



# The Co-factor of LIM Domains (CLIM/LDB/NLI) Maintains Basal Mammary Epithelial Stem Cells and Promotes Breast Tumorigenesis

Michael L. Salmans<sup>1,2</sup>, Zhengquan Yu<sup>3</sup>, Kazuhide Watanabe<sup>1</sup>, Eric Cam<sup>1</sup>, Peng Sun<sup>1</sup>, Padhraic Smyth<sup>2,4</sup>, Xing Dai<sup>1</sup>, Bogi Andersen<sup>1,2,5\*</sup>

**1** Department of Biological Chemistry, University of California, Irvine, Irvine, California, United States of America, **2** Institute for Genomics and Bioinformatics, University of California, Irvine, Irvine, California, United States of America, **3** State Key Laboratories for AgroBiotechnology, College of Biological Sciences, China Agricultural University, Beijing, PR China, **4** Department of Computer Science, University of California, Irvine, Irvine, California, United States of America, **5** Department of Medicine, University of California, Irvine, Irvine, California, United States of America

## Abstract

Mammary gland branching morphogenesis and ductal homeostasis relies on mammary stem cell function for the maintenance of basal and luminal cell compartments. The mechanisms of transcriptional regulation of the basal cell compartment are currently unknown. We explored these mechanisms in the basal cell compartment and identified the Co-factor of LIM domains (CLIM/LDB/NLI) as a transcriptional regulator that maintains these cells. Clims act within the basal cell compartment to promote branching morphogenesis by maintaining the number and proliferative potential of basal mammary epithelial stem cells. *Clim2*, in a complex with LMO4, supports mammary stem cells by directly targeting the *Fgfr2* promoter in basal cells to increase its expression. Strikingly, Clims also coordinate basal-specific transcriptional programs to preserve luminal cell identity. These basal-derived cues inhibit epidermis-like differentiation of the luminal cell compartment and enhance the expression of luminal cell-specific oncogenes *ErbB2* and *ErbB3*. Consistently, basal-expressed Clims promote the initiation and progression of breast cancer in the MMTV-PyMT tumor model, and the *Clim*-regulated branching morphogenesis gene network is a prognostic indicator of poor breast cancer outcome in humans.

**Citation:** Salmans ML, Yu Z, Watanabe K, Cam E, Sun P, et al. (2014) The Co-factor of LIM Domains (CLIM/LDB/NLI) Maintains Basal Mammary Epithelial Stem Cells and Promotes Breast Tumorigenesis. *PLoS Genet* 10(7): e1004520. doi:10.1371/journal.pgen.1004520

**Editor:** Kent W. Hunter, National Cancer Institute, United States of America

**Received:** March 12, 2014; **Accepted:** June 3, 2014; **Published:** July 31, 2014

**Copyright:** © 2014 Salmans et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files. All gene expression microarray data are available from GEO (Accession number GSE57724).

**Funding:** This work was supported by National Institutes of Health Grant AR44882 (to BA), National Institutes of Health - National Library of Medicine Biomedical Informatics Training Grant T15LM0744 (to MLS), and California Institute for Regenerative Medicine (CIRM) Training Grant TG2-01152 (to MLS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: bogi@uci.edu

## Introduction

Mouse mammary gland morphogenesis begins during mid-gestation with the development of two bilateral epithelial ridges along the ventral epidermis that form invasive, multipotent stem cell-enriched placodes migrating into the underlying mesenchyme, later branching to form a rudimentary ductal tree by birth. The structure remains relatively quiescent until hormonal stimuli at puberty initiate the formation of stem cell-enriched terminal end buds (TEBs) that rapidly proliferate and invade into the mammary fat pad, frequently bifurcating to generate a ductal network. The resulting ducts consist of luminal epithelial cells surrounded by a basal myoepithelial cell layer. Basal and luminal cells communicate signals to each other through paracrine and direct cell-cell interactions to regulate proper morphogenesis [1,2]; the nature of these complex cell-cell interactions remains to be fully defined.

Mammary stem cells (MaSCs) coordinate ductal morphogenesis and homeostasis of the luminal and basal cell compartments in the adult mammary gland. Two models have been proposed for the function of MaSCs in the mammary gland: either committed

unipotent luminal and basal epithelial stem cells maintain their respective compartment [3], or bipotent MaSCs in the basal cell compartment give rise to both lineages [4]. CD49<sup>Hi</sup>CD29<sup>Hi</sup>CD24<sup>+</sup> basal epithelial cells maintain a small population of basal stem cells (BSCs) with the potential to regenerate a functional mammary gland [5,6], while luminal stem cells (LSCs), enriched in the CD49<sup>Lo</sup>CD29<sup>Lo</sup>CD24<sup>+</sup> luminal epithelial cell population, maintain unipotent potential to preserve the luminal cell population [3,4] and cannot regenerate the mammary gland. Transcription factors that control the maintenance of stem cells and lineage specification along the mammary epithelial cell (MEC) hierarchy are best characterized in the luminal cell compartment [7]. However, the current knowledge of transcriptional regulation of BSCs and their differentiation is limited.

The LIM domain, a tandem zinc finger motif that serves as an interface for protein-protein interactions, is found in a variety of protein families, including the LIM-homeodomain (Lhx) and LIM-only (LMO) transcription factors [8]. Originally discovered as a co-activator of Lhx and LMO [9–12], the co-factor of LIM domains (CLIM/LDB/NLI) coordinates the organization of

## Author Summary

Recent advancements in mammary gland biology demonstrate conflicting models in maintenance of basal and luminal cell compartments by either unipotent or bipotent mammary stem cells. However, the molecular mechanisms underlying control of the basal cell compartment, including stem cells, remain poorly understood. Here we explore the currently unknown transcriptional mechanisms of basal stem cell (BSC) maintenance, in addition to addressing the role of the basal cell compartment in preserving luminal cell fate and promoting development of human breast tumors of luminal origin. We discover a novel function for the Co-factor of LIM domains (Clim) transcriptional regulator in promoting mammary gland branching morphogenesis and breast tumorigenesis through maintenance of the basal stem cell population. The transcriptional networks coordinated by Clims in basal mammary epithelial cells also preserve the identity of luminal epithelial cells, demonstrating a crosstalk between these two cellular compartments. Furthermore, we correlate developmental gene expression data with human breast cancer to investigate the role of developmental pathways during the initiation and progression of breast cancer. The gene regulatory networks identified during development, including those specifically coordinated by Clims, correlate with breast cancer patient outcome, suggesting these genes play an important role in the progression of breast cancer.

transcriptional complexes through two key regulatory domains: the amino-terminal homodimerization domain and the carboxy-terminal LIM-interacting domain (LID) [9,13,14]. Additionally, Clims can associate with other DNA-binding proteins, including GATA, bHLH, and Otx family members [15,16]. It is through these higher order transcriptional complexes that Clims mediate enhancer-promoter interactions to influence gene regulation [17–19]. The Clim family consists of two highly conserved members: the ubiquitously expressed *Clim2* (*Ldb1/Nli*), and the more regionally expressed *Clim1* (*Ldb2*) [10]. Germline deletion of *Clim2* in the mouse causes embryonic lethality by day 9.5 [20], while *Clim1* knockout mice display no developmental defects (as indicated by the Mouse Genome Informatics website) possibly due to compensatory effects by *Clim2*. The function of Clims in mammary gland development has not been investigated. However, in breast cancer, Clim expression correlates with estrogen receptor (ER) positivity where Clims are involved in coordinating transcriptional networks through ER $\alpha$  [21].

Because of the lethality of germline *Clim2* deletion and the possible overlap in function between *Clim1* and *Clim2*, we investigated the role of Clims in the mammary gland using the Keratin 14 (K14) promoter to express a dominant-negative Clim (DN-Clim), consisting of the highly conserved LID fused to a nuclear localization signal and Myc-tag [22], directing its expression to basal MECs. The DN-Clim molecule, which targets all Clim family members, interacts with LIM domain proteins through the LID, but lacks the dimerization domain essential for coordinating enhancer-promoter interactions [18], thereby inhibiting Clim-mediated transcriptional regulation. Using the K14-DN-Clim mouse model we discovered novel transcriptional regulatory programs coordinated by Clims in the basal cells of the mammary gland that promote branching morphogenesis through the maintenance of BSCs, in addition to controlling basal cell-derived transcriptional programs that preserve luminal cell identity. Transcriptional networks within the stem cell-enriched

TEB and differentiated duct correlate with human breast cancer gene expression and patient outcome in an opposing fashion; a significant proportion of these networks are regulated by Clims. Furthermore, Clim-mediated transcriptional machinery in the basal cell population permits tumor initiation and progression in the MMTV-Polyoma Middle T (PyMT) breast cancer model. Altogether, these data describe novel mechanisms of transcriptional regulation in BSCs and their role in breast cancer.

## Results

### Clims promote mammary gland branching morphogenesis

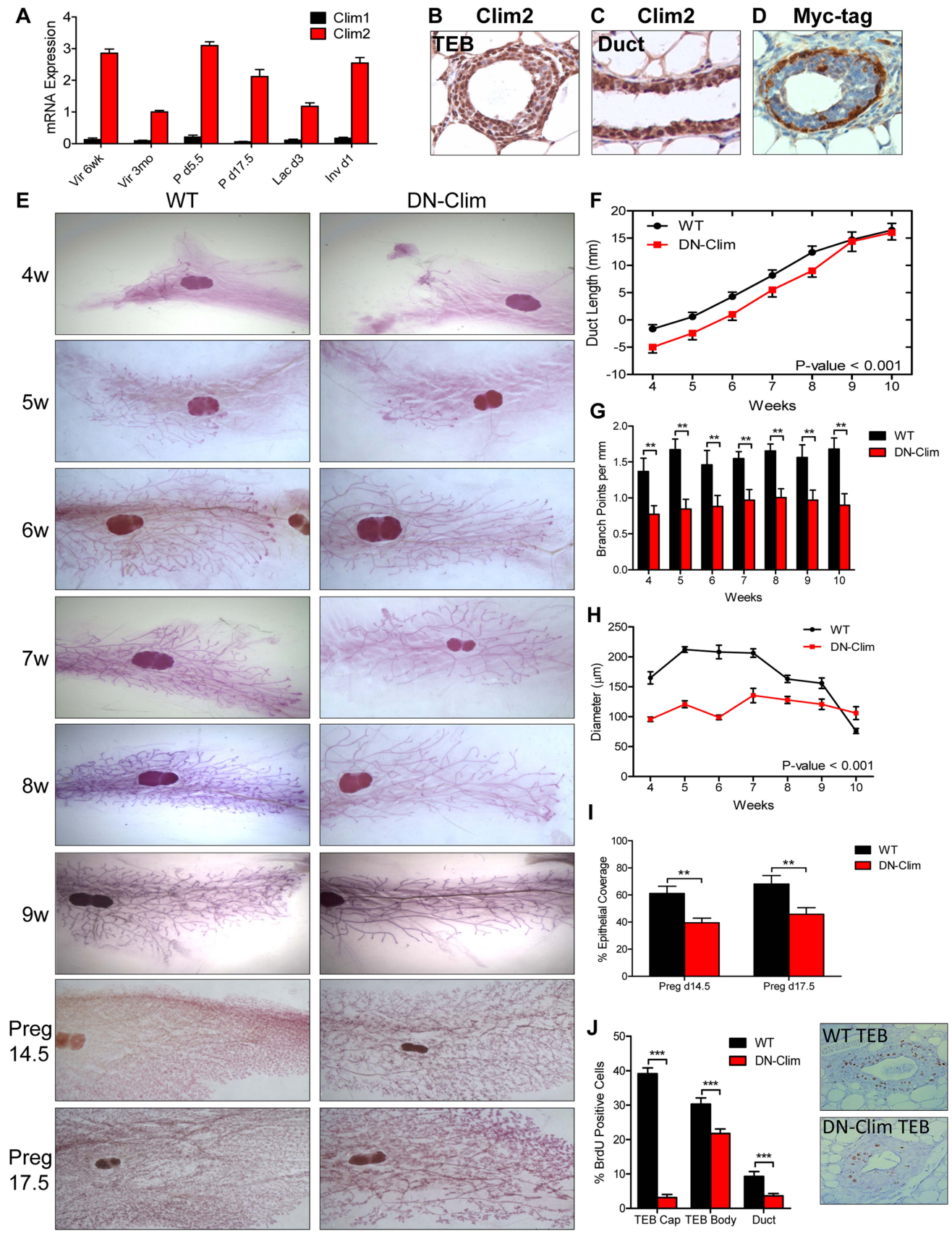
Among the two Clim family members, *Clim2* is preferentially expressed throughout mammary development (Figure 1A), localized to basal and luminal MECs (Figure 1B–C, S1A). To study the role of Clims in MECs we used the K14-DN-Clim mouse model, targeting DN-Clim expression to the K14-positive basal MECs, as indicated by Myc-tagged DN-Clim expression (Figure 1D). DN-Clim is only expressed in basal cells of ducts and TEB cap cells with rare expression in the TEB body cells, consistent with K14 expression in a subset of these cells [23].

