Cell-State Transitions Regulated by SLUG Are Critical for Tissue Regeneration and Tumor Initiation

Sarah Phillips,[1,2] Aleix Prat,[3] Maja Sedic,[1,2] Theresa Proia,[6] Ania Wronski,[1,2] Sohini Mazumdar,[1] Adam Skibinski,[1,2] Stephanie H. Shirley,[4] Charles M. Perou,[5] Grace Gill,[1] Piyush B. Gupta,[6,7,*] and Charlotte Kuperwasser[1,2,*]

[1]Department of Developmental, Molecular and Chemical Biology, Sackler School of Graduate Biomedical Sciences, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
[2]Molecular Oncology Research Institute, Tufts Medical Center, 800 Washington Street, Boston, MA 02111, USA
[3]Translational Genomics Group, Vall d'Hebron Institute of Oncology (VHIO), Barcelona 08035, Spain
[4]Department of Molecular Carcinogenesis, Science Park - Research Division, The University of Texas MD Anderson Cancer Center, Smithville, TX 78957, USA
[5]Department of Genetics and Pathology, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27514, USA
[6]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge MA, 02142, USA
[7]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: pgupta@wi.mit.edu (P.B.G.), charlotte.kuperwasser@tufts.edu (C.K.)
http://dx.doi.org/10.1016/j.stemcr.2014.03.008

SUMMARY

Perturbations in stem cell activity and differentiation can lead to developmental defects and cancer. We use an approach involving a quantitative model of cell-state transitions in vitro to gain insights into how SLUG/SNAI2, a key developmental transcription factor, modulates mammary epithelial stem cell activity and differentiation in vivo. In the absence of SLUG, stem cells fail to transition into basal progenitor cells, while existing basal progenitor cells undergo luminal differentiation; together, these changes result in abnormal mammary architecture and defects in tissue function. Furthermore, we show that in the absence of SLUG, mammary stem cell activity necessary for tissue regeneration and cancer initiation is lost. Mechanistically, SLUG regulates differentiation and cellular plasticity by recruiting the chromatin modifier lysine-specific demethylase 1 (LSD1) to promoters of lineage-specific genes to repress transcription. Together, these results demonstrate that SLUG plays a dual role in repressing luminal epithelial differentiation while unlocking stem cell transitions necessary for tumorigenesis.

INTRODUCTION

In renewable tissues such as the hematopoietic system, skin, and intestine, multipotent stem cells serve as a reservoir of cells that are called upon to maintain tissue homeostasis and function (Blanpain and Fuchs, 2006; Tesori et al., 2013; Toma et al., 2001; Barker et al., 2008; Weissman, 2000). These stem cells have been implicated as precursors to cancer, presumably due to their long-term persistence and high self-renewing capabilities (Barker et al., 2009; Bonnet and Dick, 1997). However, in other tissues such as the mammary gland, lineage-restricted progenitor cells, as opposed to multipotent stem cells, are responsible for tissue maintenance and homeostasis (Van Keymeulen et al., 2011). When called upon for tissue regeneration, as is the case upon transplantation or injury, these lineage-committed progenitor cells unlock primitive stem cell programs that are not normally required for tissue development or tissue homeostasis (Blanpain et al., 2004; Doupé et al., 2012; Kordon and Smith, 1998; Shackleton et al., 2006; Stingl et al., 2006; van Amerongen et al., 2012; Van Keymeulen et al., 2011). By doing so, these cells acquire properties that make them amenable to cancer initiation (Pacheco-Pinedo et al., 2011; Proia et al., 2011; Schwitalla

et al., 2013; Youssef et al., 2010, 2012). However, the molecular mechanism by which committed progenitor cells access latent stem cell programs is not well understood.

Previously, we showed that the transcription factor SLUG is an important regulator of mammary epithelial lineage commitment and differentiation (Proia et al., 2011). Recent studies have also shown that SLUG is necessary for the mammary stem cell state (Guo et al., 2012). However, SLUG-deficient mice develop mammary glands, and transplantation of tissue fragments from these mice were able to fully regenerate functional mammary glands; this suggests that SLUG might be dispensable for stem cell activity (Nassour et al., 2012). Thus, the precise role of SLUG in mammary stem and progenitor cell dynamics remains unclear.

The ability to study stem cell-state transitions and progenitor cell dynamics in vivo is challenging; even when cell-state markers are available, most transitions are short-lived and difficult to capture. We sought to gain insights into how SLUG controls stem cell activity in normal disease-free mammary epithelial cells by using a recently developed and validated quantitative model to predict cell-state transition rates in vitro (Gupta et al., 2011). Using this approach, we were able to (1) infer differences in cell-state transition probabilities between wild-type (WT) and

CrossMark
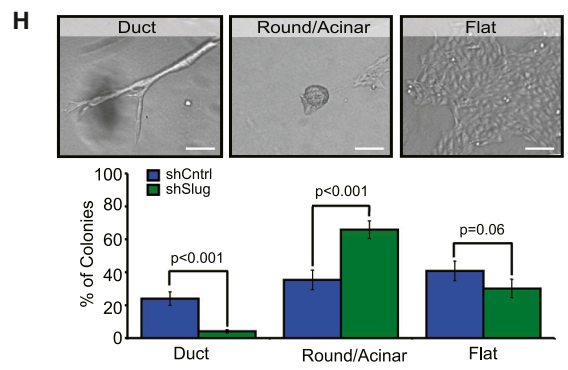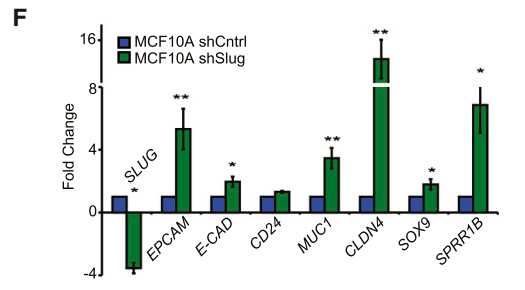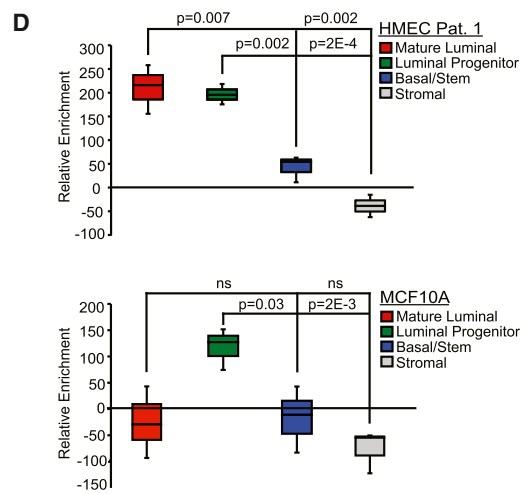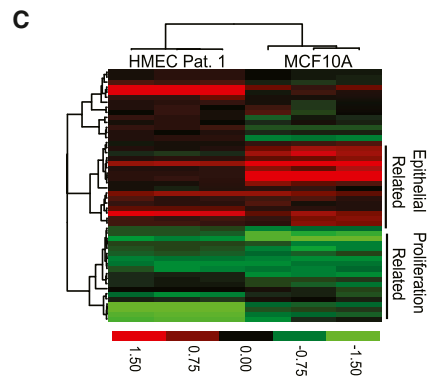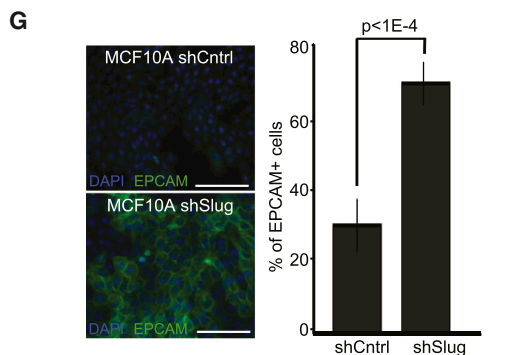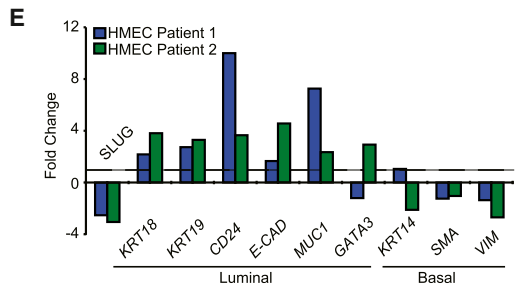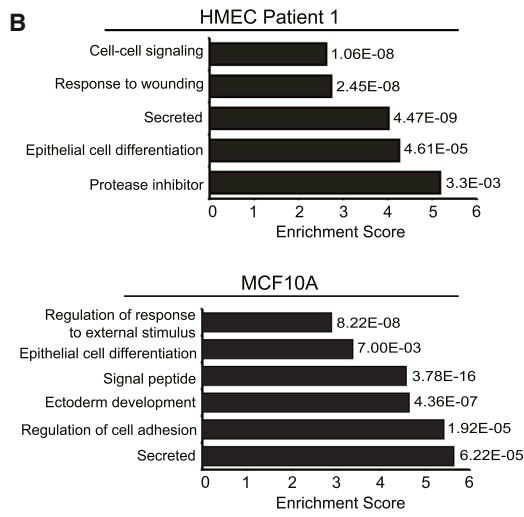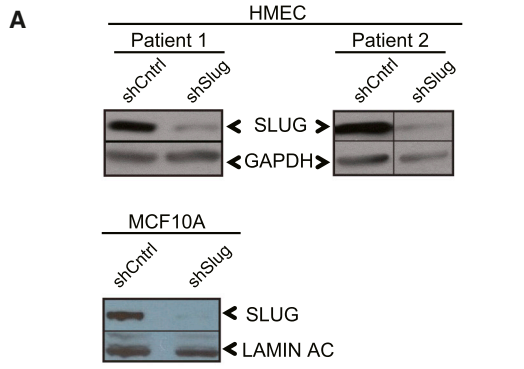
(legend on next page)

SLUG-deficient mammary epithelial cell populations, (2) accurately predict the in vivo phenotype associated with SLUG deficiency, and (3) provide insights into how SLUG inhibition influences progenitor cell dynamics to ultimately disrupt cellular differentiation as well as tissue homeostasis, regeneration, and tumor initiation.

RESULTS

SLUG Inhibits Differentiation of Breast Epithelial Cells

SLUG could be regulating stem cell activity by preventing proliferation, by inhibiting differentiation, or by affecting cell-state transitions between stem cells and lineage-committed cells. To begin to distinguish between these possibilities, we used lentiviral-mediated short hairpin RNA to knockdown *SLUG* in human basal progenitor cell lines: human telomerase reverse transcriptase (hTERT) immortalized mammary epithelial cells (HMECs) derived from two different patient samples and the spontaneously immortalized MCF10A breast epithelial cell line (Figure 1A). In agreement with our previous findings (Proia et al., 2011), *SLUG*-inhibited (shSlug) HMECs adopted a tightly packed cobblestone-like appearance compared to control cells (shControl), consistent with a more differentiated epithelial cell phenotype (Figure S1A available online).

To more comprehensively define the phenotype associated with *SLUG* inhibition in mammary progenitor cells, we performed global gene expression analysis on shControl and shSlug HMEC and MCF10A cells (Supplemental Experimental Procedures). Gene Ontology (GO)

analysis revealed significant enrichment of genes involved in epithelial cell differentiation, ectoderm development, secretion, and the regulation of cell adhesion in *SLUG*-inhibited HMECs and MCF10A cells (Figure 1B). To further assess breast epithelial differentiation, microarray data were analyzed using the recently described Genomic Differentiation Predictor (Prat et al., 2010) and the 50 genes of the PAM50 breast cancer intrinsic subtype predictor (Parker et al., 2009). As predicted, *SLUG* inhibition led to significant changes in the expression of epithelial-related genes in both HMECs and MCF10A cells (Figure 1C; Figure S1B). Notably, there was increased expression of luminal genes (e.g., *FOXA1*, *BAG1*, and *NAT1*) and decreased expression of proliferation genes (e.g., *MKI67*, *CDC6*, and *CDCA1*) in both cell lines (Figure 1C; Gene Expression Omnibus [GEO] accession number GSE54735). Furthermore, when probed for signatures of lineage-specific mammary epithelial subpopulations (Lim et al., 2009), *SLUG* inhibition resulted in a significant enrichment in the luminal progenitor signature in both cell lines, while HMECs also displayed a significant enrichment in the mature luminal profile (Figure 1D). Transcriptional analysis using quantitative RT-PCR further confirmed the strong induction of luminal differentiation genes, with a concomitant decrease in the expression of genes associated with basal/stem differentiation following *SLUG* knockdown (Figures 1E and 1F). Consistent with these data, immunofluorescence analysis for the luminal marker EPCAM revealed a significant increase in its expression upon *SLUG* knockdown in MCF10A cells (Figure 1G).

