Review

OPEN ACCESS

THE EMBO JOURNAL

Transcriptional regulation of normal human mammary cell heterogeneity and its perturbation in breast cancer

Davide Pellacani, Susanna Tan, Sylvain Lefort & Connie J Eaves*

Abstract

The mammary gland in adult women consists of biologically distinct cell types that differ in their surface phenotypes. Isolation and molecular characterization of these subpopulations of mammary cells have provided extensive insights into their different transcriptional programs and regulation. This information is now serving as a baseline for interpreting the heterogeneous features of human breast cancers. Examination of breast cancer mutational profiles further indicates that most have undergone a complex evolutionary process even before being detected. The consequent intra-tumoral as well as inter-tumoral heterogeneity of these cancers thus poses major challenges to deriving information from early and hence likely pervasive changes in potential therapeutic interest. Recently described reproducible and efficient methods for generating human breast cancers *de novo* in immunodeficient mice transplanted with genetically altered *primary* cells now offer a promising alternative to investigate initial stages of human breast cancer development. In this review, we summarize current knowledge about key transcriptional regulatory processes operative in these partially characterized subpopulations of normal human mammary cells and effects of disrupting these processes in experimentally produced human breast cancers.

Keywords breast cancer; chromatin; epigenomics; mammary; transcriptional regulation

DOI 10.15252/embj.2018100330 | Received 23 July 2018 | Revised 22 October 2018 | Accepted 8 November 2018 | Published online 11 January 2019

The EMBO Journal (2019) 38: e100330

The normal adult human mammary gland

The adult human female mammary gland is a continuous branching tree of ducts that extend radially from the nipple and terminate in expanded alveolar structures frequently called lobules (Fig 1A). This structure is encased in a basement membrane and an outer layer of fibroblasts, all of which are embedded in a collagen-rich stroma containing adipocytes, macrophages, lymphocytes, and blood and lymph vessels. The mammary gland, itself, consists of two layers of cells with different features and functions. The outer "basal" layer is made up of cells that are in direct contact with the basement membrane. These cells are also referred to as myoepithelial cells because they possess contractile, smooth muscle-like properties. The inner "luminal" layer of the gland contains cells with quite different, polarized epithelial features and an ability to produce and secrete milk upon hormone induction.

The initial stages of development of the mammary gland that take place in humans before birth are not well documented, and hence, knowledge of these has had to rely on inferences drawn from studies of mice (Veltmaat *et al,* 2003; Spike *et al,* 2012; Makarem *et al,* 2013b). In that species, the mammary gland can be seen to originate in the embryo from cells in the ventral ectoderm that invade the underlying mesoderm to form a primitive branching structure. At this stage, the rudimentary gland is composed of cells with a mixture of properties that are associated with distinct cell types found in the adult mouse mammary gland. This primitive structure then expands rapidly after the onset of puberty. Thereafter, until menopause, the entire mammary gland in humans and mice alike undergoes continuous cyclical phases of expansion and involution under the control of changing levels of estrogen (E) and progesterone (P) (Fig 1B; Ramakrishnan *et al,* 2002). Current evidence indicates that the stimulatory effects of these hormones are exerted indirectly by activating paracrine signaling mechanisms that involve an upregulated production of amphiregulin by E, an induced secretion of RANKL by P, and an enhancing effect of hormonally controlled changes by WNT-producing macrophages (Wilson *et al,* 2006; Asselin-Labat *et al,* 2010; Brisken & O'Malley, 2010; Joshi *et al,* 2010; Roarty & Rosen, 2010; Visvader & Stingl, 2014; Arendt & Kuperwasser, 2015; Chakrabarti *et al,* 2018). Other growth factors implicated in regulating mammary gland development and homeostasis include members of the epidermal growth factor (EGF), insulin-like growth factor (IGF), and fibroblast growth factor (FGF) families (Hynes & Watson, 2010).

The development of reproducible methods for isolating the different cell types that constitute the major components of the normal adult human mammary gland as separate suspensions of single viable cells was a key advance because it then enabled the further biological and molecular characterization of these different cell

Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, BC, Canada
*Corresponding author. Tel: +1 604 675 8122; Fax: +1 604 877 0712; E-mail: ceaves@bccrc.ca

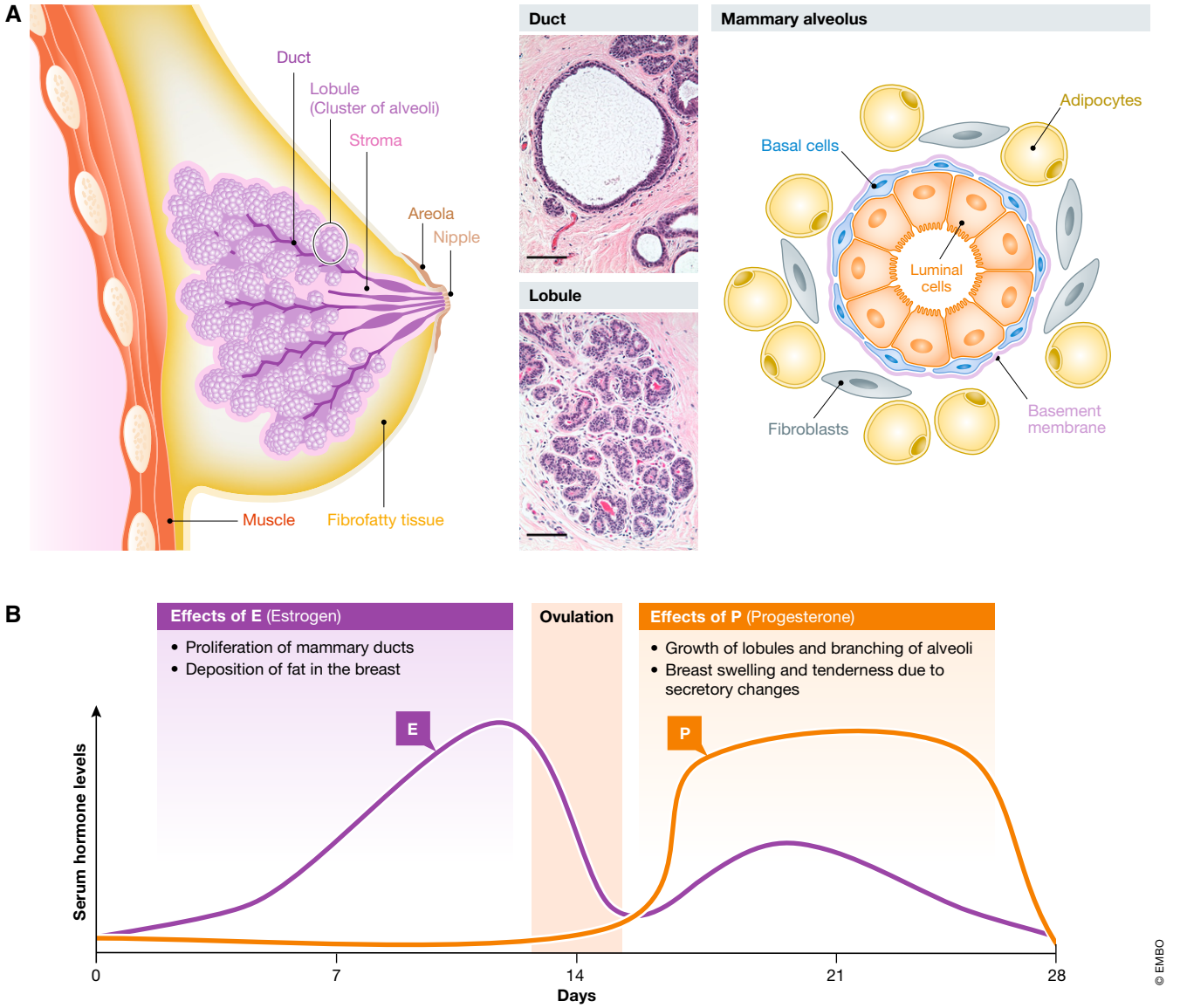

Figure 1. Macro- and microscopic structure of the normal human breast.

