MIT Libraries | DSpace@MIT

MIT Open Access Articles

BCL11B Drives Human Mammary Stem Cell Self-Renewal In Vitro by Inhibiting Basal Differentiation

Massachusetts Institute of Technology

Stem Cell Reports
Article

BCL11B Drives Human Mammary Stem Cell Self-Renewal *In Vitro* by Inhibiting Basal Differentiation

Daniel H. Miller,[1,2] Dexter X. Jin,[1,2] Ethan S. Sokol,[1,2] Janel R. Cabrera,[3,4] Daphne A. Superville,[1,2] Rebecca A. Gorelov,[1,5] Charlotte Kuperwasser,[3,4] and Piyush B. Gupta[1,2,6,7,*]

[1]Whitehead Institute for Biomedical Research, 455 Main St., Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Department of Developmental, Chemical, and Molecular Biology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
[4]Raymond & Beverly Sackler Convergence Laboratory, Tufts University School of Medicine, 145 Harrison Avenue, Boston, MA 02111, USA
[5]Department of Biology, Williams College, Williamstown, MA 01267, USA
[6]Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA
[7]Harvard Stem Cell Institute, Cambridge, MA 02138, USA
*Correspondence: pgupta@wi.mit.edu
https://doi.org/10.1016/j.stemcr.2018.01.036

SUMMARY

The epithelial compartment of the mammary gland contains basal and luminal cell lineages, as well as stem and progenitor cells that reside upstream in the differentiation hierarchy. Stem and progenitor cell differentiation is regulated to maintain adult tissue and mediate expansion during pregnancy and lactation. The genetic factors that regulate the transition of cells between differentiation states remain incompletely understood. Here, we present a genome-scale method to discover genes driving cell-state specification. Applying this method, we identify a transcription factor, BCL11B, which drives stem cell self-renewal *in vitro*, by inhibiting differentiation into the basal lineage. To validate BCL11B's functional role, we use two-dimensional colony-forming and three-dimensional tissue differentiation assays to assess the lineage differentiation potential and functional abilities of primary human mammary cells. These findings show that BCL11B regulates mammary cell differentiation and demonstrate the utility of our proposed genome-scale strategy for identifying lineage regulators in mammalian tissues.

INTRODUCTION

Adult tissues are maintained through the regulated self-renewal and differentiation of stem cells that give rise to differentiated cell types. The mammary epithelium typifies this process by undergoing cyclic expansion and contraction during the estrous cycle and further differentiating during pregnancy and lactation (Fata et al., 2001; Macias and Hinck, 2012; Schedin et al., 2000). This extensive tissue turnover demands a continuous source of newly minted differentiated cells of both the luminal and basal epithelial lineages. While the existence of both bipotent and lineage-restricted mammary stem cells (MaSCs) has been established (Davis et al., 2016; Inman et al., 2015; Pal et al., 2017; Rios et al., 2014; Scheele et al., 2017; Shackleton et al., 2006; Stingl et al., 2005, 2006; Van Keymeulen et al., 2017; Van Keymeulen et al., 2011; Visvader and Stingl, 2014; Wang et al., 2015), relatively little is known about the genes that regulate the self-renewal or differentiation of these stem and progenitor cell types.

The identification of self-renewal and lineage commitment regulators in the human mammary gland has been complicated by two main factors. First, while cell surface markers often enrich for stem and more differentiated cell states, they rarely allow investigators to isolate pure subpopulations of mammary epithelial cells (MECs). The absence of definitive markers, particularly those distin-

guishing MaSCs from differentiated cells of the basal lineage, precludes the use of transcriptomic or other profiling strategies to identify candidate regulators that commit bipotent cells to the basal lineage. Second, while powerful transplantation and lineage tracing assays are available for study of the murine mammary gland, a lack of appropriate experimental models precludes similar analyses in human cells (Visvader and Stingl, 2014). While the mouse has been indispensably valuable as an experimental model, murine mammary glands display morphological, developmental, and genomic differences when compared with the human gland (Carroll et al., 2017; Fridriksdottir et al., 2011; Visvader, 2009), suggesting that rodents may not fully replicate aspects of the biology of human mammary tissue. This latter limitation was partially addressed by the recent development of three-dimensional (3D) culture models that support the outgrowth of morphologically complex mammary tissue from primary human MECs (Miller et al., 2017; Sokol et al., 2016).

This study was motivated by the need for systematic methods to identify functional regulators of stem and progenitor cell states. We approached this problem by observing that genes whose expression is sufficient to specify a particular cell lineage would need to be stably repressed in other cell lineages to prevent aberrant differentiation (Mall et al., 2017; Schafer et al., 1990; Schoenherr and Anderson, 1995; Terranova et al., 2006). In principle,

this observation provides a means of distinguishing between downstream markers and key functional regulators of a cell state. Thus, in a mixed population of cells containing two cell lineages, genes sufficient to alter a cell's state would be stably repressed in one lineage and expressed in the other. Even though such genes would either be expressed or repressed in any given cell, profiling of a population of cells would show that such genes display epigenetic marks associated with both gene repression (histone H3K27me3) and active expression (histone H3K4me3) (Barski et al., 2007; Koch et al., 2007).

We use an additional feature to distinguish such "pseudo-bivalent" genes (Ψ-bivalent), expressed in some cell types and repressed in others, from genes that are uniformly bivalent and thus poised for expression (though not actually expressed) in all of the cells within a population. At an epigenetic level, Ψ-bivalent genes can be distinguished from truly bivalent genes by markers of transcriptional elongation (histone H3K79me2) and the presence of transcribed mRNAs. We hypothesized that Ψ-bivalent genes encompass functionally significant regulators of stem cell self-renewal and lineage commitment. Here, we identify a cohort of such Ψ-bivalent genes in a mammary stem/progenitor cell line. We validate the role of one of these genes, *BCL11B*, in driving the self-renewal of primary human multipotent MaSCs in two-dimensional (2D) and 3D assays of progenitor activity, differentiation potential, and tissue development. Furthermore, we show that BCL11B functions to maintain *in vitro* multipotency by specifically inhibiting basal lineage commitment.

RESULTS

Identification of Pseudo-bivalent TFs as Candidate Cell Lineage Specifiers
To identify candidate genes regulating cell state, we used the MCF10A cell line as a model system (Soule et al., 1990). When seeded into 3D collagen gels, single MCF10A cells can form morphologically complex ductal-lobular tissue rudiments (organoids) (Figures 1A and 1B), indicating that this line contains bipotent stem cells capable of differentiating and self-organizing into the tree-like architecture characteristic of mammary tissue. In addition to forming ductal-lobular organoids, single MCF10A cells also form either duct-only or lobule-only organoids, indicating the presence of lineage-committed progenitors (Figure 1B). Because this cell line contains the lineage-committed basal and luminal progenitor cell states required for tissue morphogenesis (Krause et al., 2008; Sarrio et al., 2012; Sokol et al., 2015), we set out to identify differentially expressed transcription factors (TFs) that may specify these states.

While many factors are differentially expressed between cell states, we were interested specifically in factors capable of reprogramming cellular lineage. We reasoned that the promoters of such factors would be actively repressed in other lineages, since, if this were not the case, stochastic fluctuations in their expression could lead to inappropriate lineage switching. Thus, a factor capable of driving cells into lineage A would be expressed in cells of that lineage while stably repressed in other cell lineages. We identified such factors using chromatin immunoprecipitation sequencing (ChIP-seq) against histone modifications marking transcriptional activation (H3K4me3), transcriptional repression (H3K27me3), and active transcriptional elongation (H3K79me2) (Figures 1C and 1D).

Based on the above reasoning, we were interested specifically in finding Ψ-bivalent TFs that appeared bivalent on the population level but were in fact either expressed or repressed in individual cells (Figure 1C). These factors would be stably activated (H3K4me3+ promoter) in a subset of cells and stably repressed (H3K27me3+ promoter) in another subpopulation. We identified a total 1,895 H3K4me3+ TFs and 1,135 H3K27me3+ TFs. We identified 55 TFs whose promoters were marked with both H3K4me3 and H3K27me3 peaks on the population level (see Experimental Procedures for details on peak calling). Of these bivalent TFs, 23 also contained H3K79me2 peaks within their gene body, indicating active elongation, suggesting that the majority of these genes were expressed in a subset of cells. However, since H3K79 methylation status is regulated in part by cell cycle status (Schulze et al., 2009), to definitively identify genes being actively transcribed we performed RT-PCR, which revealed that 48 of the bivalent TFs expressed detectable mRNAs on the population level (Table S1 and Figure 1E). We classified these 48 TFs as Ψ-bivalent candidate regulators of differentiation.

