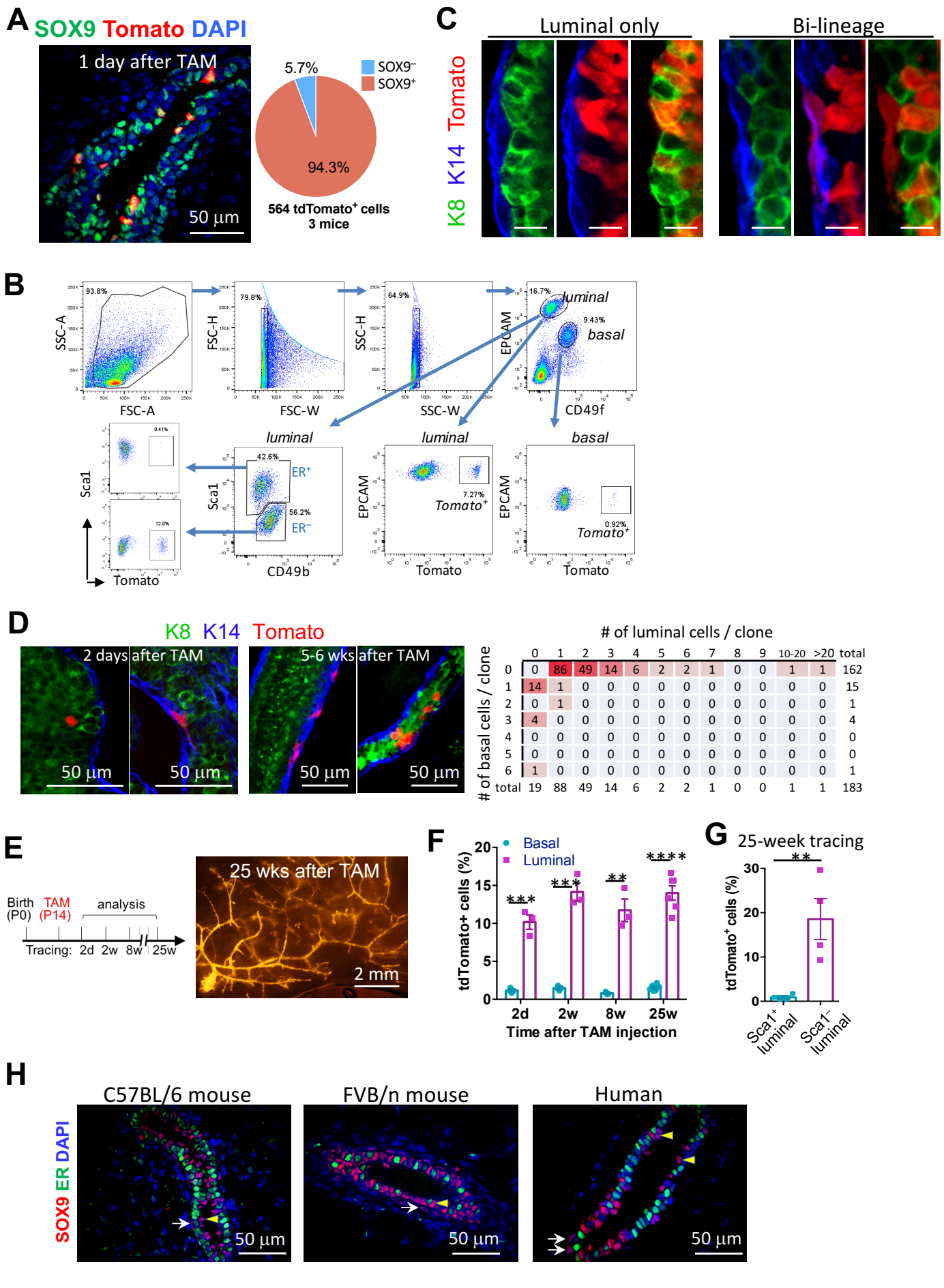
Figure S1, Related to Figure 1



# of basal cells / clone	0	1	2	3	4	5	6	7	8	9	10-20	>20	total
0	0	86	49	14	6	2	2	1	0	0	1	1	162
1	14	1	0	0	0	0	0	0	0	0	0	0	15
2	0	1	0	0	0	0	0	0	0	0	0	0	1
3	4	0	0	0	0	0	0	0	0	0	0	0	4
4	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0	0	1
total	19	88	49	14	6	2	2	1	0	0	1	1	183

Figure S1, related to Figure 1.

