Cell Genomics, Volume *2*

Supplemental information

A single-cell regulatory map of postnatal lung alveologenesis in humans and mice

Thu Elizabeth Duong, Yan Wu, Brandon Chin Sos, Weixiu Dong, Siddharth Limaye, Lauraine H. Rivier, Greg Myers, James S. Hagood, and Kun Zhang

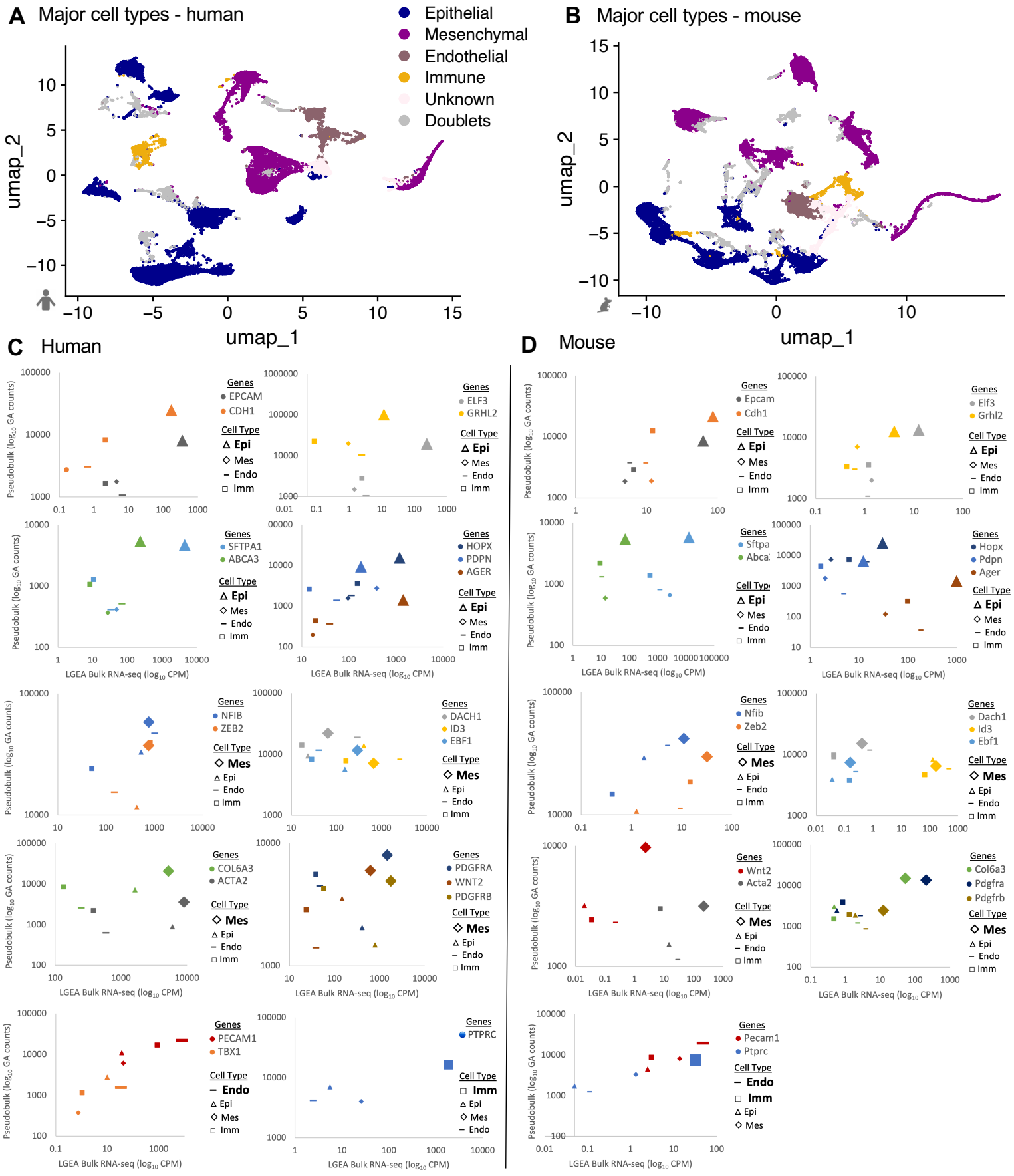


Figure S1. Correlation of marker genes between scTHS-seq and LGEA bulk RNA-seq, Related to Figure 1. **(A-B)** UMAP of human and mouse scTHS-seq in Figure 1B-C colored by major cell type. **(C-D)** Scatter plots of pseudobulk scTHS-seq gene activity counts for marker genes used to classify clusters in major cell types with corresponding expression (log10 cells counts per million) of LGEA bulk RNA-seq of sorted epithelial, mesenchymal, endothelial, and immune cells [S1-3].

See also Table S7.