(A) Diagram showing the macroscopic structure of the human breast and histological sections of ducts and alveoli (scale bar = 100 μm). (B) Effects of serum hormone levels on the human mammary epithelium during the menstrual cycle.

types. Most studies of normal human mammary cells have made use of discarded tissue obtained from women without known breast disease undergoing reduction mammoplasties. The pieces of tissue obtained are then subjected to a series of enzymatic dissociation and filtration steps, followed by removal of prevalent blood and endothelial cells using antibodies against CD45 and CD31. The three major cell types that constitute the mammary gland, plus remaining stromal fibroblasts, can then be separately isolated using flow cytometry according to their differential staining with antibodies to CD49f and EpCAM (Fig 2A). The three subpopulations of mammary cells obtained are typically referred to as basal cells (BCs), luminal progenitors (LPs), and luminal cells (LCs). Other antibody cocktails have also been used to obtain highly overlapping phenotypes with very similar biological and molecular properties (Raouf *et al,* 2008; Bachelard-Cascales *et al,* 2010; Keller *et al,* 2012; Kannan *et al,* 2014; Nguyen *et al,* 2014; Fridriksdottir *et al,* 2015; Lawson *et al,* 2015; Britschgi *et al,* 2017), and additional markers have proven useful to subdivide these three subpopulations of human mammary cells even further (Eirew *et al,* 2012; Shehata *et al,* 2012; Knapp *et al,* 2017; Morel *et al,* 2017). However, the combination of antibodies to CD49f and EpCAM has generally been the most widely utilized.

BCs are defined by their $CD49f^{+}EpCAM^{low}$ phenotype and are so-named because they express numerous markers (e.g., KRT14, TP63, ACTA2/SMA, MME/CD10, and THY1/CD90) that distinguish cells of the basal layer from those of the luminal layer in histological preparations of normal human mammary tissue. In culture media containing insulin and EGF, as well as other supplements and a

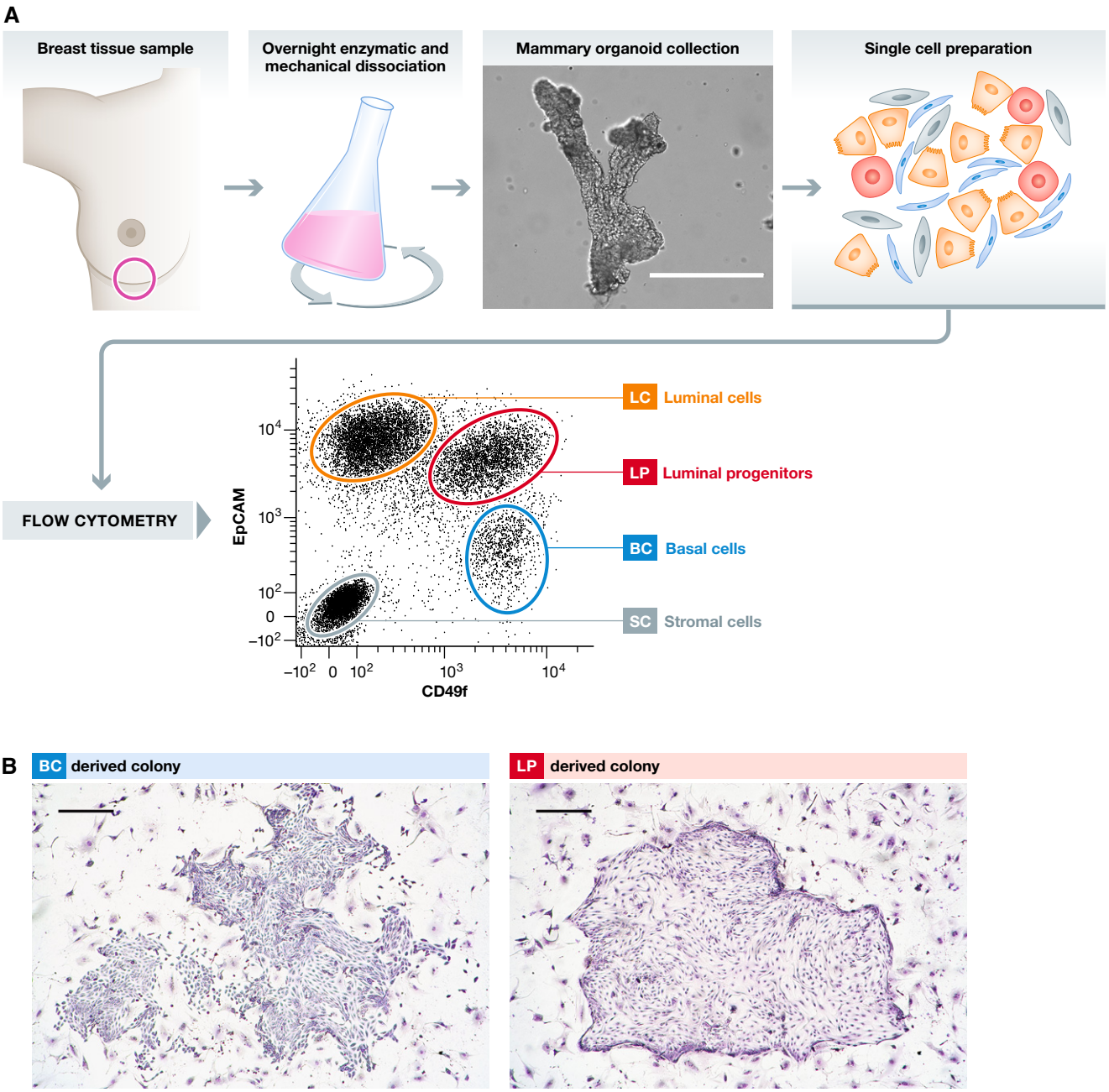

Figure 2. Subpopulations of cells within the normal adult human mammary gland.

(A) Diagram showing the workflow for separating the four main cell populations present in the breast in addition to blood cells and endothelial cells (scale bar = 400 μm). (B) Examples of typical Giemsa-stained colonies derived from BCs and LPs and assessed after 7–9 days *in vitro* (scale bar = 400 μm).

feeder layer of fibroblasts, ~ 10–20% of freshly isolated BCs plated at low density will produce readily visualized adherent colonies within 8–10 days (Fig 2B; Eirew *et al*, 2008; Kannan *et al*, 2013). Many of the individual colonies produced from BCs under these conditions will contain a mixture of cells expressing either basal or luminal markers (Stingl *et al*, 2001). A smaller fraction of the BCs (~ 0.1%) will produce bilayered epithelial structures that resemble the normal human mammary gland when injected directly into "humanized" fat pads (Kuperwasser *et al*, 2004; Proia & Kuperwasser, 2006; Lim *et al*, 2009) or when transplanted in collagen gels that are then inserted either under the kidney capsule (Eirew *et al*, 2008, 2010; Nguyen *et al*, 2015) or subcutaneously (Pellacani D and Eaves C, unpublished) in immunocompromised mice. In both of these sites, the regenerated human mammary gland-like structures contain the same spectrum of EGF-dependent *in vitro* mammary colony-forming cells (CFCs) that are present in the normal human mammary gland, as well as rarer cells that can regenerate similar bilayered mammary gland structures and

mammary CFCs upon transplantation into secondary hosts (Eirew *et al,* 2008; Lim *et al,* 2009; Nguyen *et al,* 2014). In addition, the regenerated human gland-like structures will produce human milk proteins when appropriately hormonally stimulated (Eirew *et al,* 2008).