(A) SOX9 immunostaining of *Sox9-CreER^{T2}; R26R-tdTomato* mammary glands 24 hours after TAM injection.

(B) Gating strategy for flow cytometric analysis of mammary gland epithelial cells in this study. EPCAM and CD49f were used to separate basal and luminal cells, and CD49b and Sca1 were used to further divide luminal cells into the ER$^+$ (Sca1$^+$) and ER$^-$ (Sca1$^-$CD49b$^+$) subpopulations.

(C) Magnified images of the selected areas in Figure 1D. Scale bars, 10 μm.

(D) K8 and K14 immunostaining of *Sox9-CreER^{T2}; R26R-tdTomato* lineage tracing clones at the indicated time points (left), and number of clones of the indicated cell types at 5-6 weeks after TAM injection at 25 μg/mouse (right). Scale bars, 50 μm.

(E) Experimental schedule (left) for lineage-tracing studies shown in Figure S1E to S1G and a representative tdTomato whole-mount image (right) of *Sox9-CreER^{T2}; R26R-tdTomato* mammary glands 25 weeks after the tamoxifen treatment on P14.

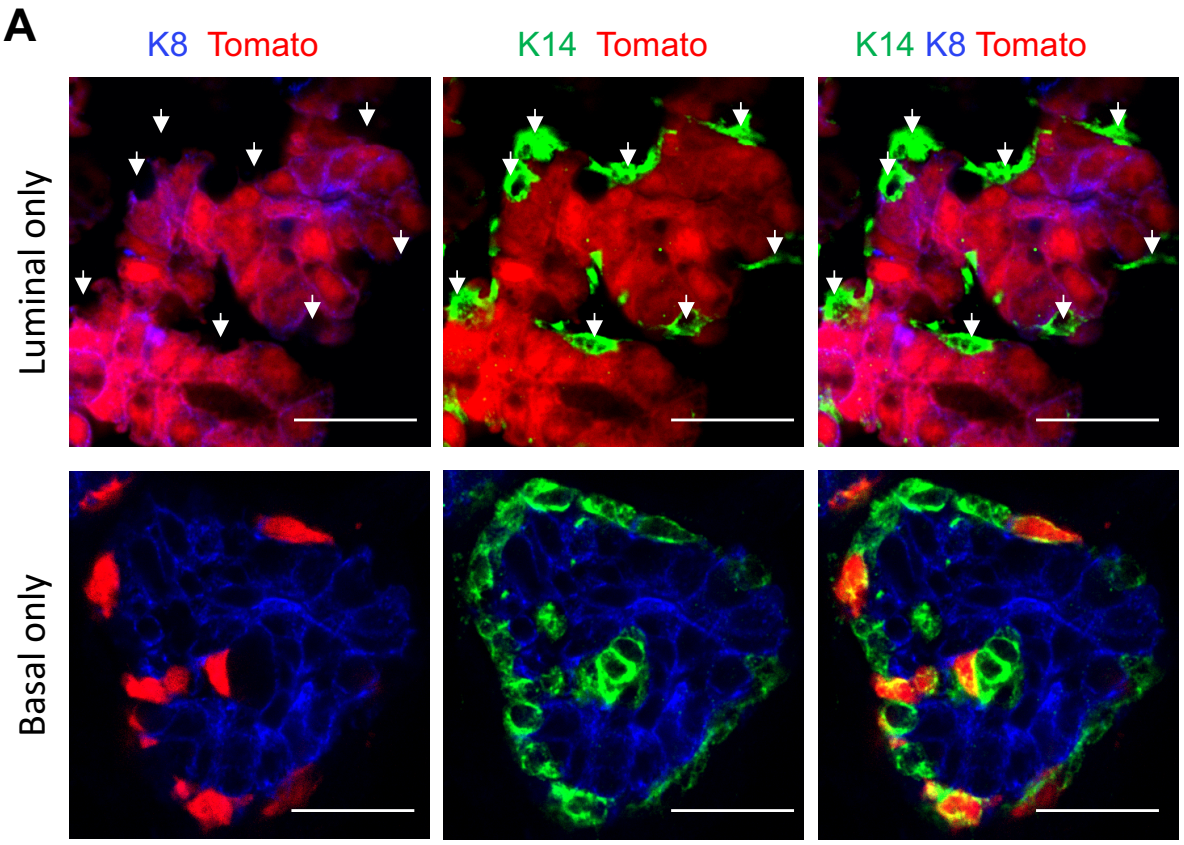
(F) Percentages of tdTomato$^+$ cells in the basal and luminal populations at indicated time points, as determined by flow cytometry (mean ± SEM, n = 3-5).