Abnormalities in branching morphogenesis of DN-Clim mammary glands appear as early as 4 weeks of age and continue through adulthood and late pregnancy (Figure 1E), featuring delayed ductal elongation, reduced branching frequency, and smaller TEBs (Figure 1F–H). While the delayed ductal penetration dissipates by early adulthood, the reduced branching frequency and smaller TEBs persist throughout development. Apart from the reduction in size, the TEBs of DN-Clim mammary glands appear to be morphologically normal. While branching morphogenesis is typically completed by week 10, as denoted by the absence of TEBs, small DN-Clim TEBs persist beyond 10 weeks of age (Figure 1H). Collectively, these data demonstrate an important role for basal cell-expressed Clims in mammary gland branching morphogenesis through enhancement of ductal penetration, branching frequency, and TEB size.

During pregnancy, DN-Clim mammary glands at day 14.5 and 17.5 exhibit reduced epithelial density (Figure 1E and 1I); however, this could be a result of the low ductal density prior to pregnancy, rather than defective side branching and alveologenesis. Additionally, while DN-Clim females produce litters comparable in size to normal mice, they can only support a limited number of pups beyond postnatal day 2 (Figure S1B), likely due to decreased milk availability from the underdeveloped mammary epithelial tree, as the surviving pups from DN-Clim females grow at a normal rate (Figure S1C) and wild type lactating females can support full litters derived from DN-Clim females. Thus, Clims are required for a fully functional lactating mammary gland.

### Clims maintain cell proliferation during branching morphogenesis

To determine the cellular mechanism for impaired mammary gland development in DN-Clim mice we measured bromodeoxyuridine (BrdU) incorporation in proliferating cells, observing significant decreases in BrdU<sup>+</sup> cells in DN-Clim TEB and ducts (Figure 1J). Consistent with the expression pattern of DN-Clim in TEB cap cells, we observed the most drastic reduction of proliferation in the cap cells, while the body cells exhibit only a moderate, yet significant, decrease in proliferation. The fact that TEB body cells, which rarely express DN-Clim, have reduced proliferative potential suggests a non-autonomous mechanism of Clim-mediated cell proliferation through signals that originate





**Figure 1. Clims promote branching morphogenesis and proliferation.** (A) *Clim2* mRNA is predominantly expressed over *Clim1* throughout mammary gland development. Gene expression is normalized to 18s mRNA. (B–C) IHC showing expression of *Clim2* in the (B) TEB and (C) duct in mammary glands from 6 week old mice. *Clim2* expression is observed in basal and luminal cells, with slightly higher expression in basal cells. (D) IHC detecting the Myc-tag of DN-Clim demonstrates expression in the K14-positive basal layer of mammary glands from 6 week old mice. (E) Whole mount analysis of mammary glands throughout branching morphogenesis and pregnancy demonstrates defective branching morphogenesis and TEB maintenance in DN-Clim mammary glands. Images are representative of at least three mice per time point. (F) Mean ductal penetration of the three longest ducts throughout branching morphogenesis indicates a delay in DN-Clim ductal penetration. Data are normalized using the center of the lymph node as the central reference point. P-value derived from repeated measures ANOVA. (G) Mean number of branch points for the three longest ducts normalized to the length of the duct (mm) indicates reduced branching frequency in DN-Clim mammary glands. \*\* P-value<0.01 determined by Student's t-test. (H) TEB diameter (μm) is reduced in DN-Clim mammary glands throughout branching morphogenesis. P-value derived from repeated measures ANOVA. (I) Proportion of pregnancy stage mammary fat pad filled with epithelial tissue demonstrates a reduced epithelial density in DN-Clim mammary glands. \*\* p-value<0.01 determined by Student's t-test. (J) Proportion of BrdU positive cells in TEB (cap and body cells) and duct cells is reduced in DN-Clim mammary glands. Representative images of the TEB are shown. \*\*\* p-value<0.01 determined by Student's t-test. Quantification for A, F–J represent mean ± SEM for at least three mice. doi:10.1371/journal.pgen.1004520.g001

from the cap cells. Thus, Clims promote proliferation in both the TEB and duct during branching morphogenesis.

### Gene sets differentially expressed in the TEB and duct represent distinct developmental functions during branching morphogenesis and are differentially expressed in breast cancer subtypes

The current knowledge of the molecular mechanisms that regulate branching morphogenesis is limited, especially regarding the transcriptional programs within the stem cell-enriched TEB and differentiated duct structures. To elucidate these transcriptional programs, and specifically those controlled by Clims, we profiled gene expression in TEB and duct cells from wild type (WT) and DN-Clim mice, using laser capture microdissection to isolate these structures from four, six, eight, and ten week old mice. Consistent with its preferential expression during branching morphogenesis, *Clim2* was expressed higher than *Clim1* in both the TEB and duct cells (Figure S2A).

To define developmentally regulated genes over the time course we applied the Bayesian Estimation of Temporal Regulation (BETR) algorithm [24] identifying 8,030 and 6,488 probe sets (7,318 and 5,870 genes) temporally regulated in TEB and duct cells, respectively, with 1,313 TEB and 815 duct genes changing 1.5-fold or greater in two or more time points (Figure 2A–C, Dataset S1 and S2). TEB genes are associated with proliferation, while the duct genes are associated with differentiation (Figure 2D–E). Because the TEB represents a stem cell-enriched population [25], and because of our direct comparison of TEB and duct cells, these gene signatures are representative of mammary stem cell and differentiated MECs, respectively.

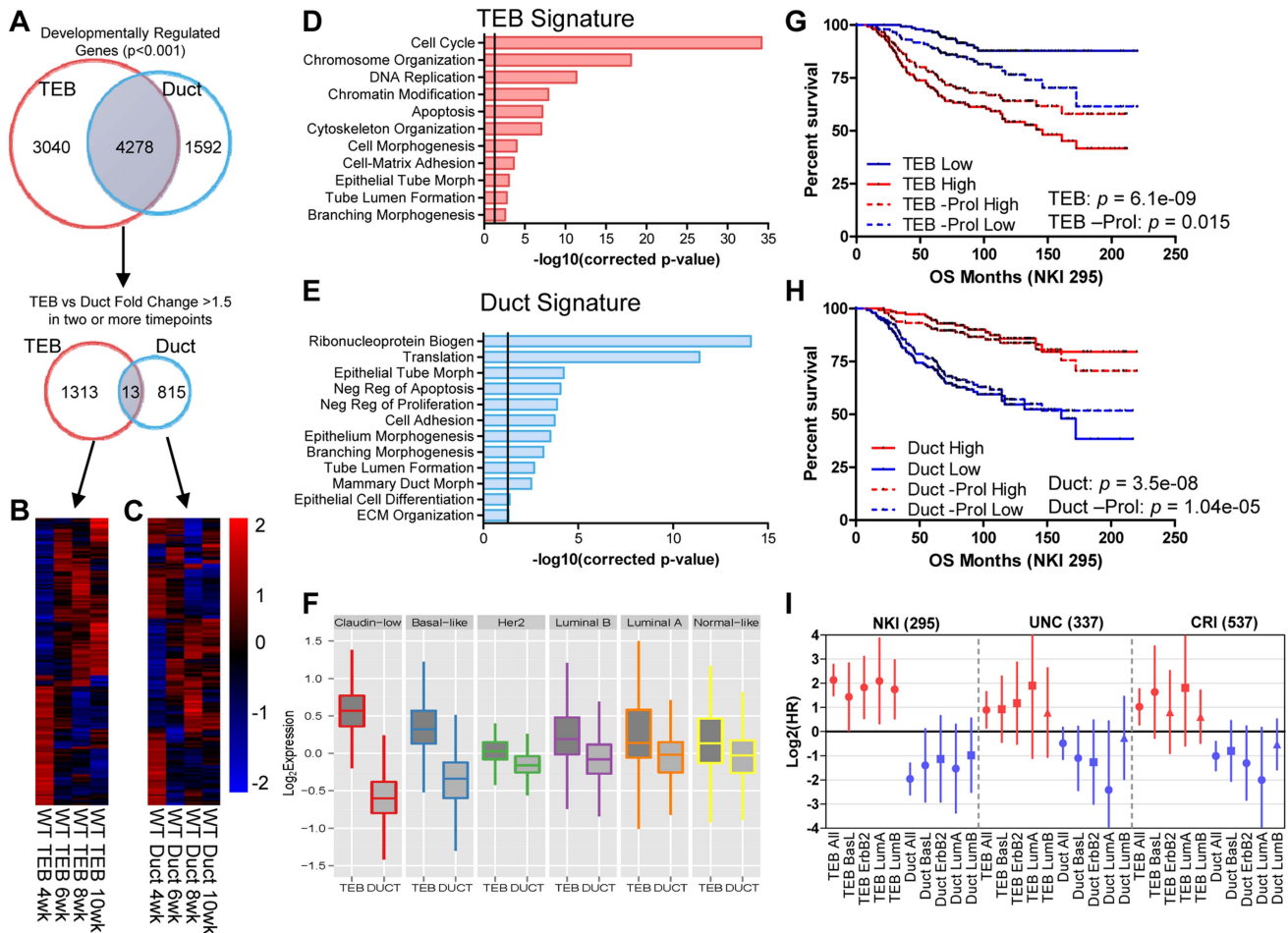
The hyperproliferative and invasive nature of the TEB bears similarity to molecular properties observed in aggressive breast and other adenocarcinomas. In addition, poorly differentiated basal-like tumors are more aggressive and possess stem cell-like transcriptional characteristics compared to the differentiated luminal subtypes [26]. This prompted us to determine whether the less differentiated TEB and fully differentiated duct gene signatures can be used as predictors of prognosis in human breast cancer. Utilizing three data sets that profile the transcriptome and classify the intrinsic subtypes of primary human breast tumors [27–29] we found the TEB signature is highly expressed in the poorly differentiated claudin-low and basal-like breast cancer subtypes, while the duct signature is more highly expressed in the differentiated luminal subtypes (Figure 2F). High TEB signature expression predicted poor overall survival in breast cancer patients (Figure 2G). An opposite trend was observed with the duct signature (Figure 2H), suggesting the expression of differentiation genes suppresses tumorigenesis. These trends are also observed within the different subtypes of breast cancer (Figure 2I). Even

after removal of proliferation genes, which significantly contribute to prognosis prediction [30], the TEB signature retains prognostic relevance, albeit with reduced power (Figure 2G). The duct signature, which is poorly enriched for proliferation genes, was relatively unaffected by removal of these genes (Figure 2H). Thus, these developmental signatures hold prognostic value in breast cancer and may be useful to identify novel genes of interest for understanding the biology of breast cancer in relationship to normal development.

### Clims regulate developmentally associated TEB and duct genes

To determine genes regulated by Clims during branching morphogenesis we used the CyberT algorithm [31] to define differentially expressed genes (DEGs) in the DN-Clim TEB and duct: 66 TEB and 137 duct genes were significantly differentially expressed in at least two time points (Figure 3A–B, Dataset S3 and S4), a sizable proportion of which are differentially expressed in both TEB and duct cells (Figure S2B). Each DEG set is significantly enriched with genes from their respective developmental signature (Figure S2C–D), indicating an important role for Clims in regulating branching morphogenesis. Using the Molecular Signature Database [32] to characterize the biological properties of the *Clim*-regulated TEB and duct gene sets revealed that both have functions in mammary stem cells, pubertal mammary gland development, several breast cancer subtypes, and Wnt-signaling in the mammary gland (Figure 3C–D). Additionally, TEB DEGs participate in metalloproteinase activity and possess serum response factor (SRF) and androgen receptor (AR) transcription factor binding motifs, while duct DEGs are enriched with smooth muscle and mammary basal epithelial cell genes, participate in FGF signaling, and possess TCF3 and SOX9 transcription factor binding motifs. Furthermore, high expression of *Clim*-regulated genes in breast cancer is an indicator of poor prognosis (Figure 3E). Collectively, these data support the involvement of Clims in gene expression networks important for both stem cell function and breast cancer.