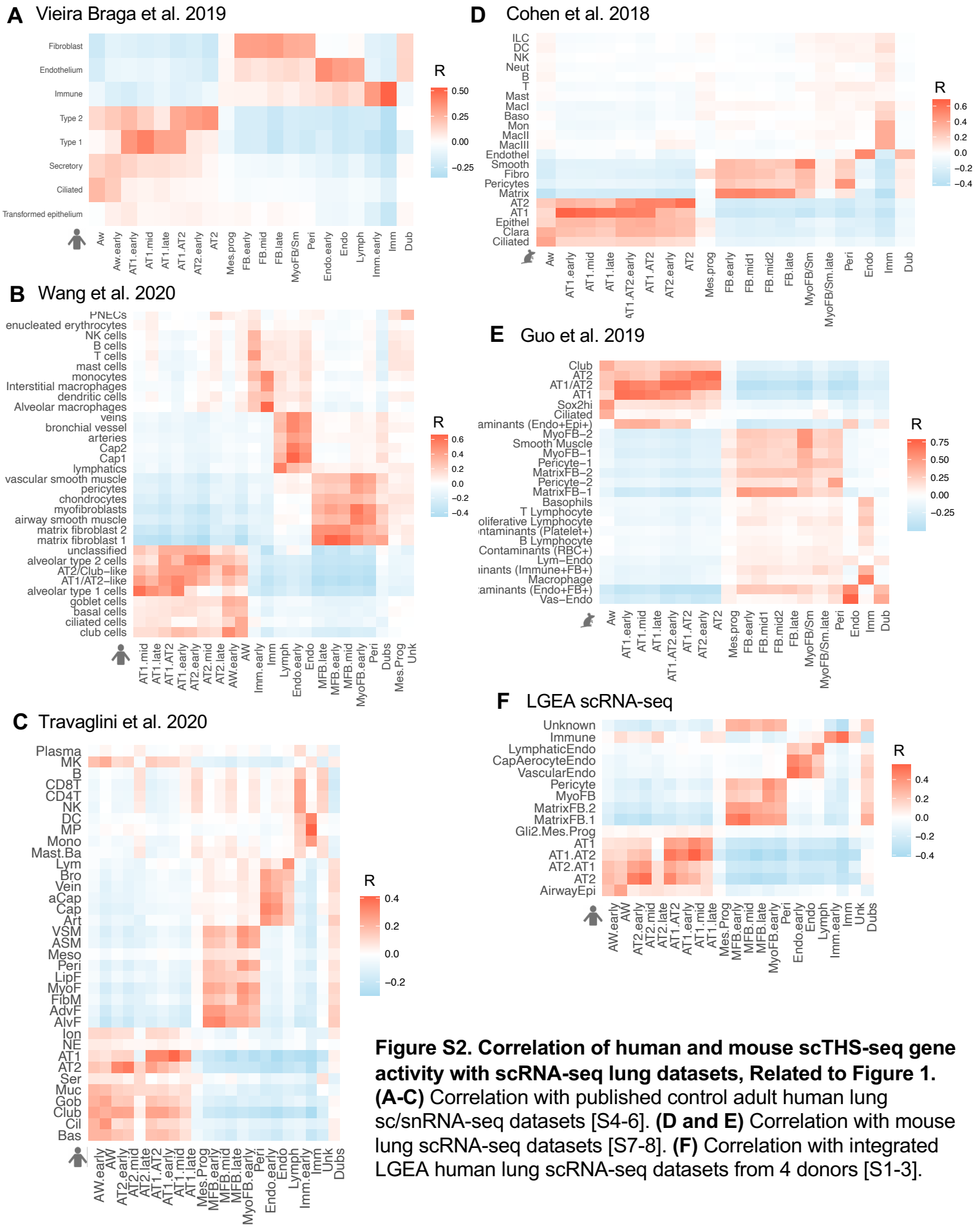


Figure S2. Correlation of human and mouse scTHS-seq gene activity with scRNA-seq lung datasets, Related to Figure 1. **(A-C)** Correlation with published control adult human lung sc/snRNA-seq datasets [S4-6]. **(D and E)** Correlation with mouse lung scRNA-seq datasets [S7-8]. **(F)** Correlation with integrated LGEA human lung scRNA-seq datasets from 4 donors [S1-3].