(G) Percentages of tdTomato$^+$ cells in Sca1$^+$ or Sca1$^-$ luminal population at 25 weeks after the tamoxifen treatment, as determined by flow cytometry (mean ± SEM, n = 4).

(H) SOX9 and ER immunostaining of mouse and human mammary gland sections. A representative of 3 human samples is shown. Note that the SOX9 and ER staining is mutually exclusive. Yellow arrowheads point to SOX9$^+$ luminal cells and arrows point to SOX9$^+$ basal cells. The SOX9 level in basal cells tend to be lower than that of luminal cells.

**P < 0.01, ***P<0.001, and ****P < 0.0001.

Figure S2, Related to Figure 2



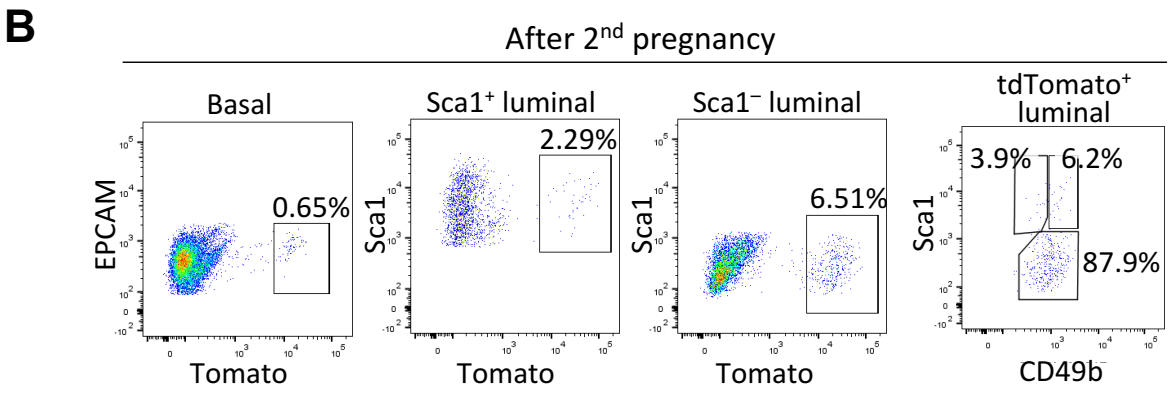


Figure S2, related to Figure 2.

(A) Magnified fields of Figure 2D to shown either luminal or basal cells were labelled with tdTomato in each alveolus. Arrows point to $K14^+$ and $tdTomato^-$ cells. Scale bard, 25 μm.

(B) Representative flow cytometric profiles of basal and luminal populations ≥3 weeks post weaning of the last litter. Most of $tdTomato^+$ luminal cells were in the $Sca1^-$ population.