To determine genes differentially expressed in basal and luminal cell compartments we sorted MECs into Lin<sup>-</sup>CD29<sup>Hi</sup>CD24<sup>+</sup> and Lin<sup>-</sup>CD29<sup>Lo</sup>CD24<sup>+</sup> populations, respectively (Figure S3A–D). Basal and luminal cell purity from WT and DN-Clim MECs was confirmed by qPCR for K8, K14, and the *K14-DN-Clim* transgene expression (Figure S3E–G); as expected, the *K14-DN-Clim* transgene is only expressed in basal cells, and not luminal cells. Basal markers (K5, K14, and smooth muscle actin) and luminal markers (K8, K18, Gata3, estrogen and progesterone receptors) are significantly overexpressed (CyberT, p-value<0.001) in their respective cell types, indicating effective basal and luminal cell separation (Dataset S5). DN-Clim basal and luminal

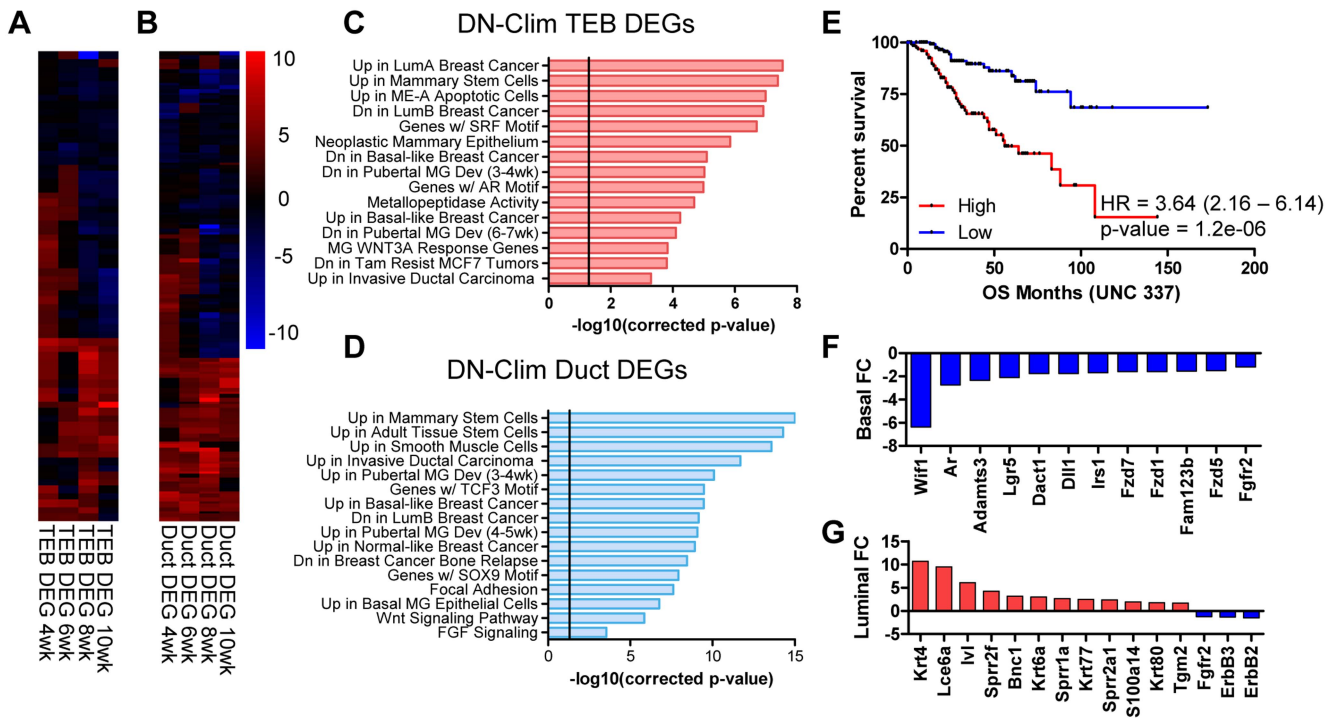


**Figure 2. Developmental TEB and duct gene signatures derived from time course expression profiling predict breast cancer prognosis.** (A) Identification of developmentally regulated genes with the BETR algorithm in TEB and duct cells over the time course of branching morphogenesis (4, 6, 8, and 10 weeks). TEB and duct signatures were further refined by selecting genes with at least 1.5-fold expression change in at least two time points for each cell type. (B–C) Heat map of resulting (B) 1,313 TEB and (C) 815 duct gene signatures expressed at least 1.5-fold or greater in their respective cell type in at least two time points. (D–E) Biological processes associated with (D) TEB and (E) duct gene signatures. The TEB and duct signatures represent a proliferation and differentiation signature, respectively. Vertical line represents p-value at 0.05. (F) TEB and Duct signatures are differentially expressed in the molecular breast cancer subtypes. The TEB and duct signatures are more highly expressed in the poorly- and well-differentiated breast cancer subtypes, respectively. (G–H) Analysis of overall survival (OS) in NKI 295 cohort of patients classified into (G) TEB and (H) duct high and low expression groups based on median expression of the respective gene signature, including the effects of removing proliferation genes (-Prol). High expression of the proliferative TEB signature confers poor prognosis, while high expression of the differentiated duct signature confers improved prognosis. P-value derived from Log-rank test. (I) Hazard ratios (HR) derived from overall survival analysis for TEB and duct signatures in the molecular subtypes of breast cancer in three separate data sets demonstrate similar trends observed in G–H. Red samples represent TEB signature and blue samples represent duct. Error bars represent 95% confidence interval; filled circles represent p-value < 0.1, squares represent trends with p-value < 0.25, and triangles represent non-significant changes with p-value > 0.25. doi:10.1371/journal.pgen.1004520.g002

cells differentially express 422 (221 up and 201 down) and 227 (139 up and 88 down) genes with at least 1.5-fold change, respectively (Dataset S6 and S7), indicating both cell and non-cell autonomous mechanisms of gene regulation coordinated by Clim; a significant proportion of basal- and luminal-specific DEGs overlap with the DEGs identified in the TEB and duct populations (Figure S3H–I). Basal-specific DEGs participate in mammary gland morphogenesis processes, including cell proliferation, adhesion, and stem cell properties (Figure 3F and S3J). Interestingly, DEGs in the luminal cell population are enriched with keratinization and squamous epithelial, specifically esophageal and epidermal, differentiation genes (Figure 3G and S3K), suggesting a transepidermal differentiation of the luminal cells in the DN-Clim mammary gland. Thus, Clim coordinate gene expression in the

basal cell compartment that ultimately help maintain the identity of luminal cells.

Several genes and pathways related to mammary gland stem cell biology are differentially expressed in the DN-Clim mammary gland throughout the developmental time course and within the basal and luminal cell compartments (Figure 3C–D, and S3J–K). Ontology analyses revealed involvement of Clim-regulated genes in mammary stem cells and the Wnt signaling pathway. Lgr5, a marker of stem cells in CD29<sup>Hi</sup>CD24<sup>+</sup> basal cells [33], is consistently downregulated in basal cells and both the TEB and duct (Dataset S3, S4, and S6). Additional Clim-regulated genes that have known roles in maintenance of the stem cell enriched TEB and branching morphogenesis are the *Fgfr2* gene [34,35] and the luminal cell-specific *ErbB2* and *ErbB3* genes (Figure 3G



**Figure 3. Clims coordinate the expression of developmental gene signatures.** (A–B) Heat map of (A) TEB (66 genes) and (B) duct (137 genes) DEGs in the DN-Clim mammary gland with CyberT p-value<0.01 in at least two time points. (C–D) Categories of enriched gene sets from the Molecular Signatures Database that significantly overlap with the (C) TEB and (D) duct DEG sets. Both gene sets are enriched with mammary stem cell and breast cancer genes. Vertical line represents p-value at 0.05. (E) Analysis of overall survival (OS) in UNC 337 cohort of patients classified into high and low expression groups based on median expression of the combined list of DEGs in the DN-Clim TEB and ducts demonstrates that high expression of Clim-regulated genes confers poor prognosis. HR: Hazard Ratio with 95% confidence interval. P-value derived from Log-rank test. (F) Fold change of Wnt-signaling factors in DN-Clim sorted basal cells suggests altered Wnt signaling. Fgfr2 is also downregulated. (G) Increased expression of epidermal differentiation genes in DN-Clim luminal cells suggests a role for basal cells in maintaining luminal cell fate. Luminal cell signaling factors ErbB2/3 and Fgfr2 are also downregulated. doi:10.1371/journal.pgen.1004520.g003

and S4) [36,37]. Given this data, we hypothesized that Clims promote branching morphogenesis through gene regulation within stem cells of the basal MEC compartment.

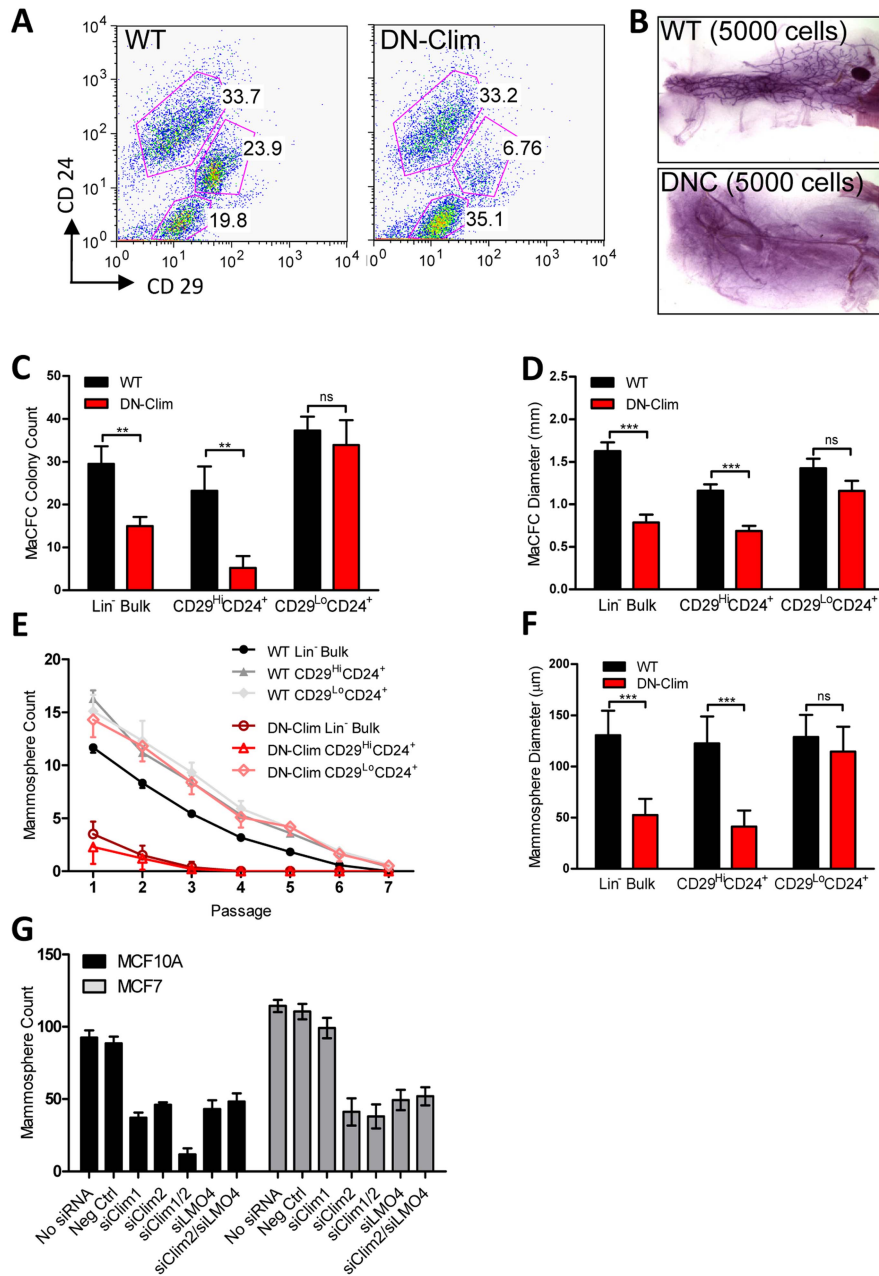
**Clims maintain the basal mammary epithelial stem cell population**

The reduced size of the stem cell-enriched TEB in DN-Clim mice and the enrichment of mammary gland stem cell genes in our DEG sets, along with previous evidence that Climb2 affects stem cells in the hair follicle [22], prompted us to examine the stem cell populations in the DN-Clim mammary gland. Quantification of the basal and luminal MEC populations revealed a reduction of the CD29<sup>Hi</sup>CD24<sup>+</sup> BSC-enriched population in the DN-Clim mammary gland (Figure 4A), while the proportion of the CD29<sup>Lo</sup>CD24<sup>+</sup> LSC-enriched population remains unchanged. No differences were observed in the luminal progenitor cell population when using CD61 as a marker for these cells (Figure S5A) [38]. Limiting dilution transplant analysis with CD29<sup>Hi</sup>CD24<sup>+</sup> cells revealed a nearly complete ablation of DN-Clim mammary repopulating units (MRU) (Table 1, Figure 4B); only two DN-Clim transplants resulted in mammary outgrowths, both of which developed poorly structured mammary trees (Figure S5B). Thus, Clims are of paramount importance in the maintenance of BSCs.

