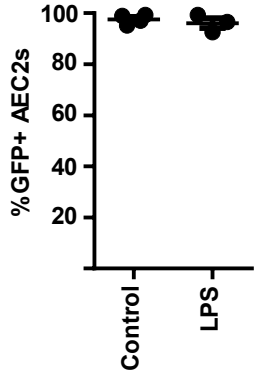
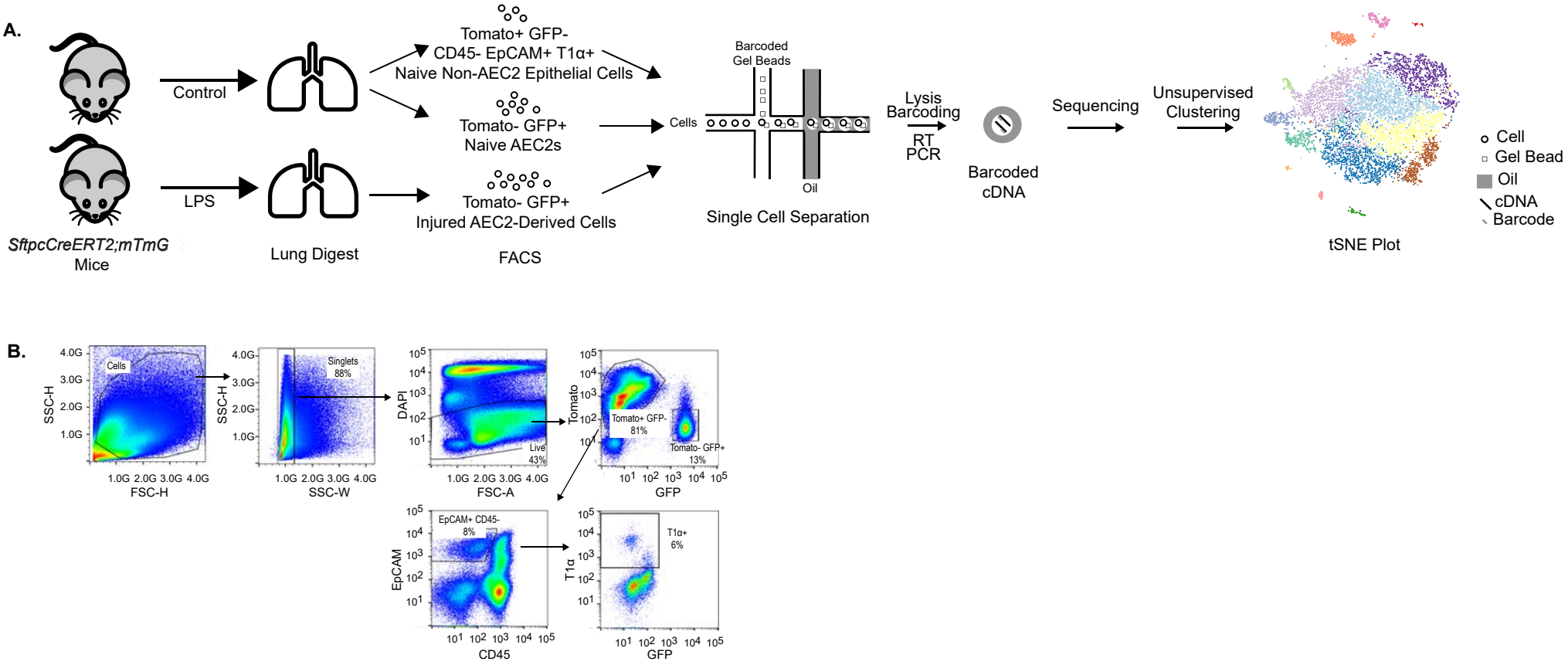
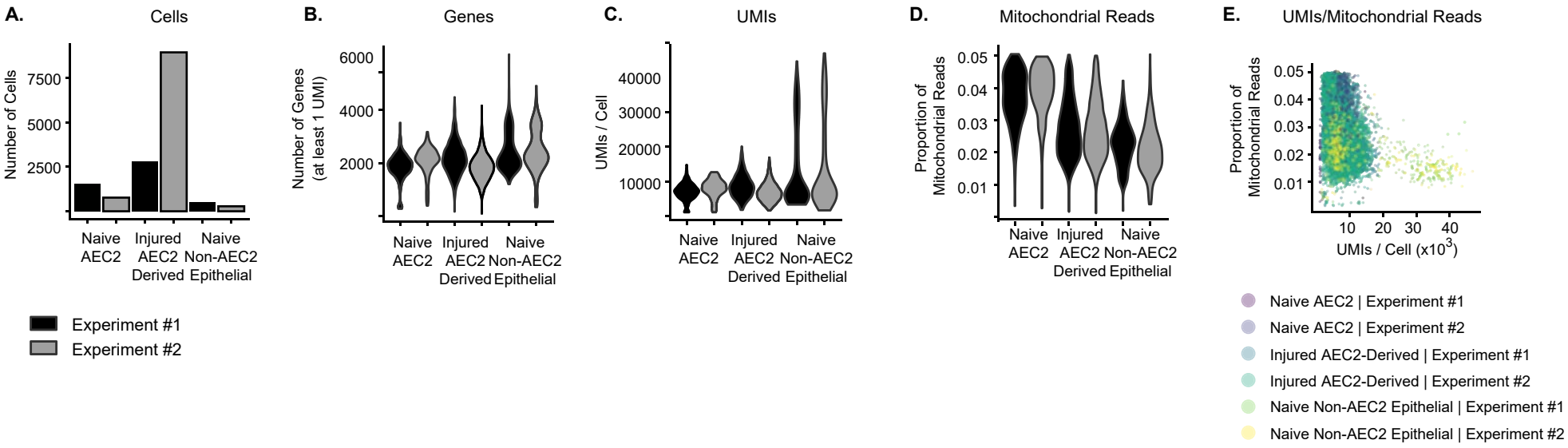


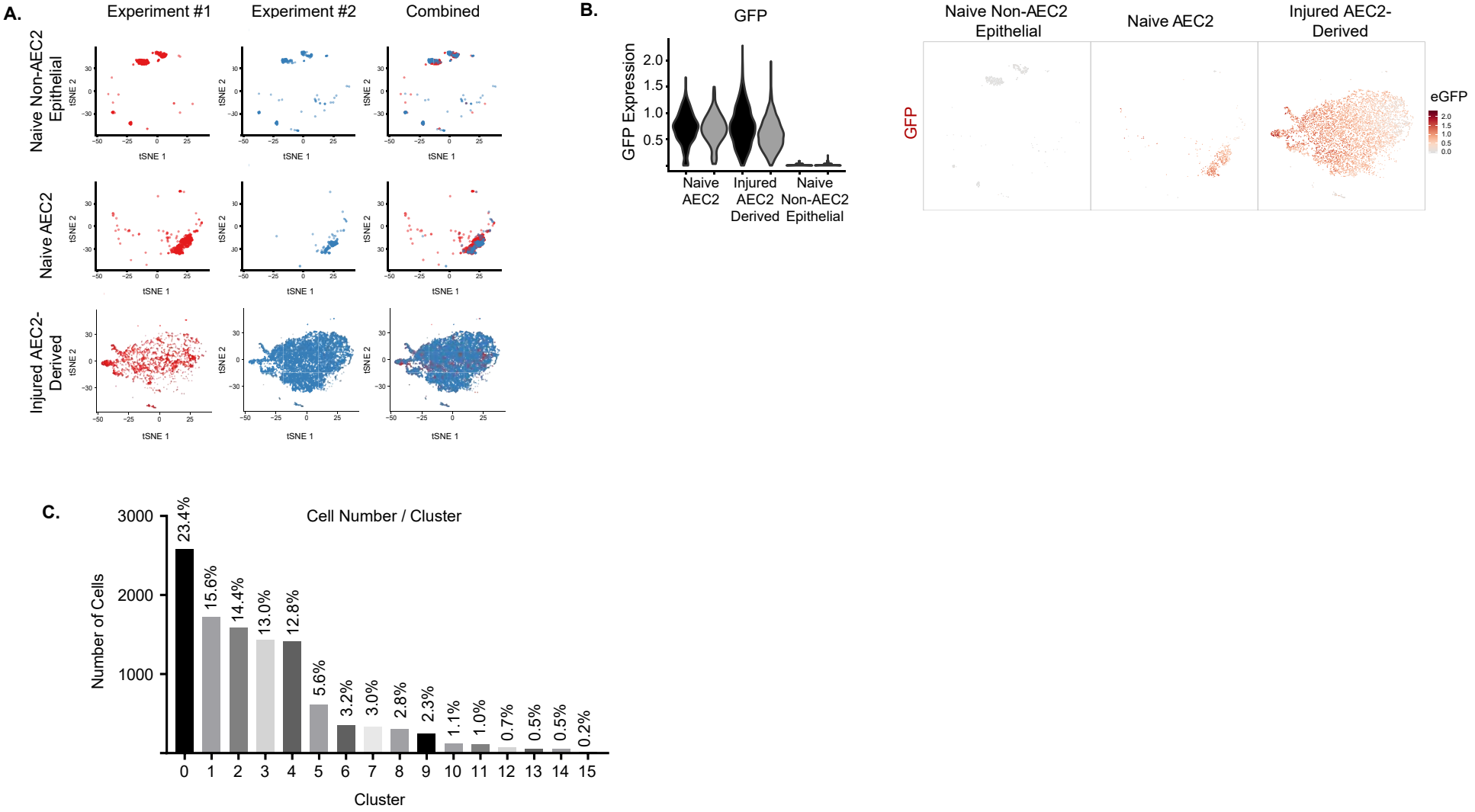
Supplementary Figure 1. AEC2s are the Principal Progenitor of AEC2s During Epithelial Regeneration in the LPS Model of Lung Injury. *SftpcCreERT2;mTmG* mice were treated with or without LPS and harvested at Day 27. The percent of proSPC+ cells (AEC2s) that were GFP+, as determined by immunostaining of lung sections, is shown. Mean ± SEM indicated.



