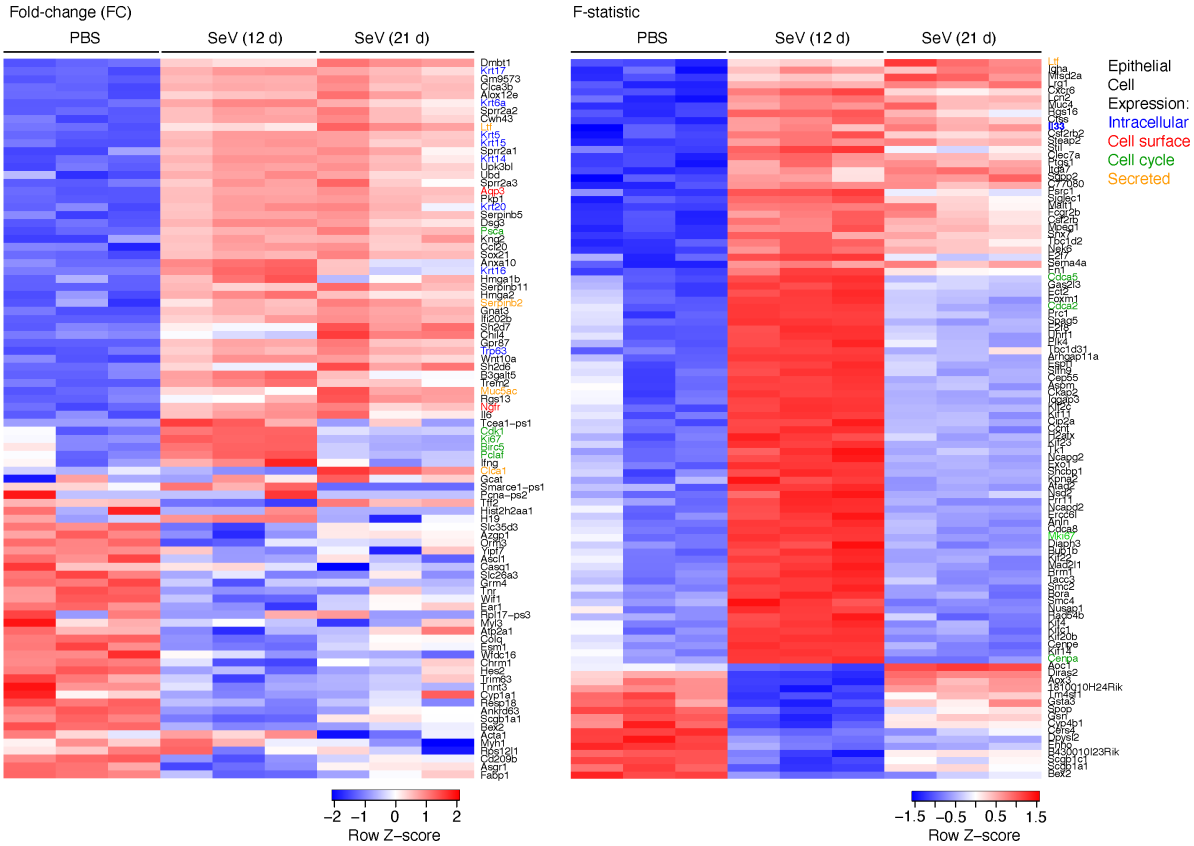
Supplemental Figures 1-10, Methods, and Tables 1-2

Basal-epithelial stem cells cross an alarmin checkpoint for post-viral lung disease

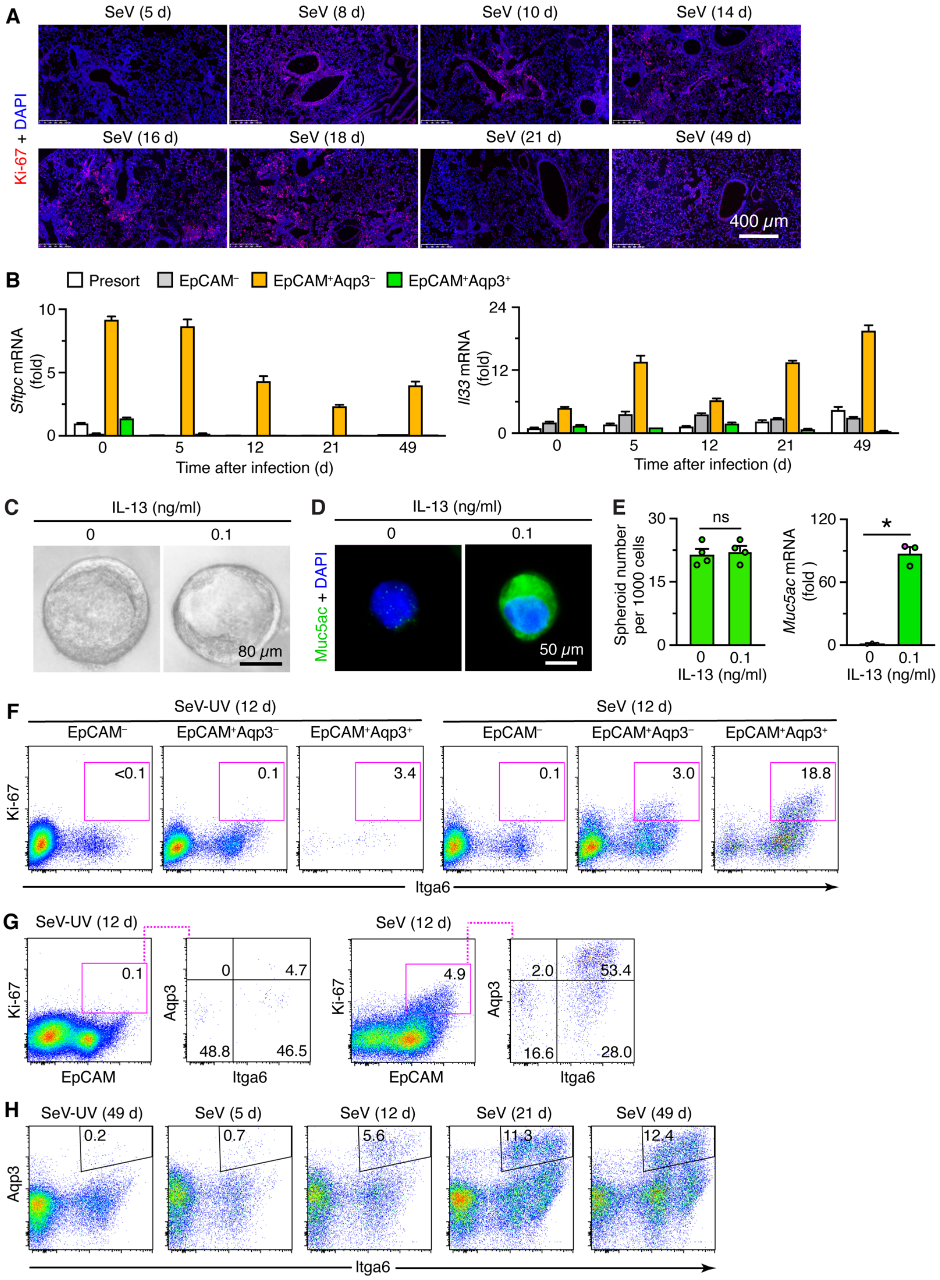
Kangyun Wu[1], Kenji Kamimoto[2,3], Yong Zhang[1], Kuangying Yang[1,4], Shamus P. Keeler[1], Benjamin J. Gerovac[1], Eugene V. Agapov[1], Stephen P. Austin[1], Jennifer Yantis[1], Kelly A. Gissy[1], Derek E. Byers, Jennifer Alexander-Brett[1,5], Christy M. Hoffmann[1], Matthew Wallace[1], Michael E. Hughes[1,2], Erika C. Crouch[5], Samantha A. Morris[2,3], and Michael J. Holtzman[1,6]*

[1]Pulmonary and Critical Care Medicine, Department of Medicine, [2]Department of Genetics, [3]Department of Developmental Biology, [4]Division of Biostatistics, [5]Department of Pathology and Immunology, [6]Department of Cell Biology and Physiology, Washington University School of Medicine, Saint Louis, MO 63110

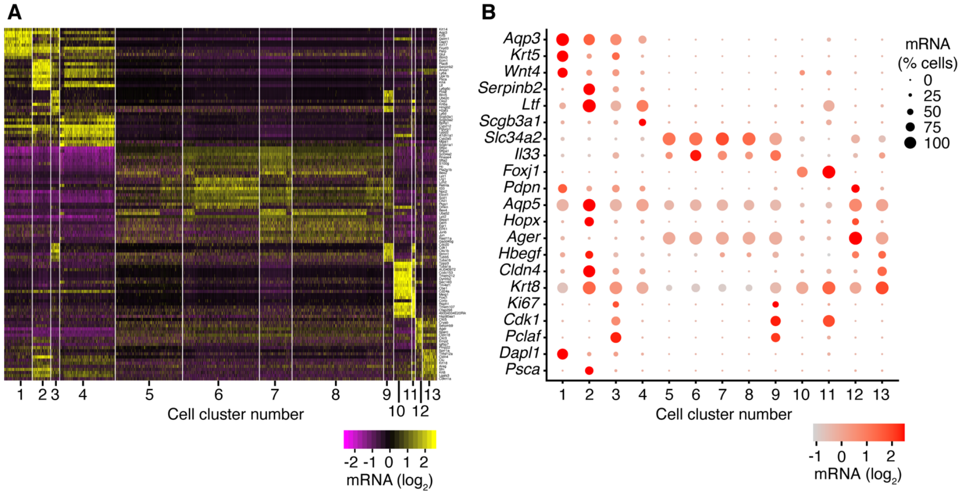
*Address correspondence to: M.J.H., Washington University School of Medicine, Campus Box 8052, 660 South Euclid Avenue, Saint Louis, MO 63110; Tel 314-362-8970; Fax 314-362-9009; E-mail mjholtzman@wustl.edu



