Supporting Information

dos Santos et al. 10.1073/pnas.1303919110

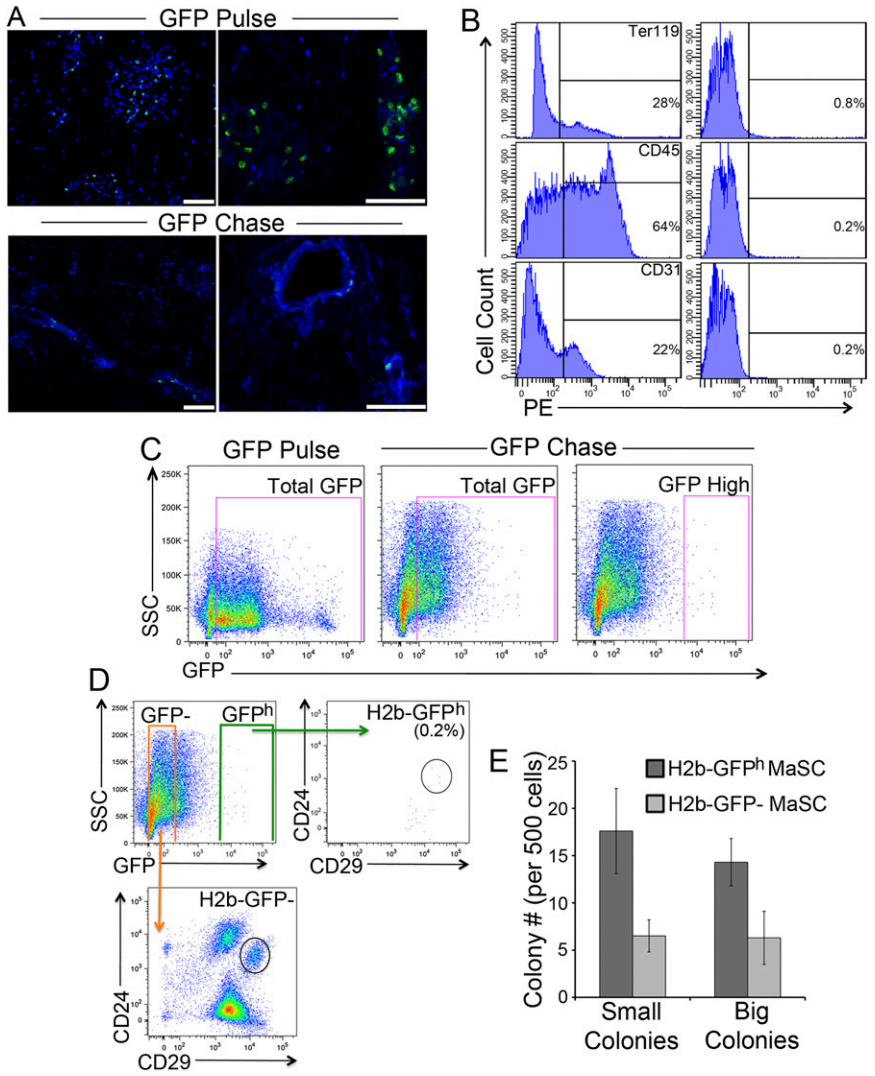

Fig. S1. Characterization of H2b-GFP mammary gland label-retaining cells. Effects of doxycycline diet on K5tTa-H2b-GFP transgenic mouse mammary glands. (*A*) Paraffin-embedded sections with DAPI nuclear staining and anti-GFP antibody. (*B*) Lineage depletion strategy. FACS analysis showing removal of PE-stained red blood cells (Ter119$^+$ cells), white blood cells (CD45$^+$ cells), and endothelial cells (CD31$^+$ cells) after magnetic bead lineage depletion. (*C*) H2b-GFP$^+$ cells gating strategy. Lin$^-$ mammary gland cells were first selected according to GFP expression (GFP$^-$ and GFP$^+$) and further analyzed according to anti-CD24 and anti-CD29 staining as displayed in Fig. 2. (*D*) FACS sorting strategy for transplantation assays. Lin$^-$ GFP chase cells, stained with 7-ADD for dead cell exclusion, were divided based on GFP expression, H2b-GFP$^-$ mammary gland stem cells (MaSCs; CD24$^+$CD29^{h}GFP$^-$), and H2b-GFP^h MaSCs (CD24$^+$CD29^{h}GFP^h) and either transplanted into cleared fat pads of prepubescent female mice or (*E*) carried through to colony-forming assays.