Candidate Regulatory TFs Mark Cell States in the Human Mammary Gland
To determine whether any of these candidate regulatory TFs play a role in human MEC identity, we asked whether their expression distinguishes mature cell types within the human gland *in vivo*. We collected single cells from elective patient reduction mammoplasty tissue and performed single-cell qPCR to quantify expression of the 48 candidate TFs, as well as established cell-state markers (Figure 2A). Using unsupervised hierarchical clustering (Figure 2B) and principal component dimensionality reduction (Figure 2C), the cells separated into three distinct clusters. Cells in two of the clusters expressed markers of the basal lineage (*CD10*, *SNAI1*, *SNAI2*, *ITGA6*) (Ballard et al., 2015; Moritani et al., 2002; Sarrio et al., 2012; Stingl et al., 2006; Villadsen et al., 2007), whereas a third cluster contained cells enriched for higher expression of luminal
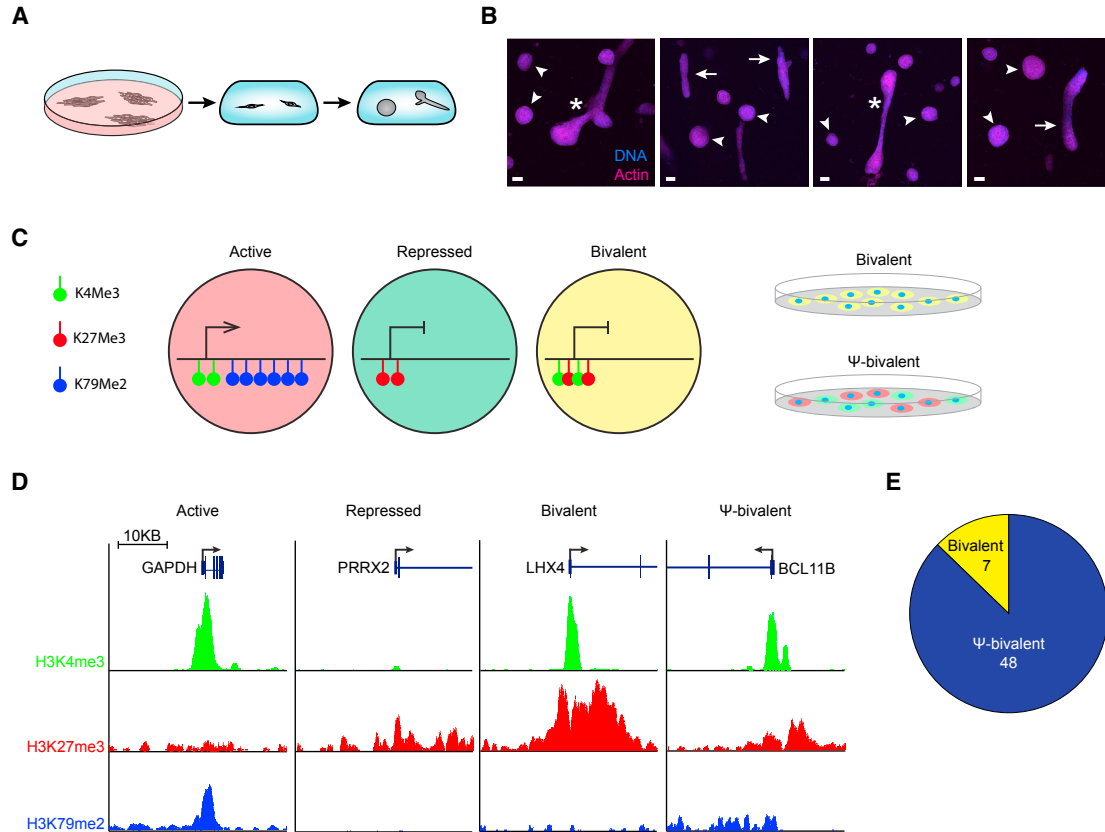
Figure 1. Discovery of Candidate Lineage Specifiers in the MCF10A Mammary Stem Cell Line

(A) Schematic showing the seeding of MCF10A cells into 3D collagen cultures, and the formation of organoids.

(B) Representative confocal microscopy images showing examples of MCF10A organoids after 8 days of 3D culture. Examples of acinar organoids are indicated with arrowheads, ductal organoids are indicated with arrows, and ductal-lobular organoids are indicated with asterisks. Scale bars, 50 μm.

(C) Schematic depiction of epigenetic marks at active, repressed, bivalent, and Ψ-bivalent genes.

(D) Representative results of ChIP-seq run for histone H3K4me3, histone H3K27me3, and histone H3K79me2, showing active, repressed, bivalent, and pseudo-bivalent genes in a mixed population of MCF10A cells.

(E) Summary of bivalent and Ψ-bivalent TF loci calls from ChIP-seq and RT-PCR results.

lineage markers (*EPCAM*, *GATA3*, *SOX9*, *KRT18*, *ETV6*) (Abd El-Rehim et al., 2004; Asselin-Labat et al., 2007; Guo et al., 2012; Kouros-Mehr et al., 2006; Shehata et al., 2012; Su et al., 1996; Taylor-Papadimitriou et al., 1989; Tognon et al., 2002) (Figures 2B and 2D).

Each of these three clusters showed statistically significant enrichment of expression of at least one of our candidate regulatory TFs (Wilcoxon rank-sum test, p < 0.01; Figure 2D). The luminal cluster was enriched for expression of *ATF3*, *BARX2*, and *ZNF175*. Basal cluster 1, which expresses high levels of the epithelial to mesenchymal transition (EMT) TF, *SNAI1*, was also enriched for the expression of *FLI1*. Basal cluster 2, which expresses high levels of the basal markers *SNAI2* and *CD10*, was also enriched for the expression of another EMT TF, *ZEB2*. Notably, a recent study by Bach et al. (2017) identified a subpopulation of

basal cells in the murine mammary gland that was also enriched for *Zeb2* expression, raising the possibility that basal cluster 2 represents a human homolog of this murine cell state. In summary, our findings raise the possibility that two distinct human basal epithelial cell types exist, which may be regulated by *FLI1* and *ZEB2*, respectively.

Candidate Regulatory TFs Are Enriched in Cultured Primary MaSCs

In order to determine whether any of these 48 candidate cell-state specifiers controlled the MaSC state, we first needed to enrich stem and progenitor cells from our patient sample. To do so, we used a recently described method (Miller et al., 2017; Sokol et al., 2016) of *in vitro* culture of freshly isolated primary cells in 3D hydrogels to expand stem and progenitor cells (Figure 3A). When grown in these
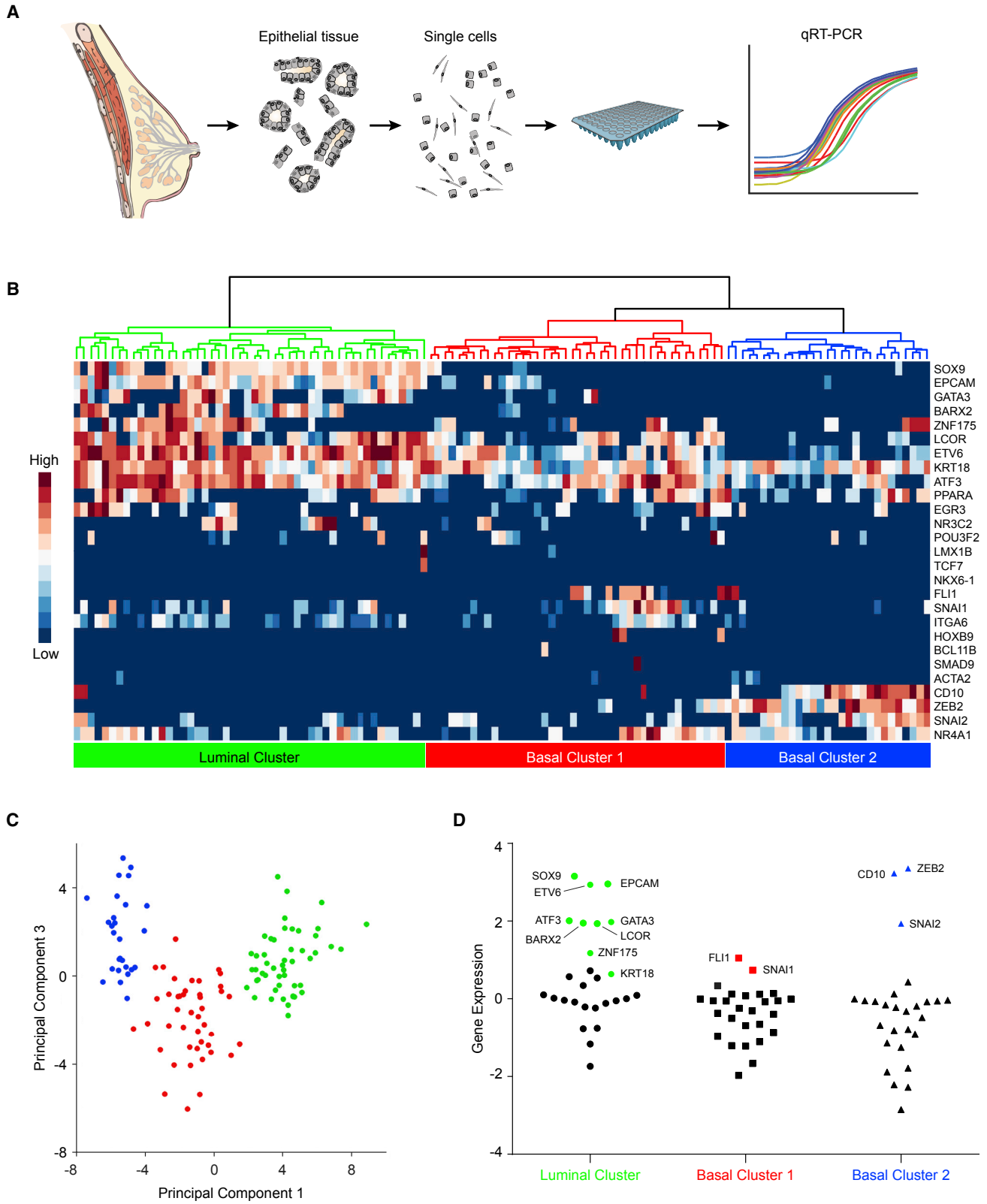
(legend on next page)
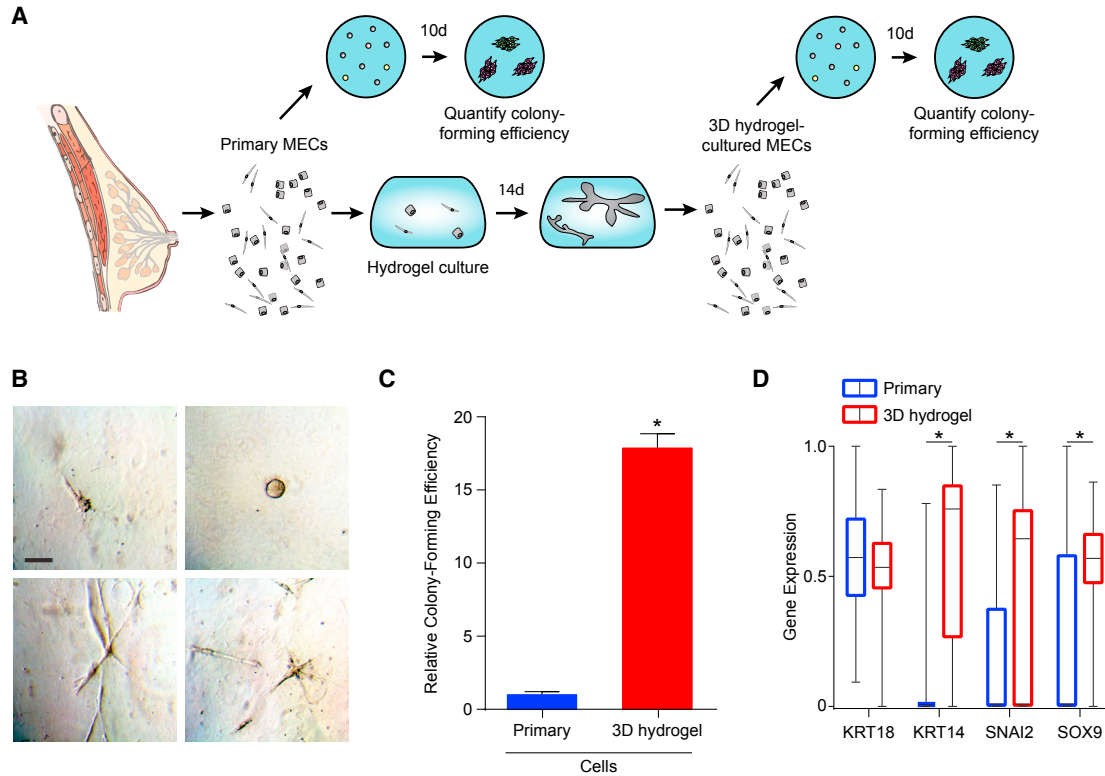
Figure 3. Expansion of Primary Human Mammary Stem/Progenitor Cells in 3D Hydrogel Culture