Mammary colony forming cell (MaCFC) and mammosphere functional assays confirmed the depleted basal stem/progenitor cell population. Unsorted Lin<sup>-</sup> bulk and sorted

CD29<sup>Hi</sup>CD24<sup>+</sup>Lin<sup>-</sup> DN-Clim primary epithelial cells have fewer MaCFCs and form smaller colonies, while a slight and insignificant decrease in MaCFCs is observed in the sorted CD29<sup>Lo</sup>CD24<sup>+</sup>Lin<sup>-</sup> population (Figure 4C–D). Consistently, Lin<sup>-</sup> bulk and CD29<sup>Hi</sup>CD24<sup>+</sup>Lin<sup>-</sup> sorted DN-Clim primary mouse MECs grown in suspension produced significantly fewer and smaller mammospheres (Figure 4E–F), while CD29<sup>Lo</sup>CD24<sup>+</sup>Lin<sup>-</sup> cells exhibit comparable sphere forming efficiency to WT cells. Serial passaging of the mammospheres every seven days resulted in enhanced depletion rates of Lin<sup>-</sup> bulk and CD29<sup>Hi</sup>CD24<sup>+</sup>Lin<sup>-</sup> DN-Clim stem/progenitor cells. Collectively, these results suggest Clims maintain the number and proliferative potential of the basal stem/progenitor cell population.

To further confirm the requirement of Clims in promoting stem-like features, Climb1 and Climb2 were knocked down by siRNA, both independently and together, in the MCF10A and MCF7 mammary epithelial cell lines (Figure S5C–D). Climb2 knockdown resulted in significantly reduced mammosphere forming efficiency in both cell lines, while Climb1 knockdown resulted in significant decreases in MCF10A cells, but not MCF7 cells (Figure 4G). Consistently, we observed a synergistic decrease in mammosphere formation efficiency in the combined siClim1/2 in MCF10A cells, but not MCF7 cells. These cell type-specific differences in promoting stem-like properties may be attributed to higher levels of Climb1 in MCF10A cells than MCF7 cells. Altogether these data further support the role of Clims in promoting stem-like features of MECs.



**Figure 4. Clims maintain basal mammary epithelial stem cells.** (A) Flow cytometry analysis of Lin<sup>-</sup> MECs with CD29 and CD24 markers reveals decreased CD29<sup>hi</sup>CD24<sup>+</sup> basal epithelial cell population. (B) DN-Clim mammary glands are absent of MRUs as demonstrated in representative whole mounts from transplants of CD29<sup>hi</sup>CD24<sup>+</sup> BSC-enriched population. See Table 1 for results from limiting dilution analysis. (C) DN-Clim basal cells, and not luminal cells, maintain fewer colony forming units determined by colony-forming cell assays with Lin<sup>-</sup> bulk and sorted primary mammary epithelial cells. (D) Reduced growth rates in colonies that form from DN-Clim basal cells determined by quantification of colony size for the corresponding MaFCF assays described in (C). (E) DN-Clim basal cells, and not luminal cells, maintain fewer mammosphere forming units determined by mammosphere assays with Lin<sup>-</sup> bulk and sorted primary mammary epithelial cells. (F) Reduced growth rates in spheres that form from DN-Clim basal cells determined by quantification of mammosphere diameter for the first passage of the mammosphere assays described in (E). (G) Clm1, Clm2 and the Clm interaction factor LMO4 are essential for maintaining stem-like features of MCF10A cells, as determined by mammosphere assays with transient siRNA knockdown of Clm1, Clm2, and LMO4. In MCF7 cells, only Clm2 and LMO4 maintain stem-like features. Data represent mean ± SEM from at least three mice (C-F) or at least three experiments (G). \*\* p-value<0.01; \*\*\* p-value<0.001, ns: not significant. doi:10.1371/journal.pgen.1004520.g004

Clim2 is known to interact with LMO4 in MECs; LMO4 promotes mammary gland morphogenesis and breast cancer [39–42]. To determine whether Clm2 may be acting through LMO4 to maintain the mammary stem/progenitor cell population we transiently knocked down *LMO4* gene expression individually and

in conjunction with *Clim2* in MCF10A and MCF7 cells (Figure S5E). We found similar reduction in mammosphere forming efficiency with individual and combined knockdowns (Figure 4G), suggesting these two transcription factors may target similar mechanisms to promote stem cell features, consistent with these



**Table 1.** Clims maintain the basal mammary epithelial stem cell population.

CD29 <sup>Hi</sup> CD24 <sup>+</sup> Basal Cells	MaSC Frequency	
	Wild Type	DN-Clim
5000	2/2	1/2
1000	2/2	0/2
500	3/3	0/3
100	2/3	1/3
50	3/5	0/5
MRU Frequency (95% CI)	1/69 (1/28 to 1/172)	1/5566 (1/1288 to 1/24,055)
<i>p</i> -value (WT vs DN-Clim)	<i>p</i> = 3.58e-09	
<i>p</i> -value (single-hit Poisson model)	<i>p</i> = 0.748	<i>p</i> = 0.222

doi:10.1371/journal.pgen.1004520.t001

two factors forming a tight complex [43–45]. Thus, Clim2 may regulate gene expression in these cells through the organization of transcriptional complexes involving LMO4.

### Clim2 regulates *Fgfr2* to maintain BSCs

*Fgfr2* is significantly downregulated in the DN-Clim mammary gland throughout the developmental time course (Figure 5A) as validated by qPCR in the TEB and duct cells from 6 week old mice (Figure 5B). *Fgfr2* expression was also decreased in the sorted basal and luminal cells from DN-Clim mice (Figure 5C). While the decrease in basal cells is likely due to cell autonomous effects of DN-Clim, the decreased *Fgfr2* expression in luminal cells is likely non-cell autonomous as the DN-Clim molecule is not expressed in these cells. Since *Fgfr2* is essential for maintaining the BSC population [46], and DN-Clim is localized to the K14-positive basal MECs, we hypothesized that Clims directly regulate the expression of *Fgfr2* to maintain BSCs. Indeed, the FGFR2 protein is decreased in DN-Clim basal cells (Figure 5D–E). Transient knockdown of *Fgfr2* in the *Fgfr2*-high MCF7 cell line impairs mammosphere forming efficiency (Figure 5F), while stable overexpression of *Fgfr2* in the *Fgfr2*-low MCF10A cell line with transient knockdown of Clims partially rescues mammosphere forming efficiency (Figure 5G). These results suggest Clims regulate *Fgfr2* expression to maintain stem-like features.

To determine if Clims directly bind the *Fgfr2* promoter, we performed ChIP assays for Clim2, DN-Clim (Myc-tag), H3K4me3 (a marker for actively transcribed genes), and the Clim-interacting partner LMO4, followed by qPCR targeting multiple sites in the 2 kb region surrounding the *Fgfr2* transcriptional start site (TSS). Clim2 binds to a region 538 bp to 708 bp upstream of the TSS in the WT mammary gland, with enriched H3K4me3 surrounding the TSS (Figure 6A), while DN-Clim binds the same region upstream of the TSS in DN-Clim mammary glands, with a drastic reduction in H3K4me3 enrichment (Figure 6B). Additionally, LMO4 is enriched at the same region in WT mammary glands, but is absent in the DN-Clim mammary gland, suggesting a necessity for Clim2 to recruit LMO4 to the promoter. To determine if Clim-binding activates transcription, we cloned the promoter-binding region (433 to 1010 bp upstream of the *Fgfr2* TSS) into the pGL3-promoter vector. In Clim2-positive MCF10A cells, this region of DNA enhances transcription of the luciferase reporter, and DN-Clim inhibits its transcription (Figure 6C). Furthermore, *Fgfr2* expression is downregulated in MCF10A and MCF7 cells upon knockdown of Clim1/2 and LMO4 (Figure 6D). Altogether, these data suggest that Clim2 promotes

the maintenance of the BSC population by directly binding the *Fgfr2* promoter to drive transcription of the gene.

