invitrogen). Competition assays were performed using passage 3 primary BCs isolated from Lef-1cKO, Lef-1cKO:ROSA-TG, Lef-1cKO:Lef-1KI, and ROSA-TG mice. Lef-1cKO, Lef-1cKO:ROSA-TG, or Lef-1cKO:Lef-1KI BCs were mixed at a 9:1 ratio with ROSA-TG BCs, treated with OH-Tam or ethanol as described above, and then co-cultured for 5 passages. Cells were split 1:5 into new culture dishes at each passage. Following Cre induction the color of each cell genotype was as follows: Lef-1cKO:ROSA-TG (GFP^+), Lef-1cKO and Lef-1cKO:Lef-1KI ($tdTomato^-/GFP^-$), and ROSA-TG ($tdTomato^+$). Cells were grown to ~85-90% confluency before passaging and quantification of populations. Quantification of GFP^+, $tdTomato^+$, and $GFP^-tdTomato^-$ cells was done using flow cytometry (Becton Dickinson LSR II). Flow cytometry was conducted at the University of Iowa Flow Cytometry Facility.

Immunofluorescence: Mouse tracheae were excised and fixed in 4% paraformaldehyde (PFA) overnight at 4°C, then washed in PBS and embedded in OCT frozen blocks. Tracheal longitudinal frozen sections were cut at 10 μm. Frozen tissue sections were post-fixed in 4% PFA, then washed in three changes of PBS. Slides were then incubated in blocking buffer containing 20% normal donkey serum, 0.3% Triton X-100, and 1 mM $CaCl_2$ in PBS for 1 hour at room temperature. Slides were incubated with primary antibodies in a diluent buffer containing 1% normal donkey serum, 0.3% Triton X-100, and 1 mM $CaCl_2$ in PBS overnight at 4°C. Slides were washed in three changes of PBS and incubated with secondary antibodies in diluent buffer for 1 hour at room temperature. Nuclei were stained using Hoechst 33342 (Invitrogen). Slides were mounted with either ProLong Gold (Invitrogen) or Aqua-Mount (Lerner Laboratories). A complete list of the antibodies used can be found in Table S2.

Airway cell cultures (both expanding BC cultures and ALI cultures) were fixed in 4% PFA for 20 min at room temperature, then washed in PBS 3 times. Primary and secondary antibody incubations were performed as described above. Detection of EdU was done using Click-iT™ Plus EdU Cell

Proliferation Kit for Imaging (ThermoFisher Scientific) and the accompanying protocol. EdU detection was performed after the completion of the secondary antibody incubation.

Fluorescent images were collected with either a Zeiss LSM 700 line-scanning confocal microscope or a Zeiss LSM 880 line-scanning confocal microscope (Carl Zeiss, Germany).

Table S2: Primary and secondary antibodies used.

Primary Antibodies		
Antibody Target	**Company**	**Catalog No.**
Keratin 5	BioLegend	905901
α-Tubulin	Cell Signaling Technology	5335
Scgb1a1	Millipore	ABS1673
Muc5B	Santa Cruz Biotechnology	sc-20119
Cyclin D1	AbCam	ab16663
Phospho-H3	AbCam	ab5176
GFP	Rockland	600-101-215
Secondary Antibodies		
Species	**Company**	**Catalog No.**
Donkey α-Chicken	Jackson ImmunoResearch	703-606-155
Donkey α-Rabbit	Jackson ImmunoResearch	711-606-152
Donkey α-Rabbit	Invitrogen	A21206
Donkey α-Goat	Invitrogen	A11055
Donkey α-Goat	Jackson ImmunoResearch	705-606-147

Fluorescent *in situ* Hybridization: Primary Lef-1cKO murine basal cells were cultured in SAGM and treated with OH-Tam or ethanol as described above. Twenty-four hours following OH-Tam treatment, cells were pulsed with 10 μM EdU for 2 hours, then lifted from the culture dish and spun onto glass microscope slides using a cytocentrifuge. Cells were spun at 700 RPM for 8 min. EdU detection was performed first as described above. Fluorescent *in situ* hybridization was then performed using a ViewRNA Cell Plus Assay Kit (ThermoFisher Scientific) with the Cytospin Module Kit (ThermoFisher Scientific), with one modification to the provided protocol: Probe incubation temperature was increased from 40°C to 44°C. The murine *Lef-1* mRNA probe used was designed, manufactured, and tested by ThermoFisher Scientific.