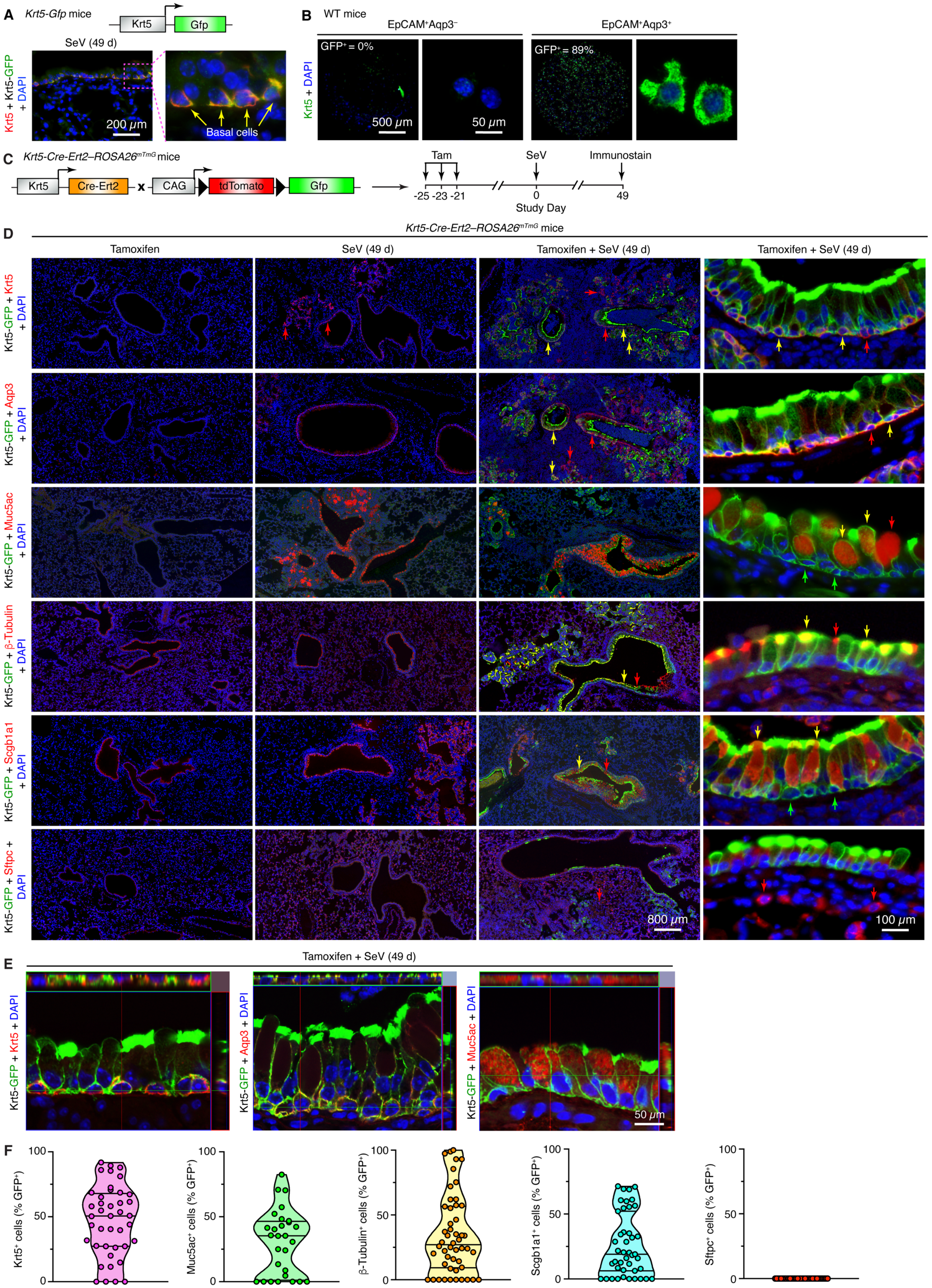
Supplemental Figure 1. Heat maps for RNAseq analysis of gene expression in lung tissue after SeV infection. Conditions correspond to Figure 1 (PBS control and 12 and 21 d after SeV infection in WT mice). Gene list is based on greatest fold-change (FC) or F-statistic difference from PBS control condition. Basal epithelial cell or basal epithelial cell lineage genes of interest to post-viral lung disease are annotated by colors for indicated categories. Data represent results from a single experiment with n=3 mice per condition.



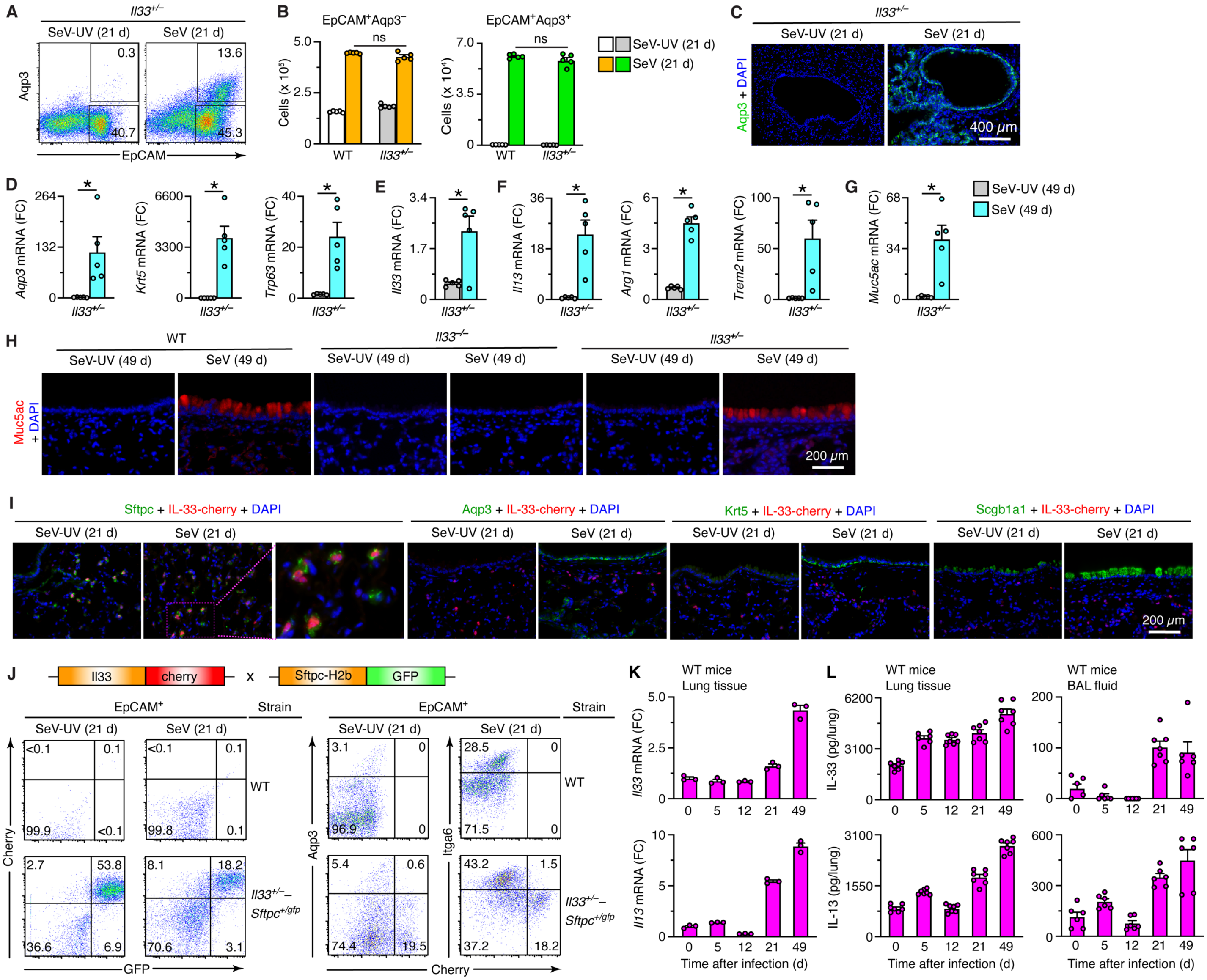
Supplemental Figure 2. Levels of proliferating Aqp3+Itga6+ basal epithelial cells and Sftpc^{+} AT2 cells after SeV infection. **A**, Immunostaining for Ki-67 in lung sections from WT mice after SeV infection. **B**, Levels of *Sftpc* and *Il33* mRNA in indicated FACS-isolated cell populations. **C**, Phase-contrast microsopy images of lung spheroids formed from EpCAM^{+}Aqp3^{+} cells and cultured with and without IL-13. **D**, Corresponding immunostaining for Muc5ac of lung spheroid cells for conditions in (D). **E**, Levels of lung spheroids and *Muc5ac* mRNA in spheroids for conditions in (D). **F**, Flow cytograms for Ki-67 in CD31^{-}CD45^{-} cells relative to Aqp3 and Itga6 expression at 12 d after SeV or SeV-UV. **G**, Flow cytograms of lung epithelial (CD31^{-}CD45^{-}) cells based on Ki-67, EpCAM, Aqp3, and Itga6 expression in wild-type mice at 12 d after SeV or SeV-UV. Values indicate % of cells within each gate. **H,** Flow cytograms of lung epithelial (CD31^{-}CD45^{-}EpCAM^{+}) cells based on Aqp3 and Itga6 expression in WT mice after SeV or SeV-UV. Values indicate % of cells within each gate. Data represent results from a single experiment with n=3 mice per condition and were replicated twice. **P*<0.05 using ANOVA with Bonferroni correction. **P*<0.01 using ANOVA with Bonferroni correction.