Figure S3, Related to Figure 3

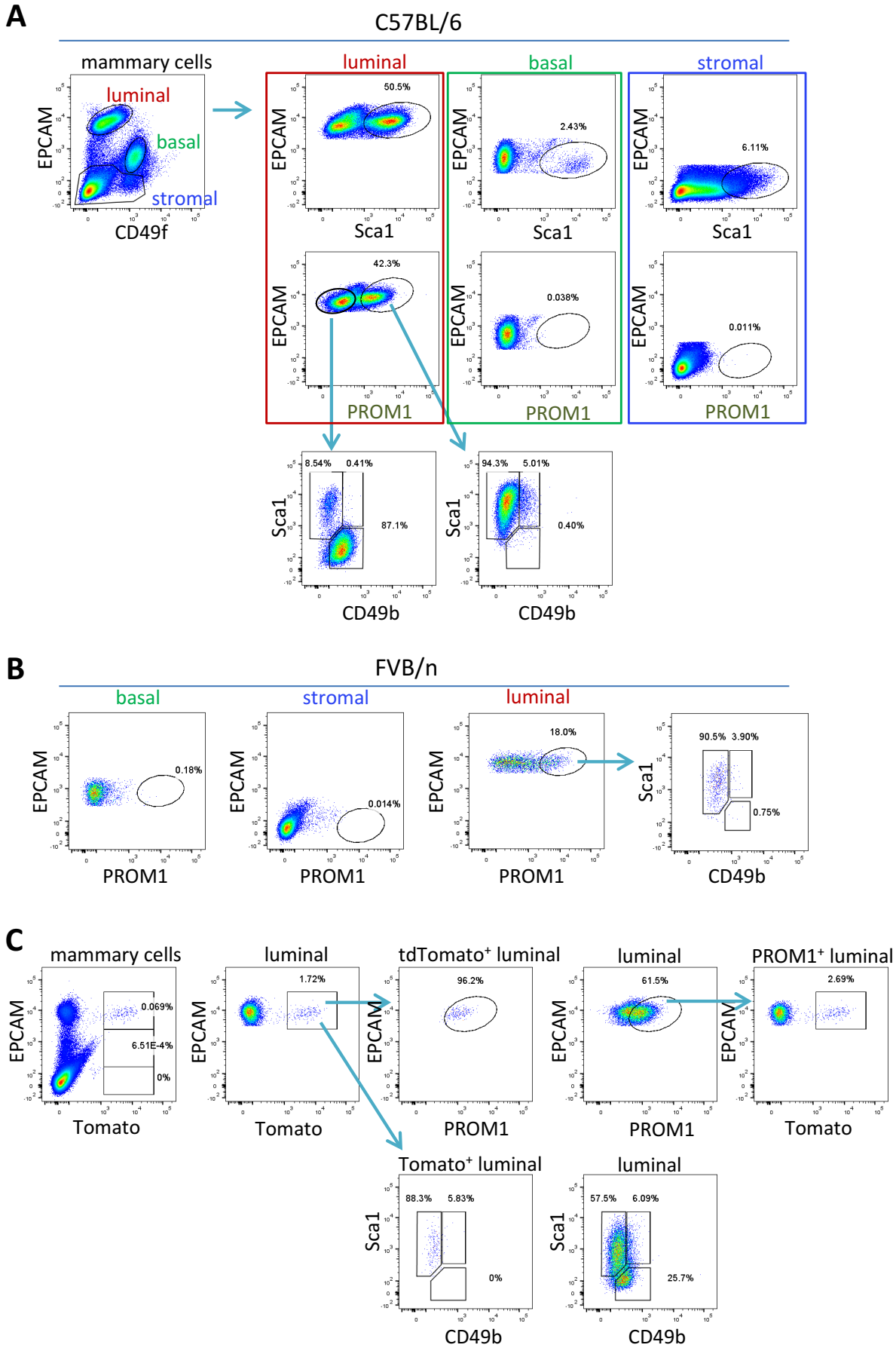


Figure S3, related to Figure 3.

(A) Flow cytometric profiles showing the expression patterns of Sca1 and PROM1 in different mammary cell types of C57BL/6 mice. PROM1 is specifically expressed in the $Sca1^{+}$ (ER^{+}) luminal cells, while Sca1 is also expressed in subsets of basal and stromal cells.

(B) PROM1 expression profiles in mammary epithelial cell populations of FVB/n mice.