(A) Schematic showing the seeding of primary MECs in 2D colony-forming assays, before and after culture in 3D hydrogels, to assay stem/progenitor activity. 10d, 10 days; 14d, 14 days.

(B) Representative brightfield microscopy images of tissue rudiments formed from single primary MECs. Scale bar, 100 μm.

(C) Quantification of colony-forming efficiency in uncultured primary cells and 3D hydrogel-cultured cells. Data plotted are normalized to primary cells, and demonstrate a roughly 18-fold increase in stem/progenitor activity in cells cultured in 3D hydrogels (n = 3 independent experiments).

(D) Gene expression, as determined by single-cell qPCR, in primary cells and 3D hydrogel-cultured cells. Expression values are gene normalized and set to a range from 0 to 1. Box and whiskers plot depicts the median, interquartile range, and minimum and maximum values.

Error bars represent mean ± SD. *p < 0.05 by Student's t test.

hydrogels, single cells form simple 3D structures over the course of 14 days (Figure 3B). The cells within these immature structures are strongly enriched for stem/progenitor cell states, as demonstrated by a roughly 18-fold increase in colony formation when plated in 2D culture (Figure 3C). These cultured cells express higher levels of *KRT14*, indicating an expansion of the basal population, which contains all of the bipotent stem cell activity in the mammary

epithelium. They also express higher levels of *SNAI2* and *SOX9*, whose co-expression is required for the stem cell state in the mouse mammary epithelium (Guo et al., 2012) (Figure 3D).

To determine whether any of our 48 candidate cell-state regulatory TFs were enriched in these cultured stem/progenitor cells, we performed single-cell qPCR. This analysis revealed two clusters of cultured cells: a basal cluster


Figure 2. Candidate Lineage Specifiers Distinguish Luminal and Basal Cell Types in Primary Human Mammary Epithelial Cells

(A) Schematic depicting the isolation of patient mammary epithelial cells and single-cell qPCR.

(B) Single-cell qPCR of established cell-state markers and candidate lineage-specifying TFs results in one luminal cluster and two basal clusters of cells.

(C) Principal component analysis of single-cell gene expression data shows three distinct clusters of cells.

(D) Enrichment of expression of each gene is plotted for each of the three clusters. Colored and labeled points indicate genes significantly enriched in a cluster, relative to cells outside of that cluster (p < 0.01 by Wilcoxon rank-sum test).
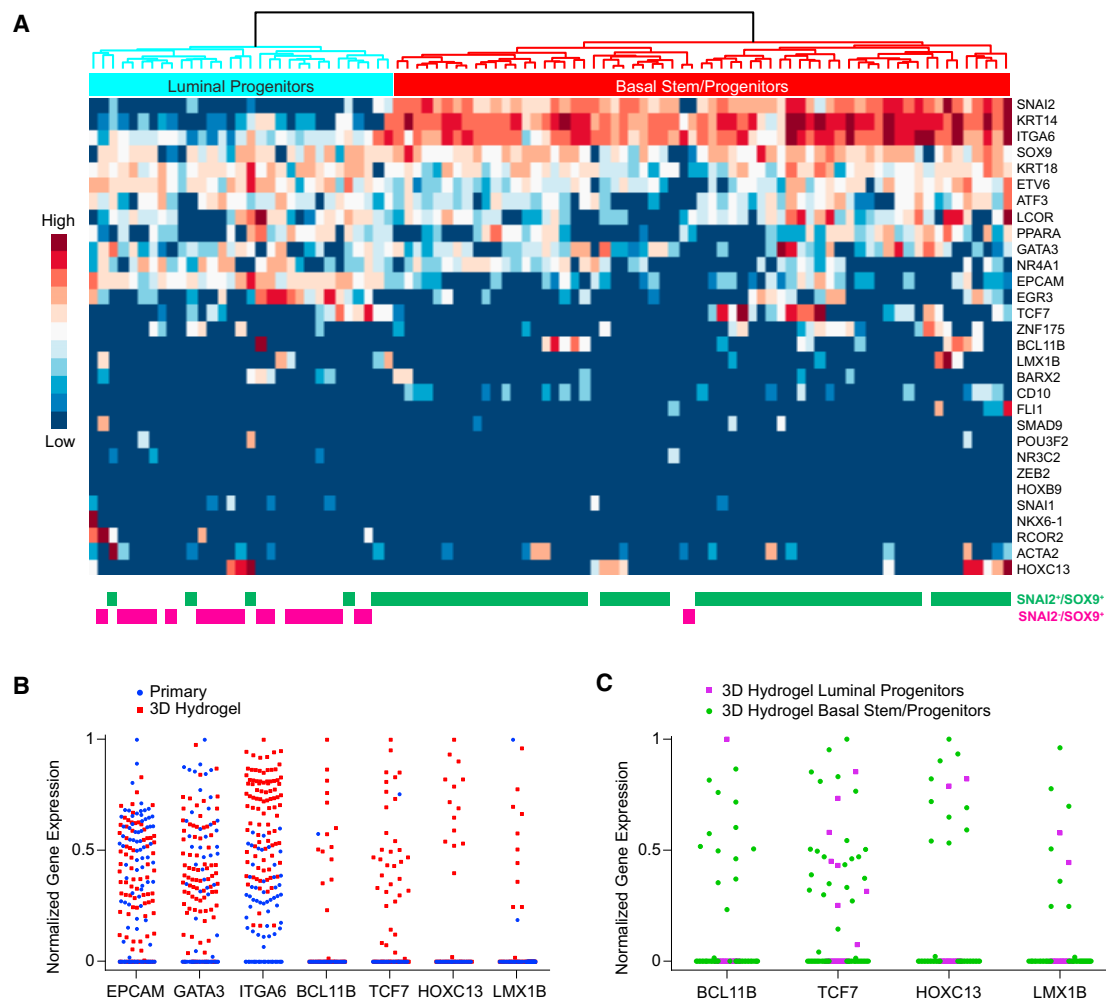
Figure 4. Expression of *BCL11B* Is Enriched in Human Mammary Stem/Progenitor Cells

(A) Results of single-cell qPCR of established cell-state markers and candidate lineage-specifying TFs in primary MECs cultured in 3D hydrogels for 14 days. Cells separate into a cluster composed of basal stem/progenitor cells (indicated by co-expression of *SNAI2* and *SOX9*) and luminal progenitor cells (expressing *SOX9*, but not *SNAI2*).