Differentiation and morphogenesis of control and shSlug MCF10A cells was also evaluated using a 3D

Figure 1. SLUG Inhibits Differentiation of Breast Epithelial Cells

(A) Western blot analysis of SLUG expression in shControl and shSlug HMECs and MCF10A cells.

(B) Gene Ontology biological process categories upregulated following *SLUG* inhibition in mammary epithelial cells: HMECs (patient 1) and MCF10A cells. The DAVID Functional Annotation Tool (Huang da et al., 2009) was used to identify categories with an enrichment score >2; the enrichment score and p value of genes differentially expressed in the microarray are shown.

(C) Hierarchical clustering heatmap of shSlug HMEC (patient 1) and shSlug MCF10A cells compared to shControl cells (n = 3 for each cell line) using the 50-gene set of the PAM50 breast cancer intrinsic subtype predictor. No gene centering was performed.

(D) Relative enrichment of mature luminal, luminal progenitor, basal/stem, and stromal signatures (defined by Lim et al., 2009) in shSlug HMEC (patient 1) and MCF10A cells compared to shControl cells.
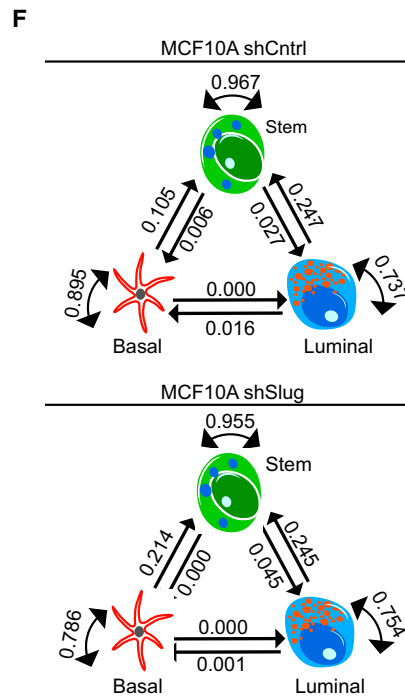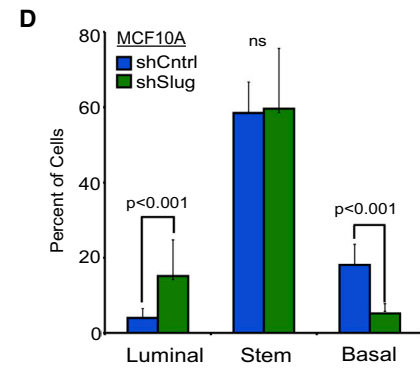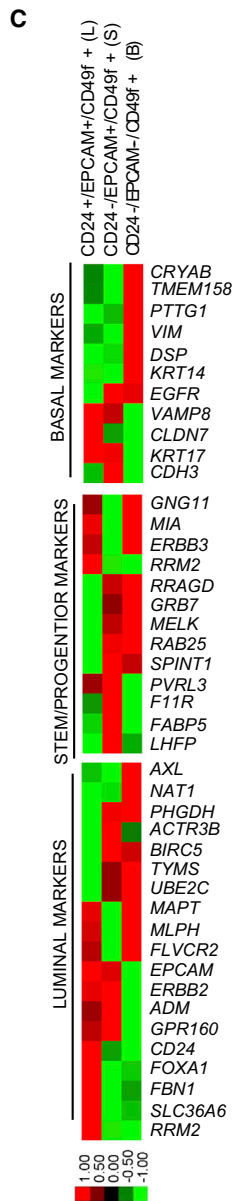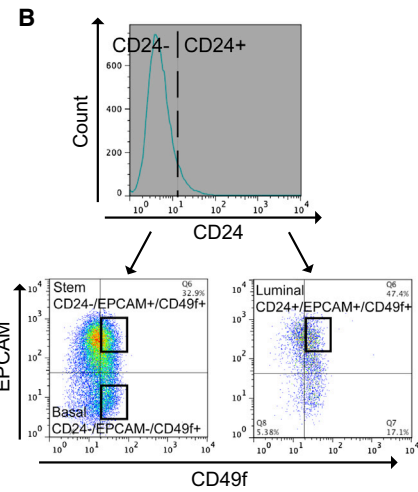
(E) Relative mRNA expression levels (normalized to *GAPDH*) of luminal and basal markers in two different patient-derived HMEC lines following *SLUG* inhibition. Genes differentially expressed in the shSlug cells compared to the control cells (dashed line) are plotted.

(F) Quantitative real-time PCR analysis of luminal marker expression (normalized to *GAPDH*) in MCF10A cells following *SLUG* inhibition. Genes differentially expressed in shSlug cells compared to shControl cells are plotted. Bars represent the fold change ± SD of three independent experiments, and p values were calculated by Student's two-tailed t test. *p < 0.05, **p < 0.005.

(G) Immunofluorescent staining for the luminal marker EPCAM in MCF10A shControl and shSlug cells. Representative images are shown and quantification is depicted in the graph below. Data represent the mean ± SEM of three independent experiments, and p values were calculated by Student's two-tailed t test. Scale bars represent 100 μm.

(H) Phase-contrast images of MCF10A cells grown in 3D culture. Quantification of the percentage of round, ductal, and flat colonies formed by shControl and shSlug MCF10A cells. Data represent the mean ± SD of three independent experiments, and p values were calculated by Student's two-tailed t test.

See also Figure S1.

(legend on next page)

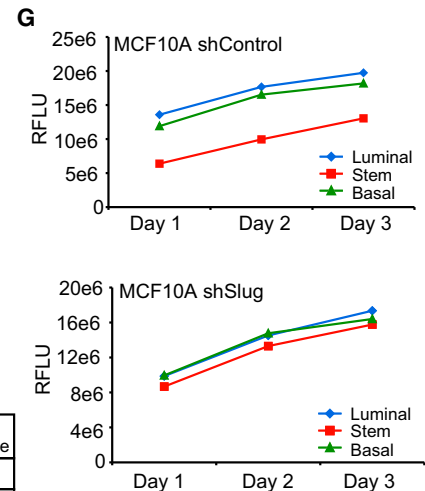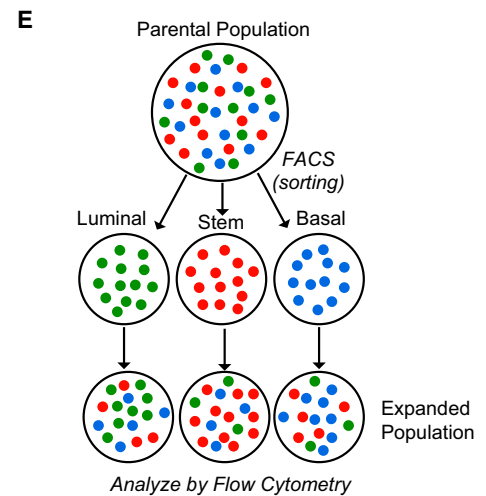collagen/Matrigel assay. In these 3D cultures, luminal progenitor cells primarily differentiate into round alveolar colonies while basal progenitor cells form distinct flat colonies and branching ductal colonies (Keller et al., 2012). As expected, *SLUG* inhibition led to a significant increase in the formation of round alveolar colonies and a significant decrease in the number of flat and ductal structures following growth on collagen (Figure 1H). Together, these data fortify SLUG as a critical repressor of breast epithelial differentiation.

Modeling the Effect of SLUG on Cell-State Transition Rates

While the above findings clearly highlight SLUG's influence on inhibiting mammary epithelial cell (MEC) differentiation, they do not divulge the dynamic cellular mechanisms that may be affecting cell-state transitions between stem cells and lineage-committed cells. To begin to address this, we examined breast epithelial cell-state proportions by fluorescence-activated cell sorting following *SLUG* inhibition. In cultured cells, luminal cells display EPCAM$^+$ and CD24$^+$ immunophenotypes, while basal cells exhibit EPCAM$^-$/CD49f$^+$ and CD24$^-$/CD49f$^+$ immunophenotypes (Eirew et al., 2008; Keller et al., 2010, 2012; Lim et al., 2009). Consistent with the gene expression data, *SLUG* inhibition resulted in increased proportions of CD24$^+$ and EPCAM$^+$ luminal cells and concomitant decreased proportions of CD24$^-$/CD49f$^+$ and EPCAM$^-$/CD49f$^+$ basal/progenitor cells (Figure 2A; Figures S2A–S2E).

To provide dynamic insights into how *SLUG* deficiency impacts these cell-state shifts, we employed a quantitative Markov model that has recently been used to study the dynamics of cell-state proportions in human breast cancer cell lines (Gupta et al., 2011). With this model, data gener-ated from short-term in vitro studies can be used to deduce the probability of cellular transitions between any two phenotypic (e.g., differentiation) states. The model can also be used to predict how genetic perturbations affect cell-state interconversion rates; this allows us to infer, on a dynamic level, how changes in gene function disrupt cellular equilibrium.

We characterized the three main mammary epithelial cell states isolated from MCF10A cells (Keller et al., 2010): luminal (L) (CD24$^+$/EPCAM$^+$/CD49f$^+$), basal (B) (CD24$^-$/EPCAM$^-$/CD49f$^+$), and stem (S) (CD24$^-$/EPCAM$^+$/CD49f$^+$) (Figure 2B). As expected, expression analysis of 105 selected genes that define different cell-differentiation states confirmed that this sorting technique successfully isolates populations of cells in each of the indicated cell-differentiation states (Figure 2C). In addition, the shSlug cells exhibited a higher percentage of cells in the luminal state (shControl = 4% versus shSlug = 15%) and a lower percentage of cells in the basal state (shControl = 23% versus shSlug = 7%), with no appreciable change in the stem state (shControl = 73% versus shSlug = 78%) compared to the control population (Figure 2D).

Based on these findings, L, B, and S sorted subpopulations were allowed to expand in culture for 4 days, at which time the proportion of cells in the luminal, basal, and stem states were reanalyzed by flow cytometry (Figure 2E). For each sorted subpopulation, there was an evident shift in cell-state proportions, suggesting that cells were interconverting between states (Figure 2F). In fact, *SLUG* inhibition reduced the plasticity of cell transitions into the basal state (L→B reduced 16-fold, S→B no longer present). In contrast, *SLUG* inhibition increased the plasticity of cell transitions from the stem state into the luminal state (S→L increases 1.7-fold) and from the basal state into the

Figure 2. Loss of SLUG Alters Stochastic Cell-State Interconversions

(A) Flow cytometry analysis of luminal (EPCAM and CD24) and basal (CD49f) cell-surface markers in shControl (n = 3) and shSlug (n = 3) MCF10A cells. Representative dot plots for EPCAM versus CD49f expression are shown. Data represent the mean ± SD, and p values were calculated by Student's two-tailed t test.

(B) Flow cytometry gating strategy for isolating luminal (CD24$^+$/EPCAM$^+$/CD49f$^+$), stem (CD24$^-$/EPCAM$^+$/CD49f$^+$), and basal (CD24$^-$/EPCAM$^-$/CD49f$^+$) subpopulations from a heterogeneous parental MCF10A population.

(C) Expression of various differentiation-related genes in luminal, basal and stem subpopulations of MCF10A cells using the nCounter platform (Nanostring Technologies) (Geiss et al., 2008). The heatmap displays selected genes that showed the greatest differential expression across the three cell states. Data presented have been gene median centered.

(D) Flow cytometry analysis showing the proportion of cells in the luminal, basal, and stem states in MCF10A shSlug and shControl parental populations immediately prior to sorting. Data represent the mean ± SD of six independent experiments, and p values were calculated by Student's two-tailed t test.

(E) Schematic of the experimental procedure used to calculate cell-state transition probabilities in MCF10A shSlug and shControl cells.

(F) Predicted cell-state transition probabilities for shControl and shSlug MCF10A cells. The table below summarizes the frequency of cell-state transitions in shControl and shSlug cells. Each transition's fold change following *SLUG* inhibition is also displayed; only transitions that changed >1.5-fold following *SLUG* inhibition are shown. NA, not applicable.