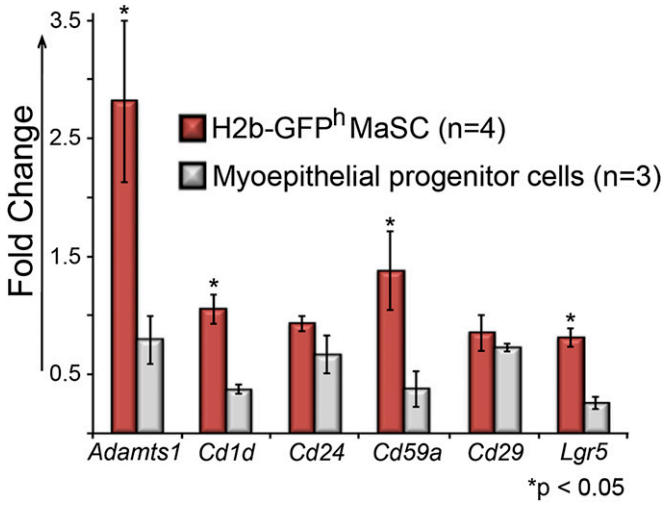

Fig. S2. Quantitative RT-PCR validation of mammary gland transcriptome. Lin$^-$ mammary gland cells (n = 10) were sorted into lyses buffer for quantitative PCR using Taqman probes. *Cd24* and *Cd29* mRNAs were included as controls. Samples were normalized to the levels of *Gapdh* expression. Error bars represent SD among replicated. *P < 0.05 by the t test.

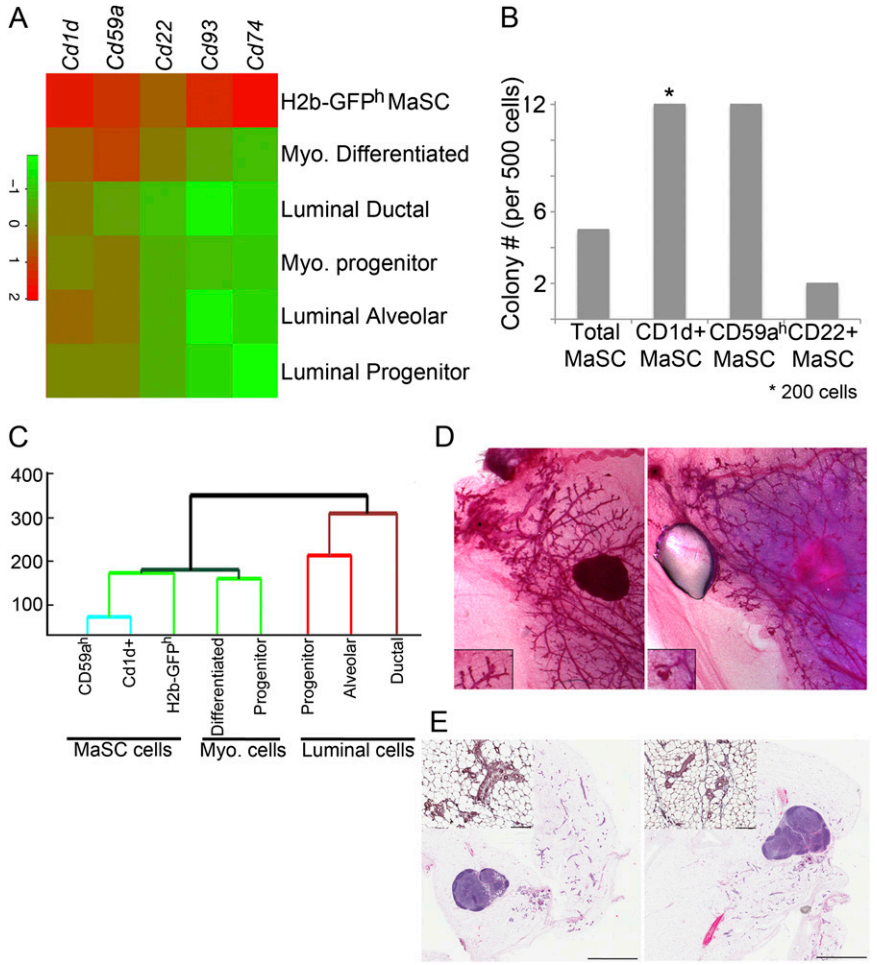

Fig. S3. Identification of MaSC cell surface markers. (*A*) Heat map of cell surface markers expression across all mammary gland cell types profiled. Those cell surface markers shown are the most abundantly expressed within the H2b-GFP^h MaSCs. (*B*) MaSC markers colony-forming assay. Cells were sorted using Cd24$^+$ Cd29^h alone (total MaSC) or Cd24$^+$Cd29^h plus one of three markers: CD1d (CD1d MaSC), CD59a (CD59a^h MaSC), or CD22 (CD22 MaSC). *Two hundred cells. (*C*) Expression dendogram for the top most abundantly expressed genes across all mammary gland cells, including CD1d MaSCs and CD59a^h MaSCs. (*D*) Representative whole-mount images from mammary glands injected with Cd1d$^+$ MaSCs. *Scar tissue from cell injection. (*E*) Representative images from paraffin-embedded sections of Cd1d$^+$ MaSC-injected glands stained with H&E and anti-GFP immunohistochemistry (IHC) (*Left Inset*). (Scale bar: 2 mm; *Left Inset*, 100 μm.)