LPs and LCs are defined by their shared high expression of EpCAM, a well-established marker of cells that constitute the luminal layer of mammary glands. Both LPs and LCs also express other markers histologically associated with the luminal layer (e.g., KRT8, KRT18, and MUC1). However, these $EpCAM^+$ mammary cells can be readily subdivided according to their differential expression of CD49f (and CD117, c-KIT). LC is the term assigned to the $CD49f^-$ cells within the $EpCAM^+$ fraction, and they include most of the cells that express E and P receptors (ER/ESR1 and PR/PGR) and express low to undetectable levels of EGFR (Lim *et al,* 2009). Not surprisingly, LCs do not mount a significant direct signaling response to EGF (Knapp *et al,* 2017) and do not proliferate when exposed to EGF *in vitro* (Kannan *et al,* 2013, 2014). They are also incapable of reconstituting epithelial structures *in vivo* that contain clonogenic progeny (Eirew *et al,* 2008). However, it was recently reported that a small proportion (~ 0.4%) of $EpCAM^+CD271^-CD166^{high}CD117^{low}$ human mammary cells, a phenotype expected to overlap with $CD49f^-EpCAM^{hi}$ LCs, will form colonies in cultures containing inhibitors to the TGF-β pathway (Fridriksdottir *et al,* 2015). Interestingly, cultures established from these cells could be expanded for 15 population doublings and their progeny continued to express ER and respond to E stimulation. In mice, similar evidence of the proliferative activity *in vivo* of non-clonogenic LCs has also been obtained from BrdU incorporation studies (Giraddi *et al,* 2015). Together, these findings raise the possibility that at least some human mammary cells with a LC phenotype can proliferate when appropriately stimulated. Nevertheless, the relevance of these *in vitro* findings to events that underpin the cellular dynamics within the mammary gland of normal adult women remains obscure as, *in situ*, very few ER^+ or PR^+ mammary cells appear to be proliferating (Clarke *et al,* 1997; Stingl, 2011).

LPs are defined as the $EpCAM^+$ cells that co-express CD49f, suggesting that they might be an intermediate stage between BCs and LCs. However, these cells express other markers specific to the luminal layer of the epithelium assessed histologically, although only a minority express ER or PR (Lim *et al,* 2009). LPs are also distinct in their expression of high levels of CD117, a marker often used for their differential isolation (Fridriksdottir *et al,* 2015; Lawson *et al,* 2015). Approximately 50% of LPs also express KRT5/6 (Lim *et al,* 2009), a type of cytokeratin known to be expressed by cells in the basal layer of many types of epithelia (Purkis *et al,* 1990; Böcker *et al,* 1992). On average, 20–30% of LPs will generate colonies *in vitro* under the same conditions as BCs (Fig 2B). But, in this case, only cells with luminal features are produced (Stingl *et al,* 2005). A small proportion of LPs have also been reported to regenerate epithelial structures *in vivo* (Shehata *et al,* 2012), but the structures produced do not contain CFCs (Eirew *et al,* 2008).

Most LPs have very short telomeres and display a pronounced telomere-associated DNA damage response, even in mammary cells obtained from women in their twenties (Kannan *et al,* 2013). Interestingly, some LPs expressing activated caspase-3 will still show considerable subsequent proliferative activity *in vitro* (Knapp *et al,* 2017). LPs are also distinguished by elevated levels of reactive oxygen species (ROS) compared to LCs and BCs. In addition, they display an innately greater resistance to oxidative stress and a higher level of associated DNA damage (Kannan *et al,* 2014), two processes that have been proposed to accelerate telomere shortening (von Zglinicki, 2002; Richter & von Zglinicki, 2007), and predispose cells to transformation.

More recently, single-cell mass cytometry (Knapp *et al,* 2017) and RNA sequencing methodologies (Nguyen *et al,* 2018) have provided further support for the segregation of normal human mammary epithelial cells into the same three main cell types. On the other hand, these studies have also highlighted their extensive molecular heterogeneity and the possible existence of new subsets within each (Shehata *et al,* 2012; Knapp *et al,* 2017; Nguyen *et al,* 2018). Nevertheless, pseudo-temporal ordering of the available single-cell transcriptional data produces a differentiation trajectory profile that separates into three main branches corresponding to the historically visualized distinction of cells produced in the normal adult human mammary gland (Nguyen *et al,* 2018).

Taken together, these findings are consistent with a hierarchically organized sequence of changes initiated in bipotent BCs that are able to generate progeny with either luminal or basal features. Cells with luminal features can then be phenotypically and biologically segregated into an intermediate, luminal-restricted but EGF-responsive state, and a state in which the capacity to proliferate in response to EGF has been lost. However, this model of a hierarchical differentiation process should not be viewed as necessarily reflecting a series of tightly co-ordinated events and may also not reflect the operation of mechanisms that maintain these subpopulations under normal homeostatic conditions. Indeed, in the mouse, where analogous populations of BCs, LPs, and LCs have been identified, some luminal cells possess or can acquire the regenerative activity originally thought to be restricted to BCs (Shehata *et al,* 2012; Makarem *et al,* 2013a). In addition, in mice, *in situ* lineage-tracing experiments suggest that both myoepithelial and luminal lineages can display self-sustaining dynamics (Van Keymeulen *et al,* 2011), despite the continued presence and activity of transplantable cells with the bipotent regenerative properties of "stem cells" (Rios *et al,* 2014). Such findings are consistent with increasing evidence of an incomplete overlap of mechanisms that control mammary cell proliferative potential and those that determine whether their differentiated state will change (or not) with sequential divisions.

At the same time, it is important to recognize the caveats and assumptions inherent in available methods for associating functional and molecular properties of individual human mammary cells or the history of their acquisition and display. Deriving these associations is necessarily limited by an inability to undertake the requisite prospective lineage-tracing experiments in humans. Accordingly, direct measurements of normal human mammary cell outputs *in situ* cannot be compared with the outputs that can be elicited from the same cells when they are exposed to highly stimulatory conditions *in vitro* or following their transplantation into mice. In addition, both flow cytometry and clonal assays have technical limitations of efficiency and specificity. They may also be compromised by the use of markers that are not co-ordinately controlled by mechanisms that regulate their functional properties. However, these caveats may be partially reduced by the use of index-sorting strategies to link molecular and functional properties

more directly (Wilson *et al,* 2015), thereby circumventing the problem of assigning functions of rare cells present in bulk isolates.

Transcriptional differences between human mammary cell subsets

A variety of technologies have been used over the past 10 years to characterize the transcriptomes of BCs, LPs, and LCs isolated from normal adult female breast tissue (Bloushtain-Qimron *et al,* 2008; Raouf *et al,* 2008; Lim *et al,* 2009, 2010; Maruyama *et al,* 2011; Shehata *et al,* 2012; Kannan *et al,* 2013; Gascard *et al,* 2015; Pellacani *et al,* 2016). These studies have revealed consistent differences in the activity of hundreds of genes in each of these phenotypically defined subsets. In turn, these studies have pointed to a number of differentially activated pathways that may regulate their different biological properties (Liu *et al,* 2005). For example, many components of the NOTCH pathway are expressed at different levels in BCs, LPs, and LCs, with some evidence of corresponding functional consequences (Dontu *et al,* 2004; Raouf *et al,* 2008). WNT pathway components also show differential patterns of expression, with biological evidence of their importance in maintaining a mammary stem cell state, at least as inferred from studies of the mouse mammary gland (Teulière *et al,* 2005; Roarty & Rosen, 2010; Zeng & Nusse, 2010; van Amerongen *et al,* 2012; Gu *et al,* 2013) with more limited, but consistent data for human cells (Arendt *et al,* 2014). Other pathways similarly implicated are the TGF-β (Moses & Barcellos-Hoff, 2011; Kahata *et al,* 2017) and the Hippo pathways (Chen *et al,* 2014; Pelissier *et al,* 2014; Skibinski *et al,* 2014; Shi *et al,* 2015; Britschgi *et al,* 2017). Importantly, all of these are variably deregulated in breast cancers (Howard & Ashworth, 2006).