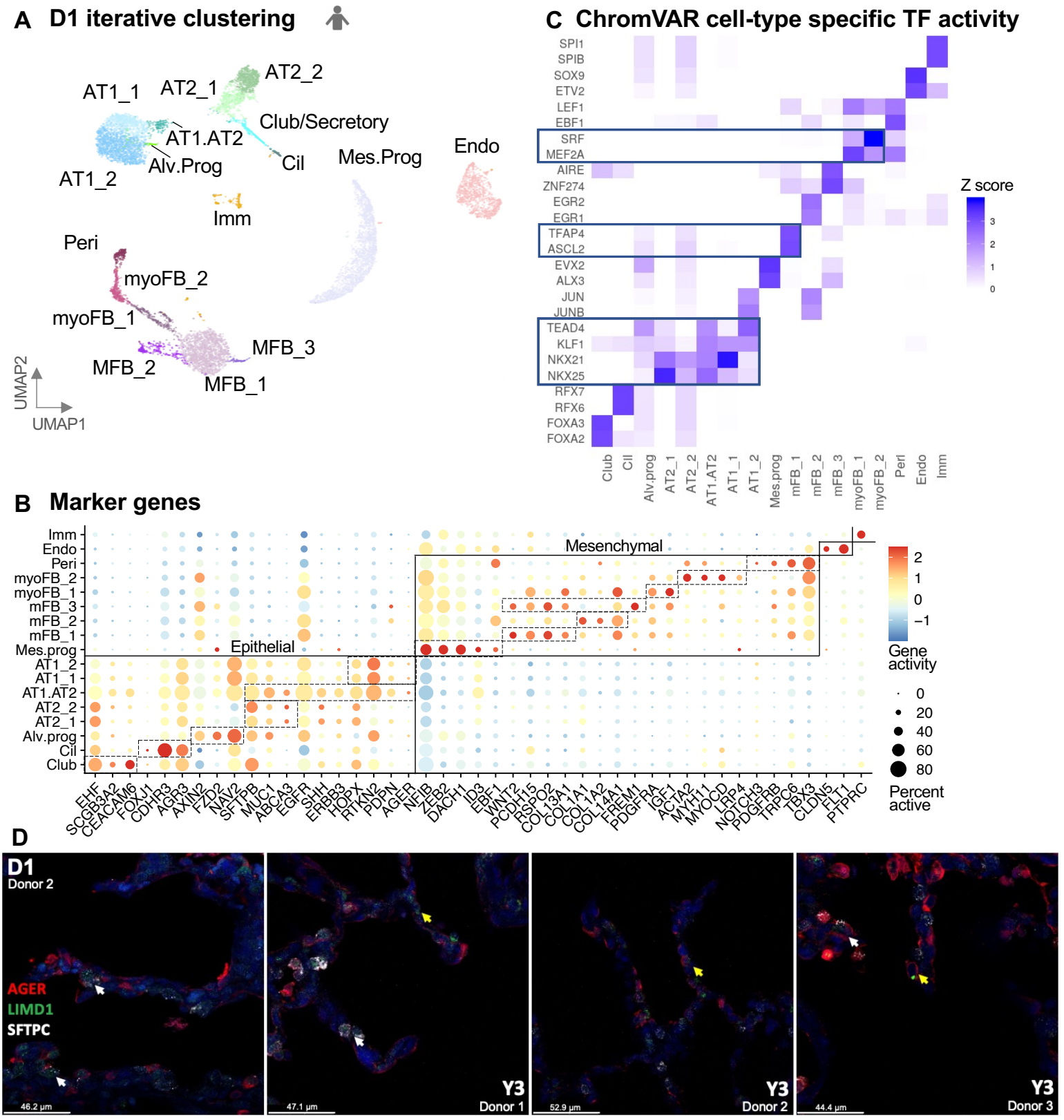
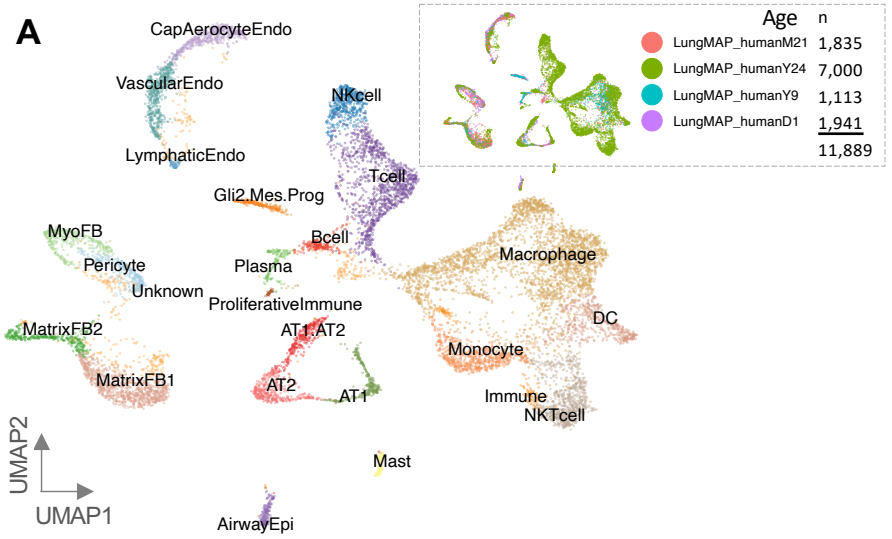


Figure S3. Iterative clustering of human day 1 lung cells and immunofluorescent staining for AT1.AT2 cells, Related to Figure 2-3, 5-6. (A) UMAP of annotated D1 cell types. **(B)** Dot plot of log-normalized gene activity scores for marker genes in each cell state. **(C)** Heatmap of top 2 differential TFs for each cluster using ChromVAR TF activity Z scores. Rectangles highlight key TFs explored in downstream construction of regulatory modules. **(D)** *LIMD1, AGER*, and *SFTPC* staining in alveolar region of human tissue in additional D1 and Y3 donors. White arrows indicate triple positive AT1.AT2 cells. Yellow arrows indicate *AGER* and *LIMD1* in AT1 cells. scTHS-seq data collected from donor 1 in each age group. See also Table S4.



