

Supplemental Information Guide

Supplemental Figure 1:

Representative FACS plot of primary mouse tracheal epithelial cells, including EpCAM⁺, tdTomato⁺ and EGFP⁺ gating strategies. Number of cells captured in each gate indicated, as well as the percentage of the parent population (schematic, bottom).

Supplemental Table 1:

Marker genes for cell-type clusters from initial 3' droplet-based scRNA-seq data (7, 193 cells). Summary tab displays marker genes of each cell-type cluster. Thresholds: Maximum FDR: 0.05, Minimum log₂ fold-change: 0.25. The top 1,000 marker genes for each individual cell-type cluster are shown in individual tabs. The following values are provided in columns: *P* (max) Maximum p-value (computed using MAST's likelihood-ratio test, two-sided, see Methods) across all comparisons for differential expression between this group and all other cell-type clusters, FDR (*Q*, max) FDR adjusted maximum p-value, *p* (Fisher's combined) Combined p-value across all comparisons for differential expression between this group and all other cell-type clusters, log₂ fold-change (min) Minimum log₂ fold-change of means between this group and all other cell-type clusters, log₂ fold-change (lower bound) Lower bound of 95% confidence interval for log₂ fold-change of means between this group and all other cell-type clusters, log₂ fold-change (mean) Average log₂ fold-change of means between this group and all other cell-type clusters.

Supplemental Table 2:

Marker genes for cell-type clusters from initial full-length plate-based scRNA-seq data (301 cells). Summary tab displays marker genes of each cell-type cluster. Goblet cells were not detected in this data set. Thresholds: Fisher's combined FDR: 0.001, Minimum log₂ fold-change: 0.25. The top 1,000 marker genes for each individual cell-type cluster are shown in individual tabs. The following values are provided in columns: *P* (max) Maximum p-value (computed using MAST's likelihood-ratio test, two-sided, see Methods) across all comparisons for differential expression between this group and all other cell-type clusters, FDR (*Q*, max) FDR adjusted maximum p-value, *p* (Fisher's combined) Combined p-value across all comparisons for differential expression between this group and all other cell-type clusters, log₂ fold-change (min) Minimum log₂ fold-change of means between this group and all other cell-type clusters, log₂ fold-change (lower bound) Lower bound of 95% confidence interval for log₂ fold-change of means between this group and all other cell-type clusters, log₂ fold-change (mean) Average log₂ fold-change of means between this group and all other cell-type clusters.

Supplemental Table 3:

Consensus marker genes for cell-type clusters from both initial 3' droplet-based and full-length plate-based scRNA-seq data (7,193 + 301 cells, in Sup. Tabs 1 and 2, respectively) Thresholds: Fisher's combined FDR: 0.001 for plate-based and maximum FDR: 0.05 for droplet-based, Minimum log₂ fold-change: 0.25.

Supplemental Table 4:

TFs tab. Related to Extended Data Fig. 3c: Cell-type enriched TFs (this sheet) from initial 3' droplet-based scRNA-seq data (7, 193 cells) and GPCRs from full-length plate-based scRNA-seq (next sheet)

Thresholds: FDR-corrected Fisher's combined p-value < 0.001. **GPCRs tab.** Related to Extended Data Fig. 8a: Cell-type enriched GPCRs from full-length plate-based scRNA-seq. Thresholds: FDR-corrected Fisher's combined p-value < 0.001.

Supplemental Table 5:

Related to Extended Data Fig. 4a: Genes enriched in proximal club cells vs distal club cells from full-length plate-based scRNA-seq (121 club cells, 46 distal, 75 proximal). Thresholds: all genes shown. The following values are provided in columns for each gene: log₂ fold-change of means, estimated log₂ fold-change (MAST), p (MAST, likelihood-ratio test one-sided), FDR.

Supplemental Table 6:

Related to Extended Data 4d: Krt4⁺/Krt13⁺ hillock-associated genes from 3' droplet-based scRNA-seq. Thresholds: all genes shown. The following values are provided in columns for each gene: log₂ fold-change of means (hillock cells vs all other cells), estimated log₂ fold-change (MAST), p (MAST, likelihood-ratio test one-sided), FDR.

Supplemental Table 7:

Related to Extended Data Fig. 5: Cell fate transition-associated genes from 3' droplet-based scRNA-seq. **All genes tab.** Thresholds: Top 250 genes associated with the 'pseudotime' ordering along a given edge from one cell-type to another in diffusion map space (p<0.001, permutation test). **All TFs tab.** All TFs associated (p<0.001, permutation test) with the 'pseudotime' ordering. Cell-fate transitions are annotated with the cell-type at the putative origin (red) and endpoint (cyan) of pseudotime.

Supplemental Table 8:

Related to Fig. 4a and Extended Data Fig. 8l: Novel tuft and goblet cell subset markers defined using 'Pulse-Seq' dataset (66,265 cells)
Thresholds: Fisher's combined FDR: 0.001, minimum log₂ fold-change: 0.1

Supplemental Table 9:

Related to Fig. 3: Cell-type abundance summary statistics from GFP⁺ (lineage-labeled) and not using 'Pulse-Seq' dataset (66,265 cells)

Supplemental Table 10:

Abundance of pulmonary ionocytes detected by airway region from human (n=1).

Supplemental Table 11:

Related to Fig. 5k: Top 100 marker genes for human pulmonary ionocytes, ranked by difference in detection fraction

Supplemental Table 12:

Related to Fig. 5d,e and Extended Data Fig. 10a. Quantitative real-time PCR (q-RT PCR) quantification of ionocyte and other cell-type markers under homeostasis and *Foxi1*- and *Ascl3*-KO. $\Delta\Delta CT$ values are the mean of the numbers of replicate mice listed (column E).