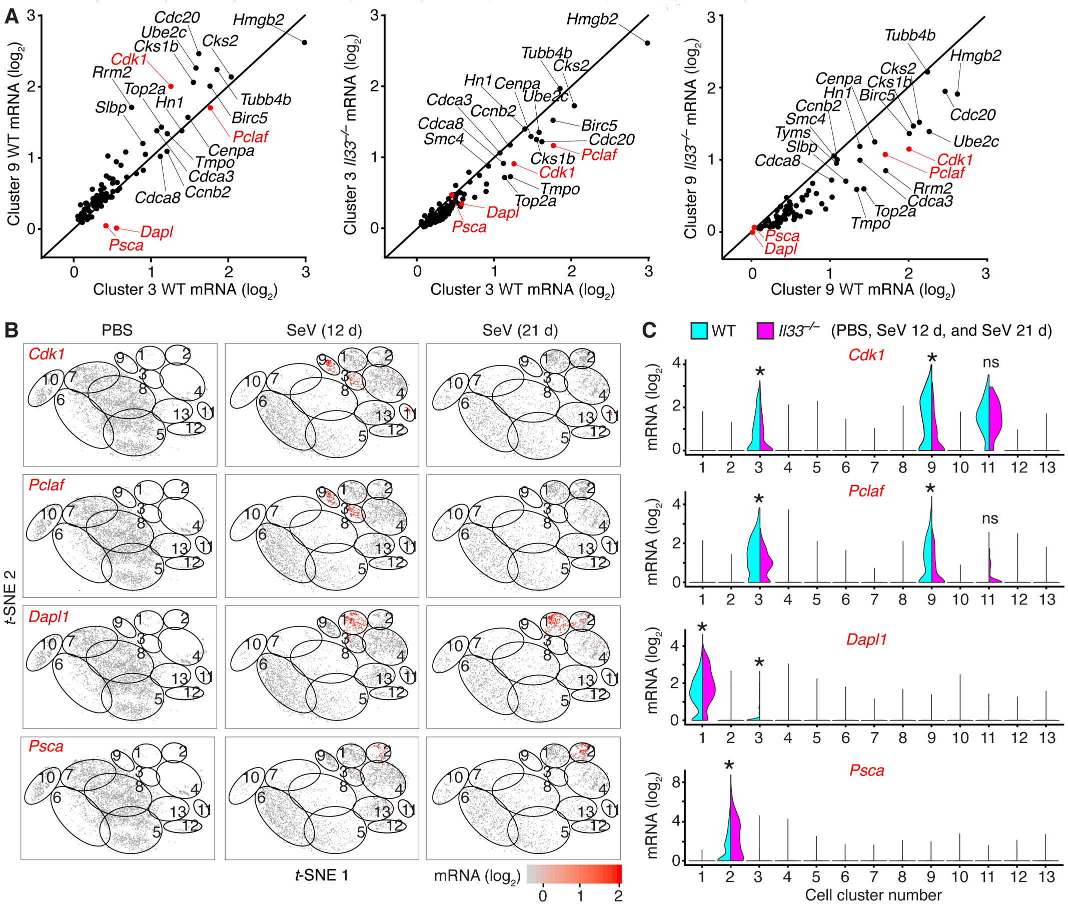
Supplemental Figure 3. Gene expression analysis for scRNAseq of lung epithelial cells after SeV infection. A, Heat map for top cluster marker genes for lung epithelial (CD45$^-$CD31$^-$ EPCAM$^+$) cells (n=18,465 single cells) from lungs of WT mice for combined sample conditions (12 and 21 d after SeV or PBS). **B,** Dot-plot of highly expressed marker genes for conditions in (**A**). Data represent results from a single experiment with n=3 mice per condition.


A
Krt5-Gfp mice
Krt5
Gfp
SeV (49 d)
Krt5 + Krt5-GFP + DAPI
200 μm
Basal cells
B
WT mice
EpCAM⁺Aqp3⁻
EpCAM⁺Aqp3⁺
GFP⁺ = 0%
GFP⁺ = 89%
Krt5 + DAPI
500 μm
50 μm
C
Krt5-Cre-Ert2–ROSA26mTmG mice
Krt5
Cre-Ert2
x
CAG
tdTomato
Gfp
Tam
SeV
Immunostain
-25 -23 -21
0
49
Study Day
D
Krt5-Cre-Ert2–ROSA26mTmG mice
Tamoxifen
SeV (49 d)
Tamoxifen + SeV (49 d)
Tamoxifen + SeV (49 d)
Krt5-GFP + Krt5 + DAPI
Krt5-GFP + Aqp3 + DAPI
Krt5-GFP + Muc5ac + DAPI
Krt5-GFP + β-Tubulin + DAPI
Krt5-GFP + Scgb1a1 + DAPI
Krt5-GFP + Sftpc + DAPI
800 μm
100 μm
E
Tamoxifen + SeV (49 d)
Krt5-GFP + Krt5 + DAPI
Krt5-GFP + Aqp3 + DAPI
Krt5-GFP + Muc5ac + DAPI
50 μm
F
Krt5⁺ cells (% GFP⁺)
Muc5ac⁺ cells (% GFP⁺)
β-Tubulin⁺ cells (% GFP⁺)
Scgb1a1⁺ cells (% GFP⁺)
Sftpc⁺ cells (% GFP⁺)
100
50
0

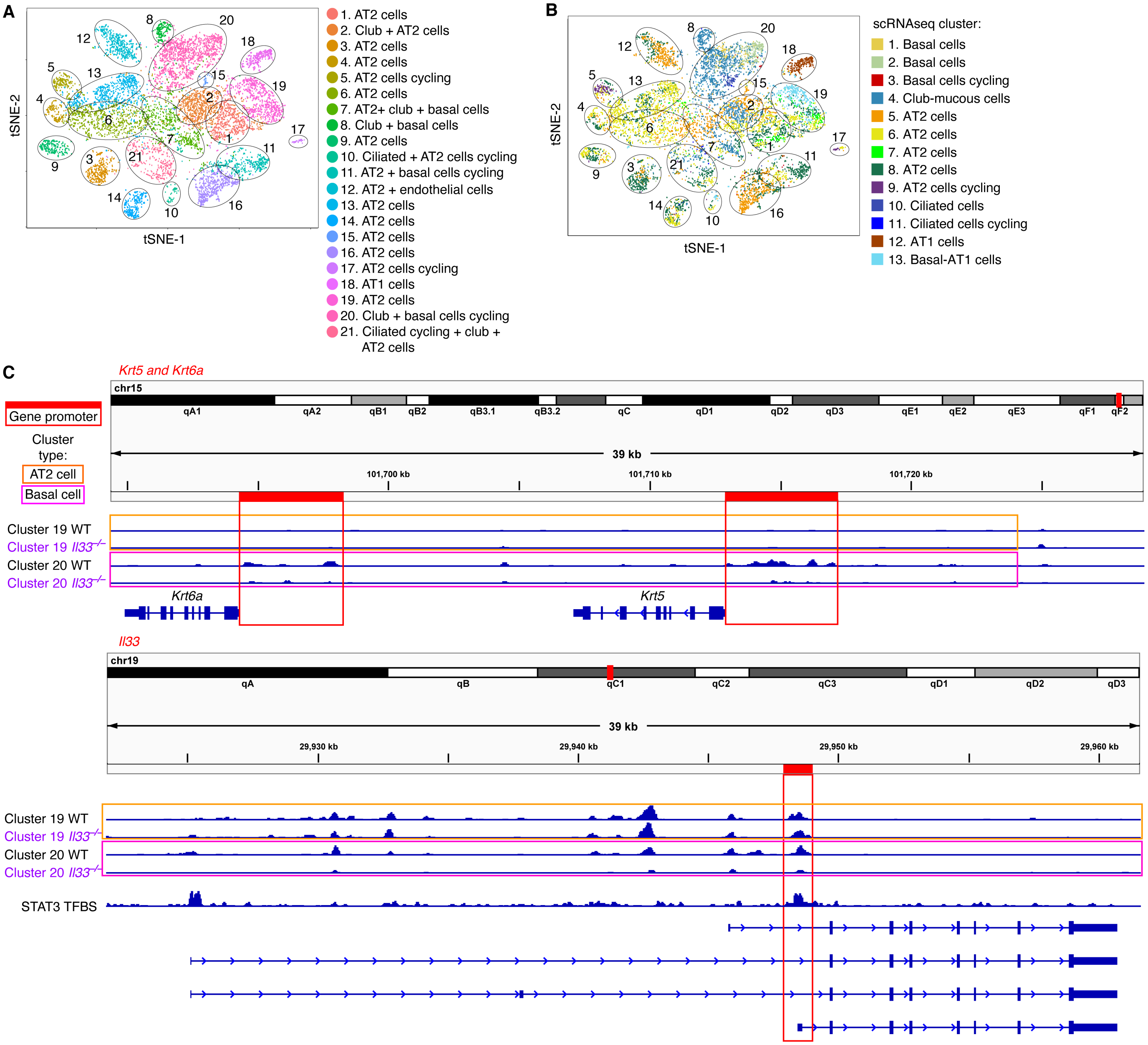
Supplemental Figure 4. Basal-ESC lineage restriction to airway (including mucous) cells. A, Immunostaining for Krt5 and GFP with DAPI counterstaining in lung sections from *Krt5-Gfp* mice at 49 d after infection with SeV. Yellow arrows, double-positive (Krt5$^+$Krt5-GFP$^+$) basal epithelial cells. **B**, Immunostaining for Krt5 with DAPI counterstaining in cells prepared from WT mice using FACS for CD31$^-$CD45$^-$EpCAM$^+$Aqp3$^-$ and CD31$^-$CD45$^-$EpCAM$^+$ Aqp3$^+$ lung cells. **C,** Scheme for lineage tracing using *Krt5-CreERT2–ROSA26mTmG* mice precisely timed for tamoxifen (Tam) treatment and SeV infection. **D**, Immunostaining for GFP and indicated marker (Krt5, Aqp3, Muc5ac, β-Tubulin IV, Scgb1a1, and Sftpc) with DAPI counterstaining in lung sections from *Krt5-CreERT2–ROSA26mTmG* mice. Yellow arrows, double-positive (Krt5-GFP$^+$epithelial-marker$^+$) cells; red arrows, single-positive (Krt5-GFP$^-$Marker$^+$) cells. **E,** Immunostaining for GFP and Krt5 or Muc5ac marker with DAPI counterstaining in lung sections imaged with confocal laser scanning microscopy to identify Krt5-GFP-trace in Krt5$^+$ and Aqp3$^+$ basal cells and Muc5ac$^+$ mucous cells for conditions in (**D**). **F**, Quantification of GFP tracing into indicated epithelial cell subset populations using violin plots for conditions in (**D**). Data represent results from a single experiment with n=3 mice per condition and were replicated twice. *$P<0.05$ using ANOVA with Bonferroni correction.