(G) In vitro proliferation curves of purified subpopulations (luminal, stem, and basal) of MCF10A shControl and shSlug cells. RFLU, relative fluorescent light unit.

See also Figure S2.
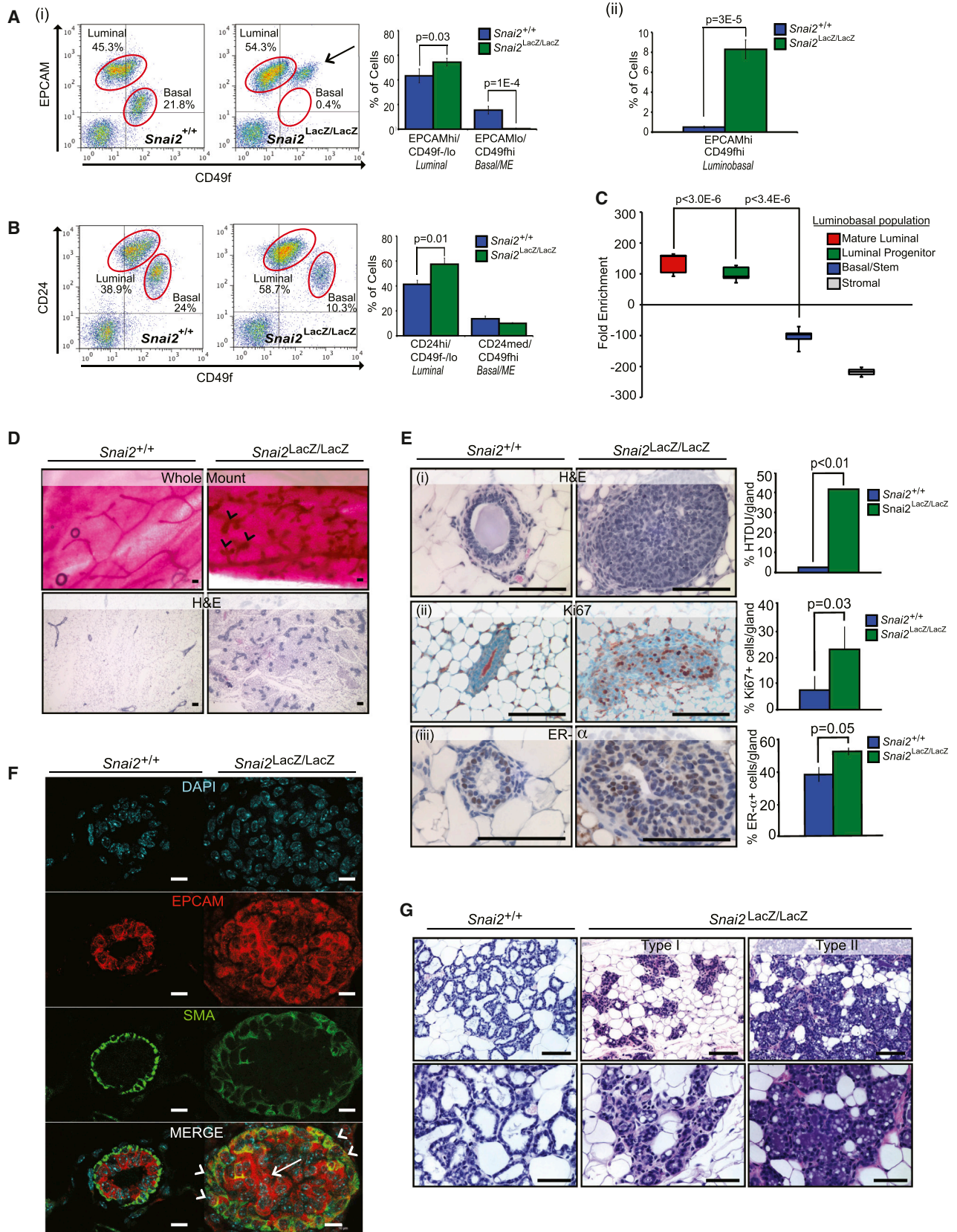
(legend on next page)

stem state (B→S increases 2-fold; Figure 2F). Because the S→B transitions were eliminated in shSlug cells, the increased proportion of luminal cells was effectively due to increased S→L transitions. Given that proliferation rates did not vary between the luminal, basal, and stem states of either parental line (Figure 2G), we could be confident that the observed transitions were not a consequence of differential growth rates inherent to each differentiation state, but rather were due to changes in cellular plasticity. Collectively, these findings, along with those above, reveal that *SLUG* inhibition affects cellular differentiation in addition to cellular plasticity, predominantly by altering transitions arising from stem and basal cells; together, these changes result in an unexpected accumulation of luminal epithelial cells.

Loss of SLUG Promotes Aberrant Differentiation and Mammary Progenitor Cell-State Transitions In Vivo

To determine if the above in vitro predictions of SLUG deficiency are observed in vivo, we examined the mammary glands of 16-week-old $Snai2^{LacZ/LacZ}$ mice; at this age, steady-state homeostasis of luminal and basal progenitor cells is well established. In these mice, the zinc-finger region of the *Slug* gene has been replaced by a β-galactosidase gene, resulting in the production of a SLUG-β-galactosidase fusion protein. The SLUG portion of this protein is nonfunctional, as it lacks the zinc-finger region; thus, mice homozygous for the $Snai2^{LacZ/LacZ}$ allele are functional SLUG knockout mice (Jiang et al., 1998; Parent et al., 2004).

Consistent with the increased proportions of luminal cells observed in vitro, $Snai2^{LacZ/LacZ}$ mammary glands exhibited a significant increase in the percentage of $CD24^{hi}/CD49f^{-/lo}$ and $EPCAM^{hi}/CD49f^{-/lo}$ luminal cells compared to WT controls (Figures 3Ai and 3B). This increase in luminal cells by flow cytometry manifested histologically as the formation of abnormal hyperplastic terminal ductal units, with intraductal micropapillary and solid growth patterns (Figures 3D and 3Ei). Notably, luminal cells, which were both Ki67 and estrogen receptor-α (ER-α) positive, accumulated within the lumens of $Snai2^{LacZ/LacZ}$ glands, consistent with a luminal progenitor state (Figures 3Eii and 3Eiii).

Also consistent with the in vitro findings, flow cytometry analysis of mammary epithelial cells from $Snai2^{LacZ/LacZ}$ mice revealed a significant reduction in the number of $EPCAM^{lo}/CD49f^{hi}$ basal/myoepithelial (ME) cells compared to age-matched WT mice (with a similar, although not statistically significant, downward trend in the $CD24^{med}/CD49f^{hi}$ basal/ME population; Figures 3Ai and 3B). Moreover, the $EPCAM^{lo}/CD49f^{hi}$ basal cell population of $Snai2^{LacZ/LacZ}$ mammary glands exhibited a significant shift toward an $EPCAM^{hi}/CD49f^{hi}$ state; termed the luminobasal state, this population reflected an altered basal state that acquired distinct features of luminal differentiation (Figure 3Aii). Consistent with this, genome-wide expression analysis on sorted luminobasal cells from $Snai2^{LacZ/LacZ}$ mammary glands revealed a significant enrichment of luminal-related epithelial genes, including several luminal cytokeratins and claudin tight junction proteins (GEO accession number GSE55620), compared to WT basal cells ($EPCAM^{lo}/CD49f^{hi}$). Additionally, when probed for signatures of lineage-specific mouse mammary epithelial subpopulations (Lim et al., 2009), the

Figure 3. SLUG Loss Promotes Aberrant Differentiation and Compromises Normal Tissue Function In Vivo
(A and B) Flow cytometry analysis of luminal (CD24hi/CD49f−/lo and EpCAMhi/CD49f−/lo), basal/myoepithelial (ME) (CD24med/CD49fhi and EPCAMlo/CD49fhi), and luminobasal (EPCAMhi/CD49fhi) MEC populations from 16-week nulliparous $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mammary glands. Representative dot plots from a $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mouse are shown. Luminal and basal/ME populations are circled in red. Arrow points to the luminobasal population. Bar charts represent the percent of cells in each population. (A) EPCAM versus CD49f luminal and basal/ME (i) and luminobasal (ii). (B) CD24 versus CD49f. Error bars are ±SEM, and p values were calculated by Student's two-tailed t test. CD24 versus CD49f: $Snai2^{+/+}$, n = 11; $Snai2^{LacZ/LacZ}$, n = 8. EPCAM versus CD49f: $Snai2^{+/+}$, n = 5; $Snai2^{LacZ/LacZ}$, n = 8.
(C) Relative enrichment of mouse mature luminal, luminal progenitor, basal/stem, and stromal signatures (defined by Lim et al., 2009) in the $Snai2^{LacZ/LacZ}$ luminobasal population (EPCAMhi/CD49fhi).
(D) Representative whole-mount and hematoxylin and eosin staining (H&E) images of nulliparous $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mammary glands (mice were 16 weeks). Arrowheads indicate hyperplastic terminal ductal units (HTDUs). Scale bars represent 100 μm.
(E) High-magnification H&E images of nulliparous $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mammary glands (a HTDU structure is depicted) (i). Quantification indicates the percent of HTDUs observed in mammary whole mounts. $Snai2^{+/+}$, n = 5; $Snai2^{LacZ/LacZ}$, n = 5. Immunohistochemical (IHC) staining for Ki67 (ii) and ER-α (iii) in nulliparous $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mammary glands. Staining is quantified to the right and indicates the percent of ER-α- or Ki67-positive cells per gland. $Snai2^{+/+}$, n = 4; $Snai2^{LacZ/LacZ}$, n = 4. All analyses were performed on 16-week-old mice. Scale bars represent 100 μm. Error bars are ±SD, and p values were calculated by Student's two-tailed t test.
(F) Confocal immunofluorescent images of $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mammary glands stained for the luminal marker EPCAM (red) and the myoepithelial marker SMA (green). Nuclei were stained with DAPI (cyan). Arrowheads represent basal cells that are aberrantly expressing EPCAM. The arrow highlights the aberrant filling of the ductal lumen with luminal cells. Scale bars represent 10 μm.
(G) Representative H&E images of mammary glands from $Snai2^{+/+}$ and $Snai2^{LacZ/LacZ}$ mice on day 1 of lactation. Type I and type II phenotypes from $Snai2^{LacZ/LacZ}$ mammary glands are shown. Mice were mated at 16 weeks. Scale bars represent 100 μm.
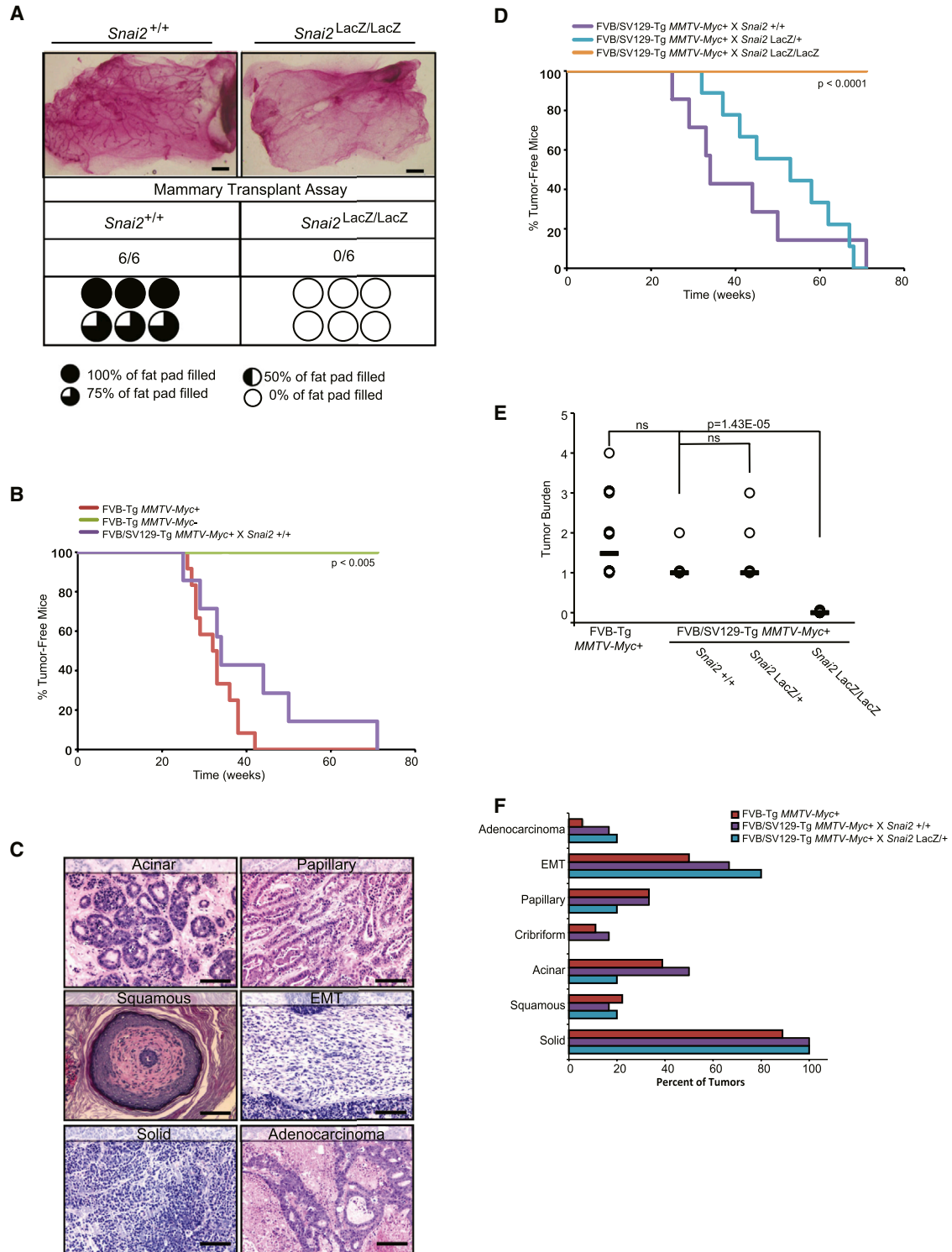
Figure 4. SLUG Loss Impairs Tissue Regeneration and Confers Resistance to Tumorigenesis