C

scTHS-seq			LGEA scRNA-seq		Wang snRNA-seq	
Proportion	n	Cluster	n	Proportion	n	Proportion
5%	1859	Aw	169	1%	2889	6%
24%	9653	AT1	270	2%	3618	8%
1%	585	AT1.AT2	315	3%	794	2%
16%	6389	AT2	355	3%	7226	16%
8%	3115	Mes.prog	224	2%	0	0%
19%	7560	FB	1159	10%	7467	16%
2%	971	myoFB	330	3%	1206	3%
1%	530	Peri	330	3%	803	2%
6%	2399	Endo	1022	9%	2252	5%
1%	581	Lymph	67	1%	887	2%
4%	1621	Imm	7148	60%	17152	37%
13%	5164	Unknown	500	4%	2206	5%
	40427		11889		46500	

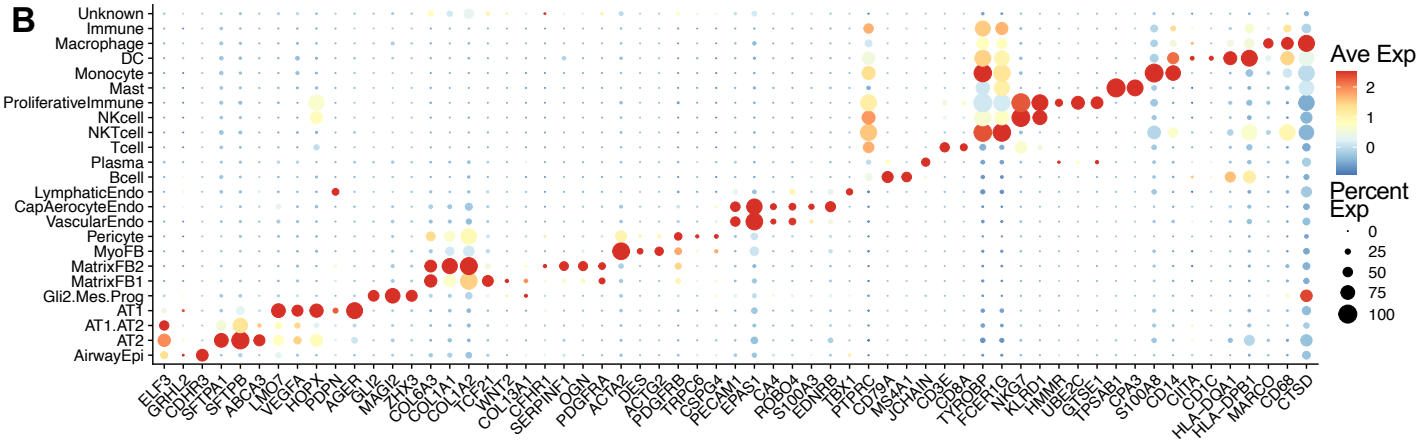
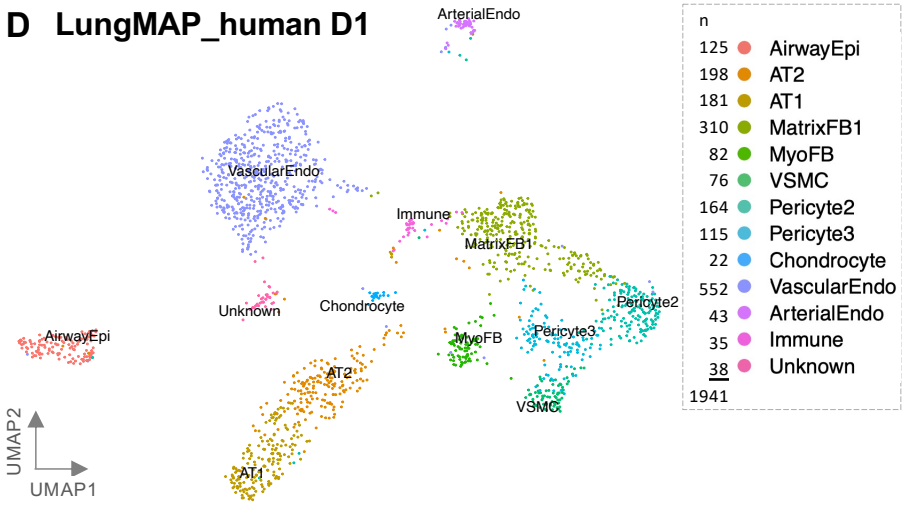
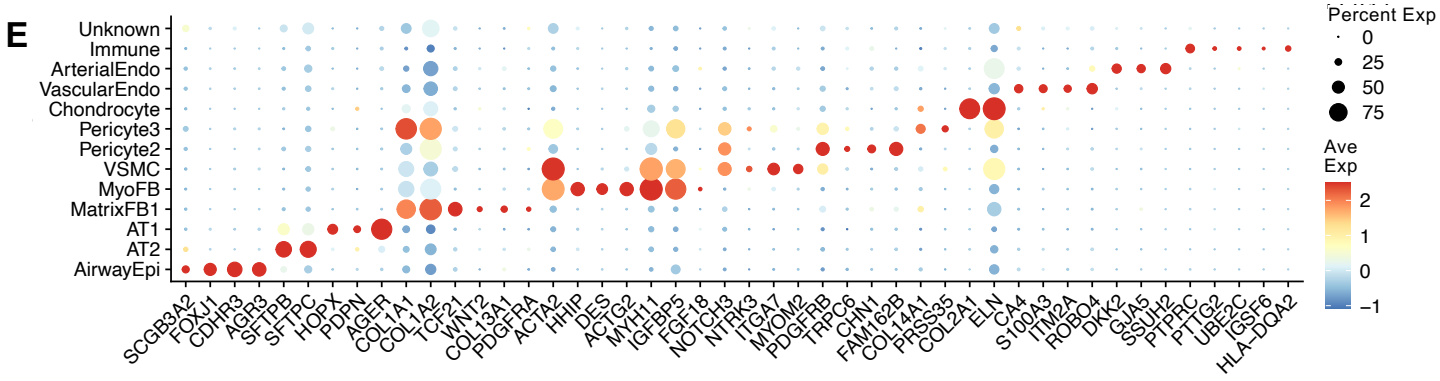