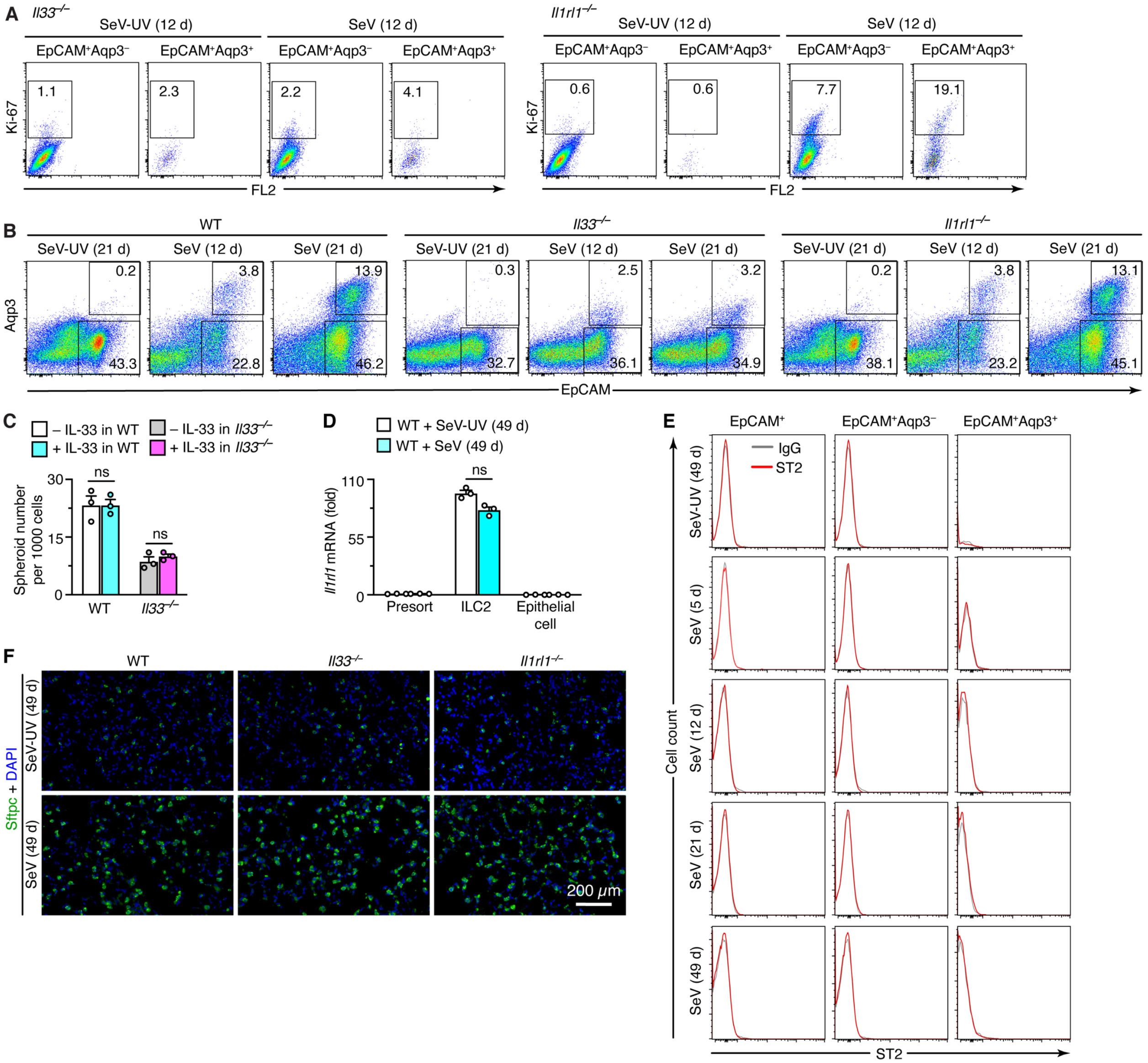
Supplemental Figure 5. IL-33-cherry-reporter mice track IL-33 expression to alveolar epithelial type 2 (AT2) cells. A, Flow cytograms of lung epithelial (CD31$^-$CD45$^-$) cells based on Aqp3 and EpCAM expression for heterozygous IL-33-cherry reporter (*Il33*$^{+/-}$) mice at 21 d after SeV or SeV-UV. Values indicate % of cells within each gate. **B**, Numbers of indicated cell populations for conditions in (**A**). **C**, Aqp3 immunostaining of lung sections from *Il33*$^{+/-}$ mice at 21 d after SeV or SeV-UV. **D-G**, Lung levels of mRNA markers for basal cells (D), IL-33$^+$ cells (E), type 2 immune response (F), and mucus production (G) for *Il33*$^{+/-}$ mice at 49 d after SeV infection or SeV-UV. **H,** Immunostaining for Muc5ac in lung sections from *WT*, *Il33*$^{-/-}$, and *Il33*$^{+/-}$ mice at 49 d after SeV or SeV-UV. **I,** Immunofluorescence for cherry-reporter and immunostaining for Sftpc, Krt5, Scgb1a1, and Aqp3 in lung sections from *Il33*$^{+/-}$ mice at 21 d after SeV or SeV-UV. **J**, Using *Il33*$^{+/-}$ mice crossed to *Sftpc*$^{+/gfp}$ mice (*Il33*$^{+/-}$*–Sftpc*$^{+/gfp}$) and WT mice, flow cytograms for cherry and GFP reporter fluorescence from lung epithelial (CD31$^-$CD45$^-$EpCAM$^+$) cells from WT and *Il33*$^{+/-}$*–Sftpc-gfp* mice at 21 d after SeV or SeV-UV and for corresponding cherry-reporter fluorescence and Aqp3 staining. **K**, Levels of *Il33* and *Il13* mRNA in lung tissue from WT mice at 0-49 d after SeV. **L**, Levels of IL-33 and IL-13 protein in lung tissue and BAL fluid for conditions in (K). Data represent results from a single experiment with n=3-5 mice per condition and were replicated twice. *$P<0.05$ using ANOVA with Bonferroni correction.



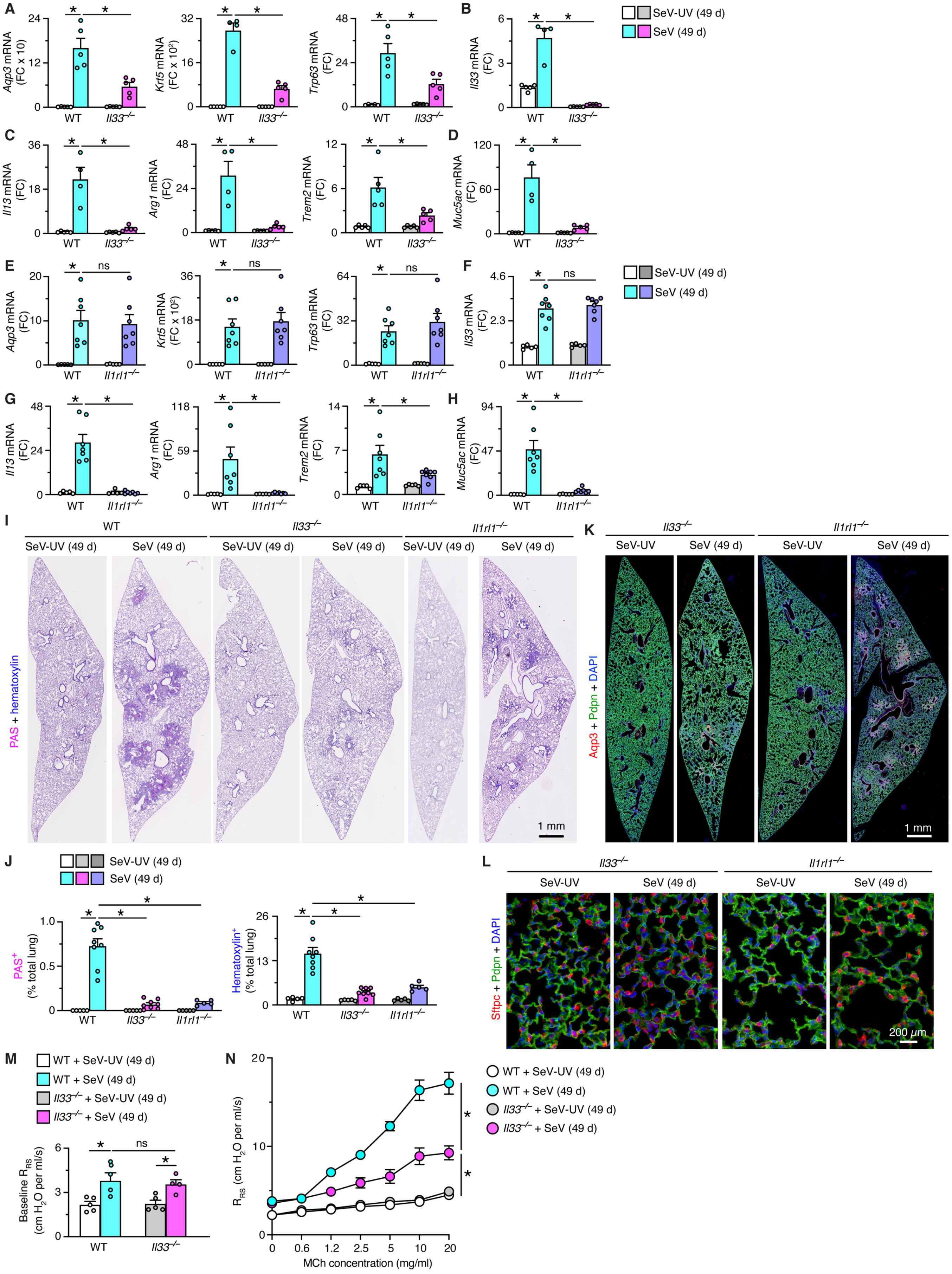
Supplemental Figure 6. Gene expression analysis for scRNAseq of lung epithelial cells and cell-cycle activation. A, Gene-gene plots for expression of cell-cycle proliferation and survival genes within and across cell clusters 3 and 9 in lungs of WT versus *Il33*$^{-/-}$ mice for combined sample conditions. **B**, *t*-SNE plots for expression of selected cell-cycle genes (*Cdk1*, *Pclaf, Dapl1, Psca*) for each sample condition in WT mice. **C**, Corresponding violin plots for expression of selected cell cycle genes for combined sample conditions from (**B**). Data represent results from a single experiment with n=3 mice per condition. **P*<0.05 using Mann-Whitney *U* test with Bonferroni correction.