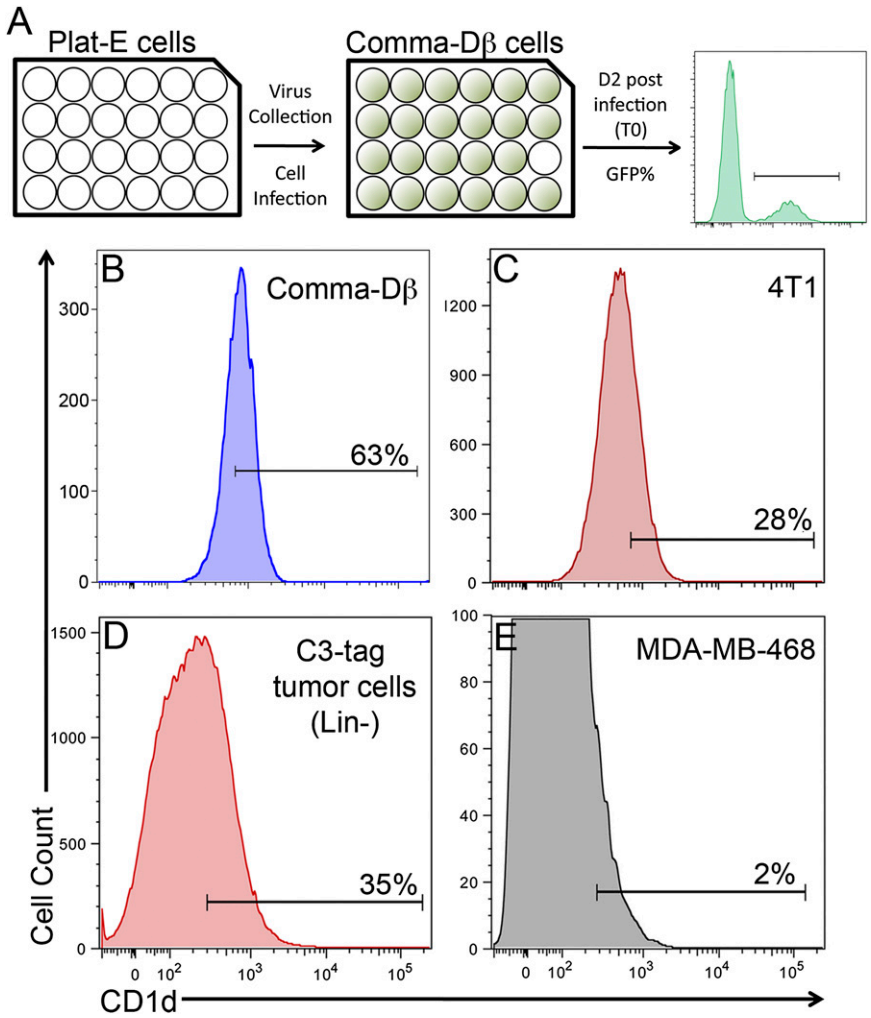

Fig. S4. Mammary gland-focused screen. (*A*) Screen strategy scheme. Plat-E cells were transfected with shRNAs as described in *Materials and Methods*. Comma-Dβ cells were infected with the virus supernatant for 20 h; 48 h postinfection, GFP percent was quantified using the MACSQuant Cell Analyzer (Miltenyi Biotech). T0 represents the GFP percent on day 2 postinfection, and T3 represents the GFP percent on day 12 d postinfection. $CD1d^+$ cells in mouse cell lines (*B*) Comma-Dβ cells, (*C*) 4T1 mouse breast cancer cells, and (*D*) C3-tag breast cancer model primary cells. (*E*) $CD1d^+$ in the human cell line MDA-MB-468.

Dataset S1. Mammary reconstitution unit (MRU) frequency in H2b-GFP^h MaSC

Dataset S1

Reconstituted mammary glands harvested 12 wk postinjection of either H2b-GFP^h MaSCs or H2b-GFP^- MaSCs. A minimum of 25 outgrowths is required to be considered a reconstituted gland. MRU frequency was estimated using the ELDA algorithm.

Dataset S2. Mammary gland pathway analysis

Dataset S2

Top differentially expressed genes of all mammary gland cell types were analyzed for pathway enrichment and molecular functions using Ingenuity Pathways Analysis (Ingenuity Systems). A minimum of 50 genes per cell type was analyzed.

Dataset S3. MRU frequency in CD1d+ MaSC

Dataset S3

Total MaSC cells ($CD24^{+}CD29^{h}$) and CD1d MaSC cells ($CD24^{+}CD29^{h}CD1d^{+}$) were isolated from the H2b-GFP transgenic mouse off doxycycline diet (GFP pulse). Reconstituted mammary glands harvested 12 wk postcell injection. A minimum of 25 outgrowths is required to be considered a reconstituted gland. MRU frequency was estimated using the ELDA algorithm.