Figure S4. Human LGEA scRNA-seq analysis, Related to Figure 3 and 6. **(A)** UMAPs of integrated LGEA human scRNA-seq data from D1, M24, Y9, and Y24 donors [S1-3] with annotated cell types based on signature expression of genes. Inset recolors UMAP by donor. **(B)** Dot plot of DE genes used to annotate cell types in A. **(C)** Table of approximate annotated cell type proportions in scTHS-seq, LGEA scRNA-seq, and Wang et al. 2020 [S5] snRNA-seq datasets. Annotated cell types were grouped by best fit into cluster labels listed. **(D)** UMAP of annotated cell types in LGEA D1 human scRNA-seq dataset. **(E)** Dot plot of DE genes used to annotate cell types in D.





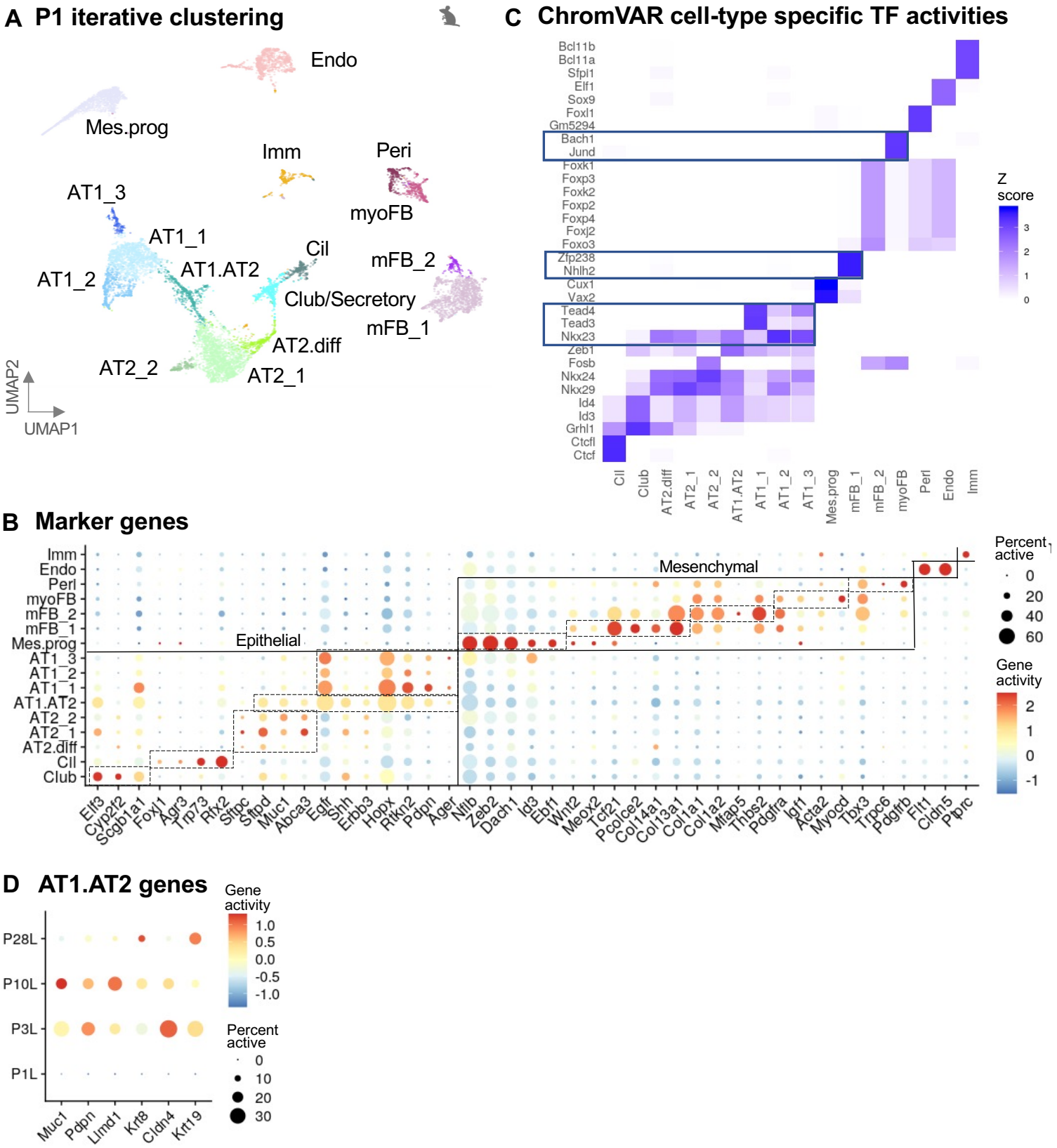


Figure S5. Iterative clustering of mouse postnatal day 1 lung cells, Related to Figure 2, 5, and 6.
(A) UMAP annotated P1 cells. **(B)** Dot plot of gene activity scores for P1 annotated clusters. **(C)** Top 2 TFs for each cluster using ChromVAR differential TF activity Z scores. **(D)** Dot plot of DA genes in AT1.AT2 cluster by donor.
See also Table S4.