(A) Whole-mount analysis of mammary epithelial outgrowths 12 weeks after transplantation. A total of 50,000 MECs isolated from $Snai2^{+/+}$ or $Snai2^{LacZ/LacZ}$ mammary glands (nulliparous, 16 weeks) were transplanted into the cleared fat pads of 3-week-old nonobese diabetic/ severe combined immunodeficiency (NOD/SCID) mice. Scale bars represent 200 μm. The extent to which each outgrowth filled the fat pad is indicated by the "pie chart" symbols.

(legend continued on next page)

luminobasal cells showed significant enrichment in the luminal progenitor and mature luminal signatures (Figure 3C).

Confocal dual-immunofluorescence microscopy analysis for the luminal marker EPCAM and the basal/ME marker SMA was also performed to validate the cellular differentiation states of luminal and basal cells in $Snai2^{LacZ/LacZ}$ mammary glands in situ. The analysis revealed two striking details: (1) the population of aberrant cells filling the ductal lumen of $Snai2^{LacZ/LacZ}$ mammary glands was predominantly of the luminal lineage (EPCAM+), but some aberrantly expressed SMA, and (2) the outer basal/ME layer of $Snai2^{LacZ/LacZ}$ mammary glands contained an uncharacteristic number of basal/ME cells coexpressing both basal and luminal lineage markers, further supporting the notion that SLUG deficiency causes basal cells to acquire luminal features (Figure 3F).

Loss of SLUG Impairs Tissue Function and Regeneration and Confers Resistance to Tumorigenesis

To determine whether the defects in luminal and basal progenitor cells might alter tissue function, 16-week-old female $Snai2^{LacZ/LacZ}$ mutant mice were examined for their ability to undergo pregnancy-induced differentiation and lactation. Although $Snai2^{LacZ/LacZ}$ mutant mice were fertile and could give rise to viable pups with similar efficiencies as WT females, all of the pups nursed by the majority of $Snai2^{LacZ/LacZ}$ mutant mice (five out of seven) died, with pups appearing to have empty stomachs. Surprisingly, examination of mammary tissues from $Snai2^{LacZ/LacZ}$ mutant mice that were unable to nurse their pups revealed two prominent phenotypes: either (1) a block in alveologenesis and failure to undergo lactational differentiation, suggestive of luminal progenitor exhaustion (type I), or (2) apparent normal alveolar differentiation and milk production but evidence of premature involution (type II; Figure 3G). Because $Snai2^{LacZ/LacZ}$ females presented no obvious abnormal maternal nurturing behavior, this latter phenotype suggested that when milk was produced in the $Snai2^{LacZ/LacZ}$ mutant mice, it was not available to neo-nates. Given that basal/ME cells exhibited increased luminal differentiation, we reasoned this failure of milk ejection might be due to impaired function and contractile activity of basal/ME cells. Consistent with this notion, morphological and histological analyses revealed that mammary glands from $Snai2^{LacZ/LacZ}$ dams contained large, dilated ducts engorged with milk compared to WT glands in which the ducts were empty due to efficient milk ejection. Collapsed alveoli and gland fragmentation, an indication of complete secretory tissue regression, were also found in $Snai2$ mutant mammary epithelium (Figure 3F), similar to those resulting from milk stasis when milk ejection is impaired in other contexts (Quarrie et al., 1996). Taken together, our in vivo analysis of $Snai2^{LacZ/LacZ}$ mammary epithelium confirms the in vitro predicted phenotype associated with altered cellular dynamics due to SLUG inhibition. Furthermore, these findings demonstrate that SLUG loss in mammary epithelial cells leads to defects in basal/ME cell differentiation.

Given the defects in progenitor differentiation as well as tissue function during pregnancy/lactation, we speculated that stem cell activity necessary for tissue regeneration following transplantation might also be compromised in SLUG-deficient mammary epithelium. The mammary transplantation assay activates committed progenitors to adopt a more primitive state with extensive bipotent and regenerative potential (Kordon and Smith, 1998; Shackleton et al., 2006; Stingl et al., 2006; Van Keymeulen et al., 2011). Consistent with the in vitro prediction that stem cells can no longer transition into basal cells following SLUG inhibition, transplantation of 50,000 $Snai2^{LacZ/LacZ}$ MECs failed to generate any mammary outgrowths, while nearly complete mammary repopulation was observed in 100% of the transplants from WT MECs (Figure 4A). This suggested that SLUG might be necessary for mammary progenitor cells to transition back into stem/basal states during tissue regeneration in vivo.

Because tumor formation also requires activation of primitive stem/basal cell programs (Youssef et al., 2012), we crossed $Snai2^{LacZ/LacZ}$ mice onto the MMTV-Myc mouse

(B) Kaplan-Meier tumor-free survival curves of FVB-Tg MMTV-Myc+ (n = 12; red), FVB-Tg MMTV-Myc− (n = 6; green) and FVB/SV129-Tg x $Snai2^{+/+}$ mice (n = 7; purple). p values were calculated using the log-rank test.
(C) MMTV-Myc tumor histopathology. Representative H&E staining of tumors from FVB-Tg MMTV-Myc+, and FVB/SV129-Tg MMTV-Myc+ mice. Scale bars represent 100 μm.
(D) Kaplan-Meier tumor-free survival curves of FVB/SV129-Tg MMTV-Myc+ mice: $Snai2^{+/+}$ (n = 7; purple), $Snai2^{LacZ/+}$ (n = 9; blue), and $Snai2^{LacZ/LacZ}$ (n = 6; orange). p values were calculated using the log-rank test.
(E) Tumor burden curve for FVB-Tg MMTV-Myc+ (parental, n = 14), FVB/SV129-Tg x $Snai2^{+/+}$ (n = 7), FVB/SV129-Tg × $Snai2^{LacZ/+}$ (n = 9), and FVB/SV129-Tg x $Snai2^{LacZ/LacZ}$ (n = 6) mice. Black bars represent the median tumor burden for each cohort. p values were calculated by Student's two-tailed t test.
(F) Histologic features from tumors arising in FVB-Tg MMTV-Myc+ (n = 14), FVB/SV129-Tg MMTV-Myc+ × $Snai2^{+/+}$ (n = 6), and FVB/SV129-Tg MMTV-Myc+ × $Snai2^{LacZ/+}$ (n = 5) mice were examined. The tumors were scored for the presence of the various dominant histological patterns.
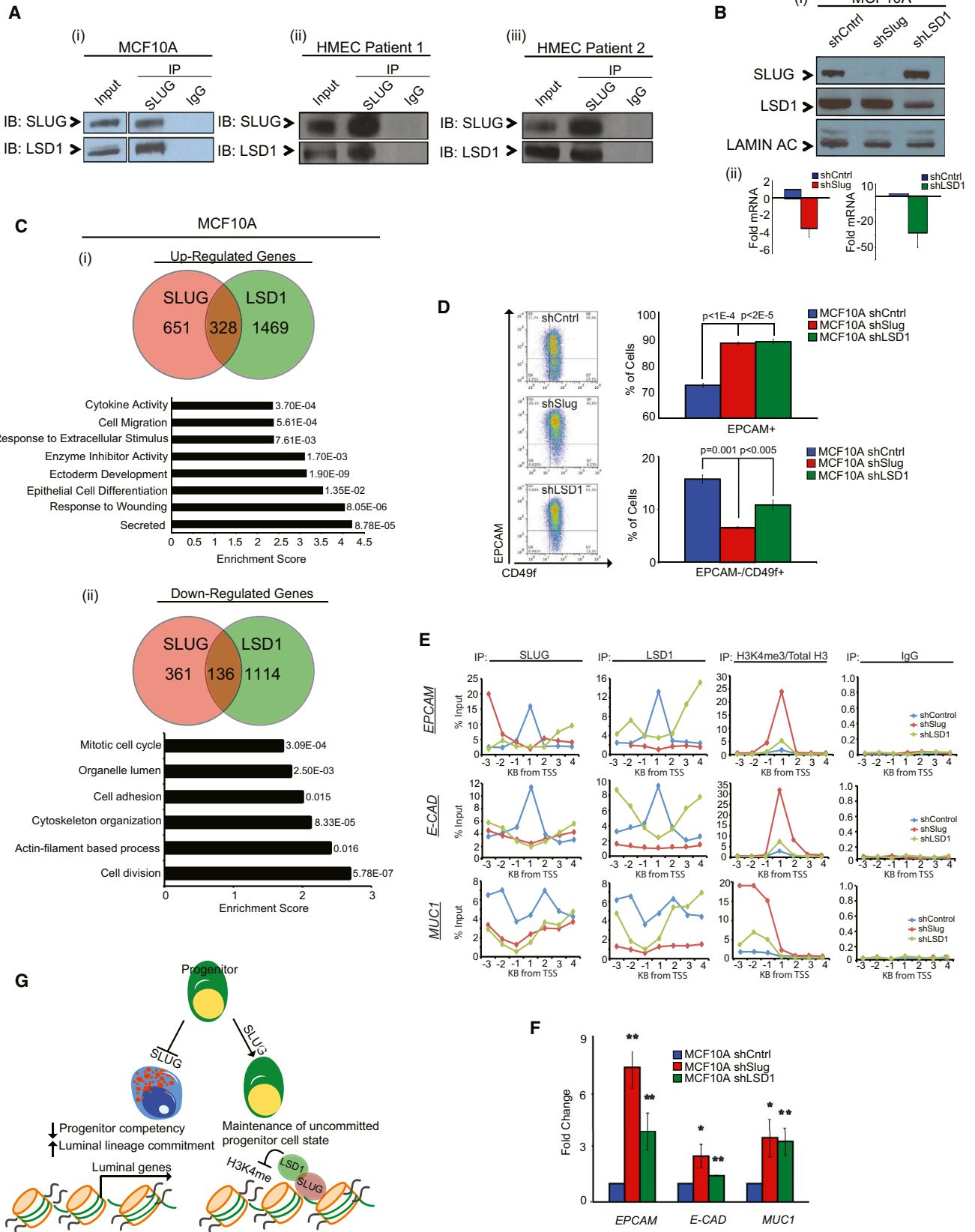
(legend on next page)

mammary tumor model background to determine whether defects in mammary progenitor cell-state transitions might also affect tumorigenesis. *MMTV-Myc*[+] mice develop heterogeneous tumors with various histologies and phenotypes associated with both luminal and basal features, including dominant luminal type tumors (microacinar and papillary), dominant basal type tumors (squamous and epithelial-to-mesenchymal transition), and mixed solid and adenocarcinomas with features of both basal and luminal differentiation (Andrechek et al., 2009; Sinn et al., 1987). *Snai2*[+/+]/*MMTV-Myc*[+] and *Snai2*[LacZ/+]/*MMTV-Myc*[+] compound mice also developed mammary adenocarcinomas with similar histologies and frequencies as the parental *MMTV-Myc*[+] mice (Figures 4B–4F). Strikingly, however, *Snai2*[LacZ/LacZ]/*MMTV-Myc*[+] compound mice completely failed to develop tumors by 70 weeks and thus were resistant to mammary tumorigenesis (Figure 4D). Taken together, these findings indicate that SLUG is necessary for mammary epithelial cell plasticity and that transitioning back into more primitive stem-like states is a prerequisite for both tissue regeneration and tumor initiation.