Human mammary cell epigenomes reflect their transcriptional profiles

Several studies have now characterized the epigenomic features of human as well as mouse mammary cells (Maruyama *et al,* 2011; Choudhury *et al,* 2013; Dos Santos *et al,* 2015; Gascard *et al,* 2015; Huh *et al,* 2015; Pellacani *et al,* 2016; Shin *et al,* 2016; Lee *et al,* 2017). Early studies reported an association of differences in the H3K27me3 and DNA methylation of genes that are differently expressed in luminal and basal subsets (Maruyama *et al,* 2011). These genes include several that encode transcriptional regulators and/or other members of pathways of reported activity in the mammary gland. Subsequent analyses revealed DNA methylation to be a stable mark of exonic and intronic usage, with evidence of intron retention events specific to each subpopulation and linked to differences in protein expression (Gascard *et al,* 2015). The latter study also found many more hypo-methylated enhancer elements in luminal cells (LPs + LCs) than in BCs and these were commonly associated with binding sites for *FOXA1*, *GATA3,* and *ZNF217*. These studies also indicated a higher overall transcriptional activity in the luminal cells. More extensive epigenomic characterization of highly purified human BCs, LPs, LCs and their associated stromal cells has now been derived from ChIP-seq analyses of H3K4me1, H3K4me3, H3K27me3, H3K27ac, H3K36me3, and H3K9me3 marks on histones and accompanying whole-genome bisulfite sequencing, with matching mRNA-seq and miRNA-seq data for the same cells (Pellacani *et al,* 2016). From these datasets, the chromatin landscape at putative enhancer sites of these different mammary cell types has been derived. Comparisons of these have also shown LPs to be intermediate between BCs and LCs, consistent with their different biological properties. Analysis of transcription factor binding sites (TFBS) and derived TF networks for each subpopulation has also enabled novel TFs to be identified as potential regulators of each subpopulation, in addition to others previously reported. Analysis of our more recently accrued epigenomic data has also provided new evidence of a bipartite TF network in LPs that includes elements of those operative in BCs and LCs (Fig 3A). In addition, this study showed that the epigenomic and transcriptional profiles of primary sources of normal human mammary cells are very different from those of established lines of immortalized but non-tumorigenic mammary cells (Fig 3B; Pellacani *et al,* 2016). This latter finding highlights the caveats of relying on data from such immortalized cell lines to infer mechanisms controlling the biological properties of normal human mammary cells, and, conversely, the importance of analyzing primary isolates for this purpose.

Epigenomic and transcriptional changes related to aging and reproductive history

Aging and pregnancy are associated, respectively, with an increase and decrease in breast cancer risk. Several groups have therefore started dissecting the molecular changes evident in mammary cells obtained from donors of different ages or different reproductive histories. These include a report of an expansion with aging of defective multipotent progenitors that show altered interactions with extracellular matrix elements and in $KRT14^+$ and $CD49f^+$ luminal cells (Garbe *et al,* 2012; Pelissier *et al,* 2014). Accompanying transcriptome changes suggested an aging-associated epigenomic deregulation, potentially mediated by changes in the microenvironment of the mammary gland (Miyano *et al,* 2017). Comparison of the transcriptomes of purified mammary cell subsets isolated from breast tissue of parous and nulliparous women has shown differences between the $CD44^+$ cells from these two sources, with *CDKN1B* (p27) as one of the most differentially expressed genes (Choudhury *et al,* 2013). More extensive studies in mice have shown pregnancy to be associated with long-lasting alterations in DNA methylation profiles at sequences enriched for STAT5 binding sites (Dos Santos *et al,* 2015).

Figure 3. Transcriptional regulation of normal human mammary cell subpopulations.

(A) TF regulatory networks constructed from the chromatin profiles at enhancers of BCs, LPs, and LCs. (B) Genome browser plots showing the differences in chromatin states defined for normal human mammary cell subpopulations and non-tumorigenic mammary cell lines around the *PROM1* and the *NT5E* genes.

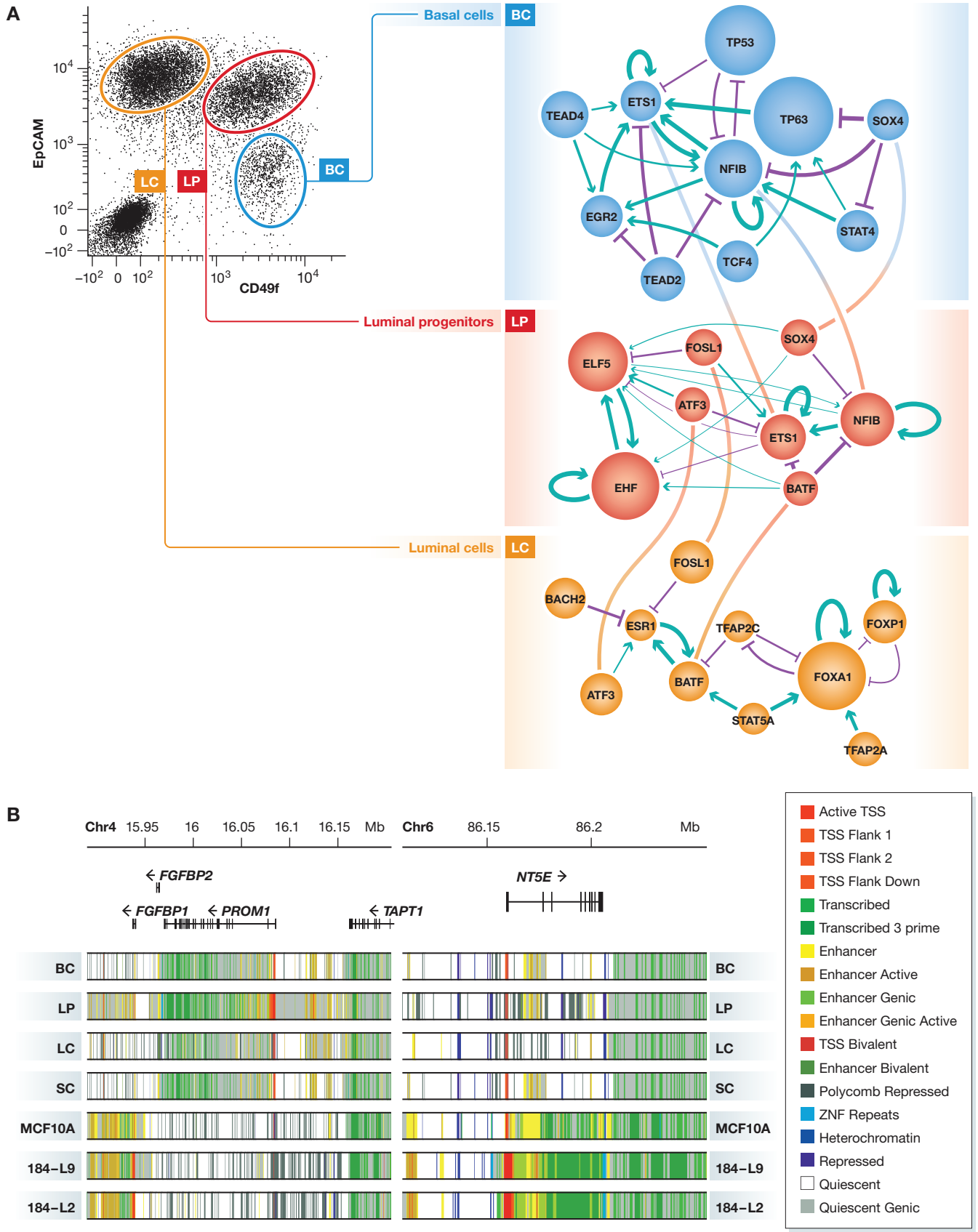

Figure 3.