RNA Sequencing of Cultured Basal Cells: Cultured basal cells (P4) were treated once with OH-Tam or ethanol as mentioned above. Cells were lifted from the plates at the following timepoints following the OH-Tam or ethanol treatment: 0 hrs (untreated), 6 hrs, 12 hrs, 18 hrs, 24 hrs, and 30 hrs. Four replicates were done for each treatment group per timepoint with each replicate consisting of an independent pool of primary cells harvested from different mice. Cells were centrifuged into pellets and lysed using Lysis Buffer from a RNeasy Plus Mini Kit (QIAGEN). RNA was then extracted following the kit's protocol. Transcription profiling using RNA-Seq was performed by the University of Iowa Genomics Division using manufacturer recommended protocols. Initially, 500 ng of DNase I-treated total RNA was used to enrich for polyA containing transcripts using oligo(dT) primers bound to beads. The enriched RNA pool was then fragmented, converted to cDNA and ligated to sequencing adaptors containing indexes using the Illumina TruSeq stranded mRNA sample preparation kit (Cat. #RS-122-2101, Illumina, Inc., San Diego, CA). The molar concentrations of the indexed libraries were measured using the 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) and combined equally into pools for sequencing. The concentrations of the pools were measured using the Illumina Library Quantification Kit (KAPA Biosystems, Wilmington, MA) and sequenced on the Illumina HiSeq 4000 genome sequencer using 150 bp paired-end SBS chemistry.

RNA-Seq Analysis: The bowtie indexed ensemble mm10 reference genome was used to align sequenced reads. Sequence alignment and transcript abundance quantification was performed using RSEM software[3]. Differential gene expression, gene enrichment and clustering analysis was performed using PyMINEr[4]. PC analysis was done using the ggplot2 package for the statistical analysis software R Studio[5]. Ingenuity Pathway Analysis (IPA; Qiagen) software was used to perform various downstream analyses.

Image Analysis: Tile-scanned fluorescent images were quantified using the Multi Wavelength Cell Scoring Application Module of MetaMorph Software. For images of longitudinal tracheal sections, areas of interest (i.e. the surface airway epithelium) were first "cut out" of the original image using ImageJ. At

least 2 sections separated by 50 μm or more were analyzed for each animal. Quantification of area performed on ALI membrane images was done using ImageJ.

SUPPLEMENTAL REFERENCES

1. Lynch TJ, Anderson PJ, Rotti PG, et al. Submucosal Gland Myoepithelial Cells Are Reserve Stem Cells That Can Regenerate Mouse Tracheal Epithelium. Cell stem cell. 2018;22:779.
2. Mou H, Vinarsky V, Tata PR, et al. Dual SMAD Signaling Inhibition Enables Long-Term Expansion of Diverse Epithelial Basal Cells. Cell stem cell. 2016;19:217-231.
3. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.
4. Tyler SR, Rotti PG, Sun X, et al. PyMINEr Finds Gene and Autocrine-Paracrine Networks from Human Islet scRNA-Seq. Cell Rep. 2019;26:1951-1964.e1958.
5. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2009.
6. Adam RC, Yang H, Ge Y, et al. Temporal Layering of Signaling Effectors Drives Chromatin Remodeling during Hair Follicle Stem Cell Lineage Progression. Cell stem cell. 2018;22:398-413.e397.

SUPPLEMENTAL TABLE LEGENDS

Table S1. Genetic mouse lines and abbreviated names. This table list the genetic names of mouse lines used in the study and the abbreviated names used in the text. Table can be found within the Supplemental Methods.

Table S2: Primary and secondary antibodies used. This Table list the antibodies used for immunostaining and the source and catalog numbers. Table can be found within the Supplemental Methods.

Table S3A. Bulk RNA-Seq of WT and Lef-1cKO murine basal cells (BCs). Columns D-AU are the individual replicate Log2(TPM) values for each gene at the timepoint indicated. Columns AV-CM list the individual replicate expression enrichment Z-scores for each gene at the indicated timepoint. Column CN lists Benjamini-Hochberg (BH) adjusted *p*-values comparing Log2(TPM) values of WT and Lef-1cKO groups for each gene across all timepoints. Differentially regulated genes are highlighted in blue (BH-adjusted $p \leq 0.05$).