(B) Normalized gene expression of three cell-state markers (*EPCAM*, *GATA3*, and *ITGA6*) is not significantly changed, but four candidate lineage-specifying TFs (*BCL11B*, *TCF7*, *HOXC13*, and *LMX1B*) are significantly enriched (p < 0.01 by Mann-Whitney test) in cultured MECs (3D hydrogel) relative to uncultured (primary) cells (n = 121 primary cells and 94 3D hydrogel-cultured cells).

(C) Normalized gene expression of *BCL11B*, *TCF7*, *HOXC13*, and *LMX1B* in cultured cells from the two clusters is plotted (n = 64 basal cells and 30 luminal cells).

marked by enriched *ITGA6* and *KRT14* expression and a luminal cluster marked by enriched *EPCAM* expression (Figure 4A). Comparing expression of our candidate TFs in stem/progenitor cells with mature primary cells, we found four TFs with significantly increased expression in stem/progenitor cells: *BCL11B*, *HOXC13*, *LMX1B*, and *TCF7* (Figure 4B). Of these four genes, *BCL11B* was most restricted to the basal population, which contains the bipotent MaSCs (Figure 4C). We therefore set out to test the role of BCL11B in the differentiation and self-renewal of human MaSCs.

BCL11B Is Required for the Self-Renewal of 3D Cultured MCF10A Cells

First, we examined the expression of BCL11B protein in MCF10A cells. As predicted by its epigenetic Ψ-bivalence, we found its expression to be heterogeneous, with some cells not detectably expressing protein and other expressing high levels of protein (Figures 5A and 5B). Next we set out to test the functional role of BCL11B in MCF10A cell self-renewal and differentiation. We used two independent short hairpin RNAs (shRNAs) to inhibit expression of BCL11B (Figure 5C), neither of which affected the growth
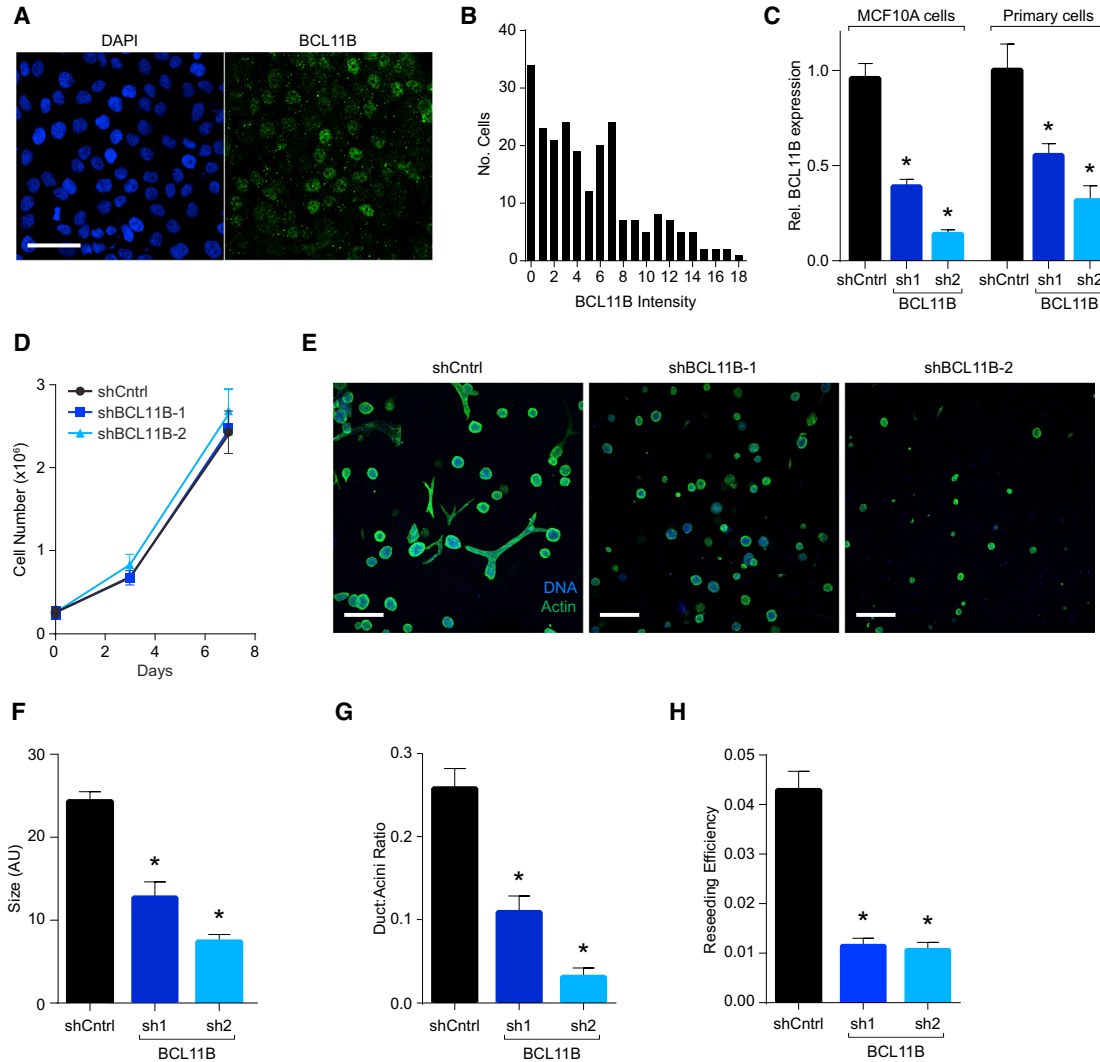
Figure 5. BCL11B Is Required for MCF10A Self-Renewal and Duct Formation

(A) Immunofluorescent staining of BCL11B protein expression. Representative of three independent replicates. Scale bar, 50 μm.

(B) Histogram showing quantification of BCL11B from (A) (n = 228 cells).

(C) Demonstration of knockdown of *BCL11B* expression, by qPCR, in MCF10A cells (left) or primary patient MECs (right) transduced with two shRNAs. Plotted is expression relative to uninfected control, normalized to GAPDH (n = 4 independent experiments).

(D) Growth curves showing no difference in growth rate of MCF10A cells transduced with shRNAs targeting *BCL11B* or *Luciferase* (shCntrl) (n = 3 independent experiments).

(E) Representative images of MCF10A cells transduced with either shRNAs targeting *BCL11B* or *Luciferase* (shCntrl), grown in 3D culture for 7 days and stained with DAPI (blue) and phalloidin (green). Scale bars, 200 μm (n = 4 independent experiments).

(F) Quantification of the size of organoids from (E) (n = 4 independent experiments).

(G) Quantification of the ratio of ductal organoids to acinar organoids from (E) (n = 4 independent experiments).

(H) Quantification of the fraction of MCF10A cells from the three lines that are capable of giving rise to secondary organoids after being removed from primary 3D culture and reseeded (n = 3 independent experiments).

Error bars represent mean ± SD. *$p < 0.05$ by Student's t test.