A SLUG/LSD1 Histone-Modifying Complex Regulates Gene Expression Programs in Basal Cells

Our findings have shown that SLUG is a critical regulator of breast epithelial cell plasticity, cell-state transitions, and differentiation. As a transcriptional repressor that mediates sequence-specific interactions with DNA (Hemavathy et al., 2000; Nieto, 2002), we hypothesized that SLUG might be interacting with other proteins to transcriptionally control these various differentiation programs. To identify interacting proteins that could cooperate with SLUG to regulate gene expression, we performed immunoprecipitation-coupled mass spectrometry by overexpressing a FLAG-tagged SLUG in 293T cells. Consistent with a recent report also using this approach, we found that lysine-specific demethylase 1 (LSD1) coimmunopurifies with SLUG (Figure S3A) (Wu et al., 2012). We confirmed this endogenous interaction in both MCF10A and HMEC basal progenitor cells (Figure 5A).

LSD1 is a histone demethylase whose transcriptional repression activity has been correlated with enzymatic removal of mono- and dimethyl groups from histone H3 Lysine 4 (H3K4; Shi et al., 2004; Wang et al., 2009). LSD1 has been shown to interact with the SNAG domain of SNAIL to regulate its activity and stability (Lin et al., 2010; Christofori, 2010). Based on this, we hypothesized that SLUG may be recruiting LSD1 to specific sites on luminal and stem cell gene promoters where it modifies histone marks associated with active transcription to suppress lineage commitment and differentiation and promote plasticity. In support of this hypothesis, genome-wide expression analysis of HMECs and MCF10A cells in which *LSD1* had been inhibited (Figure 5B; Figures S3B and S3C) identified a common set of epithelial-related genes whose expression was similarly regulated upon *SLUG* knockdown (Figure 5C; Figure S3D). Additionally, similar to shSlug cells, inhibition of *LSD1* led to an increased proportion of EPCAM[+] luminal cells and a decreased proportion of EPCAM[−]/CD49f[+] basal cells (Figures 2A and 5D; Figures S3E and S3F).

Figure 5. A SLUG/LSD1 Complex Regulates Gene Expression in Basal Cells

(A) Coimmunoprecipitation of endogenous SLUG and LSD1 from MCF10A cells (i), HMEC patient 1 (ii), and HMEC patient 2 (iii).

(B) Western blot analysis of SLUG and LSD1 expression in shControl, shSlug, and shLSD1 MCF10A cells (i). Quantitative real-time PCR analysis of *SLUG* and *LSD1* mRNA levels (normalized to *GAPDH*) in shSlug and shLSD1 MCF10A cells relative to control cells (ii). Data represent the mean ± SD of three independent experiments, and p values were calculated by Student's two-tailed t test.

(C) Top: Venn diagrams showing the overlap of microarray-identified genes commonly upregulated (i) or downregulated (ii) in shSlug and shLSD1 MCF10A cells compared to control cells. Bottom: Gene Ontology biological process categories commonly upregulated (i) or downregulated (ii) in shSlug and shLSD1 MCF10A cells compared to control cells. The DAVID Functional Annotation Tool was used to define categories with an enrichment score >1.5; the enrichment score and the p value of genes differentially expressed in each category are shown.

(D) Flow cytometry analysis of luminal (EPCAM) and basal (CD49f) cell-surface marker expression in shControl, shSlug and shLSD1 MCF10A cells. Representative dot plots for EPCAM versus CD49f expression are shown. Data represent the mean ± SD of three independent experiments, and p values were calculated by Student's two-tailed t test.

(E) ChIP analysis showing profiles for SLUG and LSD1 localization along the transcriptional start site (TSS; −3 kb to 4 kb) of common SLUG/LSD1 target genes (*EPCAM*, *E-CAD* and *MUC1*). Also depicted is relative enrichment of H3K4me3 (compared to total H3 levels) along these sites. Data are plotted as percent of input. ChIP was performed in shControl, shSlug, and shLSD1 MCF10A cells. For each cell line, DNA from two separate ChIP experiments was pooled, and enrichment was tested at the indicated promoter regions using a customized quantitative RT-PCR array.

(F) Quantitative real-time PCR analysis of luminal marker expression (normalized to *GAPDH*) in MCF10A shSlug and shLSD1 cells. Genes differentially expressed in the shSlug and shLSD1 cells compared to the scrambled control cells are plotted. Bars represent the fold change ± SD of three independent experiments, and p values were calculated by Student's two-tailed t test. *p < 0.05, **p < 0.01.

(G) Schematic diagram showing the mechanism by which SLUG and LSD1 regulate cell-state dynamics in mammary epithelial cells.

See also Figure S3.

To show that the SLUG/LSD1 complex transcriptionally regulates MEC differentiation, a core set of SLUG/LSD1 target genes, including *EPCAM*, *E-CADHERIN* (*E-CAD*), and *MUC1*, were used to evaluate SLUG recruitment of LSD1 to their promoters. In addition, these genes were used to examine whether the SLUG/LSD1 complex was modifying histone marks to affect chromatin and gene transcription. Chromatin immunoprecipitation (ChIP) confirmed SLUG and LSD1 occupancy at the proximal promoters of the indicated luminal epithelial genes (Figure 5E). Upon *SLUG* knockdown, SLUG and LSD1 occupancy at these promoter regions was lost. This loss of SLUG and LSD1 occupancy was accompanied by a significant increase in the LSD1-associated activating H3K4me3 histone mark (Saleque et al., 2007) at these promoters (Figure 5E) and increased gene expression (Figure 5F). Interestingly, upon *LSD1* knockdown, SLUG and LSD1 occupancy at these promoter regions was also lost and gene expression was increased; this suggested that LSD1 is necessary for SLUG recruitment to these promoters. In addition, *LSD1* knockdown also led to increased H3K4me3 histone marks at these promoters, although it was not as robust as compared to *SLUG* knockdown. This suggests that SLUG is necessary to recruit LSD1 in a sequence-specific manner but that LSD1 is also necessary for effective occupancy and histone demethylase activity of the SLUG/LSD1 complex. Together, these findings show that SLUG interaction with LSD1 is important to modify chromatin and repress luminal gene expression (Figure 5G).

DISCUSSION

In this study, we uncover the complex connection between cellular differentiation and cellular plasticity regulated by SLUG, thus revealing how it impacts stem cell activity during tissue homeostasis, function, regeneration, and even the genesis of cancer. Our findings support a model in which the transcription factor, SLUG, plays a dual role in regulating MEC lineage identity; on the one hand, it represses luminal epithelial differentiation, but on the other it promotes stem cell-state transitions necessary for transplantation and tumorigenesis. As a transcriptional repressor of luminal differentiation, we showed that SLUG interacts with LSD1 and recruits it to specific luminal gene promoters where it demethylates H3K4 (Figure 5E). In the absence of SLUG, LSD1 is no longer recruited to these genes to modify chromatin; this results in the expression of otherwise-repressed genes. Consistent with this, our in vivo and in vitro observations showed that loss of SLUG resulted in aberrant EPCAM expression in basal/ME cells. As a regulator of progenitor cell dynamics, we found

that SLUG also actively influences stem cell transitions. In the absence of SLUG, cellular transitions from and into the basal and stem cell states were compromised. This was supported in vivo, where SLUG-deficient MECs were unable to regenerate a mammary gland following transplantation and were unable to form oncogene induced mammary tumors; both processes require the transition of basal progenitor cells into primitive stem-like states not normally present during development (van Amerongen et al., 2012; Van Keymeulen et al., 2011). Future studies will be needed to determine whether LSD1 is also necessary for repressing SLUG target genes that regulate stem cell activity.

The observations that transplantation of *Snai2*[LacZ/LacZ] mouse tissue fragments, rather than dissociated cells, results in mature, highly branched mammary ductal trees that infiltrate the entire fat pad (Nassour et al., 2012) and that *Snai2*[LacZ/LacZ] mice exhibit normal embryonic and pubertal mammary stem cell activity suggest that SLUG is dispensable for normal mammary stem cell activity. However, our findings and those of others clearly show that SLUG is necessary for mammary stem cell activity during transplantation of dissociated cells (Figure 4A) (Guo et al., 2012). These conflicting observations can be reconciled by the differences in stem cell activity required during tissue regeneration following transplantation of dissociated cells from those used during ductal elongation of already-established structures. Mammary tissues are regenerated by unipotent lineage-restricted progenitor cells that can expand to give rise to mature luminal or basal/ME cells (Keller et al., 2011; van Amerongen et al., 2012; Van Keymeulen et al., 2011). Therefore, dissociated cells must adopt a bipotent fate upon transplantation, unlocking a regenerative potential that is not normally used during development (van Amerongen et al., 2012; Van Keymeulen et al., 2011). Thus, our findings suggest that the cell-state transitions necessary for this process require SLUG.

The dedifferentiation of lineage-committed cells into more primitive stem-like states during transplantation is similar to the cell-state transitions that occur during tumor initiation (Keller et al., 2010). For example, preceding the onset of basal cell carcinoma, adult tumor-initiating cells are reprogrammed into a fate resembling embryonic hair follicle progenitors (Youssef et al., 2012). Similarly, in a model of intestinal cancer, epithelial non-stem cells have been shown to dedifferentiate into tumor-initiating cells (Schwitalla et al., 2013). These changes in cell-state transitions are genetically and epigenetically controlled (Chaffer et al., 2013; Pacheco-Pinedo et al., 2011), and here we show that in the mammary gland, they require the transcription factor SLUG.

Although epithelial cell-state conversions have been shown to occur in vivo (Doupé et al., 2012; Wang et al.,

2013), quantitative modeling of this dynamic process has been challenging. We speculated that a Markov model previously used to study cell-state transitions of cancer cells in vitro (Gupta et al., 2011) could also be used to uncover important cell-state conversions in normal MEC populations; we also wondered whether the model's predictions could be tested in vivo. Accordingly, we applied the model to normal, disease-free mammary epithelial cells in vitro and found that it could accurately predict epithelial plasticity in vivo. Remarkably, the model predicted that upon loss of SLUG, there would be an accumulation of luminal cells due to the increased probability that stem cells would transition into luminal cells. Indeed, this phenotype was observed in vivo in $Snai2^{LacZ/LacZ}$ mice. Thus, this approach opens new possibilities for assessing the consequence of genetic perturbations on cell-state transitions in normal cells that can be used to predict and understand complex phenotypes in vivo.

EXPERIMENTAL PROCEDURES

Cell Lines and Tissue Culture

HMECs and MCF10A cells were cultured in mammary epithelial growth medium (MEGM Bullet Kit; Lonza Corporation) comprising mammary epithelial cell basal medium supplemented with bovine pituitary extract (52 μg/ml), hydrocortisone (0.5 μg/ml), human epidermal growth factor (10 ng/ml), and insulin (5 μg/ml). MCF10A cells were further supplemented with cholera toxin (100 ng/ml). For proliferation assays, cells were plated in 96-well plates at 1,000 cells per well. On days 1, 2, and 3, cells were lysed with CellTiter-Glo Luminescent Cell Viability Assay reagent (Promega) and luminescence was read using the Sirius/FB15 Luminometer (Berthold).

Mouse Strains

The generation of $Snai2^{LacZ/LacZ}$ mice was described previously (Jiang et al., 1998) and generously provided by Donna F. Kusewitt. The *MMTV-Myc* mouse (FVB-Tg *MMTV-Myc*) used in this study has been described elsewhere (Sinn et al., 1987).

Animal Surgery

All animal procedures were conducted in accordance with a protocol approved by the Tufts University IACUC committee. A colony of $Snai2^{LacZ/LacZ}$, FVB-Tg *MMTV-Myc*[+], and FVB/SV129-Tg *MMTV-Myc*[+] × $Snai2^{LacZ/LacZ}$ mice were maintained in-house under aseptic sterile conditions. Mice were administered autoclaved food and water ad libitum. Surgeries were performed under sterile conditions, and animals received antibiotics in the drinking water up to 2 weeks after all surgical procedures.