Table S3B. Differentially expressed genes in Lef-1cKO basal cells (BCs). A total of 6,406 out of 21,390 genes were differentially expressed (BH-adjusted $p < 0.05$) between WT and Lef-1cKO BCs in at least one time point. Columns D-AU list the individual replicate Log2(TPM) values of each gene at the indicated timepoint. Columns AV-CM list the individual replicate expression enrichment Z-scores of each gene at the indicated timepoint. Column CN lists the BH-adjusted *p*-values of each gene. Genes are sorted from lowest to highest BH-adjusted *p*-value.

Table S3C. Average expression of differentially regulated genes in Lef-1cKO basal cells (BCs). Average Log2(TPM) values were calculated from the four replicates for each timepoint and treatment condition in Table S1B. Columns D-N list these averages. Columns O-Y list the expression enrichment Z-scores calculated from these average values. Column Z lists the BH-adjusted *p*-value for each gene. Genes are sorted from lowest to highest BH-adjusted *p*-value.

Table S4A. Fold change in expression of differentially regulated genes in Lef-1cKO basal cells (BCs). Log2(fold change/FC) was determined for genes that were differentially expressed (BH-adjusted *p*-value < 0.05) at each timepoint in Lef-1cKO BCs. Log2(FC) values > 0 are shaded red, while Lef-1cKO BCs values < 0 are shaded blue.

Table S4B. Genes with a significant fold change in expression in Lef-1cKO basal cells (BCs). Of 6,406 differentially expressed genes, 1,028 had a Log2(FC) ≥ 1 or ≤ -1 (absolute FC ≥ 2 or ≤ -2). These genes are depicted in the heatmap in Figure 3.

Table S5A. Significantly altered biological processes and functions in Lef-1cKO basal cells (BCs). Ingenuity Pathway Analysis (IPA) was performed on the Log2(FC) of differentially expressed genes between WT and Lef-1cKO BCs (Table S1B) to identify differentially regulated biological processes and functions. Shown are the processes and functions that were statistically significant. Column D displays

the predicted activated state of the corresponding process/function, as determined by IPA. Blank values indicate that the activation Z-score was not significant enough to make a prediction (>2 or <-2). Column E displays the overall activation Z-score for each process. The gene names and total number of genes in each pathway, which were differentially expressed in Lef-1cKO BCs, are listed in column F and G, respectively. Yellow highlighted rows indicate pathways that were included in the graph in Figure 4.

Table S5B. Significantly altered canonical pathways in Lef-1cKO basal cells (BCs). Ingenuity Pathway Analysis (IPA) was performed on the Log2(FC) of differentially expressed genes between WT and Lef-1cKO BCs (Table S1B) to identify differentially regulated biological processes and functions. Shown are the canonical pathways that were statistically significant ($p < 0.05$). Column C displays the overall activation Z-score for each pathway, as determined by IPA. Positive values indicate that a majority of the molecules (genes) within that pathway were upregulated, while negative values indicate that a majority were downregulated. Values marked N/A could not calculate a Z-score. The differentially expressed genes in each pathway are included in column D. Yellow highlighted rows indicate pathways that were included in the graph in Figure 4.

Table S5C. DNA Damage Response (DDR) genes differentially expressed in Lef-1cKO basal cells (BCs). The list of differentially expressed genes with Log2(FC) from Table S1B was submitted to Ingenuity Pathway Analysis (IPA). 139 of those genes were identified as being linked to the DDR pathway, and are listed in Column A. Columns B-K list the expression enrichment Z-scores for each of these genes at the indicated timepoints. Columns L-P list the Log2(FC) between time-matched ethanol and OH-Tam treated BCs for each gene.

Table S6A. Potential direct Lef-1 target genes differentially expressed in Lef-1cKO basal cells (BCs). The list of genes generated from a previously published Lef-1 ChIP-seq dataset on hair follicle stem cells [6] was cross referenced to our list of differentially expressed genes (DEGs) in Lef-1cKO BCs (Table S1B). This resulted in a list of 1,138 genes that are DEGs in Lef-1cKO BCs and potential direct

Lef-1 targets (Column A). Columns B-K lists the expression enrichment Z-score of each gene listed at indicated timepoints. Columns L-P list the Log2(FC) between time-matched ethanol and OH-Tam treated BCs for each gene.