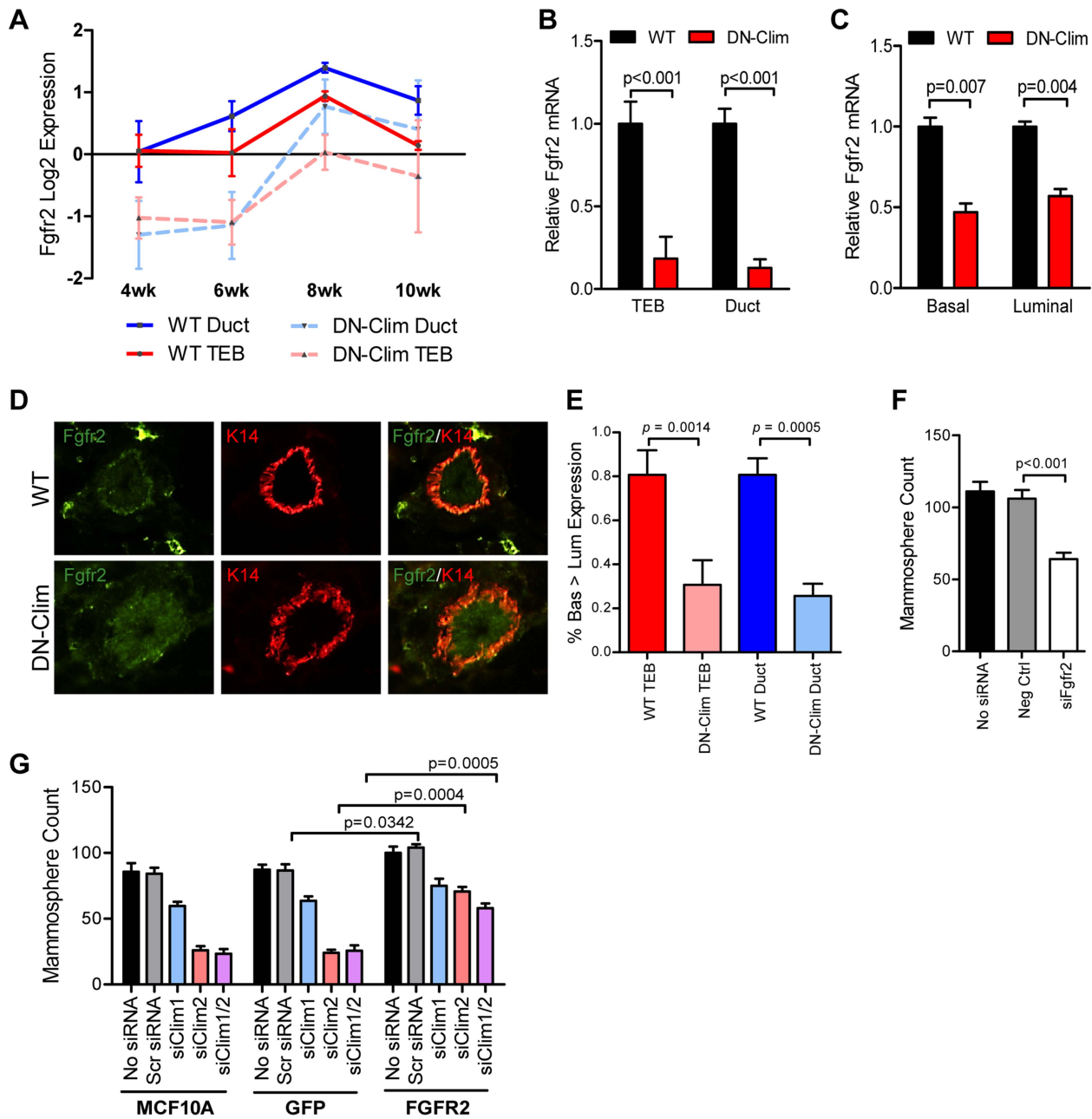
### Clim2 promotes breast tumorigenesis

The CLIM protein is expressed in ER-positive breast tumors [21], and we show here that the Clim-regulated gene network is associated with poor prognosis in breast cancer (Figure 3E). Furthermore, its interacting partner LMO4 is upregulated in breast cancer and promotes breast tumorigenesis [40,42,47,48]. We examined the expression of Clims in the molecular subtypes of breast cancer and observed highest expression of Clim2 in the less differentiated claudin-low and basal-like breast cancer subtypes, while Clim1 is most highly expressed in the differentiated luminal cell types (Figure 7A–B). Accordingly, Clim2 expression alone predicts poor prognosis in breast cancer, while Clim1 expression displays an opposite trend (Figure S6). The combined expression level of both Clim genes is a strong predictor of poor outcome (Figure 7C), suggesting Clim2 is a more powerful predictor of breast cancer outcome than Clim1.

To determine if Clims play a role in breast tumorigenesis, we bred the *K14-DN-Clim* gene into the MMTV-PyMT mouse model. PyMT mice develop palpable tumors in nearly every mammary gland, whereas DN-Clim/PyMT mice only develop tumors in no more than six mammary glands (Figure 7D). Additionally, DN-Clim/PyMT mice exhibit a significant delay in the development of palpable tumors (Figure 7E) and the growth rate of these tumors is drastically impaired (Figure 7F). While tumor development is rapid and frequent in PyMT mice, with consecutive palpable tumors developing within days of each other, there was a significant delay in the time required for the second palpable DN-Clim/PyMT tumor to develop and the growth rate of the second tumor was significantly slower than the first (Figure 7E–F). Taken together, these data demonstrate that Clim promotes the initiation and progression of breast cancer in the PyMT model.

To determine the stage at which Clim may be promoting breast cancer, we collected whole mounts of mammary glands from as early as 3 weeks to as late as 3 months. Hyperplastic lesions appear in the mammary glands of 3-week-old PyMT and DN-Clim/PyMT mice (Figure 7G). However, by 7 weeks entire arms of the mammary tree display hyperplastic lesions in PyMT mice, while the DN-Clim/PyMT mice only exhibit sparse lesions (Figure 7G). Consistently, after removal of the primary tumor from mammary glands of 3 month old mice we observed a decrease in the area of the epithelial tree covered in tumor tissue. The drastic difference in

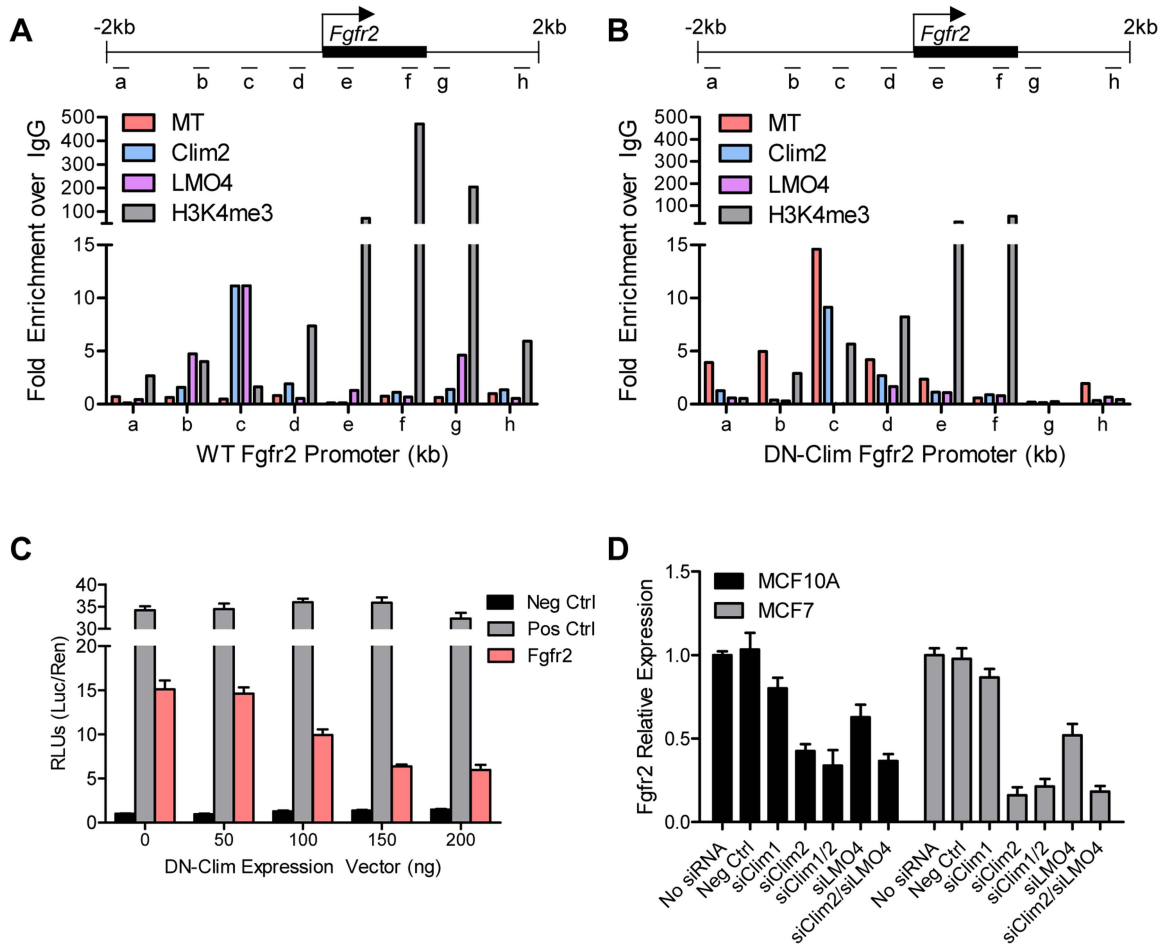




**Figure 5. *Fgfr2* downregulation in the mammary gland leads to loss of stem/progenitor cell activity.** (A) *Fgfr2* is downregulated in TEB and duct cells throughout branching morphogenesis. The data are derived from the time course gene expression microarray analysis. (B) qPCR validation of decreased *Fgfr2* expression in DN-Clim TEB and duct cells. Gene expression is normalized to 18s mRNA. (C) qPCR validation of decreased *Fgfr2* expression in sorted basal and luminal cells. Gene expression is normalized to 18s mRNA. (D) Immunofluorescence analysis of FGFR2 protein expression reveals decreased protein in the basal layer of DN-Clim mammary glands. (E) Semi-quantitative analysis of FGFR2 protein expression in TEB and luminal structures suggests reduced protein levels in basal cells of the DN-Clim mammary gland. WT mammary glands exhibit more protein in basal cells when compared to luminal cells, but DN-Clim mammary glands exhibit equal FGFR2 protein levels in basal and luminal cells. The percentage of TEB and duct structures that exhibit higher FGFR2 protein in basal cells was quantified by counting at least 10 structures from at least 10 frozen sections from three WT and three DN-Clim mammary glands. The data represent the mean  $\pm$  SEM from three biological replicates. (F) Transient siRNA knockdown of *Fgfr2* in MCF7 reduces sphere forming efficiency in mammosphere assays. (G) *Fgfr2* overexpression partially rescues reduced mammosphere forming efficiency upon transient siRNA knockdown of Clims in the MCF10A cell line. Data represent mean  $\pm$  SEM from at least two littermate mice (A–C, E), or three replicate experiments (F–G). doi:10.1371/journal.pgen.1004520.g005

the number of hyperplastic nodules developing from the mammary tree suggest that Clims act during the early stages of tumorigenesis. IHC staining of 6 week old mammary glands demonstrated that these tumors primarily express K8, and not

K14 (Figure 7H), confirming they are luminal tumors [49]. PyMT tumors exhibit a rare population of K14-expressing cells at the leading, basal edge of the tumor, which also express DN-Clim in the DN-Clim/PyMT tumors (Figure 7H). These data demonstrate



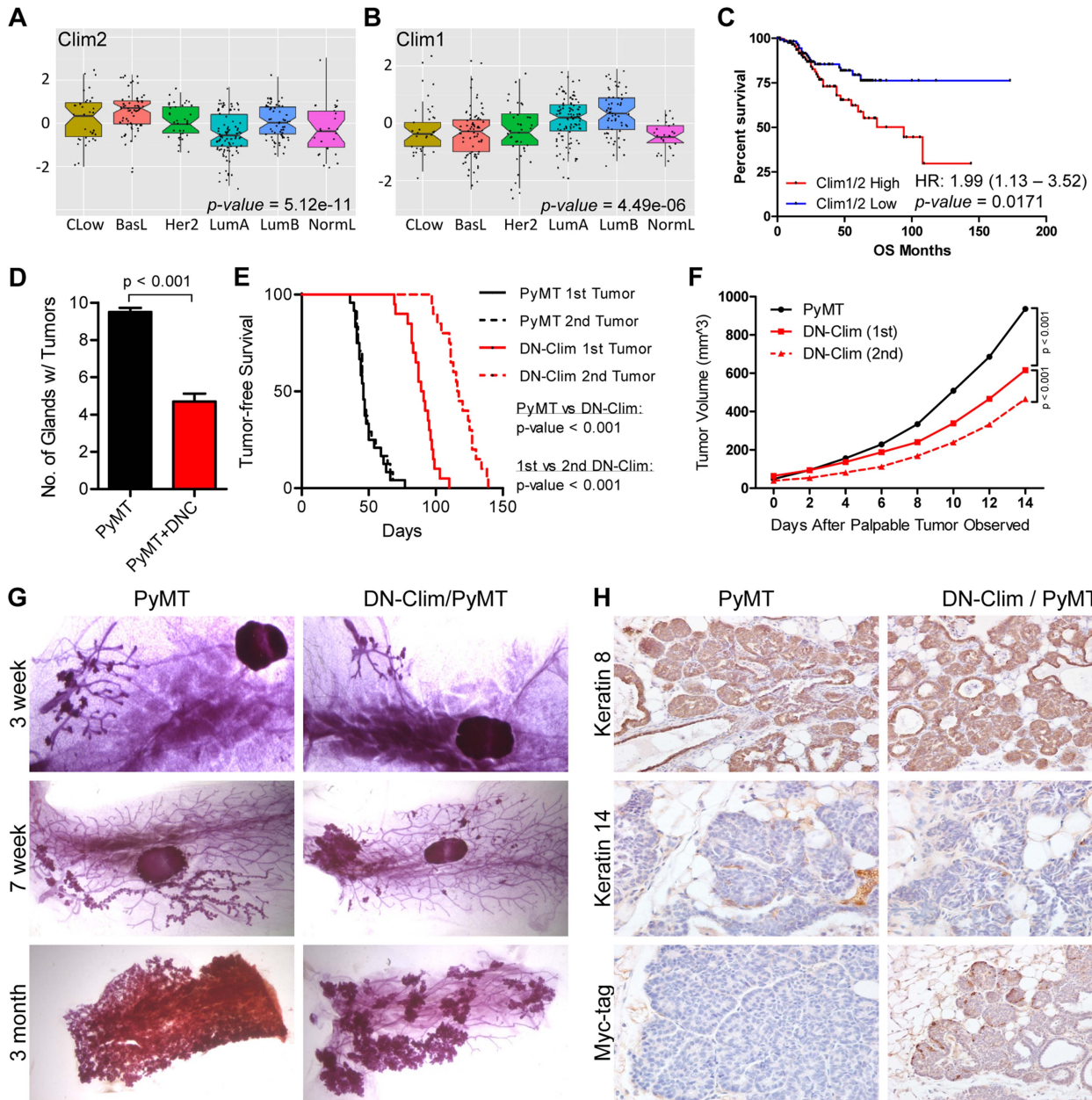
**Figure 6. Climbs and LMO4 directly target and the *Fgfr2* promoter to induce gene expression.** (A–B) ChIP-qPCR assays of Myc-tagged DN-Clim, Clim2, LMO4, and H3K4me3 enrichment at the indicated locations of the  $\pm 2$  kb surrounding the *Fgfr2* promoter (top panel) in (A) WT and (B) DN-Clim MECs. MECs were collected from three littermate mice then pooled for ChIP-qPCR experiments. The data suggest Clim2 binds approximately 0.5–1.0 kb upstream of the *Fgfr2* TSS. When DN-Clim is bound to this region, LMO4 binding is lost and H3K4me3 levels are decreased. (C) Luciferase assays from MCF10A cells transiently expressing the Clim-binding promoter region of *Fgfr2* in the pGL3-promoter luciferase reporter vector indicate Clim acts as an inducer of *Fgfr2* expression. The DN-Clim expression vector was transiently titrated to demonstrate reduction of luciferase activity with increasing DN-Clim expression. (D) Expression of *Fgfr2* in the MCF10A and MCF7 cells is reduced upon transient siRNA knockdown of Clim1, Clim2, and LMO4. Expression is normalized to GAPDH. Data represent mean  $\pm$  SEM from three replicate experiments (C–D). doi:10.1371/journal.pgen.1004520.g006