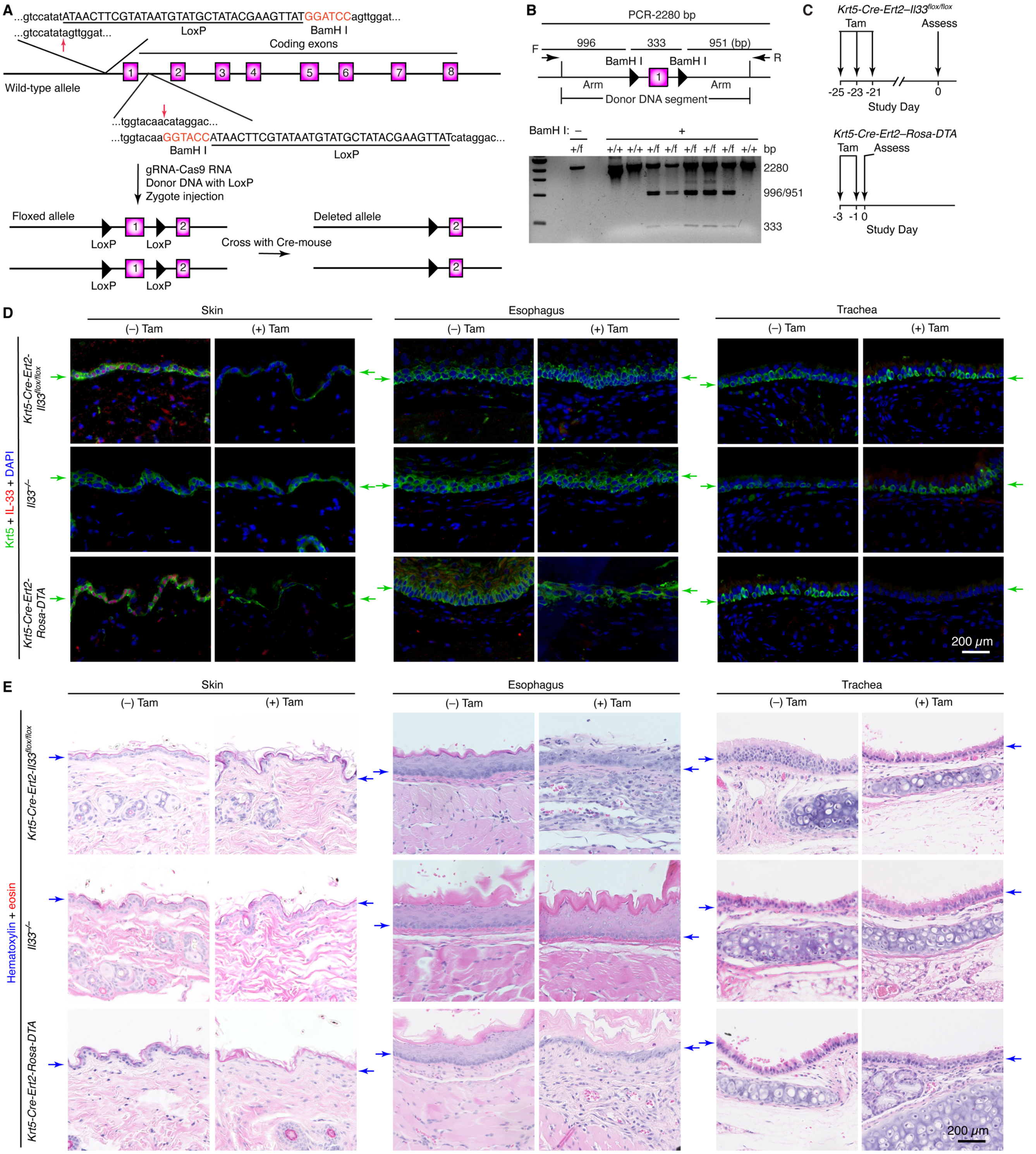
Supplemental Figure 7. Single-cell chromatin accessibility linked to IL33 expression in a basal-ESC subset. A, *t*-SNE plot for scATACseq analysis of lung epithelial (CD31$^-$CD45$^-$ EpCAM$^+$) cells from WT and *Il33*$^{-/-}$ mice at 12 d after SeV infection. **B**, *t*-SNE plot for cell clusters in (a) mapped on to gene clusters identified from scRNAseq analysis at 12 d after SeV infection. **C**, Integrated genomic view (IGV) analysis for peaks of activity for the *Il33* gene for the individual cell clusters identified from (A) at 12 d after SeV infection. Orange-line box designates AT2 cells (clusters 6, 13, 16, 19) with unchanged signal in *Il33*$^{-/-}$ compared to WT mice; red-line box designates basal cells and basal-cell lineage cells (cluster 20) with a decreased signal in *Il33*$^{-/-}$ compared to WT mice. Red boxes indicate gene-promoters near the transcription start site with an associated STAT3-transcription factor binding site (TFBS). Data represent results from a single experiment with n=3 mice per condition.



Supplemental Figure 8. Effect of *Il33* and *Il1rl1* gene expression on ESC proliferation after viral infection. **A**, Flow cytograms for Ki-67 levels in lung epithelial (CD31$^-$CD45$^-$EpCAM$^+$) cells (Aqp3$^-$ and Aqp3$^+$ subsets) in *Il33*$^{-/-}$*, and Il1rl1*$^{-/-}$ mice at 12 d after SeV infection or SeV-UV. (WT data reproduced from Figure 2H). **B,** Flow cytograms of lung epithelial (CD31$^-$CD45$^-$) cells based on EpCAM and Aqp3 expression in WT*, Il33*$^{-/-}$*, and Il1rl1*$^{-/-}$ mice at 12 and 21 d after SeV infection or SeV-UV. Values indicate % of cells within each gate. **C**, Levels of lung spheroid formation from Aqp3$^+$ lung epithelial (CD31$^-$CD45$^-$ EpCAM$^+$) cells isolated from WT and *Il33*$^{-/-}$ mice at 21 d after SeV infection and then 3D-cultured with and without treatment with IL-33 (10 ng/ml for 10 d). **D,** Levels of *Il1rl1* mRNA in ILC2 (SS^{low}lin$^-$ CD90.2$^+$ST2$^+$) and epithelial cells (CD31$^-$CD45$^-$EpCAM$^+$) that were isolated by FACS from lung tissue of WT mice at 49 d after SeV or SeV-UV. **E,** Flow cytometry histograms for IL-33R (ST2) levels on indicated epithelial cell populations after SeV or SeV-UV. **F**, Immunostaining for Sftpc in lungs sections at 49 d after SeV infection for conditions in (A). Data represent results from a single experiment with n=3 mice per condition and were replicated twice. *P<0.05 using ANOVA with Bonferroni correction.


A
Aqp3 mRNA (FC x 10)
Krt5 mRNA (FC x 10²)
Trp63 mRNA (FC)
B
Il33 mRNA (FC)
SeV-UV (49 d)
SeV (49 d)
WT
Il33⁻/⁻
C
Il13 mRNA (FC)
Arg1 mRNA (FC)
Trem2 mRNA (FC)
D
Muc5ac mRNA (FC)
E
ns
F
Il1rl1⁻/⁻
G
H
I
PAS + hematoxylin
1 mm
K
Aqp3 + Pdpn + DAPI
J
PAS⁺ (% total lung)
Hematoxylin⁺ (% total lung)
L
Sftpc + Pdpn + DAPI
SeV-UV
200 μm
M
WT + SeV-UV (49 d)
WT + SeV (49 d)
Il33⁻/⁻ + SeV-UV (49 d)
Il33⁻/⁻ + SeV (49 d)
Baseline R_RS (cm H₂O per ml/s)
N
R_RS (cm H₂O per ml/s)
MCh concentration (mg/ml)
0
0.6
1.2
2.5
5
10
20
*

Supplemental Figure 9. IL-33 versus ST2 requirement for gene expression, remodeling, and functional readouts for PVLD. A-D, Lung levels of mRNA markers for basal cells (**A**), AT2 cells (**B**), type 2 immune response (**C**), and mucus production (**D**) for WT and *Il33*$^{-/-}$ mice at 49 d after SeV versus SeV-UV**. E-H**, Corresponding lung levels of mRNA markers for WT and *Il1rl1*$^{-/-}$ mice for conditions in (A-D). **I**, PAS and hematoxylin staining of lung sections for WT, *Il33*$^{-/-}$, and *Il1rl1*$^{-/-}$ mice at 49 d after SeV or SeV-UV. **J**, Quantitation for conditions in (I). **K**, Immunostaining for Aqp3 and Pdpn in lung sections from *Il33*$^{-/-}$ and *Il1rl1*$^{-/-}$ mice at 49 d after SeV or SeV-UV. **L**, Immunostaining for Sftpc and Pdpn for conditions in (K)**. M**, Respiratory system resistance (R_{RS}) at baseline for WT and *Il33*$^{-/-}$ mice at 49 d after SeV or SeV-UV. **N**, Corresponding R_{RS} at baseline and after inhaled methacholine (MCh) for conditions in (M). Data represent results from a single experiment with n=5 mice per condition and were replicated twice. *$P<0.05$ using ANOVA with Bonferroni correction.