rate of the cells in 2D culture (Figure 5D). When seeded into 3D culture, MCF10A cells with inhibited expression of BCL11B prematurely differentiate, as indicated by a reduction in organoid size (Figures 5E and 5F), failure to form acini (Figures 5E and 5G), and loss of ability to repopulate new organoids when reseeded into secondary 3D cultures (Figure 5H). These phenotypes all pointed toward a loss of MCF10A self-renewal upon inhibition of BCL11B
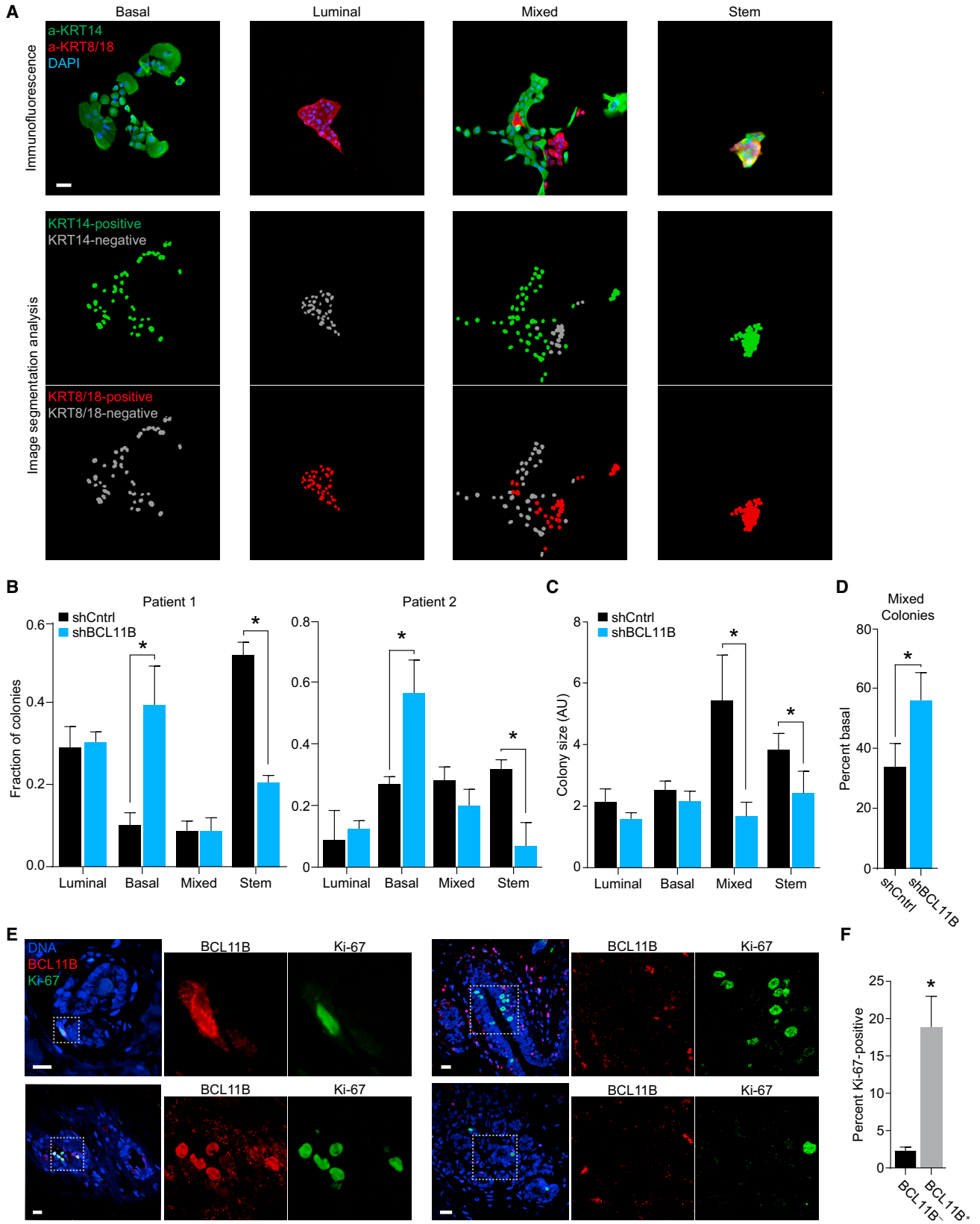
(legend on next page)

when cells are grown in differentiating conditions (3D culture).

BCL11B Drives Primary Human MaSC Self-Renewal by Inhibiting Basal Differentiation

To determine the role of BCL11B in the differentiation of primary human stem and progenitor cells *in vitro*, we performed a canonical 2D plate differentiation assay in wild-type patient cells and in cells transduced with shRNAs targeting *BCL11B*. Inhibition of BCL11B in two independent patient samples did not affect the formation of luminal colonies or mixed luminal-basal colonies. However, the fraction of bipotent stem cell colonies, identified in this assay by the presence of co-expression of luminal and basal markers (Clayton et al., 2004; Proia et al., 2011; Smalley et al., 1998; Sokol et al., 2015; Stingl et al., 2001), dropped significantly, indicative of a loss of bipotent stem cell self-renewal in these *in vitro* conditions (Lim et al., 2009; Petersen and Polyak, 2010; Santagata et al., 2014). This was accompanied by a reciprocal increase in mature basal colonies, indicating that, rather than self-renewing, stem cells were differentiating into basal progenitors in the absence of BCL11B (Figures 6A and 6B).

In addition to the shift in colony ratios, stem colonies and colonies of mixed lineage were significantly smaller when BCL11B expression was inhibited, indicating decreased self-renewal of the bipotent cells that give rise to these colonies (Figure 6C). Mixed lineage colonies also showed an increase in the fraction of basal cells within the colony, indicating a bias toward basal differentiation upon loss of BCL11B (Figure 6D).

A recent study by Cai et al. (2017) found that, in murine MECs, *Bcl11b* drives a quiescent MaSC state, inconsistent with our demonstration of a functional role of BCL11B in driving self-renewal in proliferating MaSCs. Recent advances in the study of mammary epithelial biology have raised concerns about the ability of canonical assays, such as 2D colony formation and single-cell transplantation, to drive non-physiological dedifferentiation of MECs (Chang et al., 2014; Lloyd-Lewis et al., 2017; Van Keymeulen et al., 2011). In light of these findings, we sought to interrogate the role of BCL11B in experimental models that do not involve the isolation of single cells or culture on 2D dishes. First, we sought to determine whether BCL11B expression in human mammary cells *in vivo* was predictive of proliferation or quiescence. Immunofluorescence staining of freshly fixed human tissue showed that this was not the case, as BCL11B-positive MECs were more likely to express the proliferation marker Ki-67 than BCL11B-negative MECs (Figures 6E and 6F).

Next, to examine the functional role of BCL11B in a more native tissue development context, we turned to our hydrogel culture system. When intact, patient-derived tissue fragments consisting of hundreds of epithelial cells are seeded into hydrogels and their growth is driven by stem cell expansion, followed by differentiation (Sokol et al., 2016). To determine BCL11B's functional role in this *in vitro* tissue development assay, we infected patient tissues with lentiviral constructs driving the expression of either a *BCL11B*-targeting shRNA and GFP, or a control *Luciferase*-targeting shRNA and GFP (Figure 7A). This infection strategy leads to only a fraction of the cells within a tissue being transduced with the virus, allowing for the direct comparison of infected cells and wild-type cells within a single tissue. In the shLuc-GFP control, GFP-positive cells can be seen distributed throughout the tissue, in both the outer basal layer and the inner luminal layer. In contrast, in the shBCL11B-GFP infected tissues, GFP-positive cells do not contribute to new outgrowths from the tissue, and instead differentiate into elongated basal cells, restricted to the core of the structure (Figure 7B). This indicated that, in contrast with the control shRNA, inhibition of BCL11B depleted

Figure 6. BCL11B Drives Self-Renewal of Primary Human MaSCs by Inhibiting Basal Differentiation

(A) Demonstration of automated colony-scoring pipeline and representative basal, luminal, mixed, and stem colonies derived from single primary human MECs grown for 10 days in 2D culture. Scale bar, 50 μm.

(B) Two independent patient samples show an increase in basal colony formation and subsequent decrease in stem colonies upon the inhibition of BCL11B, relative to control shRNA.

(C) Quantification of colony size shows decreased size of mixed and stem colonies upon inhibition of BCL11B, relative to control shRNA.

(D) Mixed colonies show an increase in the fraction of basal cells when BCL11B is inhibited, relative to control. For (B–D), n = 3 independent experiments per patient, with a total of 139 patient 1 shCntrl colonies, 150 patient 1 shBCL11B colonies, 115 patient 2 shCntrl colonies, and 87 patient 2 shBCL11B colonies.

(E) Four representative immunofluorescence images of primary human mammary tissue stained for BCL11B (red), Ki-67 (green), and DAPI (blue). The dotted regions are enlarged at right, and individual channels are depicted. The two fields on the left show examples of BCL11B+/Ki-67+ cells, while the two fields on the right show examples of BCL11B+/Ki-67− and BCL11B−/Ki-67+ cells (n = 3 independent patient samples). Scale bars, 20 μm.

(F) Quantification of Ki-67 staining from (E). Plotted is the percentage of BCL11B-negative cells (black bar) and BCL11B-positive cells (gray bar) that expressed Ki-67 (n = 3 independent patient samples).