Supplementary Figure 2. scRNAseq of Epithelial Cells Isolated from Naïve and LPS-injured mice. A,B) *SftpcCreERT2;mTmG* mice were treated with or without LPS. At Day 7, lungs were digested, stained for EpCAM, CD45, and T1α, and subjected to fluorescence-activated cell sorting (FACS). Tomato+ GFP- CD45- EpCAM+ T1α+ (Non-AEC2 Epithelial) and Tomato- GFP+ (AEC2) cells from naïve mice and Tomato- GFP+ (Injured AEC2-Derived) cells from LPS-treated mice were loaded into a 10X Genomics Chromium Single Cell 3' Solution. Cells and gel beads containing barcoded oligonucleotides and reagents were encapsulated in droplets. Cells were lysed, and mRNA was reverse transcribed to cDNA, which introduced cell barcodes and UMI sequences. cDNA was amplified by PCR and sequenced, and cells were subjected to unsupervised clustering based on their transcriptomes. B) Flow cytometry gating strategy. The experiment with one mouse per treatment group was performed on two separate occasions, for a total of 2 mice/group.

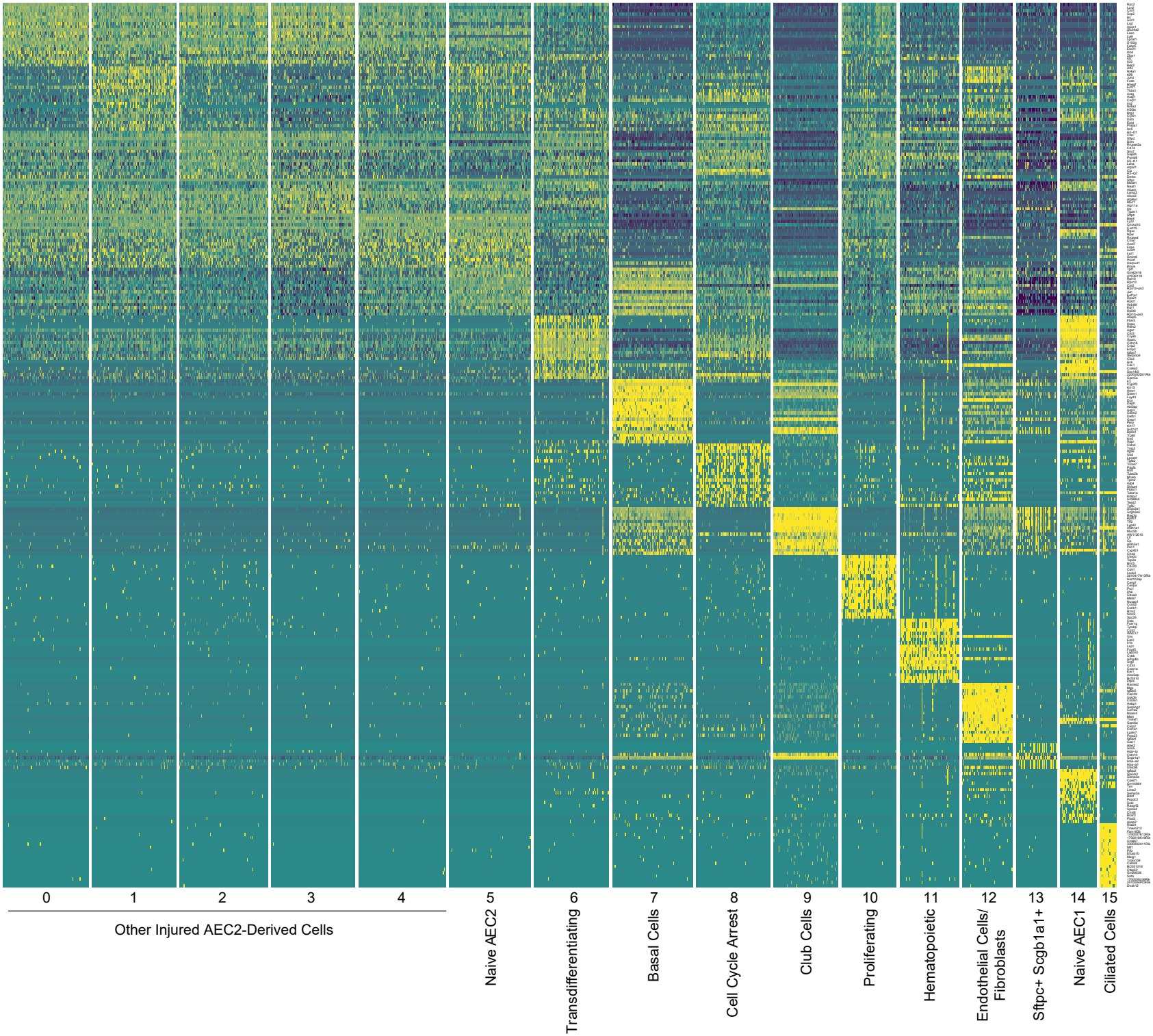


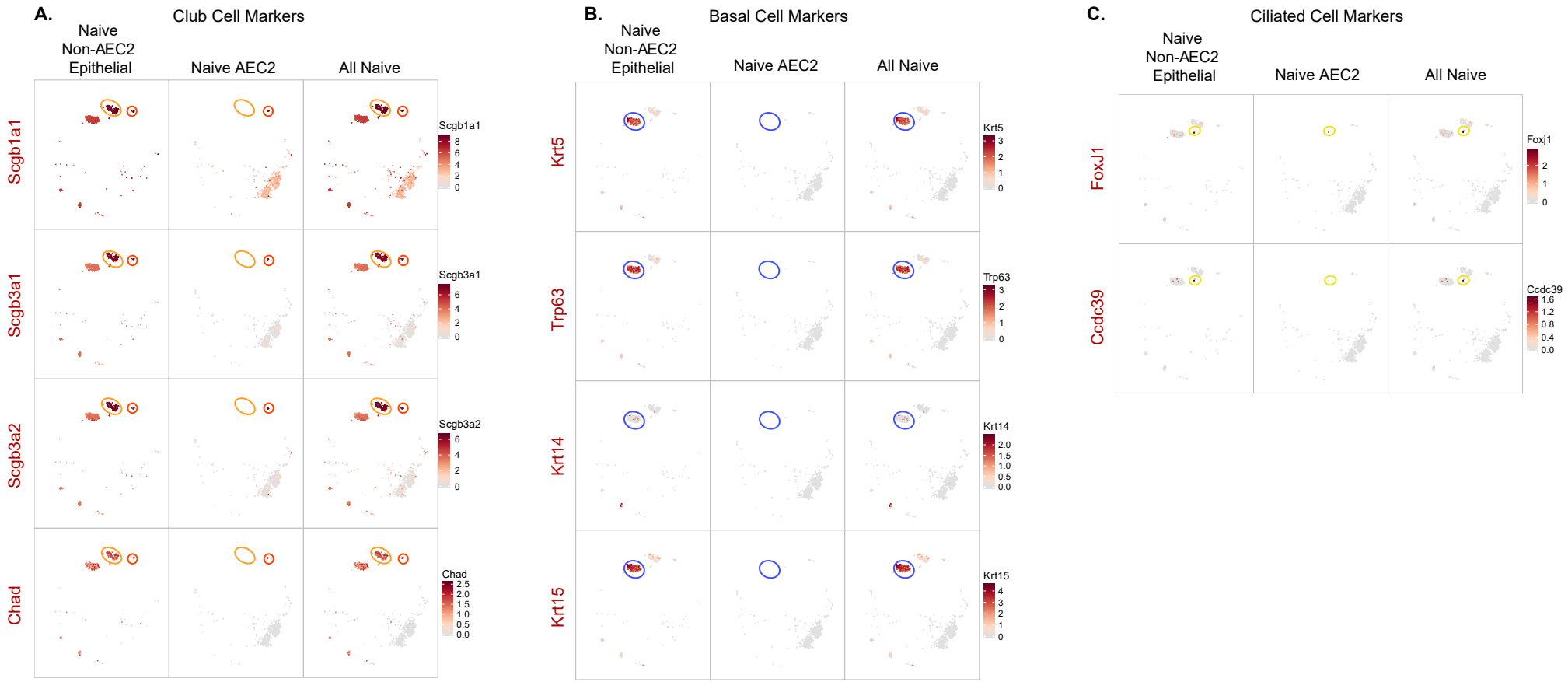
Supplementary Figure 3. scRNAseq of Epithelial Cells Isolated from Naïve and LPS-injured mice. The number of cells sequenced (A) and number of genes (B), transcripts (UMIs) (C,E), and mitochondrial reads (D,E) per cell after quality filtering. n=2 mice/group.