ACCESSION NUMBERS

The GEO accession number for the human microarray data reported in this paper is GSE54735. The GEO accession number for the mouse microarray data is GSE55620.

REFERENCES

Andrechek, E.R., Cardiff, R.D., Chang, J.T., Gatza, M.L., Acharya, C.R., Potti, A., and Nevins, J.R. (2009). Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential. Proc. Natl. Acad. Sci. USA *106*, 16387–16392.

Barker, N., van de Wetering, M., and Clevers, H. (2008). The intestinal stem cell. Genes Dev. *22*, 1856–1864.

Barker, N., Ridgway, R.A., van Es, J.H., van de Wetering, M., Begthel, H., van den Born, M., Danenberg, E., Clarke, A.R., Sansom, O.J., and Clevers, H. (2009). Crypt stem cells as the cells-of-origin of intestinal cancer. Nature *457*, 608–611.

Blanpain, C., and Fuchs, E. (2006). Epidermal stem cells of the skin. Annu. Rev. Cell Dev. Biol. *22*, 339–373.

Blanpain, C., Lowry, W.E., Geoghegan, A., Polak, L., and Fuchs, E. (2004). Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. Cell *118*, 635–648.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat. Med. *3*, 730–737.

Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell *154*, 61–74.

Christofori, G. (2010). Snail1 links transcriptional control with epigenetic regulation. EMBO J. *29*, 1787–1789.

Doupé, D.P., Alcolea, M.P., Roshan, A., Zhang, G., Klein, A.M., Simons, B.D., and Jones, P.H. (2012). A single progenitor population switches behavior to maintain and repair esophageal epithelium. Science 337, 1091–1093.

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J.T., and Eaves, C.J. (2008). A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat. Med. 14, 1384–1389.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat. Biotechnol. 26, 317–325.

Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zürrer-Härdi, U., Bell, G., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell 148, 1015–1028.

Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., and Lander, E.S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell 146, 633–644.

Hemavathy, K., Ashraf, S.I., and Ip, Y.T. (2000). Snail/slug family of repressors: slowly going into the fast lane of development and cancer. Gene 257, 1–12.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Jiang, R., Lan, Y., Norton, C.R., Sundberg, J.P., and Gridley, T. (1998). The Slug gene is not essential for mesoderm or neural crest development in mice. Dev. Biol. 198, 277–285.

Keller, P.J., Lin, A.F., Arendt, L.M., Klebba, I., Jones, A.D., Rudnick, J.A., DiMeo, T.A., Gilmore, H., Jefferson, D.M., Graham, R.A., et al. (2010). Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. Breast Cancer Res. 12, R87.

Keller, P.J., Arendt, L.M., and Kuperwasser, C. (2011). Stem cell maintenance of the mammary gland: it takes two. Cell Stem Cell 9, 496–497.

Keller, P.J., Arendt, L.M., Skibinski, A., Logvinenko, T., Klebba, I., Dong, S., Smith, A.E., Prat, A., Perou, C.M., Gilmore, H., et al. (2012). Defining the cellular precursors to human breast cancer. Proc. Natl. Acad. Sci. USA 109, 2772–2777.

Kordon, E.C., and Smith, G.H. (1998). An entire functional mammary gland may comprise the progeny from a single cell. Development 125, 1921–1930.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al.; kConFab (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. 15, 907–913.

Lin, T., Ponn, A., Hu, X., Law, B.K., and Lu, J. (2010). Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition. Oncogene 29, 4896–4904.

Nassour, M., Idoux-Gillet, Y., Selmi, A., Côme, C., Faraldo, M.L., Deugnier, M.A., and Savagner, P. (2012). Slug controls stem/

progenitor cell growth dynamics during mammary gland morphogenesis. PLoS ONE 7, e53498.

Nieto, M.A. (2002). The snail superfamily of zinc-finger transcription factors. Nat. Rev. Mol. Cell Biol. 3, 155–166.

Pacheco-Pinedo, E.C., Durham, A.C., Stewart, K.M., Goss, A.M., Lu, M.M., Demayo, F.J., and Morrisey, E.E. (2011). Wnt/β-catenin signaling accelerates mouse lung tumorigenesis by imposing an embryonic distal progenitor phenotype on lung epithelium. J. Clin. Invest. 121, 1935–1945.

Parent, A.E., Choi, C., Caudy, K., Gridley, T., and Kusewitt, D.F. (2004). The developmental transcription factor slug is widely expressed in tissues of adult mice. J. Histochem. Cytochem. 52, 959–965.

Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. J. Clin. Oncol. 27, 1160–1167.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 12, R68.

Proia, T.A., Keller, P.J., Gupta, P.B., Klebba, I., Jones, A.D., Sedic, M., Gilmore, H., Tung, N., Naber, S.P., Schnitt, S., et al. (2011). Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. Cell Stem Cell 8, 149–163.

Quarrie, L.H., Addey, C.V., and Wilde, C.J. (1996). Programmed cell death during mammary tissue involution induced by weaning, litter removal, and milk stasis. J. Cell. Physiol. 168, 559–569.

Saleque, S., Kim, J., Rooke, H.M., and Orkin, S.H. (2007). Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. Mol. Cell 27, 562–572.

Schwitalla, S., Fingerle, A.A., Cammareri, P., Nebelsiek, T., Göktuna, S.I., Ziegler, P.K., Canli, O., Heijmans, J., Huels, D.J., Moreaux, G., et al. (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. Cell 152, 25–38.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature 439, 84–88.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119, 941–953.

Sinn, E., Muller, W., Pattengale, P., Tepler, I., Wallace, R., and Leder, P. (1987). Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. Cell 49, 465–475.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature 439, 993–997.

Tesori, V., Puglisi, M.A., Lattanzi, W., Gasbarrini, G.B., and Gasbarrini, A. (2013). Update on small intestinal stem cells. World J. Gastroenterol. 19, 4671–4678.

Toma, J.G., Akhavan, M., Fernandes, K.J., Barnabé-Heider, F., Sadikot, A., Kaplan, D.R., and Miller, F.D. (2001). Isolation of multipotent adult stem cells from the dermis of mammalian skin. Nat. Cell Biol. *3*, 778–784.

van Amerongen, R., Bowman, A.N., and Nusse, R. (2012). Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. Cell Stem Cell *11*, 387–400.

Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189–193.

Wang, J., Hevi, S., Kurash, J.K., Lei, H., Gay, F., Bajko, J., Su, H., Sun, W., Chang, H., Xu, G., et al. (2009). The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. Nat. Genet. *41*, 125–129.

Wang, Z.A., Mitrofanova, A., Bergren, S.K., Abate-Shen, C., Cardiff, R.D., Califano, A., and Shen, M.M. (2013). Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. Nat. Cell Biol. *15*, 274–283.

Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. Cell *100*, 157–168.

Wu, Z.Q., Li, X.Y., Hu, C.Y., Ford, M., Kleer, C.G., and Weiss, S.J. (2012). Canonical Wnt signaling regulates Slug activity and links epithelial-mesenchymal transition with epigenetic Breast Cancer 1, Early Onset (BRCA1) repression. Proc. Natl. Acad. Sci. USA *109*, 16654–16659.

Youssef, K.K., Van Keymeulen, A., Lapouge, G., Beck, B., Michaux, C., Achouri, Y., Sotiropoulou, P.A., and Blanpain, C. (2010). Identification of the cell lineage at the origin of basal cell carcinoma. Nat. Cell Biol. *12*, 299–305.

Youssef, K.K., Lapouge, G., Bouvrée, K., Rorive, S., Brohée, S., Appelstein, O., Larsimont, J.C., Sukumaran, V., Van de Sande, B., Pucci, D., et al. (2012). Adult interfollicular tumour-initiating cells are reprogrammed into an embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. Nat. Cell Biol. *14*, 1282–1294.

Cell-State Transitions Regulated by SLUG Are Critical for Tissue Regeneration and Tumor Initiation

Sarah Phillips, Aleix Prat, Maja Sedic, Theresa Proia, Ania Wronski, Sohini Mazumdar, Adam Skibinski, Stephanie H. Shirley, Charles M. Perou, Grace Gill, Piyush B. Gupta, and Charlotte Kuperwasser

Inventory of Supplemental Information

1. **Figure S1. Inhibition of SLUG promotes epithelial cell differentiation.** Related to Figure 1:

This Figure supports Figure 1 in the main text. It provides brightfield images of the shControl and shSlug HMEC cells. It also provides additional analysis of microarray data from shControl and shSlug HMEC and MCF10A cells.

2. **Figure S2. Inhibition of SLUG results in increased proportions of luminal cells *in vitro*.** Related to Figure 2:

This Figure supports Figure 2A in the main text. It provides additional flow cytometry analysis (and representative dot plots) of luminal and basal populations in the shControl and shSlug HMEC and MCF10A cells.

3. **Figure S3. SLUG and LSD1 interact in mammary epithelial cells.** Related to Figure 5:

This Figure supports Figure 5 in the main text. It provides additional data detailing the interaction between SLUG and LSD1. Included in this figure is data showing knockdown levels of SLUG and LSD1 (by protein and mRNA) in the HMEC cells, as well as microarray and flow cytometry analysis of shSlug and shLSD1 HMEC cells. Figure 5 includes this data for the MCF10A cells.

4. **Supplemental Figure Legends**

5. **Supplemental Experimental Procedures**: This section provides a comprehensive list of the methods, reagents, etc. used to generate the data in the main text as well as the supplemental materials.

6. **Supplemental References**

A

B

Figure S1. Inhibition of SLUG promotes epithelial cell differentiation. Related to Figure 1

(A) Phase contrast images of shControl and shSlug HMECs (Patient 1 and Patient 2). Arrowheads highlight the tightly packed, cobblestone-like morphology of shSlug cells. Scale bars represent 100 μm.

(B) Genomic Differentiation Scores for HMEC (Patient 1) and MCF10A cells, calculated from microarray gene expression data (Prat et al., 2010).

Figure S2. Inhibition of SLUG results in increased proportions of luminal cells *in vitro*. Related to Figure 2

(A and B) Flow cytometry analysis of luminal (EPCAM and CD24) and basal (CD49f) cell surface markers in shControl (n=3) and shSlug (n=3) HMECs (A) Patient 1 and (B) Patient 2. Representative dot plots for EPCAM vs. CD49f expression are shown. Data represent the mean ±SD of three independent experiments and p values were calculated by Student's two-tailed t test.

(C-E) Representative dot plots for CD24 vs. CD49f expression in (C) MCF10A cells, (D) HMEC Patient 1 and (E) HMEC Patient 2 cells.

Figure S3. SLUG and LSD1 interact in mammary epithelial cells. Related to Figure 5