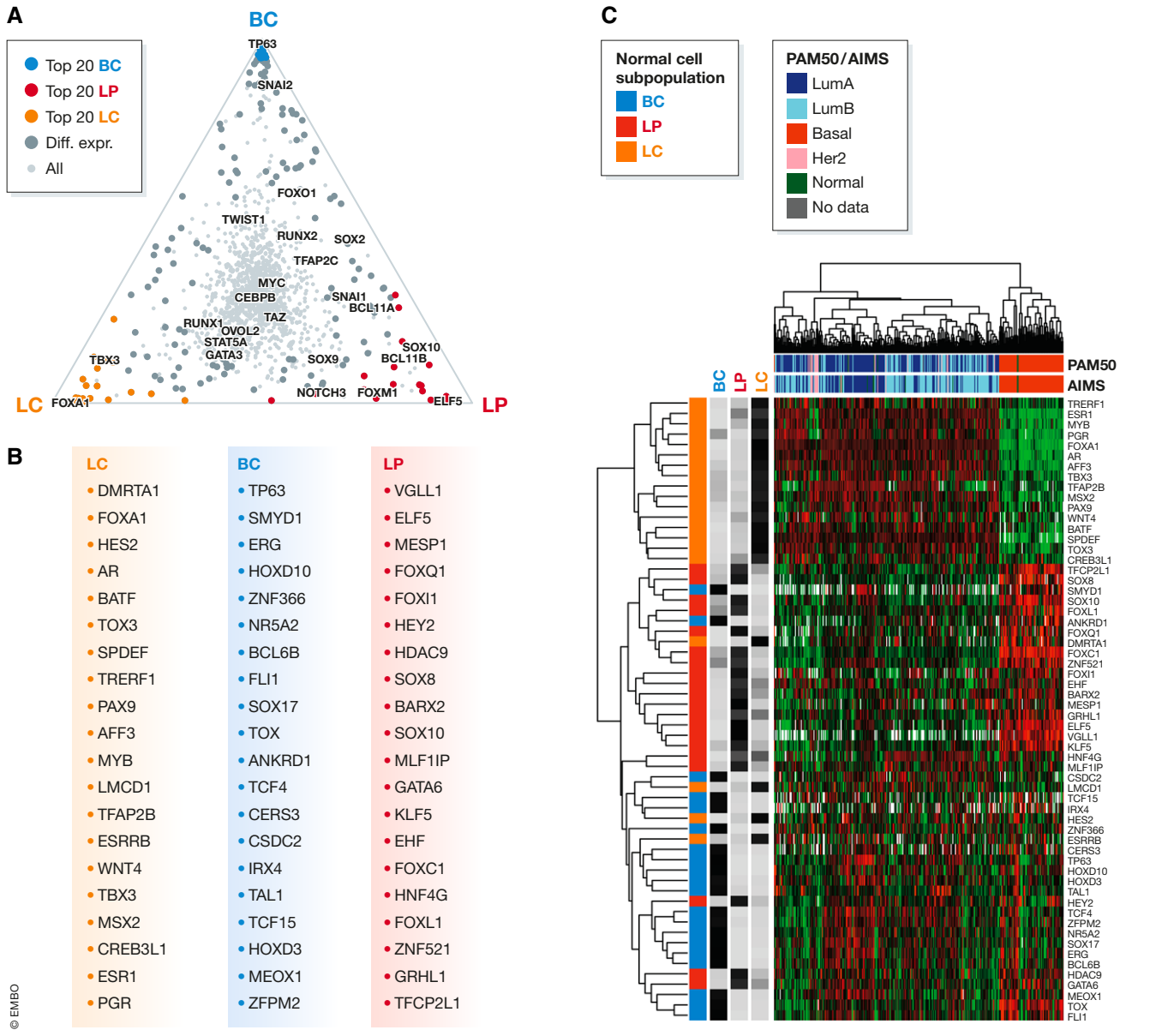

Figure 4. Transcriptional regulators active in the normal human mammary gland and in human breast cancer.

(A) Ternary plot of relative expression of all transcriptional regulators in normal human mammary cell subpopulations from a re-analysis of the RNA-seq data presented in Pellacani *et al* (2016). Transcriptional regulators discussed in the text are highlighted. (B) List of the top 20 transcriptional regulators most specific to each cell type highlighted in (A). (C) Clustering of the tumors profiled by RNA-seq in Nik-Zainal *et al* (2016) using the genes shown in (B).

Transcription factors regulating normal mammary cells

Epigenomic and transcriptional profiling of primary human mammary cells has also led to the identification of many candidate TFs that show subpopulation specificity. For example, several TFs are significantly elevated in only one of the three major subpopulations of normal human mammary cells (Fig 4A–C). *In silico* predictions further identify a differential enrichment of associated TFBSs at epigenetically defined promoter and enhancer regions in these cell types (Lim *et al,* 2010; Kannan *et al,* 2013; Gascard *et al,* 2015; Pellacani *et al,* 2016). Several studies in mice or human cell lines have also implicated a multitude of TFs to be involved in mammary cell development and differentiation. However, similar analyses of primary human cells are still very limited, although the strong correlations found between *in silico* predictions and results obtained from mice justify a brief overview of these.

One of the TFs implicated in modulating mouse mammary stem cell activity by acting directly on BCs is ΔNp63, a known regulator of normal stem cell maintenance in multiple epithelial tissues (Senoo *et al,* 2007). ΔNp63 appears to act by modulating several key pathways. These include enhancing WNT signaling by upregulating *Fzd7* expression (Chakrabarti *et al,* 2014), activating Hedgehog signaling (Li *et al,* 2008; Memmi *et al,* 2015), and partially counteracting the effects of Notch signaling (Yalcin-Ozuysal *et al,*

2010). TP63 expression in basal cells is also necessary during pregnancy and lactation: Genetic deletion of *Trp63* in keratin 14-expressing cells of the adult mouse leads to defects in luminal cell proliferation and differentiation, and failure to produce milk, due to lack of expression of the EGF family ligand NRG1 in basal cells which is required for ERBB4/STAT5A activation in luminal cells (Forster *et al,* 2014). Several SOX family TFs have likewise been implicated. For example, modulation of SOX9 expression was found to directly influence the ability of mouse mammary cells to produce organoid structures *in vitro* (Guo *et al,* 2012) and its conditional knockout impaired postnatal development of the gland (Malhotra *et al,* 2014). SOX10 is expressed specifically in mammary cells exhibiting the highest levels of stem/progenitor activity (Dravis *et al,* 2015) and SOX2 has also been implicated, albeit less directly, as its expression was induced by LGR4 downstream of WNT signaling (Wang *et al,* 2013).

Many of these studies in mice have associated expression of SOX TFs with the acquisition of features characteristic of mesenchymal cells in a process resembling an embryonic epithelial–mesenchymal transition (EMT). In fact, the possession of mesenchymal features has been frequently associated with mammary stem cell activity, both during development and subsequently throughout adulthood (Mani *et al,* 2008; Guen *et al,* 2017), although this is still controversial (Sikandar *et al,* 2017). Nevertheless, many other TFs associated with EMT have been directly linked to changes in the clonogenic or repopulating activity of mouse mammary cells. Of these, SNAI2 (SLUG) has been reported to cooperate with SOX9 (Guo *et al,* 2012) in regulating the transition of mouse mammary stem cells to short-term progenitors (Phillips *et al,* 2014). SNAI1 (SNAIL) is another member of this group, and it was found to regulate the spindle orientation machinery in mammary stem cells responding to SLIT2/ROBO1 signaling (Ballard *et al,* 2015). OVOL2, a transcriptional repressor, was likewise reported to restrict activation of EMT (Watanabe *et al,* 2014). More recently, another transcription factor, ZEB1, was shown to be expressed at high levels in a fraction of mammary BCs (Nguyen *et al,* 2018) and associated with cells expressing protein C receptor (ProCR; Wang *et al,* 2015). ZEB1 was also recently reported to have a protective role against oncogene-induced DNA damage in normal human mammary epithelial cells (Morel *et al,* 2017). Other TFs involved in mammary stem cell function include FOXO1 (Sreekumar *et al,* 2017), RUNX2 (Ferrari *et al,* 2015), MYC (Hynes & Stoelzle, 2009; Moumen *et al,* 2012), CEBPB (C/EBPβ; LaMarca *et al,* 2010), BCL11A (Khaled *et al,* 2015), and BCL11B (Miller *et al,* 2018).