(C) Flow cytometric profiles showing the immunophenotype of tdTomato-labeled cells 2 days after TAM injection at P28 days. These cells are $PROM1^{+}$ and in the $Sca1^{+}$ (ER^{+}) luminal population, demonstrating the specificity of *Prom1-CreER^{T2}*. About 2% $PROM1^{+}$ cells were labeled with tdTomato by a single dose of 1.5 mg TAM (see also Figure 3C).

Figure S4, Related to Figure 4

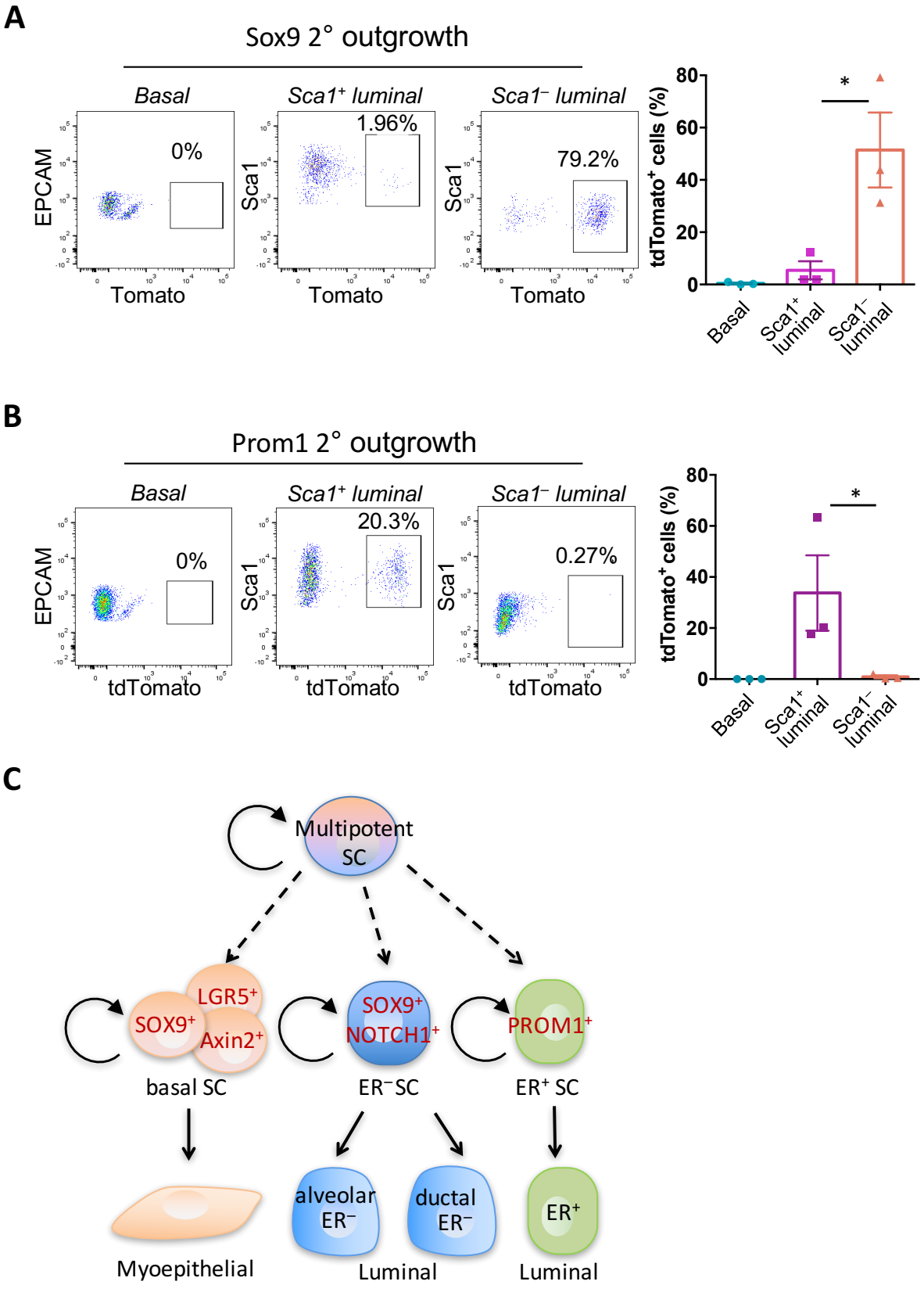


Figure S4, related to Figure 4.