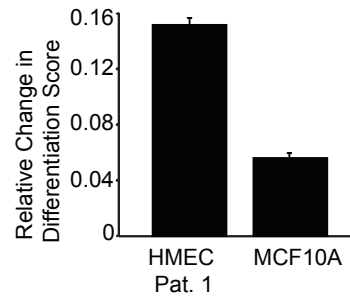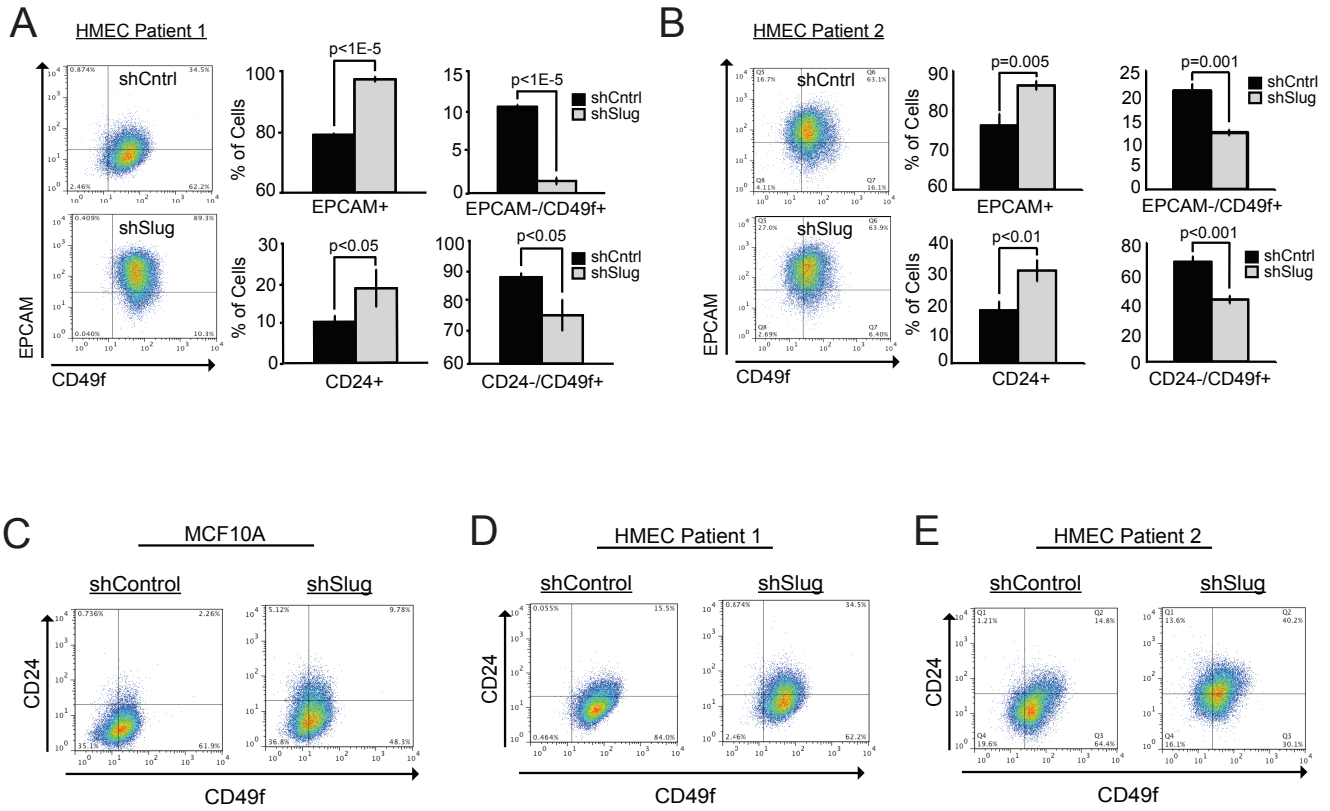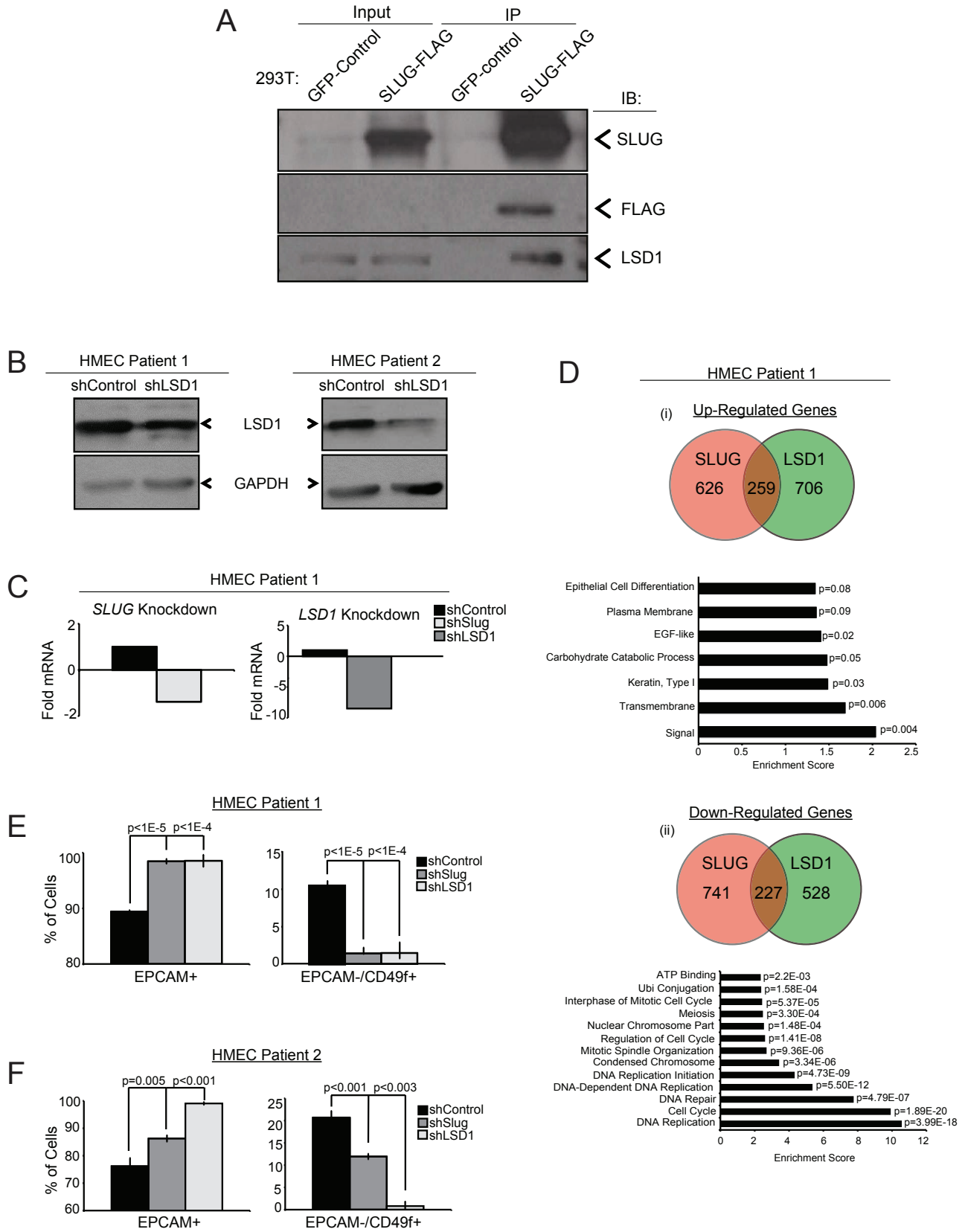(A) Immunoblot analysis of proteins immunoprecipitated with anti-FLAG beads from 293T cells transfected with FLAG-SLUG or a control GFP vector and probed with anti-FLAG, anti-SLUG and anti-LSD1 antibodies.

(B) Western Blot analysis of LSD1 expression in shControl and shLSD1 HMECs (Patient 1 and Patient 2).

(C) Quantitative Real-time PCR analysis of *SLUG* and *LSD1* mRNA levels in shSlug and shLSD1 HMEC cells (Patient 1) relative to control cells.

(D) Above: Venn Diagrams showing the overlap of microarray-identified genes commonly (i) up-regulated or (ii) down-regulated in shSlug and shLSD1 HMEC (Patient 1) cells compared to control cells. Below: Gene ontology biological process categories commonly (i) up-regulated or (ii) down-regulated in shSlug and shLSD1 HMEC (Patient 1) cells compared to control cells. The DAVID Functional Annotation Tool was used to define categories with an enrichment score >1.5; the enrichment score and the p value of genes differentially expressed in each category are shown.

(E and F) Flow cytometry analysis of luminal (EPCAM) and basal (CD49f) cell surface marker expression in shControl, shSlug and shLSD1 HMEC cells, (E) Patient 1 and (F) Patient 2. Data represent the mean ±SD of three independent experiments and p values were calculated by Student's two-tailed t test.

Supplemental Experimental Procedure

Lentiviral Infection

The VSV-G-pseudotyped lentiviral vectors were generated by transient cotransfection of the vector construct with the VSV-G-expressing construct pCMV-VSVG and the packaging construct pCMV DR8.2Dvpr (Miyoshi et al., 1998) into 293T cells together with Mirius transfection reagent (Invitrogen). Lentiviral shRNA constructs targeting *SLUG* (Addgene plasmids 10904 and 10905), *LSD1* (Sigma-Aldrich, NM_015013.1-1812s1c1) and a scramble control (plKO.1 puro, Addgene) were prepared as previously described (Gupta et al., 2005).

Western Blot Analysis

Cultured cells were harvested by trypsinization, pelleted and incubated in RIPA buffer supplemented with protease inhibitor (Roche) to obtain whole cell lysates. Cellular debris was removed by centrifugation at 13,000 rpm for 10 min. 40 μg of the whole cell lysate was used per sample. Western blotting was performed according to the manufacturers protocol (BioRad). Briefly, 4-12% pre-cast gels and XT-MOPS running buffer were used for SDS-PAGE electrophoresis. 0.45 μm nitrocellulose membranes were used for protein transfer. Membranes were incubated overnight at 4°C with primary antibodies diluted in 1% bovine serum albumin in TBS-T or 5% milk in TBS-T. Secondary antibodies (HRP-conjugated) were applied for 1 hr at room temperature (Cell Signaling; #7076 and #7074). The antibodies used for these experiments included: anti-SLUG (Cell Signaling; #9585), anti-LSD1 (Abcam; ab62582), anti-LAMIN AC (Cell Signaling; #2032) and anti-GAPDH (Millipore; MAB374).

Immunoprecipitation (IP)

HMECs and MCF10A cells were lysed in IP buffer (20 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100) supplemented with protease and phosphatase inhibitors (Roche). For immunoprecipitation assays, protein lysates (1mg) were combined with 10 µg of antibody: anti-SLUG (Cell Signaling; #9589) or mouse normal IgG (Santa Cruz; SC2025) and 25 µl of Protein A-Plus agarose beads (Repligen). Following an overnight incubation at 4°C, agarose beads were extensively washed in IP buffer, resuspended in SDS sample buffer (125 Mm Tris pH 6.8, 2.5% SDS, 10% glycerol, 2.5% 2-mercaptoethanol, 0.01% bromo-phenol-blue) and loaded onto a protein gel. Antibodies used for immunoblotting included: anti-SLUG (Cell Signaling; #9585) and anti-LSD1 (Abcam; ab62582).

293T cells were transfected with pPGS-hSlug-FLAG or a control GFP construct using FuGene 6 ® Transfection Reagent (Roche). 48 hrs after transfection, cells were harvested and lysed as described above. 1 mg of protein was incubated overnight at 4°C with 40 µl of Anti-FLAG ® M2 Affinity Gel (Sigma). Following overnight incubation, beads were extensively washed in IP buffer, resuspended in SDS sample buffer and loaded onto a protein gel. Antibodies used for immunoblotting were as follows: anti-SLUG (Cell Signaling; #9585), anti-LSD1 (Abcam; ab62582) and Anti-FLAG ® M2 (Cell Signaling; 2044).

Quantitative RT-PCR

Total RNA from cultured cells was extracted with the RNeasy Mini Kit (QIAGEN). cDNA was prepared with an iScript kit (BioRad) and PCR was carried out with SYBR Green (BioRad). The following primers (human) were used in this study:

SLUG: Forward '5-GCATTTCTTCACTCCGAAGC-3' and Reverse '5-TGAATTCCATGCTCTTGCAG-3'

LSD1: Forward '5-CAAGTGTCAATTTGTTCGGG-3' and Reverse '5-TTCTTTGGGCTGAGGTACTG-3'

GAPDH: Forward '5-GAGTCAACGGATTTGGTCGT-3' and Reverse '5-GACAAGCTTCCCGTTCTCAG-3'

KRT18: Forward '5-TGATGACACCAATATCACACGAC-3' and Reverse '5-TACCTCCACGGTCAACCCA-3'

KRT19: Forward '5-ACCAAGTTTGAGACGGAACAG-3' and Reverse '5-CCCTCAGCGTACTGATTTCCT-3'

CD24: Forward '5-TGAAGAACATGTGAGAGGTTTGAC-3' and Reverse '5-AGAGTGAGACCACGAAGAGAC-3'

MUC1: Forward '5-CGCCGAAAGAACTACGGGCAGCTG-3' and Reverse '5-CAAGTTGGCAGAAGTGGCTGCCAC-3'

GATA3: Forward '5-GCGGGCTCTATCACAAAATGA-3' and Reverse '5-GCTCTCCTGGCTGCAGACAGC-3'

KRT14: Forward '5-CATGAGTGTGGAAGCCGACAT-3' and Reverse '5-GCCTCTCAGGGCATTCATCTC-3'

SMA: Forward '5-CAGGGCTGTTTTCCCATCCAT-3' and Reverse '5-GCCATGTTCTATCGGGTACTTC-3'

VIM: Forward '5-GAGTCCACTGAGTACCGGAGAC-3' and Reverse '5-TGTAGGTGGCAATCTCAATGTC-3'

CLDN4: Forward '5-GGGGCAAGTGTACCAACTG-3' and Reverse '5-GACACCGGCACTATCACCA-3'

SOX9: Forward '5-AGCGAACGCACATCAAGAC-3'and Reverse '5-CTGTAGGCGATCTGTTGGGG-3'

EPCAM: Forward '5-AATCGTCAATGCCAGTGTACTT-3' and Reverse '5-TCTCATCGCAGTCAGGATCATAA-3'

E-CADHERIN: Forward '5-GAACGCATTGCCACATACAC-3' and Reverse '5-GAATTCGGGCTTGTTGTCAT-3'

Gene Expression Microarray Analysis – Human Cell Lines

Total RNA for microarray expression studies was isolated from MCF10As and HMECs via the RNeasy Mini kit (QIAGEN). RNA was profiled as described previously using oligo microarrays (Agilent Technologies, Santa Clara, CA, USA; Hu et al., 2006). The probes or genes for all analyses were filtered by requiring the lowest normalized intensity values in both sample and control to be > 10. The normalized log2 ratios (Cy5 shSlug sample/Cy3 shControl sample) of probes mapping to the same gene (Entrez ID as

defined by the manufacturer) were averaged to generate independent expression estimates. Differentially expressed genes between shSlug sample and shControl sample were identified after performing a one class Significance Analysis of Microarrays (SAM; Tusher et al., 2001). All microarray data are available in the University of North Carolina (UNC) Microarray Database (https://genome.unc.edu/) and have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE54735.