TFs implicated in regulating luminal cell production and maintenance have also been identified. Of these, GATA3 was found to have an essential role in controlling the morphogenesis of the mammary gland in the mouse embryo, during puberty, and in adult life (Kouros-Mehr *et al,* 2006; Asselin-Labat *et al,* 2007). In addition, GATA3 promoted differentiation of cells within the luminal lineage in mice, potentially through a positive regulatory loop with ESR1 (Eeckhoute *et al,* 2007). FOXA1 was found to be involved in hormone-induced mammary ductal invasion (Bernardo *et al,* 2010), but did not affect lobulo-alveolar maturation and milk production. ELF5 was shown to be necessary for alveologenesis during pregnancy (Choi *et al,* 2009), and its deletion led to an accumulation of cells with mixed basal/luminal molecular phenotypes (Chakrabarti *et al,* 2012b). ELF5 was found to suppress EMT by down-regulating transcription of SNAI2 (Chakrabarti *et al,* 2012a). ELF5 also acted directly in LPs (Yamaji *et al,* 2009) to influence expression of STAT5A (Choi *et al,* 2009), another TF involved in alveologenesis (Liu *et al,* 1997). Contrary to the effects of RUNX2, RUNX1 was shown to induce the appearance of ER^+ luminal cells at least partially through the modulation of *ELF5* and *FOXA1* expression (van Bragt *et al,* 2014), potentially downstream of the p38α kinase (Del Barco Barrantes *et al,* 2018).

Notably, the Hippo pathway regulator TAZ, together with many other TFs, has recently emerged as a negative regulator of luminal differentiation in primary human cells (Skibinski *et al,* 2014). Other TFs and chromatin modifiers necessary for correct human luminal cell differentiation include TFAP2C (Cyr *et al,* 2015), TBX3 (Arendt *et al,* 2014), NOTCH3 (Raouf *et al,* 2008), FOXM1 (Carr *et al,* 2012), and KDM6A (Yoo *et al,* 2016). However, many "potential" TFs identified more recently from genome-wide epigenomic analyses of both human and mouse mammary gland cells remain poorly characterized.

Cellular and molecular heterogeneity of human breast cancers

Breast cancers arise from single cells as aberrant clones of progeny that undergo a continuous process of evolution, demarcated by distributed genetic and epigenetic alterations in successive generations of daughter cells (Balani *et al,* 2017). Those that maintain and/or confer a selective growth advantage promote successive waves of subclonal expansion depending on local conditions and/or exposure to therapeutic agents. Such a complex history of subclonal evolution leading to the production of billions of genetically heterogeneous cells in human breast cancers has been dramatically revealed from genomic DNA sequence data (Nik-Zainal *et al,* 2012; Eirew *et al,* 2014). And this profound inter-tumor as well as intra-tumor heterogeneity is further exacerbated by the metastatic process in which subclones differentially populated different sites.

Breast cancers are currently classified clinically on the basis of their extent and confinement, or not, within the basement membrane that surrounds the normal mammary gland, the proliferative activity and presence of nuclear abnormalities in the malignant cells, and their expression of ER, PR, and HER2. Global gene expression profiling has led to the identification of five major subtypes (Perou *et al,* 2000) that can now be distinguished based on the measurement of transcript levels of just 50 genes (PAM50; Parker *et al,* 2009; Nielsen *et al,* 2010; Chia *et al,* 2012). Notably, many of these detect the same perturbed features that have long been recognized histologically (Table 1). The five major subtypes thus identified are referred to as follows: basal-like, luminal A, and luminal B, normal-like, and claudin-low tumors. More recently, additional subdivisions have come from analyses of both genomic sequencing data (Cancer Genome Atlas Network, 2012; Curtis *et al,* 2012) and altered epigenomic marks (Holm *et al,* 2010, 2016; Kamalakaran *et al,* 2011).

Interestingly, the expression profiles of the five main cancer subtypes are correlated with expression profiles of BCs, LPs, and LCs (Table 1). Even the PAM50 signature relies on an assessment of many gene transcripts (e.g., *FOXA1, PGR, ESR1, KRT14, KRT5, EGFR, FOXC1*, and *MIA*) that are normally present at different levels in BCs, LPs, and LCs. Generally, the transcriptional profiles of basal-like breast cancers are closest to those of LPs, those of luminal A and B cancers to LCs, and claudin-low cancers to BCs. These

Table 1. List of the genes used for the PAM50 classification.

Symbol	Histology	BC vs. LP	BC vs. LC	LC vs. LP
ACTR3B				
ANLN				
BAG1				
BCL2			DN	UP
BIRC5				
BLVRA				UP
CCNB1				
CCNE1				
CDC20				
CDC6				
CDH3				
CENPF				
CEP55				
CXXC5				UP
EGFR			UP	DN
ERBB2	✓			
ESR1	✓	DN	DN	
EXO1				
FGFR4				
FOXA1			DN	UP
FOXC1				DN
GPR160			DN	UP
GRB7				
KIF2C				
KRT14		UP	UP	
KRT17		UP	UP	
KRT5		UP	UP	DN
MAPT				
MDM2				
MELK				
MIA		DN		DN
MKI67	✓			
MLPH		DN	DN	UP
MMP11				
MYBL2				
MYC				
NAT1			DN	UP
NDC80				
NUF2				
ORC6				
PGR	✓			UP
PHGDH				DN
PTTG1			UP	
RRM2				
SFRP1			UP	DN
SLC39A6				UP
TMEM45B			DN	UP
TYMS				
UBE2C				
UBE2T				

Gene products used routinely in histological studies (✓) and transcripts increased (UP) or decreased (DN) in mammary epithelial cells are marked. Differential gene expression data are based on the RNA-seq data presented in Pellacani *et al* (2016).

findings reinforce the concept that malignancies represent perturbations of the normal tissue from which they arise and frequently retain many components of the transcriptional regulatory networks that control cell production, differentiation, and death in the normal human mammary gland.

Altered transcriptional regulation in human breast cancers

Breast cancer "drivers" is a term that has been used to refer to mutations that are found repeatedly, suggesting they contribute to the malignant properties of the cells. In contrast, "passenger mutations" is a term often assigned to mutations that are rare and do not appear to be relevant to the genesis or progression of the malignant population. It is notable that a majority of the most frequently encountered mutations affect genes linked directly or indirectly to transcriptional regulation (Nik-Zainal *et al,* 2016; Zacksenhaus *et al,* 2017).

One of the most frequently altered transcriptional regulators is *GATA3* (mutated in > 10% of cases; Cancer Genome Atlas Network, 2012), most often in ER^+ breast cancers (Fig 5; Nik-Zainal *et al,* 2016). Both clinical and experimental lines of evidence link mutations in *GATA3* directly to breast cancer development and progression. Expression of GATA3 has been associated with a favorable prognosis, although this is still debated (Chou *et al,* 2010; Takaku *et al,* 2018), and similarly, in mice and cell lines, a heightened expression reduces tumorigenesis, suppresses metastasis, and promotes expression of a luminal molecular signature. In contrast, a loss of GATA3 has been found to accelerate tumor progression (Asselin-Labat *et al,* 2011; Chou *et al,* 2013).

In > 15% of breast cancers, *MYC* is amplified. This is generally associated with an unfavorable clinical prognosis (Deming *et al,* 2000) and an ER^- breast cancer phenotype (Fig 5; Nik-Zainal *et al,* 2016). MYC is one of the most intensively studied oncogenes (Fallah *et al,* 2017). Of particular note is recent evidence that overexpression of MYC in immortalized human mammary cells triggers a reprogramming of the epigenome that confers tumor-initiating proprieties and a down-regulation of luminal-specific TFs and genes (Poli *et al,* 2018). MYC activity has also been shown recently to be influenced by its interaction with *EPIC1,* a long non-coding RNA, that is upregulated in many cancers (Wang *et al,* 2018). Interestingly, MYC amplification was also reported to be a frequent event in the genesis of transformants from primary human mammary cells