Table S6B. Differentially expressed candidate direct Lef-1 target genes associated with the DDR pathway. The list of differentially expressed genes (DEGs) generated in Table S4A was cross-referenced to a list of genes associated with the DNA damage response (DDR) pathway, according to IPA. This resulted in a list of 26 DEGs in Lef-1cKO BCs that are potential direct Lef-1 targets (Column A). Columns B-K list the expression enrichment Z-score of each gene listed at the indicated timepoints. Columns L-P list the Log2(FC) between time-matched ethanol and OH-Tam treated BCs for each gene.

Table S7A. IPA Upstream Regulator Analysis of differentially expressed genes in Lef-1cKO basal cells (BCs). IPA software was used to predict the activation state of upstream regulators based on the differential expression of their downstream target genes in our RNA-Seq dataset (Table S1B). Column A lists the upstream regulators identified in this analysis. Column C shows the predicted activation state of the corresponding upstream regulator. Regulators with activation Z-scores <-2 are listed as "Inhibited", and regulators with activation z-scores >2 are listed as "Activated". Genes with activation Z-scores (Column D) between -2 and 2 do not have an activation prediction. Column E lists the overlap p-value. The overlap p-value measures whether there is a statistically significant overlap between the dataset genes and the genes that are regulated by the upstream regulator. It is calculated using Fisher's Exact Test. Column F lists the subset of target genes controlled by the corresponding upstream regulator that had significantly altered gene expression in Lef-1cKO BCs. Upstream regulators highlighted in yellow are included in Figure S4A and S4B.

Table S7B. TGFβ1-regulated genes that are differentially expressed in Lef-1cKO basal cells (BCs). IPA upstream regulator analysis (Table S5A) identified TGFβ1 as significantly activated. This table lists the differentially expressed genes (Column A) from our Lef-1cKO BCs RNA-Seq dataset that where

classified by IPA regulated targets of TGFβ1. Column B lists the predicted state of TGFβ1 signaling, based on the Log2(FC) of the corresponding gene in our dataset. "Activated" indicates that a gene's expression pattern is consistent with an activated state of TGFβ1 signaling, while "Inhibited" indicates that a gene's expression change is the opposite of expected when TGFβ1 signaling is activated. "Affected" indicates that there isn't enough data available to determine how TGFβ1 signaling changes a gene's expression. Columns C-G list the Log2(FC) in expression of the corresponding gene in our RNA-Seq data set. Columns H-L list the Z-scores for the indicated OH-Tam treated samples. Gene expression enrichment Z-scores with an "Activated" TGFβ1 prediction (Column B) are shown in the heatmap in Figure S4C.

Table S7C. SMAD3-regulated genes that are differentially expressed in Lef-1cKO basal cells (BCs). IPA upstream regulator analysis (Table S5A) identified SMAD3 as significantly activated. This table lists the differentially expressed genes (Column A) from our Lef-1cKO BCs RNA-Seq dataset that where classified by IPA regulated targets of SMAD3. Column B lists the predicted state of SMAD4 signaling, based on the Log2(FC) of the corresponding gene in our dataset. "Activated" indicates that a gene's expression pattern is consistent with an activated state of SMAD3 signaling, while "Inhibited" indicates that a gene's expression change is the opposite of expected when SMAD3 signaling is activated. "Affected" indicates that there isn't enough data available to determine how SMAD3 signaling changes a gene's expression. Columns C-G list the Log2(FC) in expression of the corresponding gene in our RNA-Seq data set. Columns H-L list the Z-scores for the indicated OH-Tam treated samples. Gene expression enrichment Z-scores with an "Activated" SMAD3 prediction (Column B) are shown in the heatmap in Figure S4D.

Table S7D. SMAD4-regulated genes that are differentially expressed in Lef-1cKO basal cells (BCs). IPA upstream regulator analysis (Table S5A) identified SMAD4 as significantly activated. This table lists the differentially expressed genes (Column A) from our Lef-1cKO BCs RNA-Seq dataset that where classified by IPA regulated targets of SMAD4. Column B lists the predicted state of SMAD4

signaling, based on the Log2(FC) of the corresponding gene in our dataset. "Activated" indicates that a gene's expression pattern is consistent with an activated state of SMAD4 signaling, while "Inhibited" indicates that a gene's expression change is the opposite of expected when SMAD4 signaling is activated. "Affected" indicates that there isn't enough data available to determine how SMAD4 signaling changes a gene's expression. Columns H-L list the Z-scores for the indicated OH-Tam treated samples. Gene expression enrichment Z-scores with an "Activated" SMAD4 prediction (Column B) are shown in the heatmap in Figure S4D.