(A) Representative flow cytometric profiles showing frequency of tdTomato$^+$ cells in the indicated cell types of *Sox9-CreER^{T2}; R26R-tdTomato* secondary outgrowths. Quantification was shown on the right (mean ± SEM, n = 3).

(B) Representative flow cytometric profiles showing frequency of tdTomato$^+$ cells in the indicated cell types of *Prom1-CreER^{T2}; R26R-tdTomato* secondary outgrowths. Quantification was shown on the right (mean ± SEM, n = 3). *P < 0.05.

(C) A revised mouse mammary epithelial differentiation hierarchy. Our results suggest that ER$^+$ and ER$^-$ luminal cells are two independent lineages that are maintained by distinct long-lived stem cell types. PROM1$^+$ cells specifically generate ER$^+$ cells, and SOX9$^+$ cells maintain ER$^-$ ductal and alveolar luminal cells. These luminal SOX9$^+$ cells are likely to express NOTCH1 as well. Furthermore, a separate population of basal-specific SOX9$^+$ cells and possibly other basal stem cells contribute to the maintenance of basal cells. The exact relationship of these cells remains to be defined in future. Multipotent stem cells are also likely to contribute to postnatal mammary gland development and maintenance. Future studies are also needed to determine the exact relationship between multipotent and lineage-restricted stem cells.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Mammary cell flow cytometry

Mammary fat pads were minced and then digested with 300 units/ml collagenase type 3 (Worthington Biochemical) and 10 μg/ml DNase (Roche or Worthington Biochemical) in the DMEM/F12 medium at 37°C for 2 hours. The digested samples were washed with PBS and spun down at 500 g for 5 minutes to enrich mammary epithelial organoids. The organoids were further dissociated with 0.05% trypsin-EDTA at 37°C for 10 minutes, and then 1 unit/ml dispase (Worthington Biochemical) plus 100 μg/ml DNase at 37°C for 5 minutes. The dissociated cells were then filtered through 40 μm cell strainers to obtain single cell suspension. Cells were labeled with fluorophore-conjugated antibodies against CD49f (clone GoH3), EPCAM (clone G8.8), CD49b (clone HMα2), Sca1 (clone D7) and PROM1 (clone 13A4). All antibodies were from BioLegend, eBioscience or BD Pharmingen, and used at 1:100. The stained cells were then fixed in 2% paraformaldehyde. Multiparameter flow cytometric analysis were performed on a LSRII equipped with FACS Diva 6.1 software (BD Biosciences) and analyzed with the FlowJo software (Tree Star).

Mammary tissue immunostaining

Dissected mammary glands were fixed with 4% paraformaldehyde, washed with PBS, equilibrated with 30% sucrose, embedded in OCT and then cut into 10-12 μm sections. For immunofluorescence, tissue sections were stained with primary antibodies against SOX9 (Millipore AB5535), cytokeratin 8 (Developmental Studies Hybridoma Bank, clone Troma1), cytokeratin 14 (BioLegend, 905301), ER (Santa Cruz, clone MC-20) and PR (ThermoFisher, clone SP2), and then with anti-rabbit or anti-rat secondary antibodies conjugated with Alexa fluor 488 or Alexa fluor 647. Nuclei were stained with DAPI. Minimum 100 $tdTomato^+$ cells were scored to determine their ER or PR status, except Figure 3E (52 cells).

Whole-mount and Confocal microscopy

For the whole-mount analysis, freshly dissected mammary glands were squeezed between two coverslips and imaged using a Nikon Plan UW 2 × objective. Confocal images were acquired using an Axio Examiner D1 microscope (Zeiss) equipped with a Yokogawa CSU-X1 confocal scan head and analyzed using the SlideBook software (Intelligent Imaging Innovations).