the necessity of Clim-mediated, basal cell-specific gene expression for robust PyMT-mediated tumorigenesis, and suggest a functional role for Clim-dependent mechanisms in basal MECs during the initiation and progression of luminal breast tumors.

## Discussion

Transcriptional networks regulating the development and maintenance of the mammary gland luminal cell compartment are well characterized, with GATA3, Notch1, Elf5, and STAT5a being key factors [38,50–52]. However, transcriptional regulation within the basal cell compartment remains poorly understood. Here, we describe a Clim-driven developmental transcriptional network in the basal cell compartment that maintains the number and proliferative potential of BSCs. Interfering with this transcriptional network leads to defective branching morphogenesis and impaired PyMT-mediated tumorigenesis. In addition, Clim-dependent basal cell-derived transcriptional programs support the expression of key luminal growth factor receptors and maintain the identity of luminal cells (Figure 8).

The widely expressed Climbs regulate unique transcriptional programs in diverse cell types. In part, this likely depends on Clim's tissue specific interactions with LIM domains of distinct LHX and LMO family members, such as LMO2 and LHX2 in erythroid cells [12,53], LHX3 in the neural tube [54], LHX2 in hair follicles [22,55,56], and LMO4 in the mammary gland [40]. Despite this diversity in interactions, Climbs seem to play a general role in promoting stem cell maintenance as previously shown in intestinal and hair follicle epithelia [22,57], blood [58], and ES cells [59]. How Climbs maintain stem cell features is largely unexplained. We address this question in the mouse mammary gland with the K14-DN-Clim mouse model that targets both Clim family members in the basal cell population. DN-Clim, effectively disrupts Clim/LIM domain-mediated transcriptional complexes *in vivo*, as has been demonstrated in *Drosophila*, Zebrafish, and mouse [22,60]. Additionally, knockout of *Lhx2*, the Clim-interacting partner in hair follicle stem cells, gives a similar phenotype as expression of DN-Clim [22,55]. Lacking the dimerization domain, DN-Clim binds LIM domain proteins to inhibit their interactions with endogenous Climbs, thereby

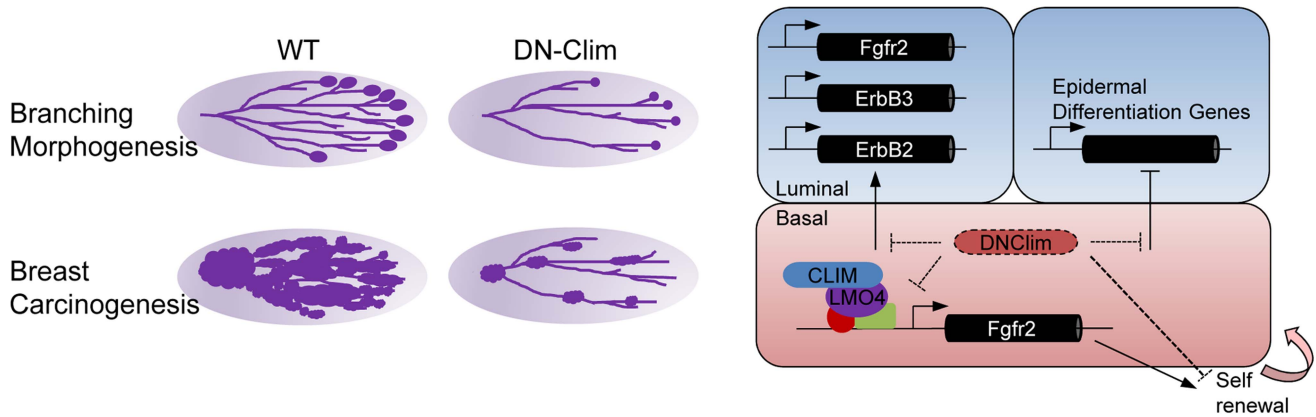


**Figure 7. Clims promote the initiation and progression of tumors in the MMTV-PyMT breast cancer mouse model.** (A–B) Expression of (A) Clim2 and (B) Clim1 in the molecular subtypes of breast cancer. Clim2 is more highly expressed in less differentiated Claudin-low (CLow) and basal-like (BasL) tumor subtypes and Clim1 is more highly expressed in the differentiated Luminal A and B (LumA and LumB) subtypes. (C) Survival analysis within the UNC 337 cohort of breast cancer patients grouped by median expression of Clim1 and Clim2 demonstrates worse prognosis in patients with highest expression of Clim1/2. (D) DN-Clim prevents tumorigenesis in the PyMT tumor model as determined by quantification of the number of mammary glands with palpable tumors. Data represent mean  $\pm$  SEM from ten mice of each genotype. (E) DN-Clim expression in the PyMT tumor model significantly delays development of palpable tumors. In PyMT mice, the second palpable tumor is detected within a few days of the first tumor, but when DN-Clim is expressed there is a significant increase in the time to detect the second palpable tumor. P-values derived from Log-rank test. (F) Decreased rate of tumor progression in DN-Clim/PyMT mice. P-values derived from repeated measures ANOVA. (G) Representative whole mounts of PyMT tumors with or without DN-Clim suggest a function for Clim in the tumor initiation stages. The primary tumor was excised from three month old mice before whole mount of the mammary gland. (H) K8, K14, and Myc-tag (DN-Clim) expression in mammary glands from 6 week old PyMT mice with or without DN-Clim. High K8 and low K14 expression demonstrate the luminal subtype features of PyMT tumors. DN-Clim expression is observed in basal cells of the tumor.  
doi:10.1371/journal.pgen.1004520.g007

preventing Clim-mediated looping involved with higher order transcriptional complexes. We still observed DN-Clim at the promoter of the *Fgf2* gene, indicating that these interactions still allow the binding of LIM domain transcription factors to their target promoters. The N-terminal dimerization domain may have

other functions not inhibited by DN-Clim, including possible interactions with other non-LIM domain transcriptional regulators [15,16] and interactions with the RLIM ubiquitin ligase that leads to degradation of associated LIM domain proteins [61]. DN-Clim has been demonstrated to stabilize nuclear LIM domain proteins





**Figure 8. Regulatory roles for Clims in normal development and cancer.** Clims promote branching morphogenesis and carcinogenesis in the mammary gland (left). Direct regulation of *Fgfr2* expression by Clims and LMO4 in the basal cell compartment facilitate basal stem cell self-renewal (right). Clims mediate a transcriptional program within the basal cell compartment to promote expression of *Fgfr2*, *ErbB2*, and *ErbB3* in the luminal cell compartment, in addition to maintaining luminal epithelial cell fate by restricting the expression of epidermal differentiation genes. doi:10.1371/journal.pgen.1004520.g008

[62], which may result in increased levels of these proteins in the basal cells of the mammary gland.

In the mammary gland, several lines of evidence argue that Clims maintain basal stem cells by stimulating expression of *Fgfr2*. Expression of a DN-Clim molecule in the mouse mammary gland and knockdown of Clims in human MECs lead to downregulation of *Fgfr2*. Also, *Fgfr2* overexpression can rescue the impaired mammosphere formation after Clim knockdown in MECs. Furthermore, we demonstrate that Clims associate with and regulate the *Fgfr2* promoter in MECs. This model is consistent with work showing that FGF signaling is essential to branching morphogenesis in several epithelial organs, including the drosophila trachea and the mammalian mammary gland, lung, kidney, and salivary gland [63]. FGF signaling promotes mammary embryonic placode formation [64], and *Fgfr2* is necessary to maintain the TEB structure and promote primary branching [34,35]. In addition, *Fgfr2* is required for maintenance of the CD29<sup>Hi</sup>CD24<sup>+</sup> basal cell population and accordingly the maintenance of regenerative BSCs within this population [46]. The defects in branching morphogenesis, basal epithelial cell frequency, and MRU capacity observed in these *Fgfr2* transgenic mouse models closely resemble the DN-Clim mammary gland phenotype. Thus, Clim acts immediately upstream of *Fgfr2* signaling in the basal cell population through direct regulation of the *Fgfr2* gene to maintain BSCs and promote branching morphogenesis.

Our data suggest that Clim associates with its known mammary gland binding partner LMO4 to regulate the *Fgfr2* gene. We find that LMO4 binds to the same region of the *Fgfr2* promoter, and knockdown of LMO4 also leads to decreased *Fgfr2* gene expression. Furthermore, we find that LMO4 knockdown leads to decreased mammosphere formation in MECs. That these proteins function as a complex is supported by experiments showing that the combined Clim/LMO4 knockdown does not lead to further decrease in mammosphere formation over individual knockdowns. Furthermore, functional studies of LMO4 in the mammary gland demonstrate its role in promoting branching morphogenesis and lobuloalveolar development by sustaining cell proliferation [39,41]. Altogether, these observations suggest Clims act through LMO4 in the mammary gland to promote BSC maintenance and development.

In addition to regulating the number and replicative potential of BSCs, we find evidence that Clims regulate a basal cell program that maintains luminal cell fate. Since one of the models of stem cell hierarchy in the mammary gland proposes that BSCs give rise to the luminal cells [4], this program may be established cell autonomously in the basal progenitors and persists in the luminal cells. Alternatively, the Clim-controlled program may act through basal-to-luminal cell signaling; there is previous evidence for crosstalk between the basal and luminal compartments through paracrine and cell adhesion-mediated mechanisms [1,2]. Irrespective of the underlying mechanisms, since DN-Clim is selectively expressed in the basal cell compartment, our gene expression data in basal and luminal cell compartments clearly reveals an effect of basal-specific Clim function on gene expression in luminal cells. In particular, DN-Clim mice exhibit downregulation of key growth factor receptors *ErbB2* and *ErbB3*, and upregulation of several epidermal differentiation genes. These observations demonstrate an essential role for Clims in the basal cell compartment to coordinate gene expression programs in the luminal cell compartment, and for the first time demonstrate the importance of basal cell factors on regulating the expression of key luminal regulators in the EGFR family. These mechanisms may contribute to the abnormal mammary development and decreased breast tumorigenesis in the DN-Clim/PyMT mice.