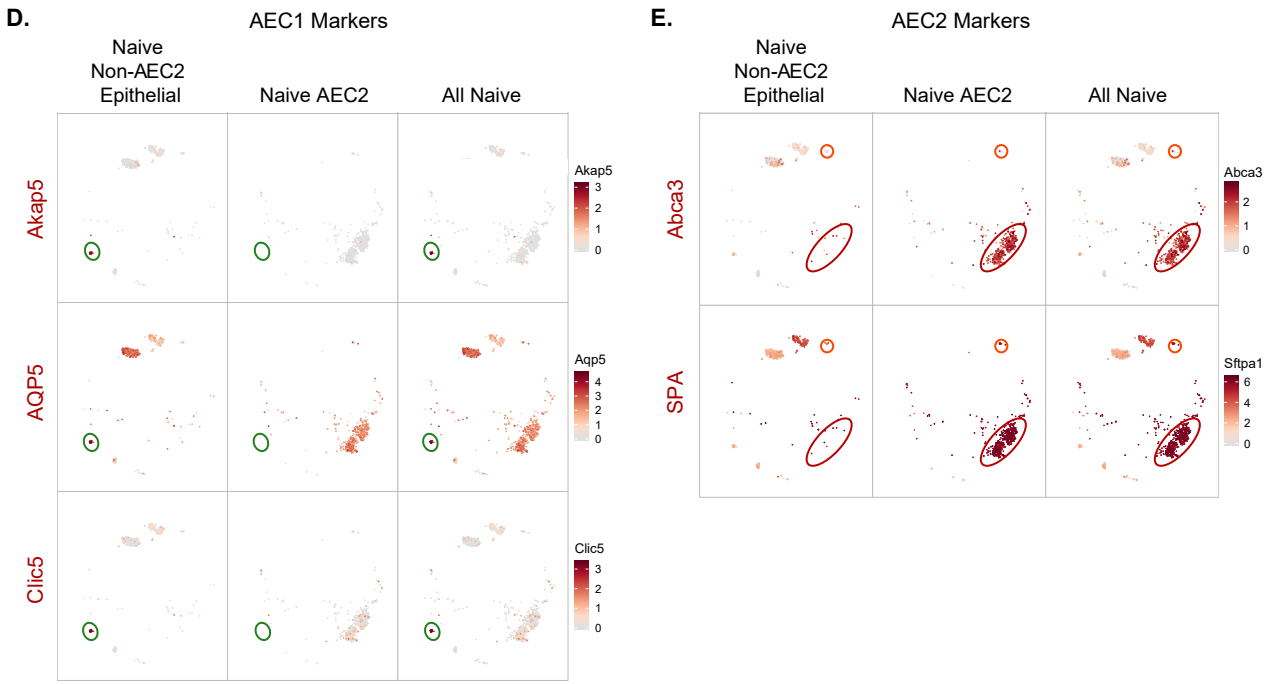
Supplementary Figure 4. Unsupervised Clustering. A) Locations within the tSNE plot of Naïve Non-AEC2 Epithelial, Naïve AEC2, and Injured AEC2-Derived cells from both experiments. B) Expression of GFP (natural log of normalized counts). C) Number and percent of cells in each cluster. D) Heatmap showing the 20 most differentially expressed genes in each cluster compared to all other clusters ranked in order of Bonferroni-corrected p-value. n=2 mice/group.

D.

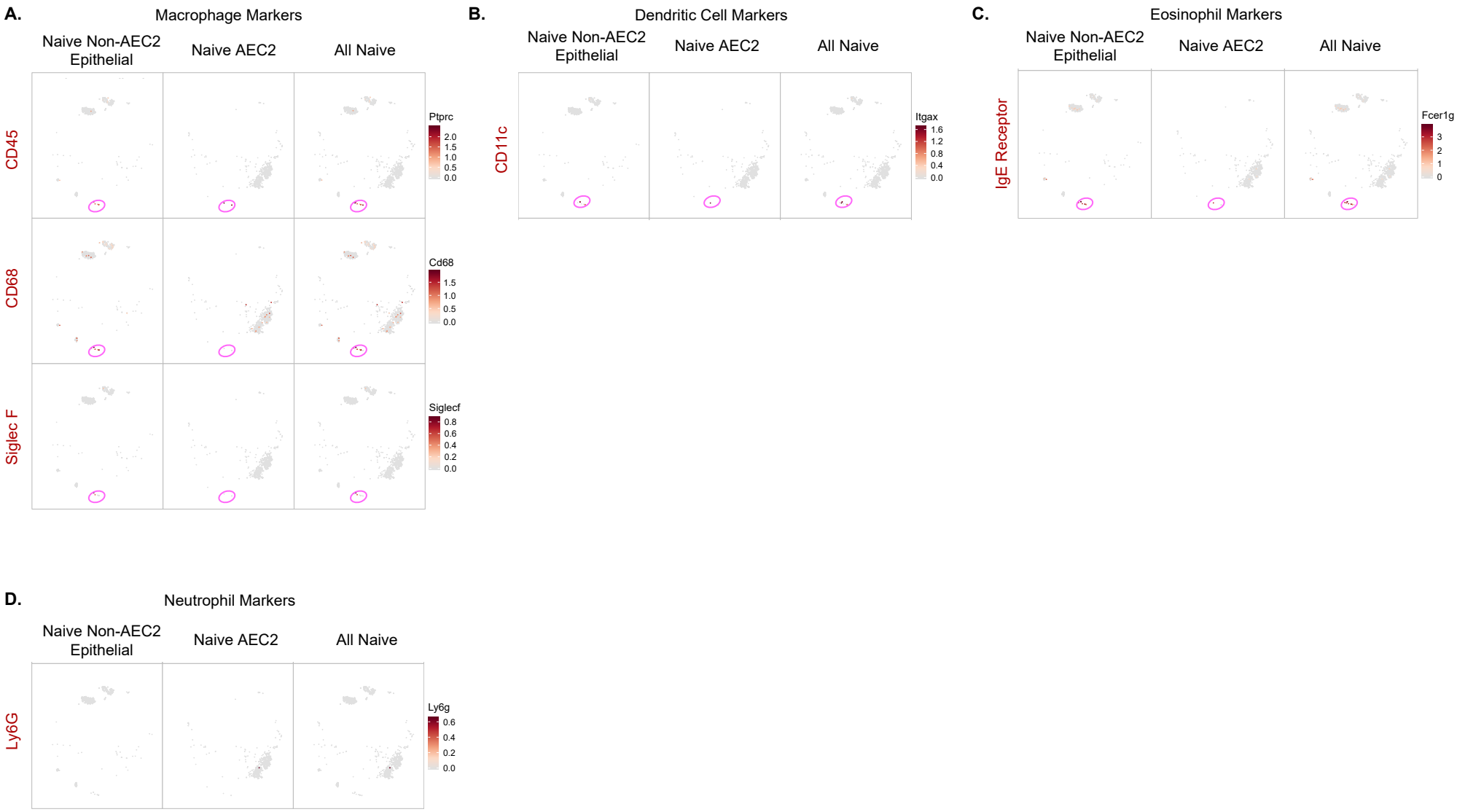

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Other Injured AEC2-Derived Cells
Naive AEC2
Transdifferentiating
Basal Cells
Cell Cycle Arrest
Club Cells
Proliferating
Hematopoietic
Endothelial Cells/ Fibroblasts
Sftpc+ Scgb1a1+
Naive AEC1
Ciliated Cells



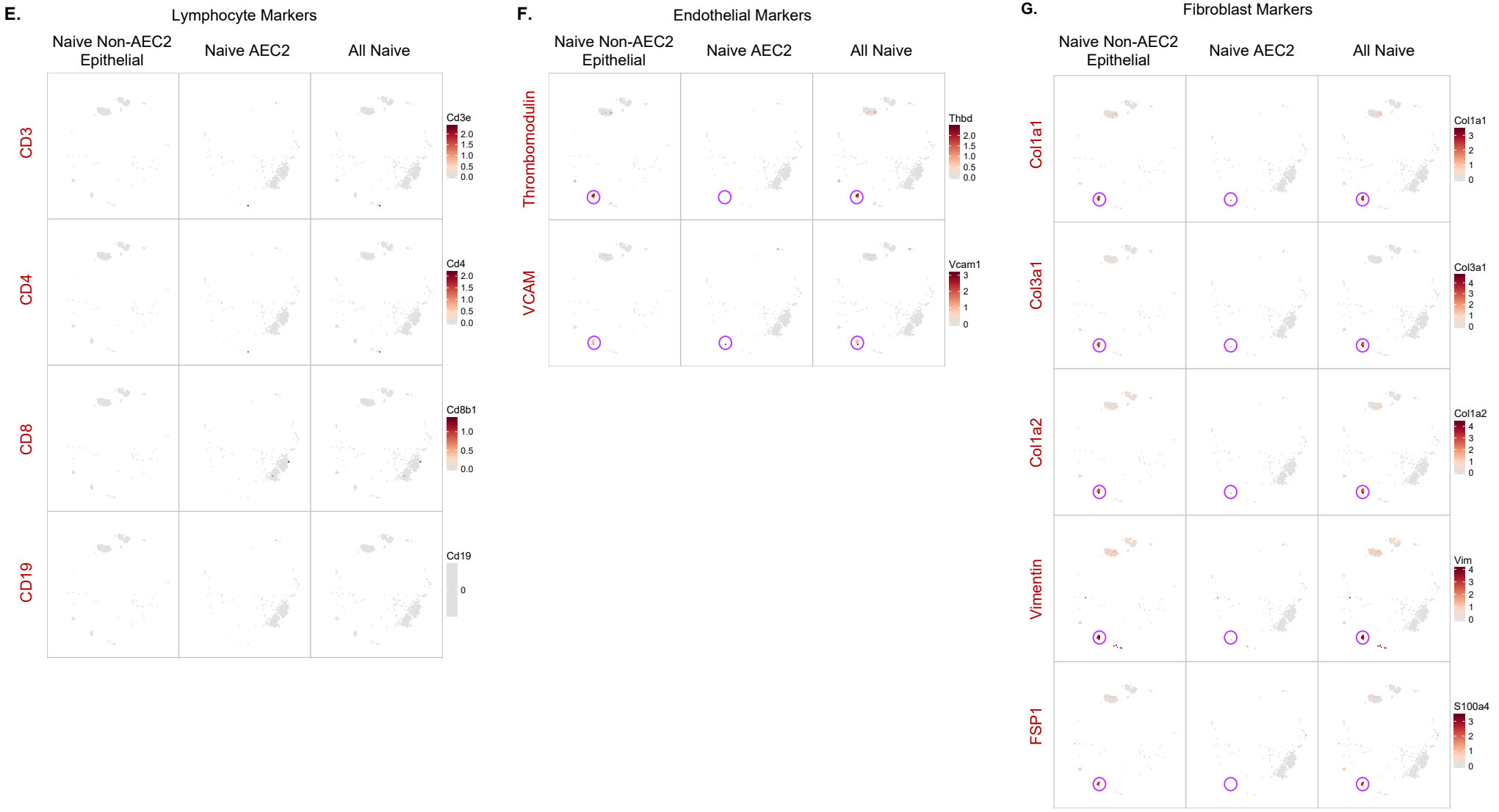
Supplementary Figure 5. scRNAseq Reveals Naïve Lung Epithelial Cell Types. Gene expression (natural log of normalized counts) of canonical Club Cell (A), Basal Cell (B), Ciliated Cell (C), AEC1 (D), and AEC2 (E) markers. Dark Orange Circles (Cluster 13): Scgb1a1+ Sftpc+ Cells, Light Orange Circles (Cluster 9): Club Cells, Yellow Circles (Cluster 15): Ciliated Cells, Blue Circles (Cluster 7): Basal Cells, Green Circles (Cluster 14): AEC1s, Red Circles (Cluster 5): AEC2s. n=2 mice/group.