Supplemental Figure 10. Conditional deletion of *Il33* reveals selective control of basal epithelial cells. A, Design scheme for mouse *Il33*-floxed allele using CRISPR/Cas9-mediated modification of the *Il33* gene. **B**, Validation of heterozygous *Il33*$^{+/flox}$ (+/f) and WT (+/+) mice based on BamH I digestion, PCR-based amplification, and agarose gel electrophoresis. **C**, Protocol scheme for analysis of *Krt5-Cre-Ert2–Il33*$^{flox/flox}$ and *Krt5-Cre-Ert2–ROSA-DTA* mice with tamoxifen (Tam) treatment. **D**, Immunostaining for Krt5 and IL-33 with DAPI counterstaining of skin, esophagus, and trachea from *Krt5-Cre-Ert2–Il33*$^{flox/flox}$, *Il33*$^{–/–}$, and *Krt5-Cre-Ert2–ROSA-DTA* mice at 10 weeks of age as diagrammed in (C). **E**, Hematoxylin and eosin staining for conditions in (D). Data represent results from a single experiment with n=5 mice per condition and were replicated twice. *$P<0.05$ using ANOVA with Bonferroni correction.

Methods

Mouse line generation and treatment. Wild-type C57BL/6J mice (000664) were obtained from Jackson Laboratories (Bar Harbor, ME). The *Il33-cherry* reporter mice were generated in the C57BL/6J strain with insertion of an mCherry-reporter cassette as described recently (1). The *Il1rl1*$^{-/-}$ mice (2, 3) were obtained from Paul Bryce (Northwestern University, Chicago, IL). The *Sftpc-gfp* reporter mice (4) and *Krt5-gfp* mice (5) were obtained from Scott Randell (University of North Carolina, Chapel Hill, NC). The *Krt5-CreERT2* mice (6) backcrossed to C57BL/6N mice (#029155) and *ROSA26*^{mTmG} reporter mice (#007676) were obtained from Jackson Laboratories and crossed to generate *Krt5-Cre-Ert2–ROSA26*^{mTmG} mice. The *Il33*$^{flox/flox}$ mice were generated in C57BL/6J mice with *LoxP* sites inserted before and after exon 1 and crossed to *Krt5-Cre-Ert2* mice to generate *Krt5-Cre-Ert2–Il33*$^{flox/flox}$ mice (Supplemental Figure 10A,B). The ROSA-DTA mice (7) were obtained from Jackson Laboratories (#009669) and crossed with *Krt5-Cre-ERT2* mice to generate heterozygous *Krt5-Cre-Ert2–ROSA-DTA* mice. Mice were inoculated intranasally with Sendai virus (SeV strain 52) or an equivalent amount of UV-inactivated virus (SeV-UV) or PBS as controls as described previously (8, 9). For lineage tracing and conditional *Il33* gene deletion experiments, *Krt5-Cre-Ert2–ROSA26*^{mTmG} and *Krt5-Cre-Ert2–Il33*$^{flox/flox}$ mice were treated with or without tamoxifen (0.25 mg/g of body weight given intraperitoneally every other day for 3 doses at 21-25 d before SeV infection) as described previously (10). All mice were maintained on a C57BL/6J background and were co-housed in a barrier facility using cages fitted with micro-isolator lids. Male mice at 5-6 weeks of age were used to minimize variability in response to infection, although male and female mice respond similarly as reported previously (8, 11, 12).

RNAseq analysis. Lung tissue RNA was purified as described previously (13) from three mice per sample condition (12 or 21 d after SeV infection or PBS control challenge) and subjected to RNAseq using a NovaSeq 6000 with S4 flow cell (Illumina, San Diego, CA) in the Genome Technology Access Center in the McDonnell Genome Institute (Washington University, St. Louis). For RNAseq analysis, gene and transcript levels were quantified using the RSEM (version 1.3.2; https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-323) pipeline, and the estimated counts were stored as the appropriate raw data to feed into the edgeR (version 3.24.3; https://bioconductor.org/packages/release/bioc/html/edgeR.html), limma (version 3.38.3; https://www.ncbi.nlm.nih.gov/pubmed/25605792; https://www.bioconductor.org/packages/release/bioc/vignettes/limma/inst/doc/usersguide.pdf), and voom (version 3.38.3; https://www.ncbi.nlm.nih.gov/pubmed/24485249) pipeline. The list of Ensemble gene ids was annotated by the biomaRt (version 2.38.0), which queried BioMart databases (https://bioconductor.org/packages/release/bioc/vignettes/biomaRt/inst/doc/biomaRt.html) to retrieve additional information including entrezgene identifiers for gene symbol and chromosome. A DGEList appropriate for our data was declared, and filterByExpr function in edgeR was used to maintain 35% of the genes in the gene list by filtering out lowly expressed genes automatically and in a principled way. The data were then normalized to account for variability in sequencing depth, library size (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0191629), and other external factors that are not of biological interest. After TMM normalization, unsupervised heirarchical clustering with the hclust function in base R (version 3.5.3) and principal component analysis with the prcomp function were conducted to identified treatment and control clusters. The correlations between replicate samples were also examined, with the lowest value of 0.94. We invoked voom software to compute precision weights and fitted linear models to the data using the lmFit function. Then, we refitted the model using the contrasts.fit function and computed moderated t- and F-statistics using the eBayes and treat functions, respectively. We chose the outputs from treat function for a more stringent definition of statistical significance to test against a fold change threshold. Two heatmaps were generated with heatmap.2 function in gplots (version 3.1.0), based on subsets of the top 100 most significant genes sorted separately by the F-statistic and fold change. We also generated scatter plots with ggplot2 (version 3.3.2) with the top genes ranked by fold change labelled of our interest.

Real-time qPCR assay. RNA was purified from lung homogenates or cellular lysates using Trizol (Invitrogen) and was converted to cDNA using a High-Capacity cDNA Archive kit (Applied Biosystems). Target mRNA was quantified by real-time PCR assay using specific fluorogenic probes and primer sets and the Fast Universal PCR Master Mix system (Applied Biosystems). Primer-probe sets for viral RNA and mouse mRNA targets are listed in (Supplemental Table 2). All probes were designed to span introns and did not react with genomic DNA. Samples were assayed with the 7300HT or QuantStudio 6 Flex Fast Real-Time PCR System and analyzed using Fast System Software (Applied Biosystems). Cycle threshold (Ct) values were normalized to *GAPDH* mRNA level and then expressed as fold-change (FC) relative to the mean Ct value for control samples using the delta-delta Ct method (14).

Immunohistochemistry and immunocytochemistry. Tissues were fixed with 10% formalin, embedded in paraffin, cut into 5-μm sections, and adhered to charged slides. Sections were deparaffinized in Fisherbrand® CitroSolv® (Fisher), hydrated, and treated with heat-activated antigen unmasking solution (Vector Laboratories, Inc). Immunostaining was performed using primary antibodies (Supplemental Table 1) followed by secondary antibodies directly conjugated to Alexa Fluor 488 nm or

594 nm fluorochromes (Invitrogen) for immunofluorescence microscopy. All sections were counterstained with DAPI and were imaged using a Leica DM5000 B microscope for conventional imaging. Staining was quantified in whole lung sections using a NanoZoomer S60 slide scanner (Hamamatsu) and ImageJ software as described previously (9, 12).

Flow cytometry and FACS. Single cell suspensions were generated from minced lung tissue that was subjected to collagenase (Liberase TM Research Grade, Roche), hyaluronidase (Sigma), DNAse I (Sigma), and Dispase II (Roche) digestion for 45 min at 37 °C and then treated with ACK buffer (Lonza) to remove red blood cells. Following FcR blockade, lung cell suspensions were incubated with labeled antibodies and were sorted using a Sony SY3200 Synergy high-speed cell sorter. The following antibodies were used: anti-mouse CD31 (clone MEC 13.3; BD Biosciences), anti-mouse CD45 (clone 30-F11; BD Biosciences), anti-mouse EpCAM (clone G8.8; BioLegend), anti-mouse/human Itga6 (Clone GoH3; eBiosciences), anti-Aqp3 (Abcam) and anti-Ki-67 (clone SolA15, eBiosciences). Anti-Aqp3 antibody is labeled by Zenon antibody labeling kit following the instruction of manufacturer (Molecular Probes). FACS results were plotted and analyzed using FlowJo software (TreeStar).