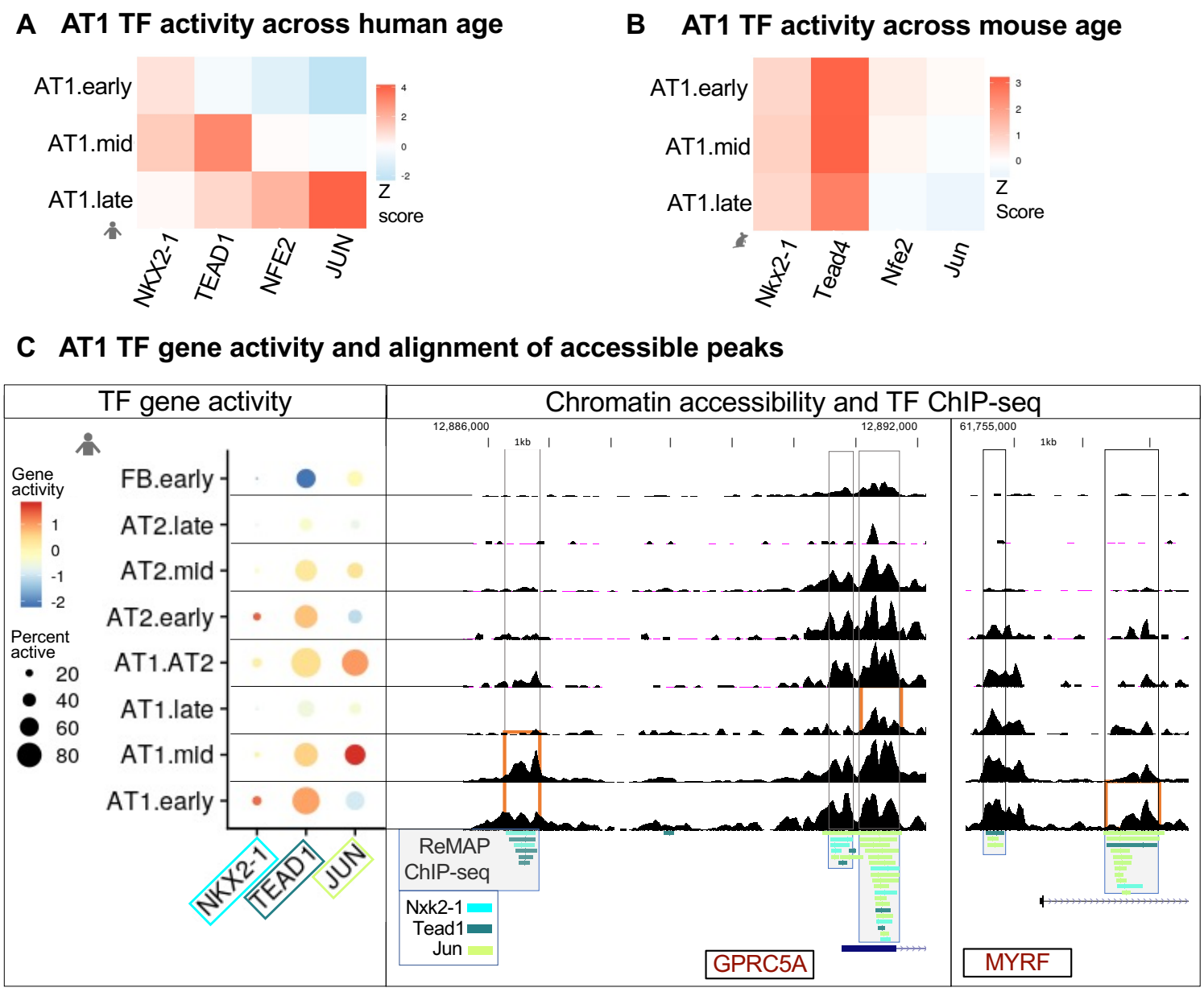


Figure S6. Dynamic AT1 cell regulation with increasing age and across species, Related to Figure 1. (A and B) TF activity Z scores for human and mouse AT1 subsets from Figure 1B and 1C. **(C)** Aligning temporal TF availability with corresponding TF binding to accessible peaks linked to AT1 marker genes. Gene activity scores for *NKX2-1, TEAD1,* and *JUN* for age-defined clusters from Figure 1B (left panel). Genome browser tracks for corresponding cell types with highlighted overlap between accessible peaks and ChIP-seq peaks for *NKX2-1, TEAD1,* and *JUN* from ReMAP 2018 [S9] near *GPRC5A* and *MYRF* TSS. Bars for ChIP-seq peaks are colored by TF to which they belong (right panel).
See also Table S2.