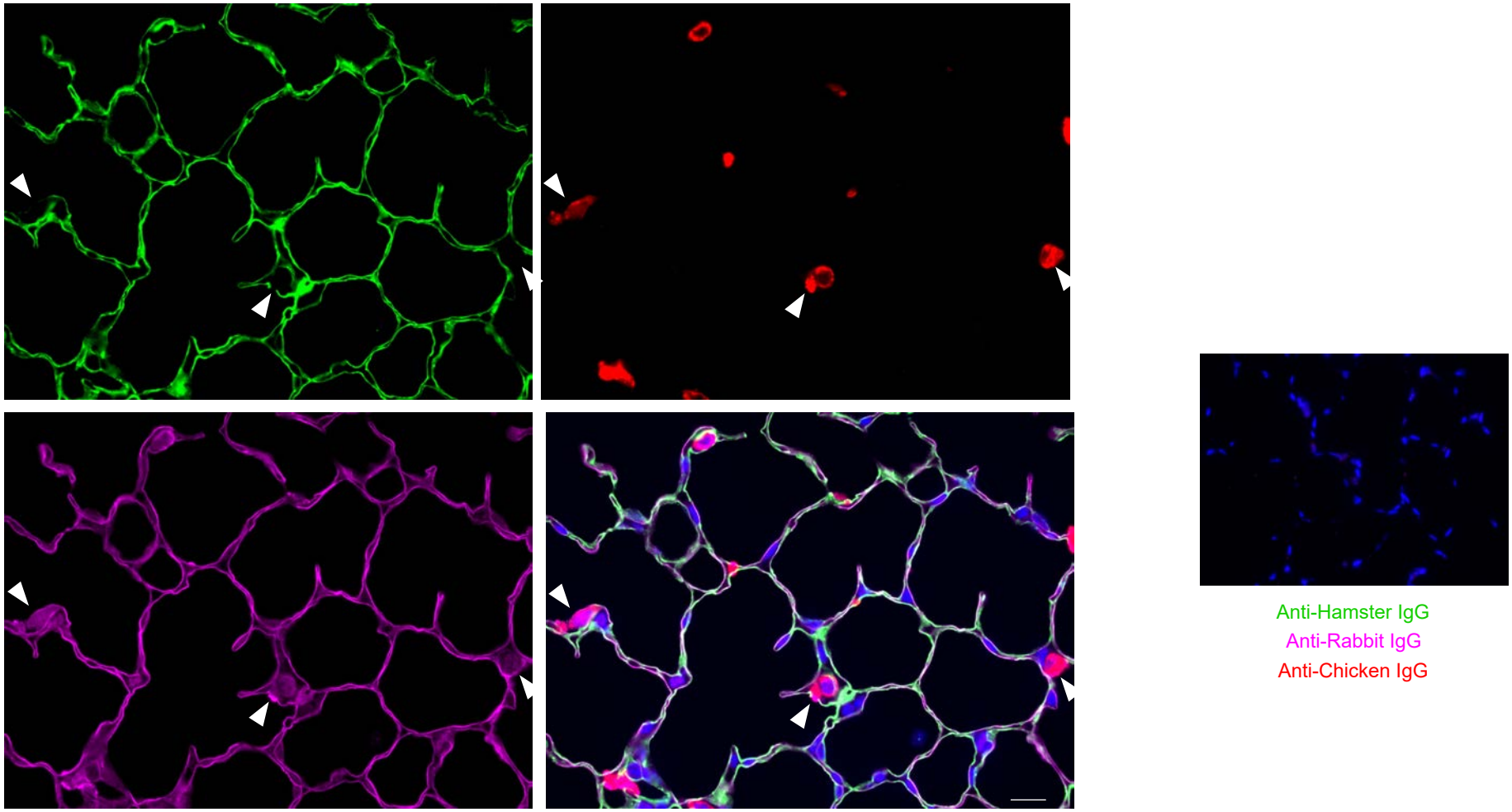

D.
AEC1 Markers
Naive Non-AEC2 Epithelial
Naive AEC2
All Naive
Akap5
AQP5
Clic5
Akap5
3
2
1
0
Aqp5
4
3
2
1
0
Clic5
3
2
1
0

E.
AEC2 Markers
Naive Non-AEC2 Epithelial
Naive AEC2
All Naive
Abca3
SPA
Abca3
2
1
0
Sftpa1
6
4
2
0

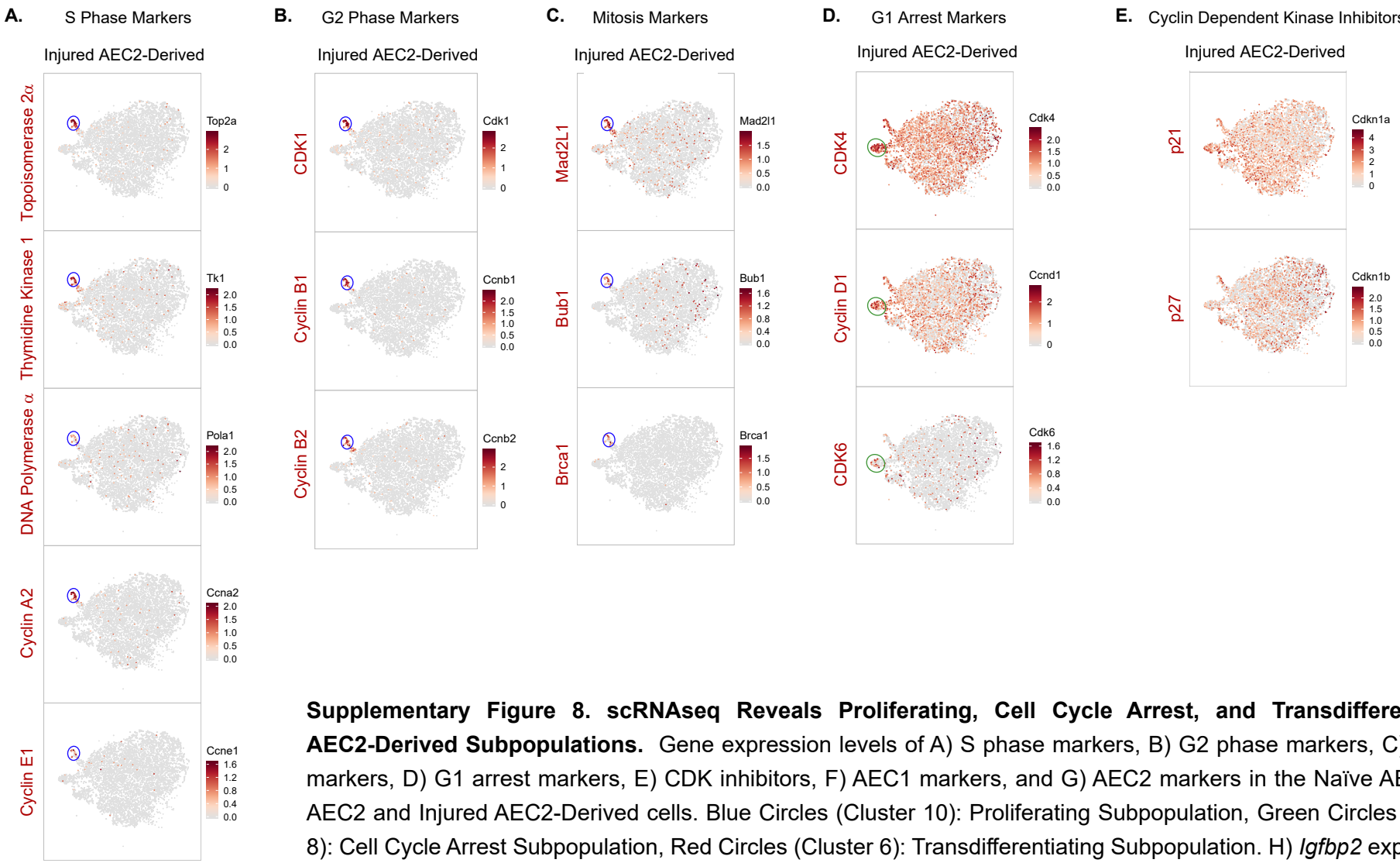


Supplementary Figure 6. scRNAseq Reveals Rare Contaminating Macrophages, Dendritic Cells, Eosinophils, Endothelial Cells, and Fibroblasts. Gene expression (natural log of normalized counts) of macrophage (A), dendritic cell (B), eosinophil (C), neutrophil (D), lymphocyte (E), endothelial cell (F), and fibroblast (G) markers. Pink Circles (Cluster 11): Hematopoietic Cells, Purple Circles (Cluster 12): Endothelial Cells/Fibroblasts.