The role of Clims is not limited to mammary gland development as our work also suggests they act within basal cells during the early stages of tumorigenesis. Expression of DN-Clim within the basal cell compartment significantly impairs the ability of PyMT to initiate neoplastic lesions. In the murine MMTV-PyMT tumor model *Fgfr2* is overexpressed in the CD29<sup>Hi</sup>CD24<sup>+</sup> tumor initiating cell population and is necessary to initiate tumorigenesis in these mice [65]. As Clims are direct regulators of *Fgfr2* in the mammary gland, the reduced tumorigenicity observed in DN-Clim/PyMT mice could in part be explained by Clim-mediated regulation of *Fgfr2* in cancer stem-like cells. The MMTV promoter directs PyMT expression to both luminal and basal cells [66], and while the histology and gene expression patterns group PyMT breast tumors with the luminal subtype of breast cancer, their cell of origin remains unknown [67]. Clims could be promoting early tumorigenesis through one of two scenarios: 1) Clims specifically maintain a basal cell-derived tumor



initiating cell, or 2) Clims coordinate the communication of basal-derived tumorigenic signals to the luminal cell compartment, such as promoting the expression of Her2 in luminal cells. These signals would have to occur during the early stages of tumorigenesis when K14-positive basal cells are still present in premalignant lesions [68]. While basal cells have previously been assumed to act as inhibitors of breast cancer progression [69], the restricted tumorigenicity observed in the DN-Clim/PyMT mammary gland suggests basal cells may permit tumorigenesis.

Corresponding to these functional mouse experiments, we observed differential expression of Clims in the molecular subtypes of breast cancer, with high expression leading to poor outcome. It is well known that Clims are expressed in luminal cells and coordinate transcriptional programs through ER $\alpha$  in human breast cancer [21]. However, our K14-DN-Clim mouse model suggests an additional function for Clims during breast tumorigenesis in the ER-negative basal cell population, either through direct regulation of the basal cell or through indirect effects on the luminal cell. Furthermore, we find that a Clim-regulated gene module correlates with poor breast cancer prognosis arguing, for relevance of the mouse experiments for human breast cancer. It is possible that Clims may in part promote breast cancer through activation of *Fgfr2* expression. Deregulation of FGF signaling is observed in breast cancer [70], with *FGFR2* amplification and overexpression observed in 5–10% of human breast cancers, correlating with poor prognosis [71,72]. In addition, single point mutations in the *FGFR2* gene are associated with increased risk in human breast cancer [73].

In conclusion, we have demonstrated that Clims are a necessary factor in mammary gland branching morphogenesis, maintaining the BSC compartment. Clim2 coordinates with LMO4 to govern these processes through direct regulation of the *Fgfr2* gene. In a breast cancer mouse model Clims act within the basal cells to permit tumorigenesis, possibly through the maintenance of cancer stem-like cells. As Clims are expressed in human breast cancer and correlate with poor differentiation of ER-positive tumors, elucidating Clim targets at a global scale may give insight into the transcriptional mechanisms that maintain primitive cancer cells. The developmental transcriptome of the TEB and duct populations reported will provide a valuable resource for delineating the molecular relationships between development and tumorigenesis, in addition to identifying novel prognostic and therapeutic targets.

## Materials and Methods

### Mouse strains

Generation and maintenance of K14-DN-Clim mice were as previously described [22]. MMTV-PyMT mice (Jackson Labs) were maintained on the K14-DN-Clim mixed background. Fox Chase SCID Beige mice (Charles River) were used as recipients for transplant experiments. All experiments conform to the regulatory guidelines approved by the International Animal Care and Use Committee of the University of California, Irvine.

### Tissue and cell preparation

Mammary glands 4 and 9 were dissected and either whole mounted, paraffin embedded, or frozen in O.C.T. For proliferation assays, mice were injected with BrdU (50  $\mu$ g/g, Sigma-Aldrich) 4 hours prior to sacrificing. Adult (8 to 12 week) mammary glands for single-cell suspensions were generated according to Stem Cell Technologies protocol. When collecting RNA, the collagenase/hyaluronidase digestion was reduced to 1.5 hours.

### Flow cytometry, sorting, and in vitro culture

Single-cell suspensions were incubated with Propidium Iodide (2  $\mu$ g/mL, Sigma-Aldrich), CD31-APC, CD45-APC, TER119-APC, CD24-PE (BD Biosciences) and analyzed by flow cytometry on FACSCalibur (BD Biosciences) or sorted on FACSARIAII (BD Biosciences). Bulk primary MECs for in vitro culture were incubated with biotinylated CD31/CD45/TER119 cocktail (Stem Cell Technologies) and magnetically separated to remove lineage cells.

### Cleared fat pad transplantation

FACS-sorted CD29<sup>Hi</sup>CD24<sup>+</sup> cells were transplanted into cleared fat pads of 3 week old immunocompromised SCID Beige recipients. Mammary glands were evaluated 8 weeks after transplantation. Limiting dilution analysis was performed on the ELDA Web-based tool (<http://bioinf.wehi.edu.au/software/elda/>).

### Microarray and bioinformatic analysis

RNA was collected from laser capture microdissected (Leica LS-AMD) or FACS sorted MECs with the RNeasy Mini Kit (Qiagen). Gene expression was measured with Affymetrix Mouse Gene 1.0ST array. Data were analyzed with PLIER algorithm. Differential expression was determined with the CyberT algorithm [31] and time course developmental gene expression analysis was performed with the BETR algorithm [24]. Detailed methods of data analysis are described in the supplemental methods accompanying this manuscript.

### Accession numbers

The gene expression microarray data has been submitted to GEO (<http://www.ncbi.nlm.nih.gov/geo/>) under the accession number GSE57724.

Additional details for the above methods are presented in Text S1, as well as methods for IHC, immunofluorescence, whole mount, ChIP, RT-qPCR, mammosphere and colony forming assays, siRNA transfection, lentiviral expression, luciferase assays and microarray analysis.

## Supporting Information

**Dataset S1** List of TEB signature genes. BETR p-values and fold change at each time point is presented for each TEB signature gene. Signature genes have BETR p-value<0.001 and  $\pm 1.5$  fold change. (XLSX)

**Dataset S2** List of Duct signature genes. BETR p-values and fold change at each time point is presented for each Duct signature gene. Signature genes have BETR p-value<0.001 and  $\pm 1.5$  fold change. (XLSX)

**Dataset S3** List of DEGs in the DN-Clim TEB. CyberT p-values and fold change at each time point for the TEB are presented. DEGs have p-value<0.01 and  $\pm 1.5$  fold change in at least two time points. NA: gene did not meet expression criteria in the specified time point. (XLSX)

**Dataset S4** List of DEGs in the DN-Clim Duct. CyberT p-values and fold change at each time point for the Duct are presented. DEGs have p-value<0.01 and  $\pm 1.5$  fold change in at least two time points. NA: gene did not meet expression criteria in the specified time point. (XLSX)

**Dataset S5** List of DEGs in the WT Basal vs Luminal cell comparison. CyberT p-values and fold change are presented. DEGs have p-value<0.001 and fold change  $\pm 1.5$ . (XLSX)

**Dataset S6** List of DEGs in the DN-Clim basal cell population. CyberT p-values and fold change are presented. DEGs have p-value<0.001 and fold change  $\pm 1.5$ . (XLSX)

**Dataset S7** List of DEGs in the DN-Clim luminal cell population. CyberT p-values and fold change are presented. DEGs have p-value<0.001 and fold change  $\pm 1.5$ . (XLSX)

**Figure S1** Specificity of Clim2 antibody and DN-Clim females fail to support full litters. (A) The Clim2 antibody specifically targets the Clim2 protein with no reactivity to Clim1, as determined by western blot on protein lysates from HEK293 cells overexpressing the Clim1 and Clim2 proteins. The Clim1/2 antibody detects Clim1 and Clim2 only in their respective overexpression lysates. Vector Ctrl Lysate = Vector transfected lysate control. (B) Average number of pups per litter from WT and DN-Clim females. DN-Clim mice are unable to support the full litter after postnatal day 2. (C) Growth rate of pups from WT and DN-Clim females. Surviving pups from DN-Clim females grow at a normal rate compared to pups from the WT mother. (PDF)

**Figure S2** Time course analysis of Clim expression and comparison of Clim-regulated genes to TEB and duct genes. (A) Expression of Clim1 and Clim2 from time course analysis of TEB and duct cells. (B) Significant overlap of differentially expressed genes from the DN-Clim TEB and duct. (C–D) DEGs in the DN-Clim (C) TEB and (D) duct are significantly enriched in their respective developmental gene set. (PDF)

**Figure S3** Gene expression profiling in sorted basal and luminal mammary epithelial cells. (A) Selection of live (PI-negative), Lin<sup>-</sup> (TER119-, CD45-, and CD31-negative) single cells. (B) Gating for basal (Lin<sup>-</sup>CD29<sup>Hi</sup>CD24<sup>+</sup>) and luminal (Lin<sup>-</sup>CD29<sup>Lo</sup>CD24<sup>+</sup>) MECs. (C) Post-sort analysis of basal MECs. (D) Post-sort analysis of luminal MECs. APC: Lin markers, PE: CD24, FITC: CD29. (E–F) qPCR validation of (E) Krt14 and (F) Krt8 in sorted cells indicates pure basal and luminal cell populations. (G) qPCR validation of DN-Clim transgene expression confirms expression of DN-Clim in basal cells. (H) DN-Clim basal and (I) DN-Clim luminal DEGs are significantly enriched in the combined list of DN-Clim TEB and Duct DEGs. Ontology analysis of (J) DN-Clim basal DEGs and (K) DN-Clim luminal DEGs. The categories represent top hits from DAVID and the Molecular Signatures Database. (PDF)

## References

- Moumen M, Chiche A, Cagnet S, Petit V, Raymond K, et al. (2011) The mammary myoepithelial cell. *Int J Dev Biol* 55: 763–771.
- Forster N, Saladi SV, van Bragt M, Sfondouris ME, Jones FE, et al. (2014) Basal Cell Signaling by p63 Controls Luminal Progenitor Function and Lactation via NRG1. *Dev Cell* 28: 147–160.
- Van Keymeulen A, Rocha AS, Ousset M, Beck B, Bouvencourt G, et al. (2011) Distinct stem cells contribute to mammary gland development and maintenance. *Nature* 479: 189–193.
- Rios AC, Fu NY, Lindeman GJ, Visvader JE (2014) In situ identification of bipotent stem cells in the mammary gland. *Nature* 506: 322–327.
- Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, et al. (2006) Generation of a functional mammary gland from a single stem cell. *Nature* 439: 84–88.
- Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, et al. (2006) Purification and unique properties of mammary epithelial stem cells. *Nature* 439: 993–997.

**Figure S4** Reduced expression of ErbB2 and ErbB3 receptor tyrosine kinases in the DN-Clim mammary gland. Expression of the (A) ErbB2 and (B) ErbB3 in the time course microarray (left panel), as determined by qPCR in 6 week old laser capture microdissected TEB and duct cells (middle panel), or in 8 week old sorted basal (Bas) and luminal (Lum) cells (right panel). Each are significantly downregulated in the TEB and duct cells. Their expression is restricted to the luminal cell compartment, and their downregulation in DN-Clim luminal cells suggests non-autonomous regulation of these genes by Clim through the basal cell population. Data represent mean  $\pm$  SEM from at least two littermate mice. \* p-value<0.05, \*\* p-value<0.01, \*\*\* p-value<0.001, ns: not significant. (PDF)

**Figure S5** Luminal progenitor cell analysis, representative whole mounts from DN-Clim transplants and validation of gene knockdown by siRNA. (A) CD61 was used as a marker for luminal progenitor cells in the Lin-CD29<sup>Lo</sup>CD24<sup>+</sup> population. No differences were observed in the quantity of these cells in the DN-Clim mammary gland. (B) Whole mounts of the two successful mammary transplants of DN-Clim CD29<sup>Hi</sup>CD24<sup>+</sup> cells. Both mammary glands exhibit defects in ductal penetration and branching morphogenesis. Inset from the fat pad transplanted with 100 DN-Clim cells shows the epithelial outgrowth indicated by the arrow. (C–E) Expression of Clim1 (C), Clim2 (D), and LMO4 (E) validates specific transient knockdown of mRNA for each respective gene. (PDF)

**Figure S6** Contribution of Clim expression to prognosis prediction. Survival analysis based on expression of (A) Clim1 or (B) Clim2. Patients were divided into high and low expressing groups based on median expression of each gene. P-values derived from the Log-rank test. (PDF)

**Text S1** Supplemental materials and methods. (DOCX)

## Acknowledgments

We thank the UCI Genomics High Throughput Facility (GHTF) and the UCI Stem Cell Core Facility for expert service.