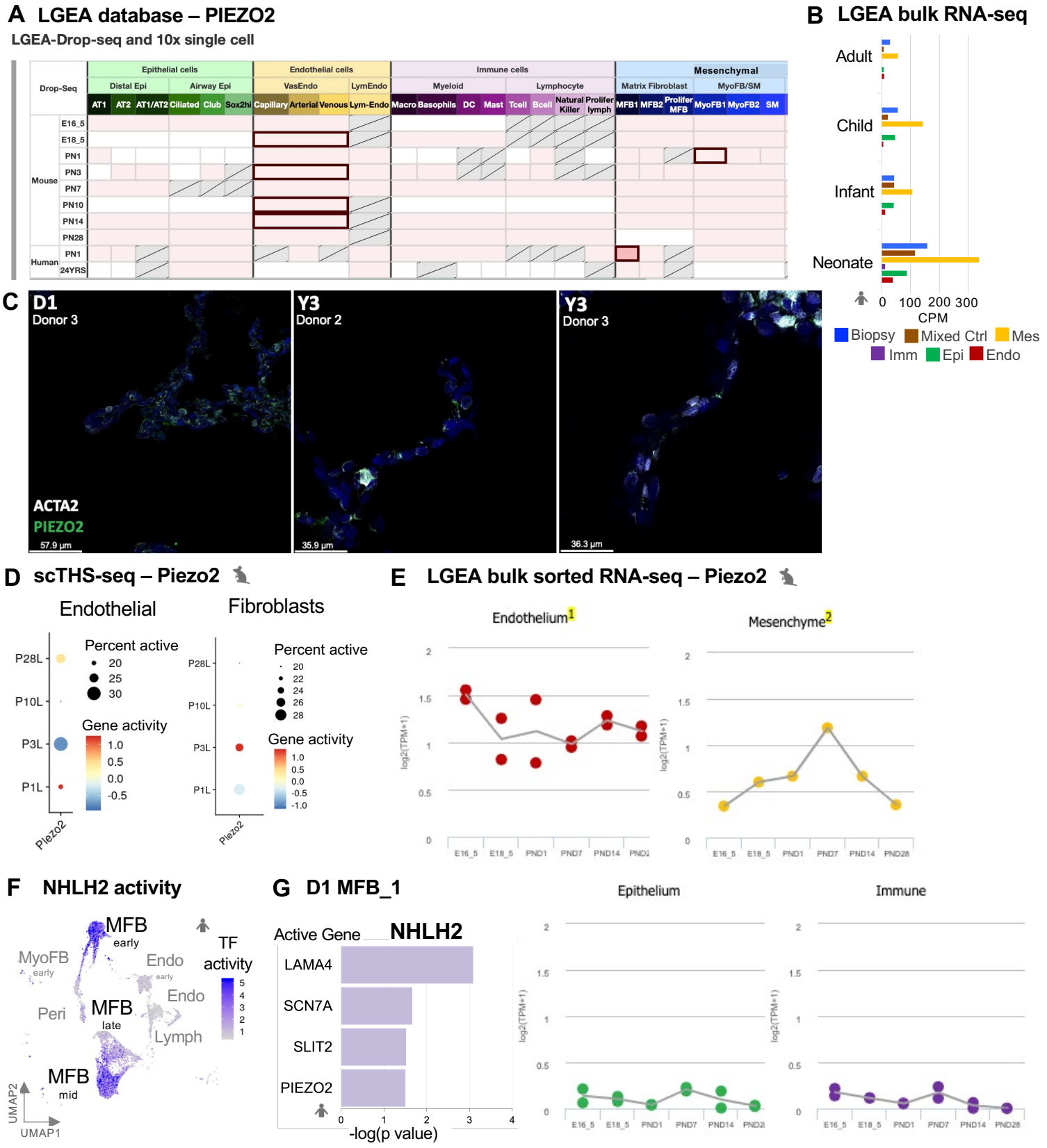


Figure S7. Accessibility, expression, and immunostaining of *PIEZO2,* Related to Figure 6.

(A) Table from LGEA with *PIEZO2* SE (red boxes) from scDrop-seq experiments in varying human and mouse cell types at different developmental ages. **(B)** Average counts per million (CPM) of *PIEZO2* expression from LGEA bulk sorted epithelial, mesenchymal, endothelial, immune, biopsy, and mixed control populations from neonatal, infant, child and adult human lung cells. **(C)** *PIEZO2* and *ACTA2* staining in alveolar region of human tissue in additional D1 and Y3 donors. **(D)** *Piezo2* gene activity for all mouse endothelial and fibroblasts grouped by age. **(E)** *Piezo2* bulk gene expression (log2(TPM+1)) in LGEA endothelium, mesenchyme, epithelium, and immune sorted lung populations from different murine developmental time points. (**F**) UMAP of stromal subset from Figure 1B highlighting *NHLH2* TF activity across development. (G) Barplot of -log$_{10}$(p values) for active genes in D1 MFB_1 cells with significant *NHLH2* motif enrichment.