Epithelial cell culture. For 3D-organoid cultures, epithelial cells from FACS were resuspended in SAGM (Lonza) mixed 1:1 with growth-factor reduced Matrigel (BD Biosciences) and plated at 3 x 10^3 cells per 100 µL in 24-well Transwells (Corning). At first 5 days, SAGM was added with 10 µM Rock inhibitor (Sigma) but without FBS. After 5 d, SAGM was added with 10 % FBS but without Rock inhibitor. The medium was changes every two or three days. Spheres were cultured for 14 d for counting and were harvested from Matrigel using cell recovery solution (BD Biosciences) and embedded in OCT (Tissue-Tek) for frozen sectioning and immunostaining.

scRNAseq analysis. $CD31^-CD45^-EpCAM^+$ cells were isolated from mouse lungs at 12 and 21 after SeV infection or control PBS challenge, and cellular mRNA was purified and subjected to scRNAseq using the 10x Genomics platform. Next-generation sequencing (NGS) data was processed with the Cell Ranger set of analysis pipelines (version 2.1.0, https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines) as follows: first, FASTQ files were generated for each sample by demultiplexing the sequencer's base call files. Then, single-cell gene count matrices were generated by running "cellranger count" with mm10 transcript reference per sample, and all sample data was aggregated with "cellranger aggr". In this process, sequence depth was normalized by subsampling reads from higher-depth data. The aggregated gene expression matrix contained 32,172 cells in total. Next, the gene expression matrix was processed with R package, Scater (1.6.3) (15). Before Scater normalization, we removed outlier cells with an extremely large or small number of UMI (UMI > 20000 or UMI < 500) or large mitochondria UMI ratio (> 0.2). Almost all cells (99.16%) passed these quality checks. We also removed genes with low average expression (mean UMI count < 0.01). Cell cycling scores were calculated for each cell using the cyclone algorithm in the R package Scran 1.6.9 (16). The gene expression matrix was then normalized with Scater that includes size-normalization and log transformation. Normalized gene expression matrix data was imported to R package Seurat 2.3.4 (17) without further normalization and was directly subjected to scaling and undesired variation removal wherein process data variations correlated with UMI count and percent of mitochondria were regressed out. All of the above analysis was performed in our high throughput computing facility in the Center for Genome Sciences and Systems Biology, Washington University in St. Louis.

Additional data analysis for clustering, dimensional reduction, marker identification, and visualization was performed with Seurat. For dimensional reduction, principal component analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) were performed. Genes with high variable expression were selected and used for PCA. The first 15 PCs were selected as important PCs based on the distribution of standard deviation for each PC. The 15 PCs were used for clustering and t-SNE. A shared nearest neighbor (SNN) modularity optimization-based clustering algorithm was used for clustering. The clusters were manually annotated by marker gene expression. Marker genes for each cluster were identified with the FindAllMarkers function in Seurat.

Cell trajectory analysis was performed using partition-based graph abstraction (PAGA) for a pseudo-temporal ordering algorithm and RNA-velocity analysis of single cells based on spliced versus unspliced mRNA abundance. The partition-based graph abstraction (PAGA) graph was constructed with the wild-type single cell dataset as described previously (18). Calculations were performed with Scanpy's implementation. A series of data preprocessing for PAGA graph construction was performed according to Scanpy's tutorial for PAGA analysis (https://scanpy-tutorials.readthedocs.io/en/latest/paga-paul15.html). The force-directed graph was constructed and used for embedding the RNA velocity trajectory graph. Normalized gene expression matrix was imported into a Python package, Scanpy 1.4.0 (19). Force-directed graph (20) was constructed according to tutorials from Scanpy (https://scanpy.readthedocs.io/en/stable/tutorials.html.). RNA velocity was analyzed with Velocyto.py (version 0.17.17) (21) using the author instructions (http://velocyto.org/velocyto.py/). The files generated with the 10x Genomics Cell Ranger pipeline were used for RNA velocity analysis. First, Velocyto command line API was used to generate gene expression matrix of spliced UMI counts and unspliced UMI counts. The output file was saved as a loom file per sample. All loom files were merged into a single file that was imported into Velocyto Python API to create the Velocyto object that was integrated with Force-directed graph and cluster label and then subjected to quality check and then filtered by the mRNA detection level (min_expr_counts=40, min_cells_express=30). After variable gene detection by a velocyto function, the UMI count of spliced and unspliced read were normalized by total molecule count. The velocyto object was subjected to a series of final data processing process; PCA, k-nn based imputation, velocity estimation, and shift calculation, and the vectors estimated by RNA velocity were projected on the Force-directed graph.

scATACseq analysis. For data preprocessing for scATACseq analysis, NGS data was processed with the Cell Ranger ATAC pipelines (Cell Ranger ATAC version 1.0.0 from 10x Genomics) to generate bam files. Bam files were subjected to the preprocessing pipeline with Snap tools (version 1.2.6). Snap files were created for each sample condition according to the developer's instructions as described at: https://support.10xgenomics.com/single-cell-atac/software/pipelines/latest/what-is-cell-ranger-atac and https://github.com/r3fang/SnapATAC/wiki/FAQs#10X_snap, respectively). Downstream analysis was performed with SnapATAC (22) according to the instructions at https://github.com/r3fang/SnapATAC/blob/master/examples/10X_brain_5k/README.md). First, WT and Il33ko data were merged into a single data file. Outlier cells were removed if they had less than 2000 UMI. In addition, cells that had a large mitochondria UMI ratio (> 0.2) were removed. A total of 15988 cells passed these quality checks. Next, cell-bin matrix (CB matrix) was calculated at 5 kb resolution (bin.size = 5000). The matrix was converted into a binary matrix (Bmat). Bmat was further converted into cell by gene matrix (Gmat) for all genes. For data analysis and visualization for scATACseq data, dimensionally reduction embeddings for diffusion map and t-SNE were calculated with Bmat (parameter: num.eigs=50). The cluster was identified by knn based Louvain clustering method (parameters: eigs.dims=1:20, k=15). For the cluster annotation, we used scRNAseq data. The Snap object was converted into the Seurat object and subjected to FindTransferAnchors, which is a data integration algorithm in Seurat. The inferred cluster name was manually modified by checking marker gene expression. Data visualization was performed with Seurat. For cell cycle phase calculation with scATACseq data, cell cycling scores were calculated for each cell with Scran (version 1.6.9). Instead of gene expression matrix, gene activity scores, which were stored as Gmat, was used cell cycling score calculation. For peak calling and differential peak identification, peak calling was done for all clusters with more than 100 cells with SnapATAC (22) and MACS2 (parameters: --nomodel --shift 100 --ext 200 --qval 5e-2 -B –SPMR). Peak information was saved as bed file per cluster and experimental condition. To find differentially accessible regions (DARs) between WT and Il33$^{-/-}$ samples, we used findDAR function in SnapATAC per each cluster (parameter: bcv=0.1). We analyzed DAR with p-value less than 0.05.

In addition, Integrative Genomics Viewer (IGV) analysis was used to visualize each track for peak calling. For this analysis, we created tdf files containing normalized coverage data for each cluster in WT and *Il33*$^{-/-}$ samples using igvtools (version 2.6.3) as described previously (23) with a user guide at: http://software.broadinstitute.org/software/igv/book/export/html/6. This normalization option multiplies each value by 1,000,000 / totalReadCount. Clusters with cell numbers less than 100 were filtered out, resulting in 11 tdf files for WT and 10 tdf files for *Il33*$^{-/-}$ samples. These tdf files were then loaded into the IGV to analyze differentially accessible peaks in the mouse *Il33* gene (preferences: normalize coverage data; genomic sequence: mm10 mouse). The four largest peaks (peak1: chr19:29,940,023-29,941,697; peak2: chr19:29,942,085-29,943,373; peak3: chr19:29,945,522-29,947,192; peak4: chr19:29,947,506-29,949,400) within the mouse *Il33* gene sequence were input to Toolkit for Cistrome Data Browser (Species: Mouse mm10; Data type in Cistrome: Transcription factor, chromatin regulator; Interval: the corresponding locus of ROI) to identify transcription factor binding sites as described previously (24, 25). BigWig format files (26) of the selected transcription factors were copied into IGV using WashU Browser (27) to match binding sites to *Il33* reference sequences.

Lung function testing. To assess lung function, we assessed blood oxygen saturation (SpO_2) and airway reactivity to methacholine (MCh). For SpO_2 monitoring, we used a MouseSTAT® pulse oximeter (Kent Scientific, Torrington, CT) applied to the paw skin under isoflurane anesthesia. To decrease variability in readings for this procedure, anesthetized mice were placed on the RightTemp warming pad, and the SpO_2 reading was recorded when the Oxiwave signal was stable. For airway reactivity, we used our protocol for MCh nebulization to anesthetized mice and measurements for R_{RS} using the FinePointe Resistance and Compliance system (Buxco Research Systems, Wilmington, NC) as described previously (9).

References

1. Wu, K., Wang, X., Keeler, S.P., Gerovac, B.J., Agapov, E., Byers, D.E., Gilfillan, S., Colonna, M., Zhang, Y., and Holtzman, M.J. 2020. Group 2 innate lymphoid cells must partner with the myeloid-macrophage lineage for long-term postviral lung disease. J Immunol 205:1084-1101.
2. Townsend, M.J., Fallon, P.G., Matthews, D.J., Jolin, H.E., and McKenzie, A.N. 2000. T1/ST2-deficient mice demonstrate the importance of T1/ST2 in developing primary T helper cell type 2 responses. J. Exp. Med. 191:1069-1076.