## Author Contributions

Conceived and designed the experiments: MLS BA. Performed the experiments: MLS ZY KW EC PSu. Analyzed the data: MLS ZY KW. Contributed reagents/materials/analysis tools: MLS BA XD PSm. Contributed to the writing of the manuscript: MLS BA.

12. Visvader JE, Mao X, Fujiwara Y, Hahm K, Orkin SH (1997) The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation. *Proc Natl Acad Sci U S A* 94: 13707–13712.
13. Jurata LW, Gill GN (1997) Functional analysis of the nuclear LIM domain interactor NLI. *Mol Cell Biol* 17: 5688–5698.
14. Matthews JM, Visvader JE (2003) LIM-domain-binding protein 1: a multifunctional cofactor that interacts with diverse proteins. *EMBO Rep* 4: 1132–1137.
15. Dawid IB, Breen JJ, Toyama R (1998) LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends Genet* 14: 156–162.
16. Bach I (2000) The LIM domain: regulation by association. *Mech Dev* 91: 5–17.
17. Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, et al. (2010) The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* 24: 277–289.
18. Deng W, Lee J, Wang H, Miller J, Reik A, et al. (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149: 1233–1244.
19. Song SH, Kim A, Ragoczy T, Bender MA, Groudine M, et al. (2010) Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. *Blood* 116: 2356–2364.
20. Mukhopadhyay M, Teufel A, Yamashita T, Agulnick AD, Chen L, et al. (2003) Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation. *Development* 130: 495–505.
21. Johnsen SA, Gungör C, Prenzler T, Riethdorf S, Riethdorf L, et al. (2009) Regulation of estrogen-dependent transcription by the LIM cofactors CLIM and RLIM in breast cancer. *Cancer Res* 69: 128–136.
22. Xu X, Mannik J, Kudryavtseva E, Lin KK, Flanagan LA, et al. (2007) Cofactors of LIM domains (Clims/Ldb/Nli) regulate corneal homeostasis and maintenance of hair follicle stem cells. *Dev Biol* 312: 484–500.
23. Sun P, Yuan Y, Li A, Li B, Dai X (2010) Cytokeratin expression during mouse embryonic and early postnatal mammary gland development. *Histochem Cell Biol* 133: 213–221.
24. Aryee MJ, Gutiérrez-Pabello JA, Kramnik I, Maiti T, Quackenbush J (2009) An improved empirical bayes approach to estimating differential gene expression in microarray time-course data: BETR (Bayesian Estimation of Temporal Regulation). *BMC Bioinformatics* 10: 409.
25. Bai L, Rohrschneider LR (2010) s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev* 24: 1882–1892.
26. Prat A, Perou CM (2011) Deconstructing the molecular portraits of breast cancer. *Mol Oncol* 5: 5–23.
27. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 27: 1160–1167.
28. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, et al. (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347: 1999–2009.
29. Guedj M, Marisa L, de Reynies A, Orsetti B, Schiappa R, et al. (2012) A refined molecular taxonomy of breast cancer. *Oncogene* 31: 1196–1206.
30. Mosley JD, Keri RA (2008) Cell cycle correlated genes dictate the prognostic power of breast cancer gene lists. *BMC Med Genomics* 1: 11.
31. Baldi P, Long AD (2001) A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. *Bioinformatics* 17: 509–519.
32. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545–15550.
33. Plaks V, Brenot A, Linnemann JR, Wong KC, Werb Z (2013) Lgr5-Expressing Cells Are Sufficient and Necessary for Postnatal Mammary Gland Organogenesis. *Cell Reports* 3: 70–78.
34. Lu P, Ewald AJ, Martin GR, Werb Z (2008) Genetic mosaic analysis reveals FGF receptor 2 function in terminal end buds during mammary gland branching morphogenesis. *Dev Biol* 321: 77–87.
35. Parsa S, Ramasamy SK, De Langhe S, Gupta VV, Haigh JJ, et al. (2008) Terminal end bud maintenance in mammary gland is dependent upon FGFR2b signaling. *Dev Biol* 317: 121–131.
36. Jackson-Fisher AJ, Bellinger G, Breindel JL, Tavassoli FA, Booth CJ, et al. (2008) ErbB3 is required for ductal morphogenesis in the mouse mammary gland. *Breast Cancer Res* 10: R96.
37. Jackson-Fisher AJ, Bellinger G, Ramabhadran R, Morris JK, Lee KF, et al. (2004) ErbB2 is required for ductal morphogenesis of the mammary gland. *Proc Natl Acad Sci U S A* 101: 17138–17143.
38. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, et al. (2007) Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol* 9: 201–209.
39. Wang N, Kudryavtseva E, Ch'en IL, McCormick J, Sugihara TM, et al. (2004) Expression of an engrailed-LMO4 fusion protein in mammary epithelial cells inhibits mammary gland development in mice. *Oncogene* 23: 1507–1513.
40. Wang N, Lin KK, Lu Z, Lam KS, Newton R, et al. (2007) The LIM-only factor LMO4 regulates expression of the BMP7 gene through an HDAC2-dependent mechanism, and controls cell proliferation and apoptosis of mammary epithelial cells. *Oncogene* 26: 6431–6441.
41. Sum EY, Shackleton M, Hahm K, Thomas RM, O'Reilly LA, et al. (2005) Loss of the LIM domain protein Lmo4 in the mammary gland during pregnancy impedes lobuloalveolar development. *Oncogene* 24: 4820–4828.
42. Sum EY, Segara D, Duscio B, Bath ML, Field AS, et al. (2005) Overexpression of LMO4 induces mammary hyperplasia, promotes cell invasion, and is a predictor of poor outcome in breast cancer. *Proc Natl Acad Sci U S A* 102: 7659–7664.
43. Sugihara TM, Bach I, Kiousi C, Rosenfeld MG, Andersen B (1998) Mouse deformed epidermal autoregulatory factor 1 recruits a LIM domain factor, LMO-4, and CLIM coregulators. *Proc Natl Acad Sci U S A* 95: 15418–15423.
44. Deane JE, Ryan DP, Sunde M, Maher MJ, Guss JM, et al. (2004) Tandem LIM domains provide synergistic binding in the LMO4:Ldb1 complex. *Embo J* 23: 3589–3598.
45. Kenny DA, Jurata LW, Saga Y, Gill GN (1998) Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis. *Proc Natl Acad Sci U S A* 95: 11257–11262.
46. Pond AC, Bin X, Batts T, Roarty K, Hilsenbeck S, et al. (2013) Fibroblast growth factor receptor signaling is essential for normal mammary gland development and stem cell function. *Stem Cells* 31: 178–189.
47. Montañez-Wiscovich ME, Seachrist DD, Landis MD, Visvader J, Andersen B, et al. (2009) LMO4 is an essential mediator of ErbB2/HER2/Neu-induced breast cancer cell cycle progression. *Oncogene* 28: 3608–3618.
48. Visvader JE, Venter D, Hahm K, Santamaria M, Sum EY, et al. (2001) The LIM domain gene LMO4 inhibits differentiation of mammary epithelial cells in vitro and is overexpressed in breast cancer. *Proc Natl Acad Sci U S A* 98: 14452–14457.
49. Smith BA, Shelton DN, Kieffer C, Milash B, Usary J, et al. (2012) Targeting the PyMT Oncogene to Diverse Mammary Cell Populations Enhances Tumor Heterogeneity and Generates Rare Breast Cancer Subtypes. *Genes Cancer* 3: 550–563.
50. Bouras T, Pal B, Vaillant F, Harburg G, Asselin-Labat ML, et al. (2008) Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. *Cell Stem Cell* 3: 429–441.
51. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardner-Garden M, et al. (2008) The Ets transcription factor Ets1 specifies mammary alveolar cell fate. *Genes Dev* 22: 581–586.
52. Yamaji D, Na R, Feuermann Y, Pechhold S, Chen W, et al. (2009) Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. *Genes Dev* 23: 2382–2387.
53. Kitajima K, Kawaguchi M, Iacovino M, Kyba M, Hara T (2013) Molecular functions of the LIM-homeobox transcription factor Lhx2 in hematopoietic progenitor cells derived from mouse embryonic stem cells. *Stem Cells* 31: 2680–2689.
54. Ostendorff HP, Tursun B, Cornils K, Schlüter A, Drung A, et al. (2006) Dynamic expression of LIM cofactors in the developing mouse neural tube. *Dev Dyn* 235: 786–791.
55. Folgueras AR, Guo X, Pasolli HA, Stokes N, Polak L, et al. (2013) Architectural niche organization by LHX2 is linked to hair follicle stem cell function. *Cell Stem Cell* 13: 314–327.
56. Mardaryev AN, Meier N, Poterlowicz K, Sharov AA, Sharova TY, et al. (2011) Lhx2 differentially regulates Sox9, Tcf4 and Lgr5 in hair follicle stem cells to promote epidermal regeneration after injury. *Development* 138: 4843–4852.
57. Dey-Guha I, Mukhopadhyay M, Phillips M, Westphal H (2009) Role of ldb1 in adult intestinal homeostasis. *Int J Biol Sci* 5: 686–694.
58. Li L, Jothi R, Cui K, Lee JY, Cohen T, et al. (2011) Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat Immunol* 12: 129–136.
59. Hwang M, Gorivodsky M, Kim M, Westphal H, Geum D (2008) The neuronal differentiation potential of Ldb1-null mutant embryonic stem cells is dependent on extrinsic influences. *Stem Cells* 26: 1490–1495.
60. Bach I, Rodriguez-Esteban C, Carrière C, Bhushan A, Kronen A, et al. (1999) RLIM inhibits functional activity of LIM homeodomain transcription factors via recruitment of the histone deacetylase complex. *Nat Genet* 22: 394–399.
61. Ostendorff HP, Peirano RI, Peters MA, Schlüter A, Bossenz M, et al. (2002) Ubiquitination-dependent cofactor exchange on LIM homeodomain transcription factors. *Nature* 416: 99–103.
62. Becker T, Ostendorff HP, Bossenz M, Schlüter A, Becker CG, et al. (2002) Multiple functions of LIM domain-binding CLIM/NLI/Ldb cofactors during zebrafish development. *Mech Dev* 117: 75–85.
63. Lu P, Sternlicht MD, Werb Z (2006) Comparative mechanisms of branching morphogenesis in diverse systems. *J Mammary Gland Biol Neoplasia* 11: 213–228.
64. Mailloux AA, Spencer-Dene B, Dillon C, Ndiaye D, Savona-Baron C, et al. (2002) Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo. *Development* 129: 53–60.
65. Kim S, Dubrovskaya A, Salamone RJ, Walker JR, Grandinetti KB, et al. (2013) FGFR2 promotes breast tumorigenicity through maintenance of breast tumorigenic cells. *PLoS One* 8: e51671.
66. Wagner KU, McAllister K, Ward T, Davis B, Wiseman R, et al. (2001) Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. *Transgenic Res* 10: 545–553.
67. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, et al. (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* 8: R76.
68. Maglione JE, Moghanaki D, Young LJ, Manner CK, Ellies LG, et al. (2001) Transgenic Polyoma middle-T mice model premalignant mammary disease. *Cancer Res* 61: 8298–8305.

69. Pandey PR, Saidou J, Watabe K (2010) Role of myoepithelial cells in breast tumor progression. *Front Biosci (Landmark Ed)* 15: 226–236.
70. Dickson C, Spencer-Dene B, Dillon C, Fantl V (2000) Tyrosine kinase signalling in breast cancer: fibroblast growth factors and their receptors. *Breast Cancer Res* 2: 191–196.
71. Adnane J, Gaudray P, Dionne CA, Crumley G, Jaye M, et al. (1991) BEK and FLG, two receptors to members of the FGF family, are amplified in subsets of human breast cancers. *Oncogene* 6: 659–663.
72. Sun S, Jiang Y, Zhang G, Song H, Zhang X, et al. (2012) Increased expression of fibroblastic growth factor receptor 2 is correlated with poor prognosis in patients with breast cancer. *J Surg Oncol* 105: 773–779.
73. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39: 870–874.