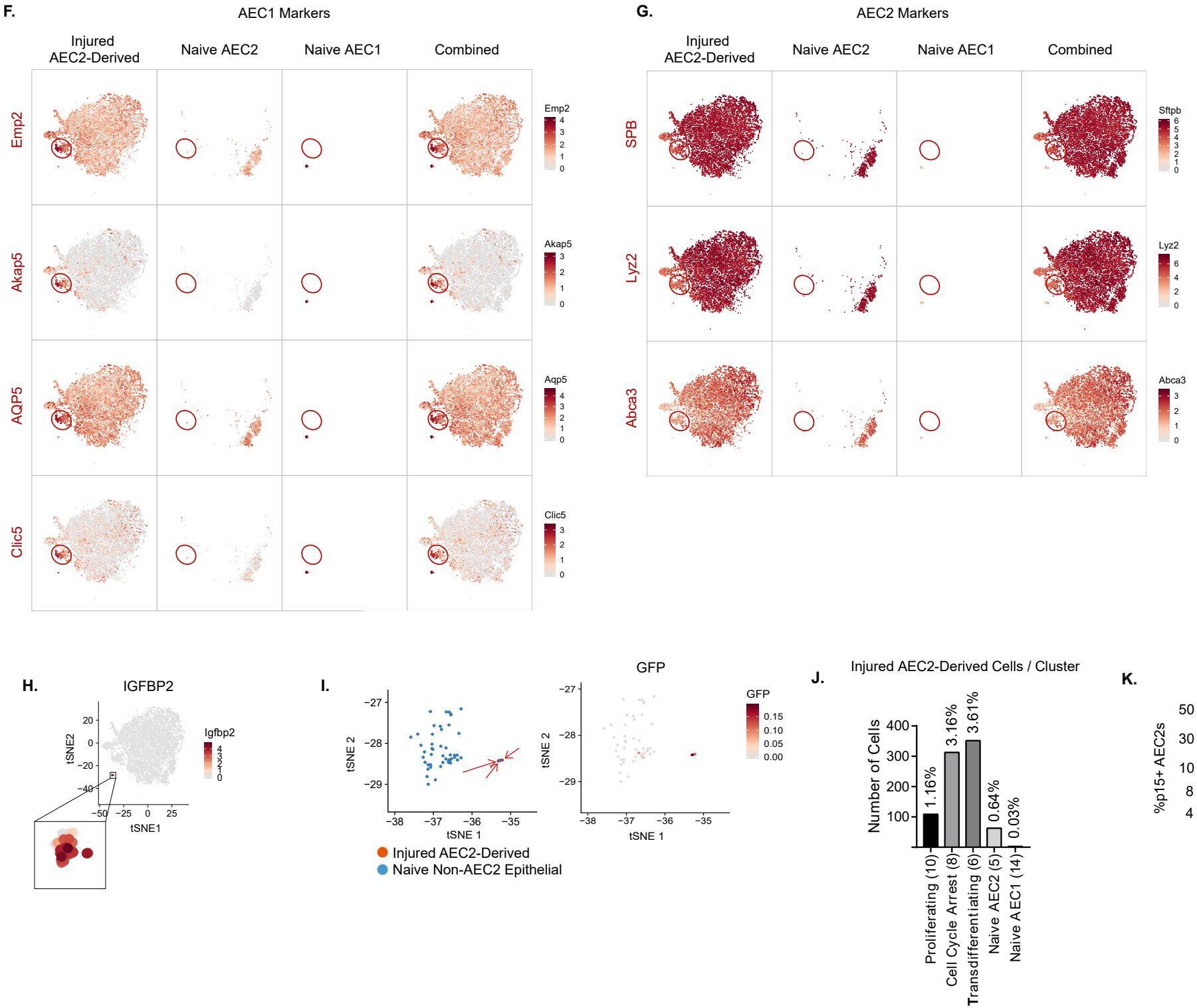

E.
Lymphocyte Markers
Naive Non-AEC2 Epithelial
Naive AEC2
All Naive
CD3
CD4
CD8
CD19
Cd3e
Cd4
Cd8b1
Cd19
F.
Endothelial Markers
Thrombomodulin
VCAM
Thbd
Vcam1
G.
Fibroblast Markers
Col1a1
Col3a1
Col1a2
Vimentin
FSP1
Vim
S100a4



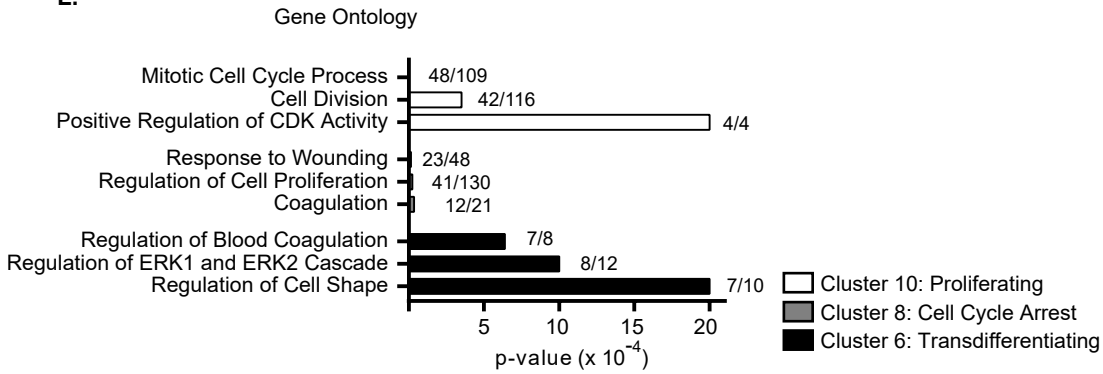
Supplementary Figure 7. AQP5 is Expressed by AEC2s. Immunostaining of lung sections from naïve *SftpcCreERT2;mTmG* mice for GFP, T1α, and AQP5. Arrowheads indicate AQP5 staining in AEC2s, particularly at the apical surface, as indicated by breaks in the T1α (AEC1) staining. Scale bar = 10 μm.