Error bars represent mean ± SD. *p < 0.05 by Student's t test.
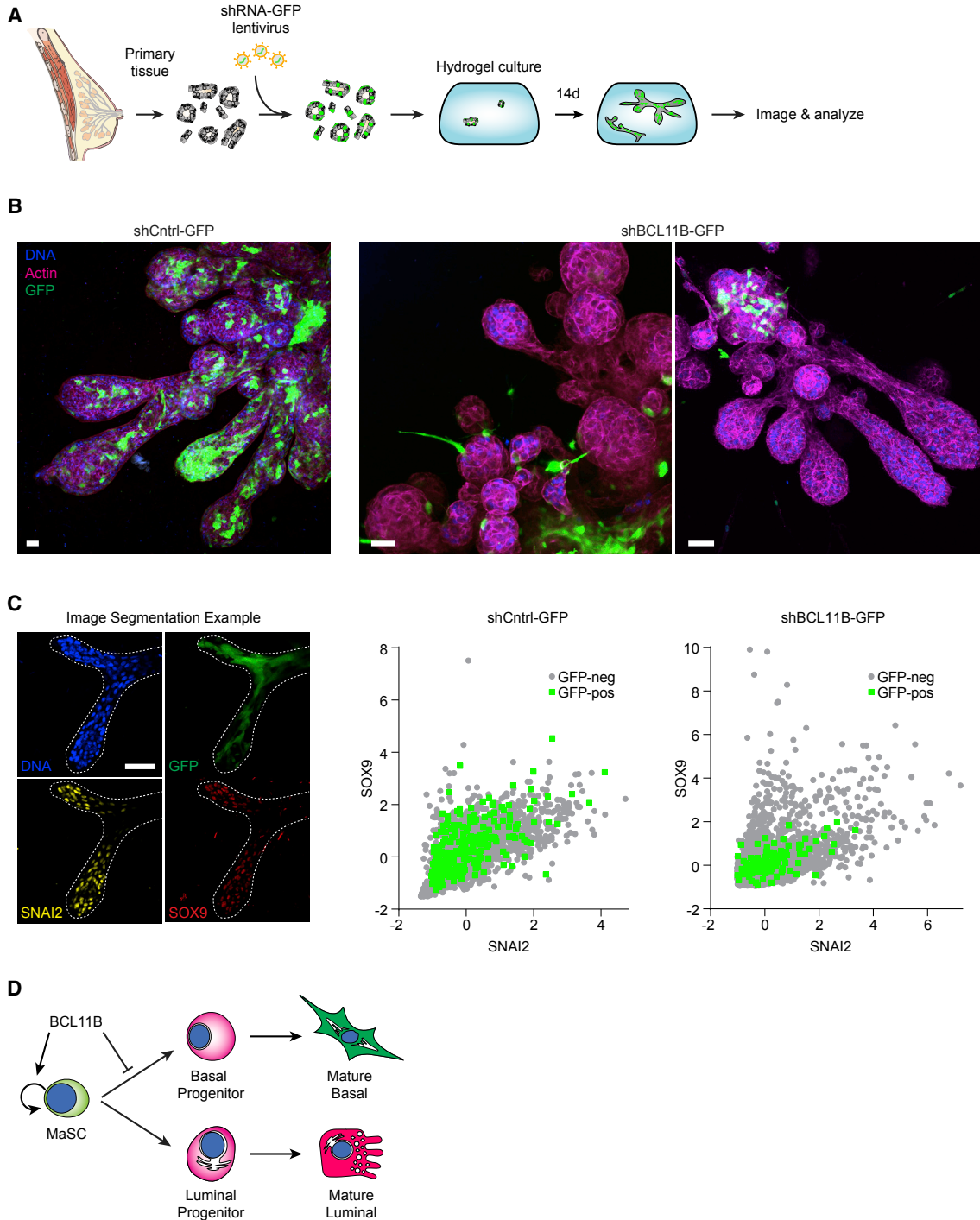
Figure 7. BCL11B Is Necessary for Growth of Primary Human Mammary Tissues in Hydrogel Culture

(A) Schematic showing the collection of primary human mammary epithelial tissue fragments, infection of tissues with lentiviral constructs, and seeding and growth of tissues in 3D hydrogel culture.

(B) Confocal microscopy images of tissues grown in culture for 14 days (representative of n = 6 independent experiments).

(C) Quantification of SOX9 and SNAI2 protein in hydrogel-cultured tissues infected with lentiviral shRNA constructs driving GFP as in (A and B). Left: example of image segmentation of a representative tissue stained for SNAI2 and SOX9; segmentation analysis allows for identification of GFP-positive and -negative cells, and quantification of SNAI2 and SOX9 expression in each cell. A shCntrl-GFP tissue is depicted in this example. Right: quantification of SOX9 and SNAI2 is plotted for individual cells from shCntrl-GFP or shBCL11B-GFP infected

(legend continued on next page)

stemness and drove basal differentiation in the infected cells.

To determine the effect of BCL11B-knockdown on stem cell self-renewal and lineage commitment, we cultured patient cells in this assay for 2 weeks and then stained for expression of SNAI2 and SOX9. In MECs, SNAI2 is a marker of the basal lineage, SOX9 is a marker of the luminal lineage, and co-expression of these factors in a single cell is indicative of the MaSC state (Guo et al., 2012). In shLuc-GFP infected control cells, we found no change in the expression of SNAI2 and SOX9, relative to uninfected control cells. However, cells infected with shBCL11B-GFP showed a marked decrease in SOX9 expression relative to tissue-matched uninfected cells (Figure 7C). This result further suggests that loss of BCL11B leads to a release from the dual-positive stem cell state and a bias toward basal differentiation, due to de-repression of basal lineage commitment of MaSCs (Figure 7D).

DISCUSSION

Here, we present an experimental strategy to identify TFs that control stem cell self-renewal and lineage commitment decisions. This approach directly addresses fundamentally important questions about how the stem cell's behavior is controlled, without the need for physically isolating distinct cell states based on marker expression. Applying this strategy to the human mammary gland identified a number of known regulators of mammary cell states (*SOX9*, *SNAI2*, and *GATA3*), as well as numerous additional candidates.

Our strategy relies on the simple assumption that factors capable of driving the specification of a cell state must be stably and actively repressed in all alternative states. While not the only way to stably repress a gene's expression, epigenetic modification (such as the trimethylation of histone H3K27) is one mechanism by which stable gene inactivation can occur. Thus, while our strategy cannot claim to identify all such factors, we were confident that the factors it did highlight would be enriched for functionally crucial genes.

Using a combination of canonical and recently developed 2D and 3D culture assays to assess human MaSC properties *in vitro*, we have demonstrated that one of these candidates, *BCL11B*, is required for self-renewal. *BCL11B's* expression is enriched in stem and progenitor cell types,

relative to mature cell types, and drives MaSC self-renewal by preventing basal lineage commitment. This role is reminiscent of the role of SNAI2, an EMT TF that inhibits luminal differentiation and is required for bipotency of MaSCs. This raises the possibility that BCL11B may act as a reciprocal differentiation inhibitor, preventing basal lineage commitment, while SNAI2 prevents luminal lineage commitment.

Interestingly, a recent study also revealed a crucial role for Bcl11b in maintenance of the MaSC state in the murine mammary gland (Cai et al., 2017). This study, however, suggested that Bcl11b expression drives quiescence in MaSCs. We find that, in human MaSCs, BCL11B-expressing cells show no indication of quiescence, as they contribute to 2D colony formation (Figures 6B–6D) and 3D tissue development (Figures 7B and 7C). In addition, BCL11B-expressing cells in fixed patient tissue sections comprise both cycling and non-cycling cells, as determined by Ki-67 staining (Figures 6E and 6F). In human MaSCs, our results suggest that BCL11B drives self-renewal by inhibiting commitment to the basal lineage and subsequent differentiation, rather than by inhibiting proliferation to drive quiescence.

While more work is needed to clarify the reason for these distinctions, it is possible that the BCL11B$^+$ MaSCs in the human and mouse glands perform distinct functions. One key developmental difference between the human and murine gland is that the murine gland only produces functional terminal ductal lobular units (TDLUs) upon pregnancy, whereas the human gland undergoes more extensive development during puberty. Cai et al. (2017) limited their study to virgin mice, but the possibility remains that Bcl11b$^+$ MaSCs could become proliferative once TDLUs form. This hypothesis suggests that the proliferation of BCL11B$^+$ MaSCs may contribute to maintenance of the gland at this more advanced stage, although further studies are required to directly test this possibility.

Previous studies of *Bcl11b* in murine T cells suggest that its expression is regulated by numerous TFs known to also have key roles in MEC biology (Li et al., 2013; Longabaugh et al., 2017). In those cells, *Bcl11b's* expression is controlled by a far downstream enhancer, which contains binding sites for Gata3 (a regulator of luminal fate commitment; Asselin-Labat et al., 2007), Runx1 (a regulator of stem cell differentiation; Sokol et al., 2015), and Tcf7 (a TF activated downstream of Wnt signaling, which drives MaSC self-renewal; Zeng and Nusse, 2010). Of note, *TCF7* was also a TF that appeared in our list of Ψ-bivalent TFs,

tissues grown in hydrogels for 14 days (n = 2,187 shLuc-GFP-negative cells, 231 shLuc-GFP-positive cells, 3,678 shBCL11B GFP-negative cells, and 121 shBCL11B GFP-positive cells; three independent experiments).
(D) Model of BCL11B's role in MaSC self-renewal and basal differentiation.
Scale bars, 50 μm.

indicating that its expression may have an important role in regulating the MaSC state. While it is unknown whether these three factors activate or repress *BCL11B* transcription in the human mammary gland, this downstream enhancer could act as a central hub for MEC lineage commitment signaling.

EXPERIMENTAL PROCEDURES

Ethics Statement
Patient reduction mammoplasty tissue samples that would otherwise have been discarded as medical waste following surgery were obtained in compliance with all relevant laws, using protocols approved by the institutional review board at Maine Medical Center. Since tissues were fully anonymized before transfer, this research was provided exemption status by the Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology. All patients enrolled in this study signed an informed consent form to agree to participate in this study and for publication of the results.

Cells and Tissues
MCF10A cells were obtained from ATCC and cultured in mammary epithelium growth medium (MEGM) (Lonza CC-3150) supplemented with 100 ng/mL cholera toxin (Sigma-Aldrich), 1× GlutaMax, and 1× penicillin and streptomycin (Gibco). Reduction mammoplasty samples were obtained from the BioBank at Maine Medical Center. Upon arrival, tissue was chopped into roughly 3 mm^3 fragments using a scalpel, and then incubated with 3 mg/mL collagenase (Roche) and 0.7 mg/mL hyaluronidase (Sigma-Aldrich) in MEGM at 37°C overnight. Epithelial tissue was separated from stroma by centrifugation; washed; and frozen in 90% fetal bovine serum (FBS), 10% DMSO for long-term storage. Following thawing, samples were depleted for fibroblasts via incubation at 37°C in DMEM with 10% FBS for 1 hr.