Table S6. Summary of key findings and supporting evidence, Related to Figure 1-7.

Finding	Evidence
Human and mouse mesenchymal progenitor (Mes.Prog) population -Signature genes: *NFIB, ZEB2, DACH1, ID3, EBF1*	Table S3 – Overlapping differential genes in similar cell populations in the following manuscripts and LGEA: 1. Human and mouse Ebf1+ fibroblasts (Liu et al., iScience 2021 PMID 34151224)[S10] 2. Mouse mesenchymal progenitor (MP) (Xie et al., Cell Rep. 2018 PMID 29590628)[S11] 3. Mouse proliferative mesenchymal progenitors (PMP) (LGEA E16.5)[S1-3]
AT1.AT2 cells -Signature genes: *LIMD1*, *KRT8, KRT19*, and *CLDN*	Corresponding LGEA [S1-3] and Wang et al., Elife 2020 [S5] sc/snRNA-seq gene expression in human lung. Figure 2C and S3D *LIMD1* protein immunostaining in human lung at D1, Y3, and adult. Three donors each for D1 and Y3. Similar cell populations: 1. Human and mouse Pre-alveolar type 1 transitional cell state (PATS) (Kobayashi et al., Nat Cell Bio 2020 PMID 32661339)[S12] 2. Mouse Krt8+ alveolar differentiation intermediate (ADI) (Strunz/Simon et al., Nat Comm 2020 PMID 32678092)[S13] 3. Mouse damage-associated transient progenitors (DATPs) (Choi et al., Cell Stem Cell 2020 PMID 32750316)[S14]
KLF and *TEAD* TF involvement in AT1 cell differentiation AT1 *PDGFA* ligand to myofibroblast *PDGFRA* receptor signaling drives postnatal alveolar septation	Corresponding TF and gene expression in LGEA [S1-3] and Wang et al., Elife 2020 [S5] sc/snRNA-seq. 1. Mouse scRNA-seq and scATAC-seq data (Zepp et al., Science 2021 PMID 33707239)[S15] – AT1 cells are a signaling hub to receptive, force-exerting myofibroblasts required for postnatal respiration.
Increase in number of myofibroblasts during human and mouse alveolar septation	LGEA ACTA2 protein immunostaining in human lung (Figure 4A). Figure 6H and S7C *ACTA2* protein immunostaining (co-staining with PIEZO2) in human lung. Three donors each for D1 and Y3.
Matrix fibroblast type 1 Signature gene: *PIEZO2*	Corresponding human and mouse single-cell expression and LGEA bulk expression of PIEZO2 (Figure 6 and S7). Figure 6H and S7C *PIEZO2* protein immunostaining in human lung. Three donors each for D1 and Y3.

Supplemental Information References:

1. Du, Y., Guo, M., Whitsett, J. A. & Xu, Y. (2015). 'LungGENS': A web-based tool for mapping single-cell gene expression in the developing lung. *Thorax* **70**, 1092–1094.
2. Du, Y., Kitzmiller, J. A., Sridharan, A., Perl, A. K., Bridges, J. P., Misra, R. S., Pryhuber, G. S., Mariani, T. J., Bhattacharya, S., Guo, M., *et al.* (2017). Lung Gene Expression Analysis (LGEA): An integrative web portal for comprehensive gene expression data analysis in lung development. *Thorax* vol. 72 481–484.
3. Du, Y., Ouyang, W., Kitzmiller, J. A., Guo, M., Zhao, S., Whitsett, J. A. & Xu, Y. (2021). Lung gene expression analysis web portal version 3: Lung-at-a-glance. *American Journal of Respiratory Cell and Molecular Biology* vol. 64 146–149.
4. Vieira Braga, F. A., Kar, G., Berg, M., Carpaij, O. A., Polanski, K., Simon, L. M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., *et al.* (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* 1 doi:10.1038/s41591-019-0468-5.
5. Wang, A., Chiou, J., Poirion, O. B., Buchanan, J., Valdez, M. J., Verheyden, J. M., Hou, X., Kudtarkar, P., Narendra, S., Newsome, J. M., *et al.* (2020). Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of sars-cov2 host genes. *Elife* **9**, 1–28.
6. Travaglini, K. J., Nabhan, A. N., Penland, L., Sinha, R., Gillich, A., Sit, R. V., Chang, S., Conley, S. D., Mori, Y., Seita, J., *et al.* (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625.
7. Cohen, M., Giladi, A., Gorki, A. D., Solodkin, D. G., Zada, M., Hladik, A., Miklosi, A., Salame, T. M., Halpern, K. B., David, E., *et al.* (2018). Lung Single-Cell Signaling Interaction Map Reveals Basophil Role in Macrophage Imprinting. *Cell* **175**, 1031-1044.e18.
8. Guo, M., Du, Y., Gokey, J. J., Ray, S., Bell, S. M., Adam, M., Sudha, P., Perl, A. K., Deshmukh, H., Potter, S. S., *et al.* (2019). Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. *Nat. Commun.* **10**, 37.
9. Chèneby, J., Gheorghe, M., Artufel, M., Mathelier, A. & Ballester, B. (2018). ReMap 2018: An updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. *Nucleic Acids Res.* **46**, D267–D275.