Gene Expression nCounter Analysis

Total RNA for microarray expression studies was isolated from Basal, Luminal and Stem MCF10A cell sorted populations via the RNeasy Mini kit (QIAGEN). For each sample, ~100 ng of total RNA was used to measure gene expression of 110 selected genes using the digital color-coded nCounter platform (Nanostring Technologies, Seatlle). Raw data was log base 2 transformed and normalized using 5 house-keeping transcripts. All data is available in GSE54735.

Gene Expression Microarray Analysis – Mouse MECs

Mammary glands from $Snai2^{LacZ/LacZ}$ mice and WT controls were dissected, resuspended as single cell suspensions and sorted into basal/ME (Lin$^-$/EPCAM^{lo}/CD49f^{hi}), luminobasal (Lin$^-$/EPCAM^{hi}/CD49f^{hi}) and unsorted fractions as described above. Pools of cells for each genotype and fraction were combined from 2-3 animals. RNA was extracted using the Qiagen RNeasy micro column with on-column DNA digestion as per manufacturer's instructions. All RNA samples were assayed on an Agilent 2100 Bioanalyzer to ensure high RNA quality (RIN > 8) prior to microarray analysis.

Microarray samples were submitted to the Tufts Center for Neuroscience Research Core. Sample preparation and analysis were conducted by the Yale Center for Genomic Analysis (Yale School of Medicine, CT) as follows: 5 ng of total RNA was amplified using the Ovation Pico WTA System v2 kit (Nugene, #3302) as per manufacturer's instructions. Amplified cDNA was purified using a QIAQuick PCR Purification kit (Qiagen) and the concentration of cDNA samples normalized to a total of 4 μg, as per Nanodrop readings (Thermo Scientific). cDNA was then labeled with the Encore BiotinIL Module Kit (Nugene, #4210) and purified with a MinElute Reaction Cleanup Kit (Qiagen) as per manufacturers' instructions. Labeled and purified cDNA was normalized to a total concentration of 750 ng and hybridized to a MouseRef v2 Expression BeadChip (Illumina) overnight at 48°C as per manufacturer's instructions. Beadchips were then washed using recommended buffers and stained with Amersham Cy3-Streptavidin (GE Healthcare #PA43001). BeadChips were scanned using the Illumina HiScan System and images analyzed using Beadstudio Software (Illumina). Quality control and data analysis were carried out according to the instructions provided by Illumina. Differential gene expression between WT and $Snai2^{LacZ/LacZ}$ samples for each fraction was computed by GenePattern Comparative Marker Selection module (Broad Institute, MIT). Gene signature comparison to the data from Lim et al. (Lim et al., 2009) was conducted as per Prat et al. (Prat et al., 2010).

3D Collagen Assay

For 3D collagen gel assays, rat-tail collagen was diluted to 1 mg/mL with 5 mM glacial acetic acid and brought to pH 7.0 with 0.01 M NaOH. Wells of 4-well chamber

slides (BD Falcon) were coated with 100 µl of collagen and allowed to set at room temperature for 20–30 min. Then, 10,000 MCF10A cells (shControl or shSlug) were plated in 1 mL of MEGM containing 2% Matrigel (BD Biosciences) per well, in triplicate. Structures were allowed to form on top of the collagen gels for 5-6 days with media changes every 3 days. Structures were quantified under a microscope at 100x magnification.

Immunostaining

For paraffin embedded tissues, samples were deparaffinized, rehydrated in graded alcohols and processed for antigen retrieval by incubation in citrate buffer (PH 6) at 95°C for 45 min. Slides were then washed in PBS before primary antibody incubation. For plated cells, samples were cultured on 8-well chamber-slides, fixed with methanol at −20°C for 10 min and washed 3x with 1x PBS before primary antibody incubation. All samples were incubated overnight at 4°C with primary antibodies: anti- (Abcam; ab8667) or anti-SMA (Vector Laboratories; VPS281) diluted in 1% bovine serum albumin in PBS. Alexa Fluor 546 and Alexa Fluor 488-conjugated secondary antibodies (Invitrogen; #A11010 and #AA11001) were applied for 1 hr at room temperature. Tissues/cells were counterstained with DAPI and mounted with the Slow-Fade mounting kit (Invitrogen). A Leica SP2 confocal microscope was used to capture these images.

Immunohistochemistry

Immunohistochemistry was performed on formalin-fixed, paraffin-embedded tissue sections with sodium citrate or Tris EDTA antigen retrieval, followed by detection with biotinylated secondary antibodies (anti-mouse, Vector; BA-202 or anti-rabbit,

Vector Laboratories; BA 1000) and visualization with the Elite ABC Peroxidase Kit and DAB substrate (Vector Laboratories). Antibodies used on paraffin-embedded tumor and mammary gland sections were as follows: anti-ER-α, (Santa Cruz; SC542, clone M20).

Staining for Ki67 was performed by the Histology Special Procedures Laboratory at Tufts Medical Center.

Tissue samples from each tumor lesion were fixed for 24 hr in 4% neutral buffered formalin. After paraffin embedding, tumor specimens were cut into 5-μm sections and stained with H&E. A Nikon Eclipse 80t microscope and SPOT camera were used for analyzing and photographing the stained sections.

Preparation of MEC Single-Cell Suspensions

The 3rd, 4th and 5th mammary glands were harvested from 8 or 16-week old mice. Glands were mechanically disaggregated and then digested with collagenase (Sigma) and Hyaluronidase (Sigma) for 1 hr at 37°C. Red blood cells were lysed in Red Blood Cell Lysis Buffer (Sigma). The resultant organoids were further digested in 0.25% trypsin-EDTA (2 min) and Dispase/DNaseI (2 min), and then filtered through a 40 μm mesh filter.

Flow Cytometry and FACS

Nonconfluent cultures of MCF10A and immortalized HMEC cells were trypsinized into single-cell suspension, counted, washed with PBS and stained with antibodies specific for human cell CD24 (PE, BD Bioscience; #555427), EPCAM (APC, BD Bioscience; #347200) and CD49f (FITC, BD Bioscience; #555735). Conjugated

isotype-matching IgGs were used as negative controls (IgG2a-PE, BD Bioscience; 559319, IgG1-APC, BD Bioscience; 3404442 and IgG2a-FITC, BD Bioscience; 555573). Antibody-bound cells were washed and resuspended at 1x 10^6 cells/ml in 2% FBS in PBS and run on a FACS Calibur flow cytometer (BD Biosciences) or sorted on a BD Influx FACS sorter (BD Biosciences). Flow cytometry data was analyzed with the Flowjo software package (TreeStar).

Freshly dissociated single cell suspensions of mouse mammary epithelial cells were stained with anti-CD24-APC (eBiosciences; #17-0242), anti-CD49f-FITC (StemCell #10111), anti-CD49f-PercpCy5.5 (Biolegend; #313618) and anti-EPCAM-APC (Biolegend; #118213) antibodies. Endothelial, lymphocytic and monocytic lineages were depleted with antibodies specific to mouse CD31 (PE, #12-0311-83), CD45 (PE, #12-0451-81) and Terr119 (PE, #12-5921) (eBiosciences). Conjugated isotype-matching IgGs were used as negative controls. They were as follows: APC-IgG2b (eBiosciences; 17-4031-81), APC-IgG2a (eBiosciences; 174321), PerCP-Cy5.5-IgG2a (Biolegend; 400531) and FITC-IgG1 (BD Biosciences; 553443).

Whole Mounts

Mammary glands were spread onto glass slides, fixed in 10% neutral buffered formalin, and stained overnight in 0.2% carmine red. Glands were subsequently dehydrated with graded ethanol solutions, cleared in xylene and mounted.

Mammary Fat-Pad Transplants:

Single cell suspensions (see above) of isolated MECs (previously harvested and frozen) were resuspended at 50,00 cells/25 μl in Matrigel:Mouse Mammary Medium

(1:3) solution, and injected into cleared inguinal fat pads of 3-week old nonobese diabetic/severe combined immunodeficient mice (Deome et al., 1959). *Snai2*$^{+/+}$ or *Snai2*$^{LacZ/LacZ}$ MECs were transplanted into contralateral inguinal fat pads. Twelve weeks after transplantation, glands were harvested and processed for whole mounting. Mouse Mammary Media: (DMEM/F12 + 2% CS, 10 μg/ml mouse-insulin, 5 ng/ml mEGF, 0.5 μg/ml hydrocortisone).

Chromatin Immunoprecipitation

The chromatin immunoprecipitation (ChIP) assay was performed by using the EZ-Magna ChIP A kit (Millipore, cat. no. 17–409) following the manufacturer's instructions. In brief, cells were grown to 60–70% confluence in 15cm dishes. Proteins were cross-linked to DNA by adding the formaldehyde directly to the culture medium to a final concentration of 1% for 10 min. The unreacted formaldehyde was quenched by incubating with 1× glycine buffer for 5 min. After rinsing twice with ice-cold PBS, the cells were resuspended in 1 ml of PBS with protease inhibitor by scraping the cells from the dishes. After brief centrifugation to pellet cells, cells were resuspended in 0.5 ml of cell lysis buffer and then nuclear lysis buffer. Sonication of cell lysate was performed four times for 25 sec each on a Branson Sonicator, followed by centrifugation at 4°C for 10 min. Supernatants were collected and diluted 1:10 with dilution buffer. A portion of diluted supernatant (1%) was kept and used as "input." Immunoprecipitation was carried out overnight at 4°C by adding the immunoprecipitating antibody and 20 μl of fully suspended protein A magnetic beads (Millipore). The antibodies used were: anti-SLUG (Cell Signaling; 9589), anti-LSD1 (Abcam; ab17721), anti-H3 (Millipore; #06-755), anti-H3K4me3 (Millipore; #NG18233938) and Mouse normal IgG (Santa Cruz; #SC2025).

After washing the protein A bead-antibody/chromatin complex, ChIP elution buffer with proteinase K was added and incubated at 62°C for 2 hrs. DNA was recovered and purified with DNA spin columns. Both the immunoprecipitated samples and the input samples were processed in the same way. After ChIP assays, we used custom-made ChIP arrays (Qiagen, EpiTect ® ChIP qPCR Array Human Custom) to assess localization of SLUG and LSD1 as well as enrichment of H3K4me3 along the proximal promoters of the *EPCAM*, *E-CADHERIN* and *MUC1* genes.

Markov Model

A description of the cell state transition Markov Model and computation of the stochastic matrix associated with a population of cells has been previously described (Gupta et al., 2011).

Supplemental References

Deome, K.B., Faulkin, L.J., Jr., Bern, H.A., and Blair, P.B. (1959). Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer Res 19, 515-520.

Gupta, P.B., Kuperwasser, C., Brunet, J.P., Ramaswamy, S., Kuo, W.L., Gray, J.W., Naber, S.P., and Weinberg, R.A. (2005). The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. Nat Genet 37, 1047-1054.

Hu, Z., Fan, C., Oh, D.S., Marron, J.S., He, X., Qaqish, B.F., Livasy, C., Carey, L.A., Reynolds, E., Dressler, L., et al. (2006). The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 7, 96.

Miyoshi, H., Blomer, U., Takahashi, M., Gage, F.H., and Verma, I.M. (1998). Development of a self-inactivating lentivirus vector. J Virol 72, 8150-8157.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America 98, 5116-5121.