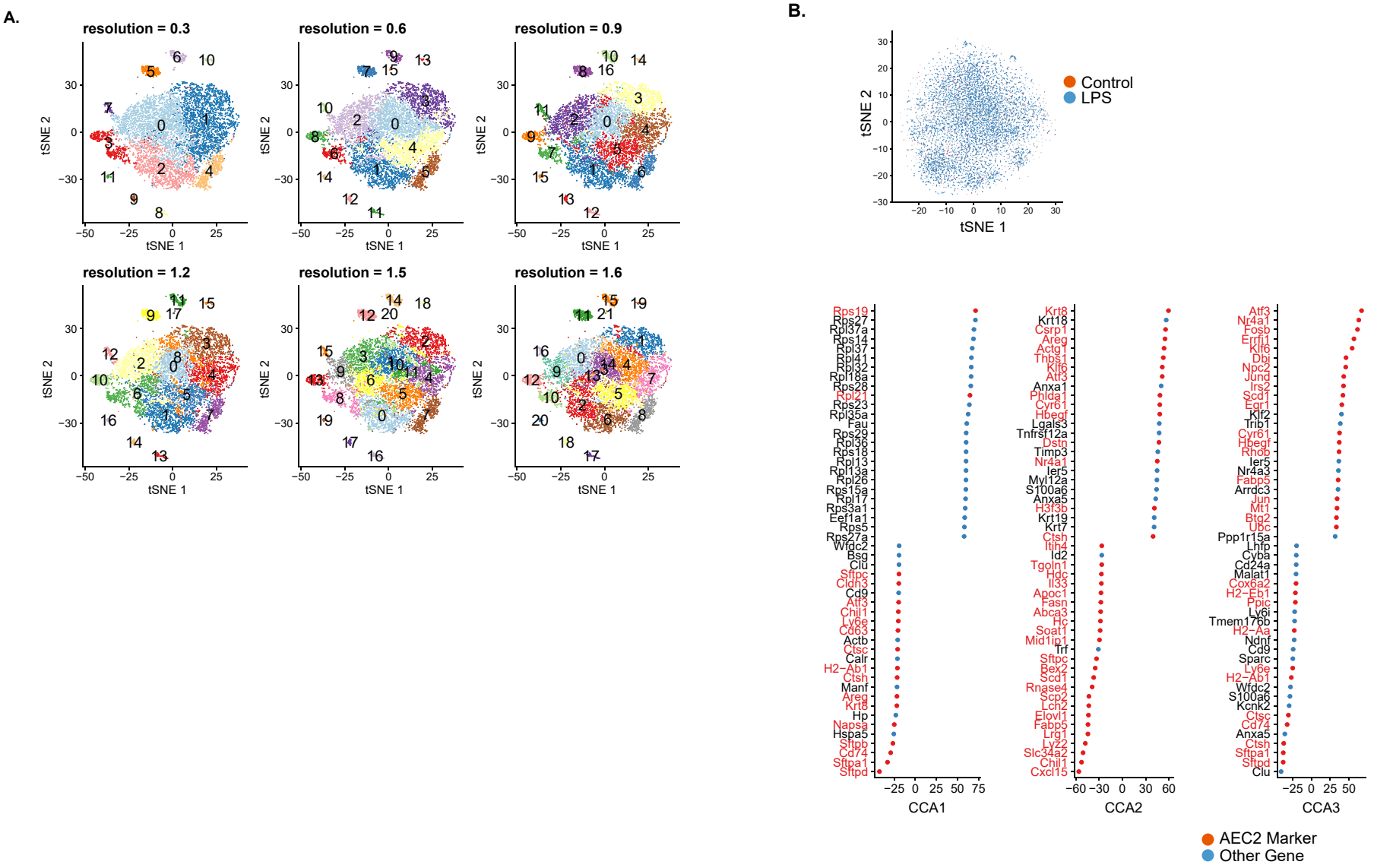


Supplementary Figure 8. scRNAseq Reveals Proliferating, Cell Cycle Arrest, and Transdifferentiating AEC2-Derived Subpopulations. Gene expression levels of A) S phase markers, B) G2 phase markers, C) Mitosis markers, D) G1 arrest markers, E) CDK inhibitors, F) AEC1 markers, and G) AEC2 markers in the Naïve AEC1 and AEC2 and Injured AEC2-Derived cells. Blue Circles (Cluster 10): Proliferating Subpopulation, Green Circles (Cluster 8): Cell Cycle Arrest Subpopulation, Red Circles (Cluster 6): Transdifferentiating Subpopulation. H) *Igfbp2* expression. A-H) Gene expression shown in natural log of normalized counts. I) Enlarged image of mature AEC1 cluster (Cluster 14) with cells derived from the Naive Non-AEC2 Epithelial Cells indicated in blue and cells derived from the Injured AEC2-Derived Cells in orange (left panel) and GFP expression (right panel). J) Number and percent of Injured AEC2-Derived Cells localized to each cluster. K) Percent of Injured AEC2-Derived Cells expressing p15 as determined by costaining with anti-GFP antibody and p15 RNA probe. Mean ± SEM shown. *p<0.05 by Mann-Whitney test. L) Gene Ontology terms for the top 3 biological processes enriched in each cluster, ranked in order of ascending p-value corrected for Elim pruning. n = 4 mice/group.


F.
AEC1 Markers
Injured AEC2-Derived
Naive AEC2
Naive AEC1
Combined
Emp2
Akap5
AQP5
Clic5
G.
AEC2 Markers
SPB
Lyz2
Abca3
Sftpb
H.
IGFBP2
tSNE2
tSNE1
Igfbp2
I.
GFP
tSNE 2
tSNE 1
Injured AEC2-Derived
Naive Non-AEC2 Epithelial
J.
Injured AEC2-Derived Cells / Cluster
Number of Cells
1.16%
3.16%
3.61%
0.64%
0.03%
Proliferating (10)
Cell Cycle Arrest (8)
Transdifferentiating (6)
Naive AEC2 (5)
Naive AEC1 (14)
K.
p15+ Cells
%p15+ AEC2s
Control
LPS

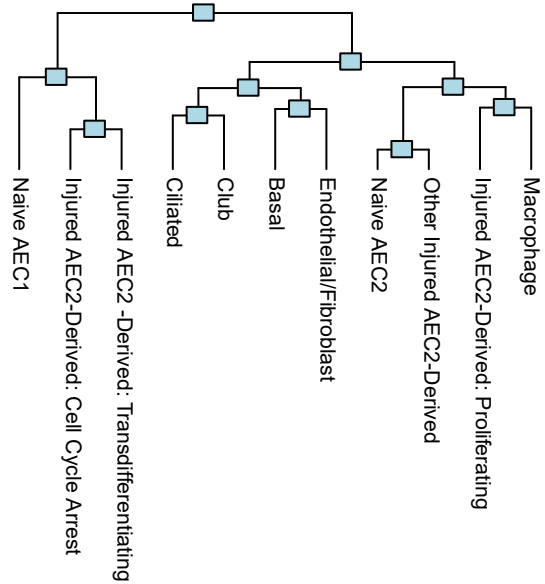
L.


Gene Ontology
Mitotic Cell Cycle Process
Cell Division
Positive Regulation of CDK Activity
Response to Wounding
Regulation of Cell Proliferation
Coagulation
Regulation of Blood Coagulation
Regulation of ERK1 and ERK2 Cascade
Regulation of Cell Shape
48/109
42/116
4/4
23/48
41/130
12/21
7/8
8/12
7/10
5
10
15
20
p-value (x 10-4)
Cluster 10: Proliferating
Cluster 8: Cell Cycle Arrest
Cluster 6: Transdifferentiating

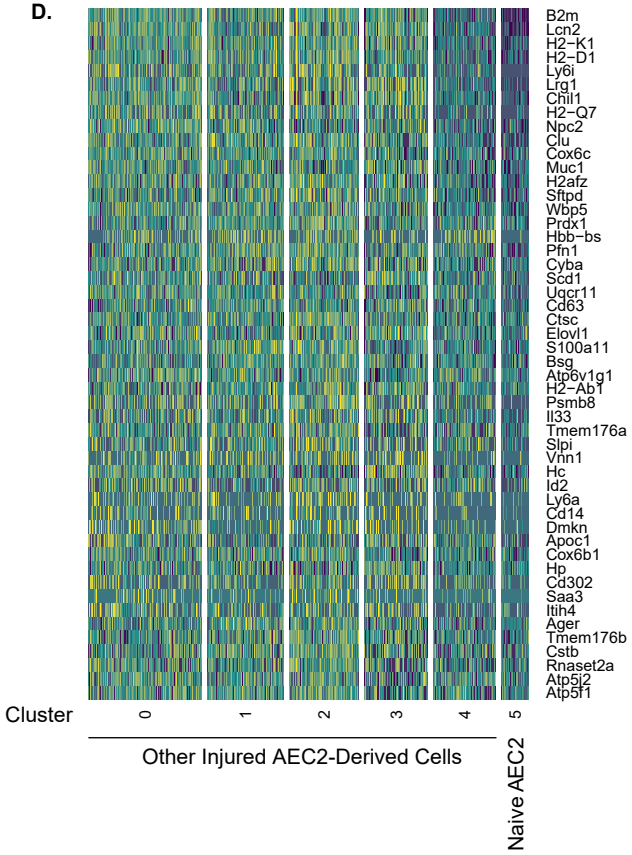


Supplementary Figure 9. scRNAseq Reveals Mildly Perturbed Gene Expression in Other Injured AEC2-Derived Cells. A) Unsupevised clustering at the indicated resolution. B) Canonical correlation analysis of the Other Injured AEC2-Derived and Naive AEC2 populations. Left Panel: tSNE plot showing the overlap of the injured (blue) and naive cells (red). Right Panel: Top genes associated with the first three CCA dimensions, with AEC2 markers highlighted in red. C) Hierarchical clustering of identified subpopulations. D) Heatmap showing the most differentially genes in Naive AEC2 cluster compared to Other Injured AEC2-Derived Cell clusters ranked in order of Bonferroni-corrected p-value. E) Gene Ontology terms for the top 10 biological processes enriched in the Other Injured AEC2-Derived Cells compared to the Naive AEC2s.

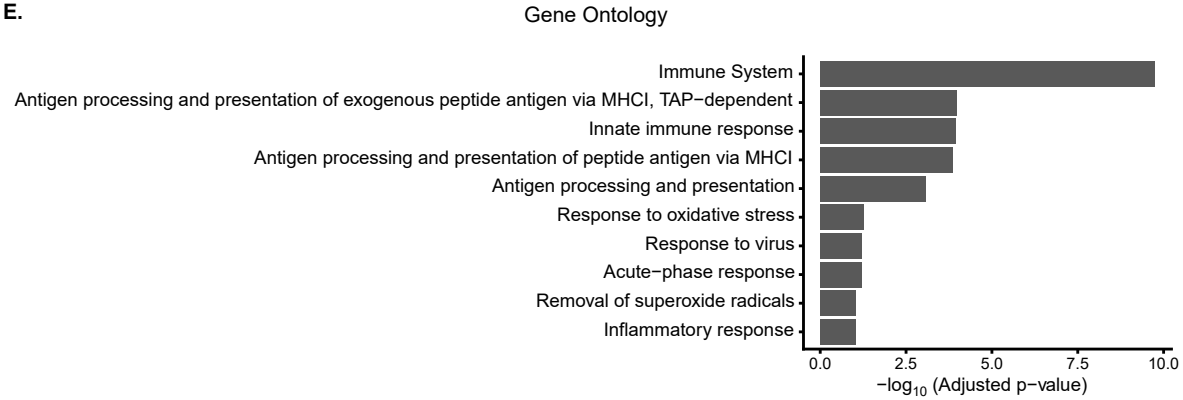
C.