Chromatin Immunoprecipitation Sequencing
Chromatin immunoprecipitation was performed as previously described (Lee et al., 2006). 1 × 10^8 MCF10A cells were fixed in 11% formaldehyde solution for 10 min at room temperature (RT), quenched with 100 mM glycine, and collected by scraping. Cells were lysed in LB1 (lysis buffer 1) for 10 min at 4°C, and resuspended in LB2 for 10 min at RT. Cells were then resuspended in LB3 for sonication. Samples were sonicated for 5 min total (30 s × 10 with 1 min gaps; 18–21 W; QSonica microtip sonicator). Following sonication, samples were spun at 20,000 g and split into three tubes (for H3K4me3, H3K27me3, H3K79me2). At this time, 50 μL was saved as an input control.

Magnetic beads (Dynabeads; Invitrogen) were prepared by washing in block (0.5% BSA in PBS) followed by an overnight incubation at 4°C in the appropriate antibody ([H3K4me3; Millipore #07-473; lot JBC 1888194; 1:50], [H3K27me3; Abcam Ab6002; lot GR 19843-1; 1:20], or [H3K79me2; Abcam Ab3594-100; lot 803369; 1:50]). In addition, 25 μL of antibody-bound beads was

added to the sonicated sample and incubated overnight at 4°C on a rotator.

The samples were washed with buffers B, C, and I, and eluted in 200 μL of elution buffer at 65°C for 15 min. Crosslinks were reversed overnight at 65°C. Samples were then digested for 2 hr with 400 μg/mL RNase A at 37°C followed by 400 μg/mL Proteinase K digestion at 55°C for 2 hr. DNA was then extracted with phenol:chloroform:isoamyl and pelleted with ice-cold ethanol. Samples were resuspended in water and analyzed on a 1% agarose gel.

Samples were sequenced at the Whitehead Genome Technology Core using the Illumina HiSeq 2000 platform and TruSeq adapters. Reads were aligned to the hg38 reference genome using Bowtie2 (version 2.3.1) (Langmead and Salzberg, 2012) using default parameters except that up to one mismatch was accepted in the seed sequence. Peaks were identified by MACS2 (version 2.1.1) (Zhang et al., 2008) using the input sample as a control while bypassing the shifting model and with the broad peak setting with a broad cutoff of 0.05 for H3K4me3 and H3K79me2 conditions. To determine bivalence, peaks were first associated with their closest gene, where antisense genes were collapsed onto their respective genes, using the HOMER software suite (version 4.9.1) (Heinz et al., 2010). Next, peaks were filtered to keep only peaks within 20 kb upstream or 5 kb downstream of a transcriptional start site. Genes that retained a H3K4me3 peak and a H3K27me3 peak were identified as bivalent. These genes were then visually inspected to eliminate false-positives, yielding a final set of 55 bivalent genes.

Immunofluorescence
Immunofluorescence was performed as previously described for 2D (Sokol et al., 2015) and 3D (Miller et al., 2017; Sokol et al., 2016) culture. Antibodies and stains used were BCL11B (Cell Signaling, 12120; 1:200), KRT14 (Life Technologies, RB-9020-P; 1:300), KRT8/18 (Vector, VP-C407; 1:500), SNAI2 (Cell Signaling, 9585; 1:400), SOX9 (Sigma, WH0006662M2; 1:50), Ki-67 (Cell Signaling, 9449; 1:800), DAPI (Life Technologies, D1306), Phalloidin-AF647 (Life Technologies, A22287; 1:100), anti-rabbit IgG- AF488 (Cell Signaling, 4,412, 1:1,000), and anti-mouse IgG-AF555 (Cell Signaling, 4409, 1:1,000).

Immunofluorescence on human tissue was performed as previously described (Skibinski et al., 2014). Primary and secondary antibodies used were BCL11B (Cell signaling, 12120; 1: 100), Ki-67 (Cell Signaling, 9585; 1: 400), anti-Rabbit IgG-AF555 (Life Technologies, A21430; 1:500), and anti-Mouse IgG-AF488 (Life Technologies, A11001; 1:500). Sections were treated with ProLong Gold antifade reagent with DAPI (Life Technologies, P36935) and mounted with coverslips.

Microscopy
Epifluorescence microscopy for imaging of 2D colony assays was done using a Nikon TE-2000 microscope. Microscopy of immunofluorescence stained hydrogel cultures and BCL11B in MCF10A cells was done using a Zeiss LSM-700 confocal microscope. Brightfield microscopy was done using a Zeiss Axiovert 25.

RT-PCR

RNA was collected from MCF10A cells in 2D culture using RNeasy kit (Qiagen) and treated with RNase-free DNase (Qiagen). RNA was reverse transcribed into cDNA using iScript (Bio-Rad). PCR was run using ExTaq (Takara) and primers listed in Table S2 for 40 cycles. PCR product was run on a 4% agarose gel and detected with ethidium bromide (Sigma).

Single-Cell Real-Time qRT-PCR

One cell per well was sorted using a flow cytometer into 96-well plates and lysed using CellsDirect Kit (Thermo Fisher). Genomic DNA was degraded using DNase I (Thermo Fisher). Reverse transcription was performed using SuperScript III (Thermo Fisher) and pooled pre-amplification primers at 50 nM each (see Table S3). Twenty cycles of pre-amplification PCR were performed using CellsDirect Kit (Thermo Fisher) and pooled primers listed in Table S3. Free primers were degraded using Exonuclease I (New England BioLabs).

Pre-amped single-cell cDNA was then used to perform real-time qRT-PCR using SsoFast EvaGreen Gene Expression (Fluidigm) protocols with 96.96 Dynamic Array plates on a BioMark HD System (Fluidigm). Primers used are listed in Table S2.

Computational Analyses and Statistics

Clustering and principal component analysis were performed using MATLAB. All statistics were calculated using GraphPad Prism.

Colony Assay

Primary mammary tissue from elective reduction mammoplasties was dissociated to single cells and plated at clonogenic density in tissue culture plates with the addition of either shLuc (shCntrl) or shBCL11B lentivirus. Colonies were grown for 10 days and then fixed and stained using immunofluorescence for expression of KRT14 and KRT8/18. Colonies were classified as luminal, basal, mixed, or stem using CellProfiler. Briefly, cells were first identified using DAPI. Corresponding KRT14 and KRT8/18 signal was then scored for each cell. Colonies in which all cells expressed KRT14 only were classified as basal. Colonies in which all cells expressed KRT8/18 only were classified as luminal. Colonies with a mixture of KRT14-only and KRT8/18-only cells were classified as mixed. Colonies with cells that co-expressed KRT14 and KRT8/18 were classified as stem.

3D Culture

3D culture of primary tissue fragments using hydrogels was performed as previously described (Miller et al., 2017; Sokol et al., 2016). Culture of MCF10A cell lines in collagen gels was performed as previously described (Sokol et al., 2015).

Lentiviral Production and Infection

Lentivirus production and cell line infection were performed as previously described (Gupta et al., 2005). To infect tissue fragments, tissues were cultured overnight in MEGM supplemented with 10 μg/mL protamine sulfate and concentrated lentivirus in ultra-low attachment tissue culture plates. After overnight incubation, tissue fragments were seeded into hydrogel cultures.

ACCESSION NUMBERS

The accession number for ChIP-seq data reported in this paper is GEO: GSE109721.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2018.01.036.

AUTHOR CONTRIBUTIONS

D.H.M. and P.B.G. conceived the project and designed experiments. D.H.M., D.X.J., E.S.S., J.R.C., D.A.S., and R.A.G. performed experiments and analyzed results. C.K. provided patient samples and assisted with experimental design. D.H.M. and P.B.G. wrote the manuscript with contributions, editorial review, and approval from all authors.

ACKNOWLEDGMENTS

REFERENCES

Abd El-Rehim, D.M., Pinder, S.E., Paish, C.E., Bell, J., Blamey, R.W., Robertson, J.F., Nicholson, R.I., and Ellis, I.O. (2004). Expression of luminal and basal cytokeratins in human breast carcinoma. J. Pathol. 203, 661–671.

Asselin-Labat, M.L., Sutherland, K.D., Barker, H., Thomas, R., Shackleton, M., Forrest, N.C., Hartley, L., Robb, L., Grosveld, F.G., van der Wees, J., et al. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat. Cell Biol. 9, 201–209.

Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D.J., Marioni, J.C., and Khaled, W.T. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat. Commun. 8, 2128.

Ballard, M.S., Zhu, A., Iwai, N., Stensrud, M., Mapps, A., Postiglione, M.P., Knoblich, J.A., and Hinck, L. (2015). Mammary stem cell self-renewal is regulated by Slit2/Robo1 signaling through SNAI1 and mINSC. Cell Rep. 13, 290–301.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution

profiling of histone methylations in the human genome. Cell *129*, 823–837.

Cai, S., Kalisky, T., Sahoo, D., Dalerba, P., Feng, W., Lin, Y., Qian, D., Kong, A., Yu, J., Wang, F., et al. (2017). A quiescent Bcl11b high stem cell population is required for maintenance of the mammary gland. Cell Stem Cell *20*, 247–260.e5.

Carroll, J.S., Hickey, T.E., Tarulli, G.A., Williams, M., and Tilley, W.D. (2017). Deciphering the divergent roles of progestogens in breast cancer. Nat. Rev. Cancer *17*, 54–64.