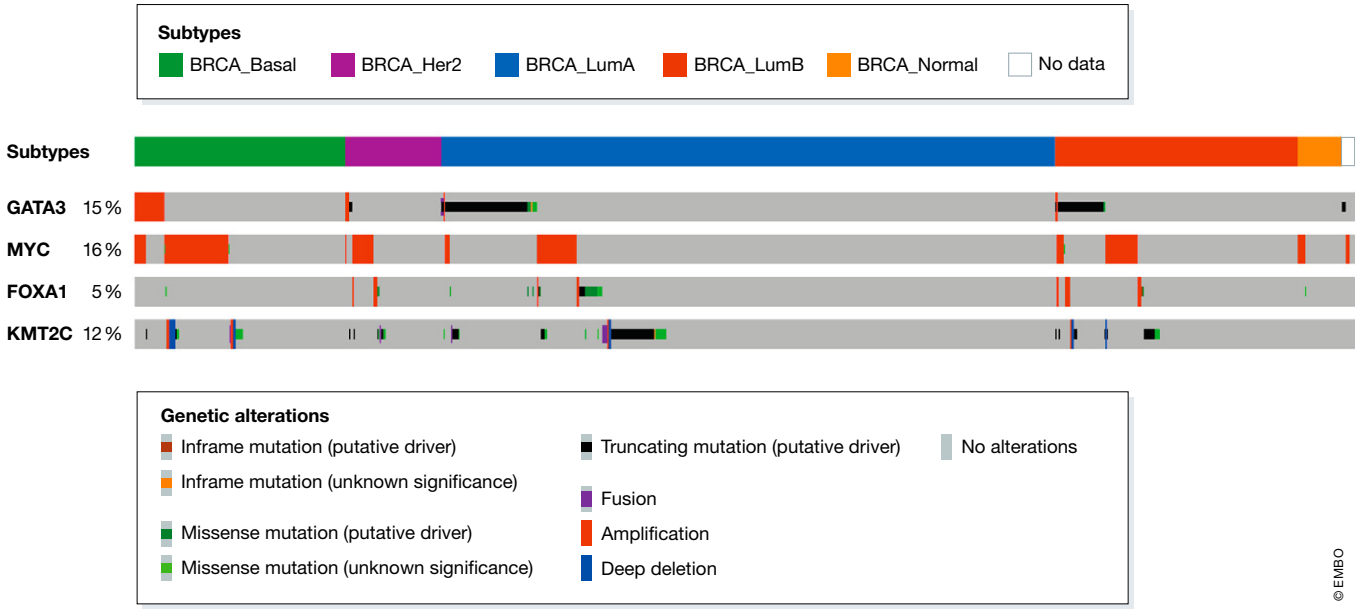

Figure 5. Frequency of genomic alteration of *GATA3*, *MYC*, *FOXA1*, and *KMT2C* in human breast cancer subtypes.
Heatmap showing the frequency of genomic alterations detected in *GATA3*, *MYC*, *FOXA1*, and *KMT2C* in human breast cancer subtypes. Data are drawn from the 993 breast cancer cases in the TCGA PanCancer Atlas study analyzed and plotted via cBioportal (http://www.cbioportal.org).

(Elenbaas *et al,* 2001) and in radiation-induced mammary cell lines (Wade *et al,* 2015).

FOXA1 is a TF that is mutated or amplified less frequently in human breast cancers (~ 2% mutated and ~ 1% amplified; Cancer Genome Atlas Network, 2012; Robinson *et al,* 2013) and usually found to be altered in ER^+ tumors (Fig 5; Nik-Zainal *et al,* 2016). FOXA1 is a main regulator of steroid receptor function in cancer (Augello *et al,* 2011), and it regulates ER signaling in breast cancer (Carroll *et al,* 2005; Lupien *et al,* 2008). FOXA1 mediates ESR1 binding and transcriptional activity (Hurtado *et al,* 2011), and its expression is associated with superior breast cancer outcomes (Shou *et al,* 2016). Molecularly, FOXA1 can recruit KMT2C (MLL3) to deposit H3K4me1 on FOXA1-bound enhancers (Jozwik *et al,* 2016).

KMT2C is another frequently mutated transcriptional regulator in breast cancer, with a mutational spectrum consistent with a loss-of-function role (Wang *et al,* 2011; Ellis *et al,* 2012; Cancer Genome Atlas Research Network, 2015). Functionally, it is the catalytic component of a complex called COMPASS (complex of proteins associated with Set1) or ASCOM (ASC-2- and MLL3-containing complex) and responsible for the monomethylation of H3K4 (Herz *et al,* 2014). In mouse models, *Mll3* deletion in the mammary gland results in hyperplasia and expansion of cells with basal features in transplant experiments, and an acceleration of PI3K-driven tumorigenesis (Zhang *et al,* 2016), supporting its role as a tumor suppressor.

Many other histone methyltransferases are deregulated in breast cancer by genetic alteration (Michalak & Visvader, 2016) and thereby contribute to an increased emergence of epigenomic alterations in breast cancer. Consequent changes in the epigenomes of analyzed human breast cancers have revealed more than 100 frequently hyper- or hypo-methylated gene promoters and pronounced global DNA hypo-methylation (Davalos *et al,* 2017; Pasculli *et al,* 2018), and the functional implication of these changes is now starting to be investigated using CRISPR/Cas9 systems (Saunderson *et al,* 2017). These findings are particularly interesting clinically, as they may offer new biomarkers of risk, prognosis, and treatment response (Pouliot *et al,* 2015; Terry *et al,* 2016) that can be robustly measured at relatively low cost (Cheuk *et al,* 2017). However, downstream transcriptional alterations are not consistently predicted and many exceptions to the general inverse correlation between promoter methylation and gene expression exist. In addition, expression of many frequently hypermethylated genes in breast cancer cells is already repressed in normal mammary cells, usually by polycomb group proteins depositing H3K27me3 (Sproul *et al,* 2011).

Comparisons of the DNA methylation profiles of individual breast cancers have shown they are highly heterogeneous. However, when subjected to unsupervised clustering, these profiles subdivide into groups that correspond largely to established transcriptionally defined breast cancer subtypes with corresponding similarities to normal human mammary subpopulations (Holm *et al,* 2010, 2016; Kamalakaran *et al,* 2011). From these, specific DNA methylation signatures with prognostic potential have been derived for luminal B and basal-like subtypes (Stefansson *et al,* 2015).

Interestingly, DNA sequence alterations that do not occur within regions that encode protein sequences directly (non-coding mutations) represent ~ 98% of mutations in cancer and most still remain poorly characterized. Of these, mutations occurring in cis-regulatory elements (i.e., enhancers and promoters) are of particular interest, as they can directly alter expression of associated gene products, by directly or indirectly altering DNA binding of TFs (Deplancke *et al,* 2016; Shi *et al,* 2016). Such mutations are frequent in breast cancer

(Bailey *et al,* 2016; Zhou *et al,* 2016; Rheinbay *et al,* 2017; Gyorffy *et al,* 2018), but their significance is generally unclear (Nik-Zainal *et al,* 2016). However, mutations in *ESR1* enhancer sequences found in ~ 7% of breast cancers have now been shown to be responsible for altering *ESR1* expression by modulating TF binding activity (Bailey *et al,* 2016). In addition, a single-nucleotide variant in one of these enhancer sequences has been associated with increased breast cancer risk. Mutations in the promoter of *FOXA1* that cause its overexpression through increased E2F binding constitute a second documented example of a biologically relevant mutation in a cis-element in some breast cancer genomes (Rheinbay *et al,* 2017). Variants linked to increased breast cancer risk have been found in distal regulatory elements of genes whose expression is modulated by FOXA1 (Cowper-Sal Lari *et al,* 2012).

Breast cancers also contain many cell types that are not part of the malignant population but, nevertheless, interact with them and co-evolve with them, adding further to the complexity and heterogeneity of breast tumors (Hanahan & Weinberg, 2011). These additional cell types include components of the blood and lymph vasculature, tissue macrophages and lymphocytes, and various stromal fibroblasts and their derivatives. Both the infiltrating leukocytes and resident cancer-associated fibroblasts (CAFs) are now well established as playing significant roles in modulating breast cancer cell growth and plasticity through direct interactions as well as through their secretion of growth factors, cytokines, and extracellular matrix components (Allinen *et al,* 2004; Aboussekhra, 2011; Place *et al,* 2011; Esquivel-Velázquez *et al,* 2015; Qiao *et al,* 2016).