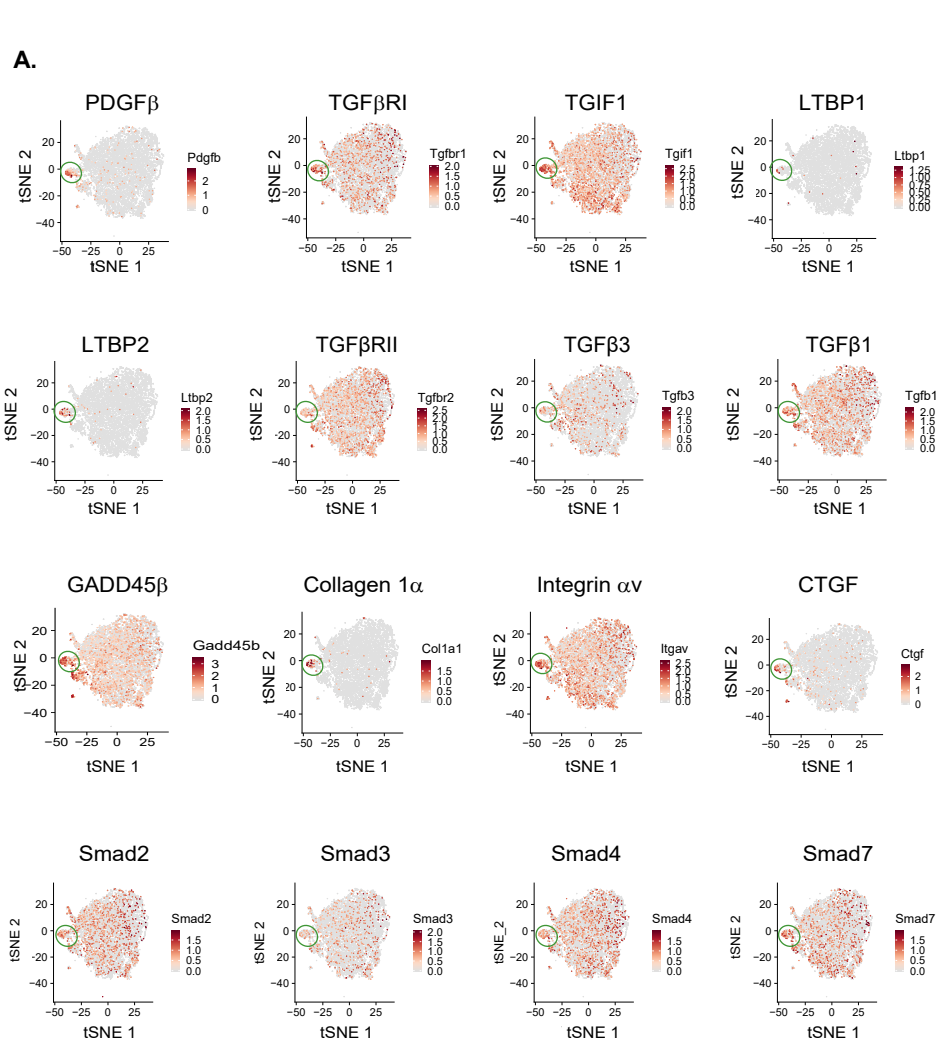

Naive AEC1
Injured AEC2-Derived: Cell Cycle Arrest
Injured AEC2 -Derived: Transdifferentiating
Ciliated
Club
Basal
Endothelial/Fibroblast
Naive AEC2
Other Injured AEC2-Derived
Injured AEC2-Derived: Proliferating
Macrophage

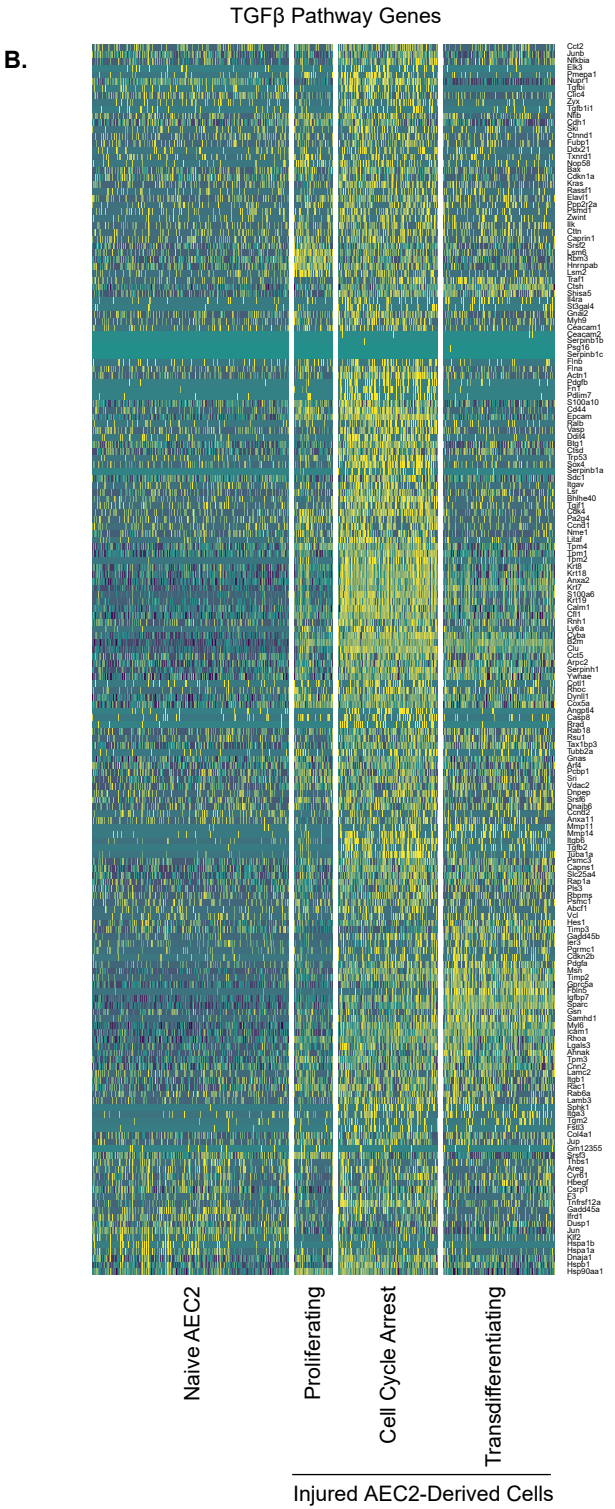
D.


B2m
Lcn2
H2−K1
H2−D1
Ly6i
Lrg1
Chil1
H2−Q7
Npc2
Clu
Cox6c
Muc1
H2afz
Sftpd
Wbp5
Prdx1
Hbb−bs
Pfn1
Cyba
Scd1
Uqcr11
Cd63
Ctsc
Elovl1
S100a11
Bsg
Atp6v1g1
H2−Ab1
Psmb8
Il33
Tmem176a
Slpi
Vnn1
Hc
Id2
Ly6a
Cd14
Dmkn
Apoc1
Cox6b1
Hp
Cd302
Saa3
Itih4
Ager
Tmem176b
Cstb
Rnaset2a
Atp5j2
Atp5f1
Cluster
0
1
2
3
4
5
Other Injured AEC2-Derived Cells
Naive AEC2

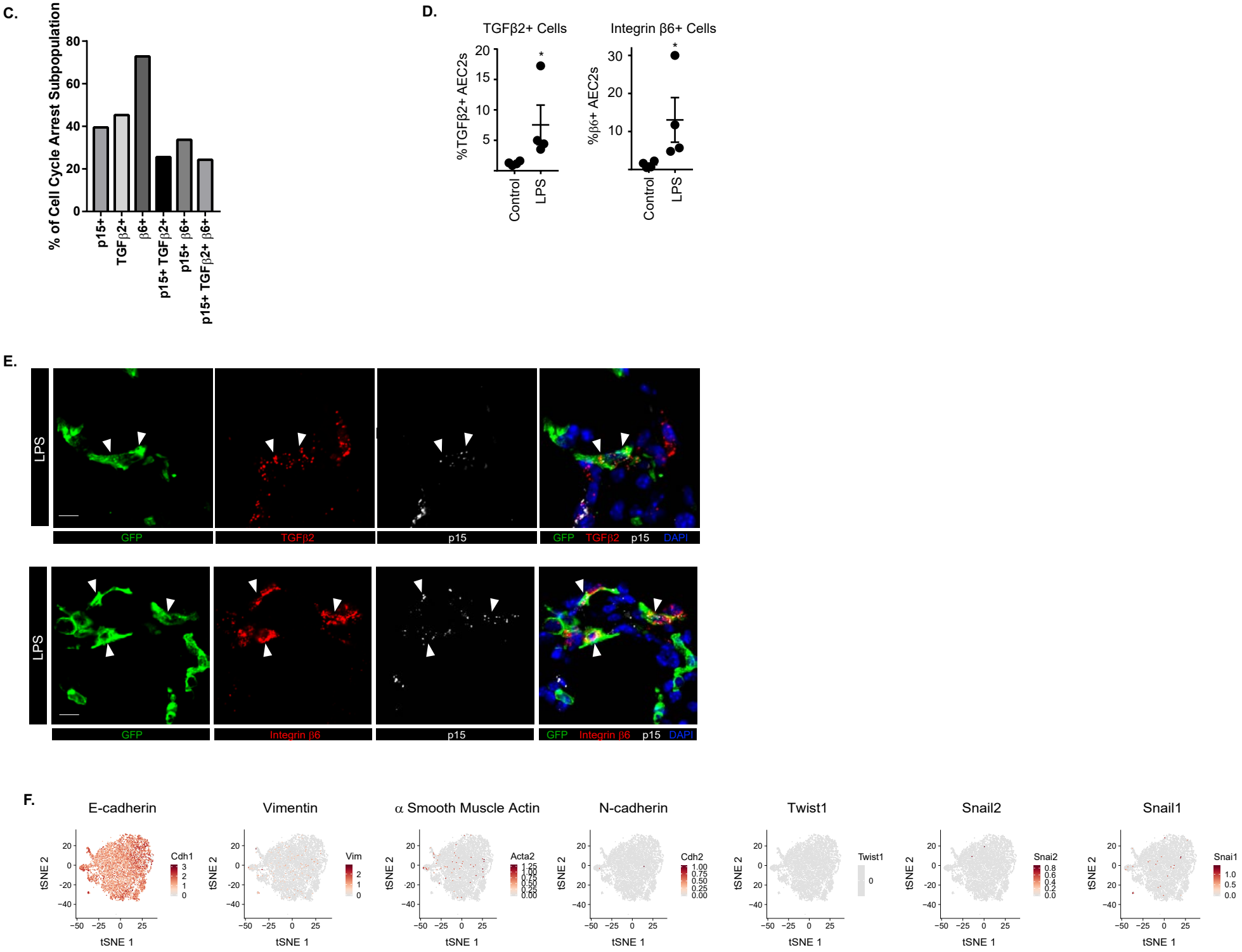
E.