Chang, T.H., Kunasegaran, K., Tarulli, G.A., De Silva, D., Voorhoeve, P.M., and Pietersen, A.M. (2014). New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. Breast Cancer Res. *16*, R1.

Clayton, H., Titley, I., and Vivanco, M. (2004). Growth and differentiation of progenitor/stem cells derived from the human mammary gland. Exp. Cell Res. *297*, 444–460.

Davis, F.M., Lloyd-Lewis, B., Harris, O.B., Kozar, S., Winton, D.J., Muresan, L., and Watson, C.J. (2016). Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. Nat. Commun. *7*, 13053.

Fata, J.E., Chaudhary, V., and Khokha, R. (2001). Cellular turnover in the mammary gland is correlated with systemic levels of progesterone and not 17beta-estradiol during the estrous cycle. Biol. Reprod. *65*, 680–688.

Fridriksdottir, A.J., Petersen, O.W., and Ronnov-Jessen, L. (2011). Mammary gland stem cells: current status and future challenges. Int. J. Dev. Biol. *55*, 719–729.

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

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.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Inman, J.L., Robertson, C., Mott, J.D., and Bissell, M.J. (2015). Mammary gland development: cell fate specification, stem cells and the microenvironment. Development *142*, 1028–1042.

Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., et al. (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome Res. *17*, 691–707.

Kouros-Mehr, H., Slorach, E.M., Sternlicht, M.D., and Werb, Z. (2006). GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell *127*, 1041–1055.

Krause, S., Maffini, M.V., Soto, A.M., and Sonnenschein, C. (2008). A novel 3D in vitro culture model to study stromal-epithelial interactions in the mammary gland. Tissue Eng. Part C Methods *14*, 261–271.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. *1*, 729–748.

Li, L., Zhang, J.A., Dose, M., Kueh, H.Y., Mosadeghi, R., Gounari, F., and Rothenberg, E.V. (2013). A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. Blood *122*, 902–911.

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. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913.

Lloyd-Lewis, B., Harris, O.B., Watson, C.J., and Davis, F.M. (2017). Mammary stem cells: premise, properties, and perspectives. Trends Cell Biol. *27*, 556–567.

Longabaugh, W.J.R., Zeng, W., Zhang, J.A., Hosokawa, H., Jansen, C.S., Li, L., Romero-Wolf, M., Liu, P., Kueh, H.Y., Mortazavi, A., et al. (2017). Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proc. Natl. Acad. Sci. USA *114*, 5800–5807.

Macias, H., and Hinck, L. (2012). Mammary gland development. Wiley Interdiscip. Rev. Dev. Biol. *1*, 533–557.

Mall, M., Kareta, M.S., Chanda, S., Ahlenius, H., Perotti, N., Zhou, B., Grieder, S.D., Ge, X., Drake, S., Euong Ang, C., et al. (2017). Myt1l safeguards neuronal identity by actively repressing many non-neuronal fates. Nature *544*, 245–249.

Miller, D.H., Sokol, E.S., and Gupta, P.B. (2017). 3D primary culture model to study human mammary development. Methods Mol. Biol. *1612*, 139–147.

Moritani, S., Kushima, R., Sugihara, H., Bamba, M., Kobayashi, T.K., and Hattori, T. (2002). Availability of CD10 immunohistochemistry as a marker of breast myoepithelial cells on paraffin sections. Mod. Pathol. *15*, 397–405.

Pal, B., Chen, Y., Vaillant, F., Jamieson, P., Gordon, L., Rios, A.C., Wilcox, S., Fu, N., Liu, K.H., Jackling, F.C., et al. (2017). Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. Nat. Commun. *8*, 1627.

Petersen, O.W., and Polyak, K. (2010). Stem cells in the human breast. Cold Spring Harb. Perspect. Biol. *2*, a003160.

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.

Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322–327.

Santagata, S., Thakkar, A., Ergonul, A., Wang, B., Woo, T., Hu, R., Harrell, J.C., McNamara, G., Schwede, M., Culhane, A.C., et al. (2014). Taxonomy of breast cancer based on normal cell phenotype predicts outcome. J. Clin. Invest. *124*, 859–870.

Sarrio, D., Franklin, C.K., Mackay, A., Reis-Filho, J.S., and Isacke, C.M. (2012). Epithelial and mesenchymal subpopulations within

normal basal breast cell lines exhibit distinct stem cell/progenitor properties. Stem Cells *30*, 292–303.

Schafer, B.W., Blakely, B.T., Darlington, G.J., and Blau, H.M. (1990). Effect of cell history on response to helix-loop-helix family of myogenic regulators. Nature *344*, 454–458.

Schedin, P., Mitrenga, T., and Kaeck, M. (2000). Estrous cycle regulation of mammary epithelial cell proliferation, differentiation, and death in the Sprague-Dawley rat: a model for investigating the role of estrous cycling in mammary carcinogenesis. J. Mammary Gland Biol. Neoplasia *5*, 211–225.

Scheele, C.L., Hannezo, E., Muraro, M.J., Zomer, A., Langedijk, N.S., van Oudenaarden, A., Simons, B.D., and van Rheenen, J. (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. Nature *542*, 313–317.

Schoenherr, C.J., and Anderson, D.J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science *267*, 1360–1363.

Schulze, J.M., Jackson, J., Nakanishi, S., Gardner, J.M., Hentrich, T., Haug, J., Johnston, M., Jaspersen, S.L., Kobor, M.S., and Shilatifard, A. (2009). Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation. Mol. Cell *35*, 626–641.

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.

Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I.A., Avril, S., Prater, M., Eirew, P., Caldas, C., Watson, C.J., et al. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. *14*, R134.

Skibinski, A., Breindel, J.L., Prat, A., Galvan, P., Smith, E., Rolfs, A., Gupta, P.B., Labaer, J., and Kuperwasser, C. (2014). The Hippo transducer TAZ interacts with the SWI/SNF complex to regulate breast epithelial lineage commitment. Cell Rep. *6*, 1059–1072.

Smalley, M.J., Titley, J., and O'Hare, M.J. (1998). Clonal characterization of mouse mammary luminal epithelial and myoepithelial cells separated by fluorescence-activated cell sorting. In Vitro Cell. Dev. Biol. Anim. *34*, 711–721.

Sokol, E.S., Sanduja, S., Jin, D.X., Miller, D.H., Mathis, R.A., and Gupta, P.B. (2015). Perturbation-expression analysis identifies RUNX1 as a regulator of human mammary stem cell differentiation. PLoS Comput. Biol. *11*, e1004161.

Sokol, E.S., Miller, D.H., Breggia, A., Spencer, K.C., Arendt, L.M., and Gupta, P.B. (2016). Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. Breast Cancer Res. *18*, 19.

Soule, H.D., Maloney, T.M., Wolman, S.R., Peterson, W.D., Jr., Brenz, R., McGrath, C.M., Russo, J., Pauley, R.J., Jones, R.F., and Brooks, S.C. (1990). Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. Cancer Res. *50*, 6075–6086.

Stingl, J., Eaves, C.J., Zandieh, I., and Emerman, J.T. (2001). Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast Cancer Res. Treat. *67*, 93–109.

Stingl, J., Raouf, A., Emerman, J.T., and Eaves, C.J. (2005). Epithelial progenitors in the normal human mammary gland. J. Mammary Gland Biol. Neoplasia *10*, 49–59.

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.

Su, L., Morgan, P.R., and Lane, E.B. (1996). Expression of cytokeratin messenger RNA versus protein in the normal mammary gland and in breast cancer. Hum. Pathol. *27*, 800–806.

Taylor-Papadimitriou, J., Stampfer, M., Bartek, J., Lewis, A., Boshell, M., Lane, E.B., and Leigh, I.M. (1989). Keratin expression in human mammary epithelial cells cultured from normal and malignant tissue: relation to in vivo phenotypes and influence of medium. J. Cell Sci. *94* (Pt 3), 403–413.

Terranova, R., Pereira, C.F., Du Roure, C., Merkenschlager, M., and Fisher, A.G. (2006). Acquisition and extinction of gene expression programs are separable events in heterokaryon reprogramming. J. Cell Sci. *119*, 2065–2072.

Tognon, C., Knezevich, S.R., Huntsman, D., Roskelley, C.D., Melnyk, N., Mathers, J.A., Becker, L., Carneiro, F., MacPherson, N., Horsman, D., et al. (2002). Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. Cancer Cell *2*, 367–376.

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.

Van Keymeulen, A., Fioramonti, M., Centonze, A., Bouvencourt, G., Achouri, Y., and Blanpain, C. (2017). Lineage-restricted mammary stem cells sustain the development, homeostasis, and regeneration of the estrogen receptor positive lineage. Cell Rep. *20*, 1525–1532.

Villadsen, R., Fridriksdottir, A.J., Ronnov-Jessen, L., Gudjonsson, T., Rank, F., LaBarge, M.A., Bissell, M.J., and Petersen, O.W. (2007). Evidence for a stem cell hierarchy in the adult human breast. J. Cell Biol. *177*, 87–101.

Visvader, J.E. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev. *23*, 2563–2577.

Visvader, J.E., and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. Genes Dev. *28*, 1143–1158.

Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.O., Yang, L., and Zeng, Y.A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. Nature *517*, 81–84.

Zeng, Y.A., and Nusse, R. (2010). Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. Cell Stem Cell *6*, 568–577.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.