3. Hsu, C.L., Neilsen, C.V., and Bryce, P.J. 2010. IL-33 is produced by mast cells and regulates IgE-dependent inflammation. PloS ONE 5:e11944.
4. Lee, J., Kim, J., Gludish, D., Roach, R.R., Saunders, A.H., Barrios, J., Woo, A.J., Chen, H., Conner, D.A., Fujiwara, Y., Stripp, B.R., and Kim, C.F. 2013. Surfactant protein-C chromatin-bound green fluorescence protein reporter mice reveal heterogeneity of surfactant protein C-expressing lung cells. Am. J. Respir. Cell Mol. Biol. 48:288-298.
5. Bruen, K.J., Campbell, C.A., Schooler, W.G., deSerres, S., Cairns, B.A., Hultman, C.S., Meyer, A.A., and Randell, S.H. 2004. Real-time monitoring of keratin 5 expression during burn re-epithelialization. J. Surg. Res. 120:12-20.
6. Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. 2011. Distinct stem cells contribute to mammary gland development and maintenance. Nature 479:189-193.
7. Voehringer, D., Liang, H.-E., and Locksley, R.M. 2008. Homeostasis and effector function of lymphopenia-induced “memory-like” T cells in constitutively T cell-depleted mice. J. Immunol. 180:4742-4753.
8. Wang, X., Wu, K., Keeler, S.P., Mao, D., Agapov, E.V., Zhang, Y., and Holtzman, M.J. 2021. TLR3-activated monocyte-derived dendritic cells trigger progression from acute viral infection to chronic disease in the lung. J. Immunol. 206:1297-1314 (selected for Top-Reads p. 1115).
9. Keeler, S.P., Agapov, E.V., Hinojosa, M.E., Letvin, A.N., Wu, K., and Holtzman, M.J. 2018. Influenza A virus infection causes chronic lung disease linked to sites of active viral RNA remnants. J. Immunol. 201:2354-2368.
10. Vaughan, A.E., Brumwell, A.N., Xi, Y., Gotts, J.E., Brownfield, D.G., Treutlein, B., Tan, K., Tan, V., Liu, F.C., Looney, M.R., Matthay, M.A., Rock, J.R., and Chapman, H.A. 2015. Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature 517:621-625.
11. van Nunen, M.C.J., and van der Veen, J. 1967. Experimental infection with Sendai virus in mice. Arch Gesamte Virusforsch 22:388-397.
12. Zhang, Y., Mao, D., Keeler, S.P., Wang, X., Wu, K., Gerovac, B.J., Shornick, L.P., Agapov, E., and Holtzman, M.J. 2019. Respiratory enterovirus (like parainfluenza virus) can cause chronic lung disease if protection by airway epithelial STAT1 is lost. J. Immunol. 202:2332-2347.
13. Zhang, Y., Mao, D., Roswit, W.T., Jin, X., Patel, A.C., Patel, D.A., Agapov, E., Wang, Z., Tidwell, R.M., Atkinson, J.J., Huang, G., McCarthy, R., Yu, J., Yun, N.E., Paessler, S.L., Lawson, T.G., Omattage, N.S., Brett, T.J., and Holtzman, M.J. 2015. PARP9-DTX3L ubiquitin ligase targets host histone H2BJ and viral 3C protease to enhance interferon signaling and control viral infection. Nat. Immunol. 16:1215-1227.
14. Livak, K.J., and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.
15. McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. 2017. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33:1179-1186.
16. Lun, A.T.L., McCarthy, D.J., and Marioni, J.C. 2019. a step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor [version 2; peer review: 3 approved, 2 approved with reservations]. F1000Research 5:2122.
17. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. 2015. Spatial reconstruction of single-cell genome expression data. Nat. Biotechnol. 33:495-502.
18. Wolf, F.A., Hamey, F., Plass, M., Solana, J., Dahlin, J.S., Gottgens, B., Rajewsky, N., Simon, L., and Theis, F.J. 2018. Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. bioRxiv:1-39.
19. Wolf, F.A., Angerer, P., and Theis, F.J. 2018. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19:15.
20. Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. 2014. ForceAtlas2, a continuous graph layout algorithm for handy netowrk visualization designed for the Gephi software. PloS ONE 9:e98679.
21. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A., Fan, J., Borm, L.E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundstrom, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., and Kharchenko, P.V. 2018. RNA velocity of single cells. Nature 560:494-498.
22. Fang, R., Preissl, S., Li, Y., Hou, X., Lucero, J., Wang, X., Motamedi, A., Shiau, A.K., Zhou, X., Xie, F., Mukamel, E.A., Zhang, K.X., Zhang, Y., Behrens, M.M., Ecker, J.R., and Ren, B. 2020. SnapATAC: a comprehensive analysis package for single cell ATAC-seq. bioRxiv.
23. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. 2011. Integrative genomics viewer. Nat Biotechnol 29:24-26.
24. Mei, S., Qin, Q., Wu, Q., Sun, H., Zheng, R., Zhang, C., Zhu, M., Wu, J., Shi, X., Taing, L., Liu, T., Brown, M., Meyer, C.A., and Liu, X.S. 2017. Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessiblity data in human and mouse. Nucl. Acids Res. 45:D658-D662.

25. Zheng, R., Wan, C., Mei, S., Qin, Q., Wu, Q., Sun, H., Chen, C.-H., Brown, M., Zhang, X., Meyer, C.A., and Liu, X.S. 2018. Cistrome Data Browser: expanded datases and new tools for gene regulatory analysis. Nucl. Acids Res. 47:D729-D735.
26. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. 2010. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics 26:2204-2207.
27. Li, D., Hsu, S., Purushotham, D., Sears, R.L., and Wang, T. 2019. WashU Epigenome Browser update 2019. Nucl. Acids Res. 47:W158-W165.

Supplemental Table 1. Antibodies for immunostaining.

Target Protein	Antibody Type	Vendor	Catalogue #
Aqp3	Rabbit pAb[1]	Abcam	ab125219
Sftpc	Rabbit pAb	Abcam	ab90716
Pdpn	Hamster mAb	Abcam	ab11936
GFP	Chicken pAb	Abcam	ab13970
Muc5ac	Mouse mAb (45M1), biotinylated	Thermo Scientific	MS-145-B
Muc5ac	Mouse mAb (45M1)	Thermo Scientifc	MS-145-P
mCherry	Mouse mAb	Abcam	ab125096
Krt5	Rabbit pAb	Abcam	ab53121
Scgba1	Goat pAb	Santa Cruz	Sc-9772
Scgba1	Mouse pAb	Santa Cruz	Sc-130411
Acetylated-α-tubulin	Mouse mAb (6-11B-1)	Sigma	T6793
Ki-67	Mouse mAb	BD Pharmingen	550609
Ki-67	Rabbit mAb	Cell Signaling	12202
Phospho-Cdk1	Rabbit mAb	Abcam	ab201008
Cdk1	Rabbit pAb	Abcam	ab234345
Active caspase-3	Rabbit mAb	Cell Signaling	9664
IL-33	Goat mAb	R&D Systems	AF3626
F4/80	Rabbit mAb	Cell Signaling	70076
Cxcl17	Sheep pAb	R&D Systems	AF4270

[1] Abbreviations: mAb, monoclonal antibody, pAb, polyclonal antibody.

Supplemental Table 2. Sequences of DNA primers and probes for real-time qPCR assays.

Target Gene		**ID/Sequence**[2]
Aqp3	F[1]	5'-AAGTTGATGGTGAGGAAGCC-3'
	R	5'-ATGCTTCACATCCGCTACC-3'
	P	5'-TCCTTGTGATGTTTGGCTGTGGC-3'
Sftpc	F	5'-ACCACAACCACGATGAGAAG-3'
	R	5'-GGACATGAGTAGCAAAGAGGT-3'
	P	5'-AGTTCCGCATCCCCTGCTGTC-3'
Krt5	F	5'-AGATCGCCACCTACAGGAA-3'
	R	5'-TCCGTAGCCAGAAGAGACA-3'
	P	5'-CCCACTCAGCCTGCACTCCTC-3'
Trp63	F	5'-GAGATGAGGAGGTGAGGAGAA-3'
	R	5'-ACAGACTGCAGCATTGTCAG-3'
	P	5'-AGATGGCATGTCGGAACTGTTCAGG-3'
Ngfr	F	5'-GCGTAGACCTTGTGATCCAT-3'
	R	5'-ACTCAGATGAAGCCAACCAC-3'
	P	5'-CCGAATGCGAGGAGATCCCTGG-3'
Ki-67	F	5'-TTCTTGGTGCATACAATGTCTTC-3'
	R	5'-AGTCTGATGTTAGGTGTTTGAGG-3'
	P	5'-ACCCTTGGCTTAGGTTCACTGTCC-3'
Il33	F	5'-TCATGTTCACCATCAGCTTCT-3'
	R	5'-GTGCTACTACGCTACTATGAGTC-3'
	P	5'-ACCGTCGCCTGATTGACTTGCA-3'
Arg1	F	5'-AGTGTTGATGTCAGTGTGAGC-3'
	R	5'-GAATGGAAGAGTCAGTGTGGT-3'
	P	5'-ACAGTCTGGCAGTTGGAAGCATCT-3'
Trem2	NA1	Mm00451744_m1
Il13	F	5'-GGTGCCAAGATCTGTGTCTC-3'
	R	5'-CCACACTCCATACCATGCTG-3'
	P	5'-AAGACCAGACTCCCCTGTGCAAC-3'
Muc5ac	F	5'-TACCACTCCCTGCTTCTGCAGCGTGTCA-3'
	R	5'-ATAGTAACAGTGGCCATCAAGGTCTGTCT-3'
	P	5'-TATACCCCTTGGGATCCATCATCTACA-3'
Cxcl17	F	5'-CCTTCTGTTGCTTCCAGTGA-3'
	R	5'-TTCCAAGAGCCACCTCCTA-3'
	P	5'-TCATGTCCATGGTCTTCAGCAGCC-3'
Lactoferrin	F	5'-CCCAAGACCACAGACATGAG-3'
	R	5'-CAGAATTTGACACAGCACACC-3'
	P	5'-TTAGCTAGACAGAGTCCAAGGGCCT-3'
Serpinb2	F	5'-GAGCATCTCATCAACATTGGC-3'
	R	5'-GTTCTCTGGGTTTCTTGTGGT-3'
	P	5'-CATCTGCTGTTCAGTGTTACCCCCA-3'
Gapdh	F	5'-GTGGAGTCATACTGGAACATGTAG-3'
	R	5'-AATGGTGAAGGTCGGTGTG-3'
	P	5'-TGCAAATGGCAGCCCTGGTG-3'
SeV-NP	F	5'-GGCGGTGGTGCAATTGAG-3'
	R	5'-CATGAGCTTCTGTTTCTAGGTCGAT-3'
	P	5'-AGCTCTAGACAATGCC-3'

[1] Abbreviations: F, forward primer; R, reverse primer; P, MGB probe.

[2] Primers and probes for all the genes were obtained from IDT, except for Trem2, which were from ABI that does not make sequences available.