Gene Ontology
Immune System
Antigen processing and presentation of exogenous peptide antigen via MHCI, TAP−dependent
Innate immune response
Antigen processing and presentation of peptide antigen via MHCI
Antigen processing and presentation
Response to oxidative stress
Response to virus
Acute−phase response
Removal of superoxide radicals
Inflammatory response
0.0
2.5
5.0
7.5
10.0
−log10 (Adjusted p−value)



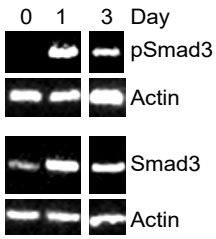


Supplementary Figure 10. Expression of TGFβ Pathway and EMT Genes. A) Gene expression (natural log of normalized counts) of TGFβ signaling molecules. B) Heatmap of TGFβ pathway gene expression in the indicated subpopulations. C) Percent of Cell Cycle Arrest subpopulation expressing indicated genes. D) Percent of AEC2-Derived Cells expressing at least 3 RNA molecules of *TGFβ2* or Integrin β6 as determined by costaining with anti-GFP antibody and RNA probes. Mean ± SEM shown. *p<0.05 by Mann-Whitney test. E) Immunofluorescence staining for GFP and in situ hybridization for *TGFβ2,* Integrin β6, and p15. Arrowheads indicate double positive cells. Scale bar 10 μm. n ≥ 4 mice/group. F) Gene expression (natural log of normalized counts) of epithelial and mesenchymal markers. G) Heatmap of expression of mesenchymal markers.

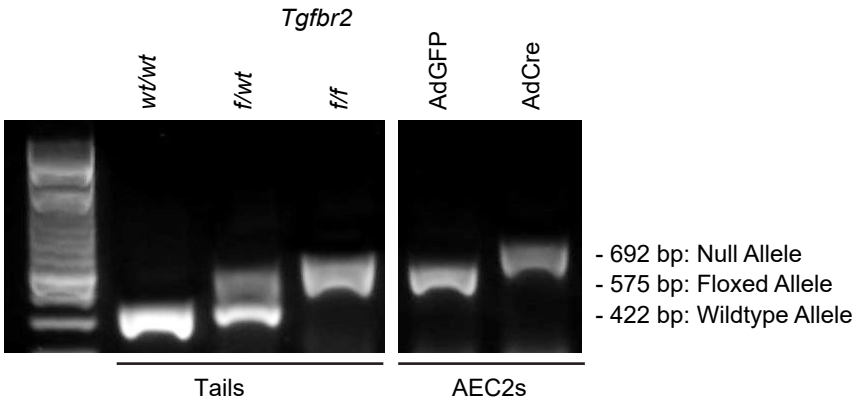

C.
% of Cell Cycle Arrest Subpopulation
80
60
40
20
0
p15+
TGFβ2+
β6+
p15+ TGFβ2+
p15+ β6+
p15+ TGFβ2+ β6+
D.
TGFβ2+ Cells
Integrin β6+ Cells
%TGFβ2+ AEC2s
%β6+ AEC2s
Control
LPS
E.
LPS
GFP
TGFβ2
p15
GFP TGFβ2 p15 DAPI
Integrin β6
GFP Integrin β6 p15 DAPI
F.
E-cadherin
Vimentin
α Smooth Muscle Actin
N-cadherin
Twist1
Snail2
Snail1
tSNE 1
tSNE 2
Cdh1
Vim
Acta2
Cdh2
Twist1
Snai2
Snai1

G.

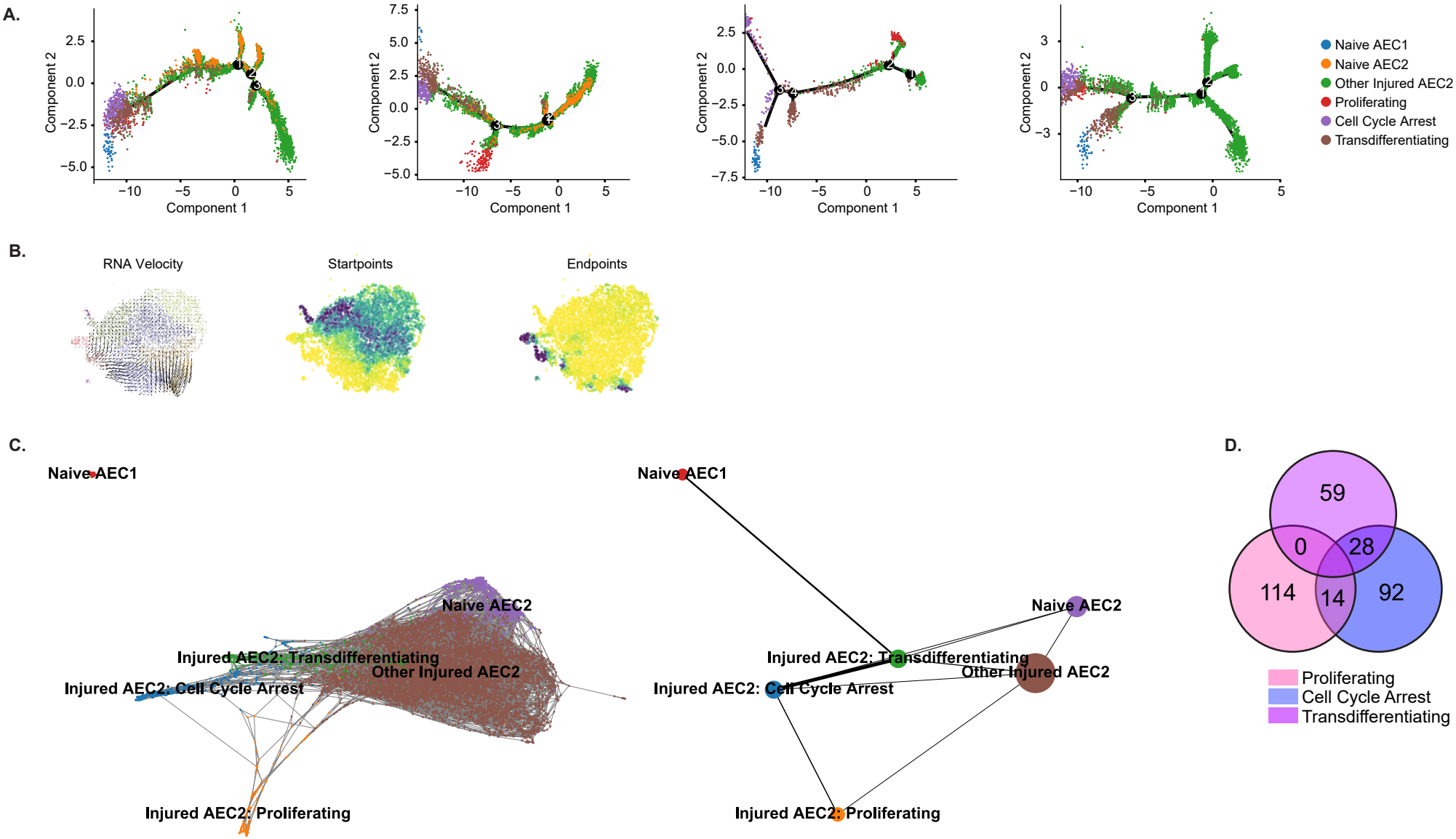

Mesenchymal Markers
Rbp1
Pixdc2
Prelp
Aldh1a1
Rarres2
Serping1
Selm
Mgp
Clec3b
Igfbp6
Ptgis
Gstm1
Sod3
Col1a2
Col3a1
Fxyd1
Col1a1
Loxl1
Ltbp4
Cth
Ogn
Lrp1
Heg1
Fn1
Zbtb20
Dusp1
Timp2
Ntn1
Ifitm3
Serpinh1
Col6a1
Cxcl14
Pdgfra
Itga8
Prex2
Wisp2
Hsd11b1
Ppp1r14a
Lbh
Gpx3
Mfap4
Pcolce2
Inmt
Crispld2
Dpep1
Spon1
G0s2
Gng11
Col6a3
Eln
Sparcl1
Fmo2
Bgn
Cdo1
Mxra8
Nr2f2
Ntib
Ptx1
Sparc
Igfbp7
Cd81
Gsn
Nupr1
Fn1
Cald1
Csrp1
Plpp3
Apoe
Cst3
Fth1
Srrhg18
Mylk
Npnt
Tns1
Cryab
Dpysl2
Timp3
Limch1
Macf1
Endothelial/Fibroblast
Cell Cycle Arrest
Proliferating
Transdifferentiating
Injured AEC2-Derived Cells



Supplementary Fig. 11. Total and Phosphorylated Smad3 Levels Increase in Cultured AEC2s. Primary rat AEC2s were cultured. Western blotting of cell lysates for pSmad3, Smad3, and actin. Vertical white lines indicate that the lanes were run on the same gel but were noncontiguous.



Supplementary Figure 12. Conditional knockout of TGFβRII in Mouse Tails or Cultured AEC2s. PCR performed on gDNA extracted from tails of *Tgfbr2*$^{wt/wt}$, *Tgfbr2*$^{f/wt}$, *or Tgfbr2*$^{f/f}$ mice or AEC2s isolated from *Tgfbr2*$^{f/f}$ mice and transduced with AdCre or AdGFP. Expected size of each allele is shown. Representative of n=3.



Supplementary Figure 13. Lineage Relationship Between Subpopulations. A) Pseudotime analysis was performed on scRNAseq data using Monocle2 with various input parameters: top 1000 most differentially expressed genes (1st panel), top 2000 differentially expressed genes (2nd panel), excluding Naive AEC2 populations (3rd, 4th panel), or with downsampling to reduce the number of injured naive AEC2 cells (3rd panel). B) RNA velocity projections overlayed on tSNE (left) and inferred trajectory start points (middle) or endpoints (right) based on Markovchain process. Dark blue indicates the highest density of end or start points. C) PAGA analysis to infer branch points, force directed graph layout (left) or PAGA representation (right). D) Venn diagram showing the overlap of genes highly differentially expressed by each subpopulation. n=2 mice/group.