One of the best characterized mechanisms of CAF modulation of human breast cancer cells is mediated by their secretion of TGF-β. Recently, this has been updated to include the suppression of adjacent normal mammary cells (Chatterjee *et al,* 2018) and the promotion of EMT in a xenografted breast cancer cell line through the transactivation of a HOX transcript antisense RNA (Ren *et al,* 2018). A third recently described role of CAFs is their induction of a FOXA1-mediated creation of a hormone-sensitive, luminal gene regulatory program in basal-like breast cancers in response to PDGF secretion by the tumor cells (Roswall *et al,* 2018). Loss of TP53 in stromal fibroblasts has also been shown to promote breast tumor development *in vivo* through the production of SDF-1 (Addadi *et al,* 2010). Additional reported mechanisms include the altered expression in CAFs of non-coding RNAs and microRNAs (Verghese *et al,* 2013; Shah *et al,* 2015; Ren *et al,* 2018). Other components of the tumor microenvironment, including tumor-associated macrophages, have been implicated in tumor promotion through the expression of *TFEB* (Fang *et al,* 2017).

Transcriptional deregulation during the initiation of breast cancers

Early events important to the genesis of human breast cancer are still limited and largely extrapolated from transgenic mouse models. Information derived from studies of human cancers has been largely limited to retrospective analyses of prevalent changes in established tumors (Futreal *et al,* 2004; Nik-Zainal *et al,* 2012), or a few analyses of preneoplastic mammary cells were obtained from carriers of *BRCA1* mutations (Lim *et al,* 2009; Proia *et al,* 2011; Choudhury *et al,* 2013) or from samples of ductal carcinoma *in situ* (DCIS; Yeong *et al,* 2017). Events that accompany the acquisition of malignant properties by immortalized, but non-tumorigenic, human mammary cell lines have also been described (Debnath *et al,* 2003; Leung & Brugge, 2012). More recently, experimental models initiated directly with primary human mammary cells have been reported.

Transgenically controlled overexpression of potential culprit genes in mice, including overexpression of MYC and HER2, was important in providing the first experimental evidence that oncogene overexpression alone could induce the formation of malignant tumors (Stewart *et al,* 1984; Muller *et al,* 1988; Bouchard *et al,* 1989). Since then, derivative approaches are now able to model metastasis due to expression of co-operating oncogenes (Sinn *et al,* 1987; Guy *et al,* 1992; Podsypanina *et al,* 2008; Adams *et al,* 2011) and assess mechanisms of pathway perturbation including TGF-β and WNT (Pierce *et al,* 1995; Li *et al,* 2000). The introduction of conditional and inducible systems to drive the expression of transgenes has enabled these models to be further refined (Sandgren *et al,* 1995; Moody *et al,* 2002; Podsypanina *et al,* 2008; Menezes *et al,* 2014; Rutkowski *et al,* 2014), including a model in mice of invasive lobular breast cancer created using CRISPR/Cas9-mediated disruption of PTEN (Annunziato *et al,* 2016).

However, a major criticism of these mouse models of breast cancer is the very ease with which the tumors can be generated. They also frequently lack the genetic complexity of human breast cancers, and their similarities to their human counterparts are often restricted to specific sites within the tumors produced (Cardiff *et al,* 2000; Hollern *et al,* 2018). In addition, their pathology may be highly dependent on the promoters used to drive expression of the oncogenic transgene and few display highly invasive properties (Cardiff *et al,* 2000). Gene expression differences in mice are also notable (Pfefferle *et al,* 2013), and some types of human breast cancer have not yet been possible to model in mice. For example, although ER^+ tumors account for the majority of all human breast cancers, stably ER^+ mouse mammary tumors have been difficult to obtain and the genetic changes that lead to ER expression in mouse tumors are frequently not characteristic of patients' ER^+ tumors (Mohibi *et al,* 2011).

Immortalized cell lines, and the MCF10A line in particular, have also been used for modeling the human mammary cell transformation process also because of their ease of use and manipulation and their availability in virtually unlimited numbers. MCF10A cells were generated by immortalizing human mammary cells obtained from a donor with benign fibrocystic disease (Soule *et al,* 1990). Forced expression of multiple cancer genes in these cells has been found to induce some features of transformation (recently reviewed in Balani *et al,* 2017). Notably, aggressively tumorigenic lines have been derived from MCF10A cells forced to overexpress HRAS and passaged *in vivo,* and their extensive characterization has revealed the presence of a number of predicted driver mutations (Maguire *et al,* 2016). However, their controlled modification has not recapitulated the phenotypic, genomic, and functional heterogeneity found in most spontaneously arising human breast cancers (Kaur & Dufour, 2012).

Analysis of DCIS has been another strategy used to investigate early events leading to invasive breast cancer. Initial transgenic mouse models of DCIS were obtained by driving expression of the

SV40 large tumor antigen with the mouse WAP promoter that becomes highly active in terminal lobular luminal cells in pregnant mice (Schulze-Garg *et al,* 2000). More recently, *in vivo* models of human DCIS have also been developed by the intraductal injection of mice with experimentally transformed human cell lines (Behbod *et al,* 2009) or primary DCIS samples from patients (Valdez *et al,* 2011). These models generally recapitulate the histology and heterogeneity of the human disease, including occasional examples of disease progression indicated by cellular invasion into the surrounding stroma.

Experimental models of *de novo* mammary tumorigenesis starting from isolated primary cells from normal tissues are particularly attractive because they avoid species differences and concerns of extrapolating from human immortalized cell lines. However, there are very few reports of genetic perturbations that consistently yield fully malignant human mammary cells in transplanted female immunodeficient mouse hosts (either NOD/SCID or NRG—NOD-*Rag1*$^{-/-}$-*IL2R*γ*c*$^{-/-}$ mice). Interestingly, most of those that have been reported have used different combinations of oncogenes, cell types, and sites of injection, with or without added fibroblasts. Immunohistological analyses of tumors produced from human EpCAM^{+} luminal cells transduced with either TP53^{R175H} + CCND1 + myristoylated PIK3CA + KRAS^{G12V} or SV40 T antigen + KRAS^{G12V} transplanted into "humanized" fat pads of NOD/SCID mice (obtained by added injection of human fibroblasts) suggested the tumors most closely resemble ductal carcinomas with predominant luminal features, including expression of ERα, CK8/18, and CK19. In contrast, the same manipulation of CD10^{+} (basal) cells caused them to acquire squamous and metaplastic features with reduced ERα and CK19 expression and robust expression of the basal marker, CK14 (Keller *et al,* 2012). On the other hand, we have found that transduction of either normal human BCs or LPs (but not LCs or SCs from the same mammoplasty samples) with just a *KRAS*^{G12D}-encoding vector produces serially transplantable invasive ductal carcinomas rapidly and at high efficiency in mice using injection sites under the kidney capsule or subcutaneously (Nguyen *et al,* 2015). These *KRAS*^{G12D}-derived tumors are also highly heterogeneous with variable proportions of cells positive for ERα, Ki67, EGFR, CK14, and CK8/18, independently of their BC or LP cell of origin (Nguyen *et al,* 2015).

Use of a DNA barcoding strategy, to track the clonal dynamics of the primary and secondary *KRAS*^{G12D}-derived tumors, showed them to be consistently and highly polyclonal, regardless of the initial cell type transduced (Nguyen *et al,* 2015). The median size of the few clones found in both primary and secondary tumors derived from the same initial inoculum was larger than most of the clones appearing only after a first passage. Interestingly, normal human mammary cells transduced with the same tracking vector also showed a delayed appearance of new and larger clones in the "normal" structures obtained in secondary as compared to primary recipients of the same original cells (Nguyen *et al,* 2014). The invasive nature of the primary clones but their general lack of perpetuation in secondary implants contrasts with the conventional concept of the oncogenic process, in which the control of invasive properties by human mammary cells is usually modeled as property that is acquired *after* deregulated growth has created a large "premalignant" population from which a more advanced derivative then arises. Taken together, these findings thus challenge previous assumptions of a requirement for a multi-step selective process during which the genetic and/or epigenetic changes needed to obtain a continuously growing invasive tumor are successively accrued.

Transcriptional profiling of the polyclonal *KRAS*^{G12D}-induced primary tumors we have described has shown they are characterized by a global deregulation of gene expression that is largely but not completely independent of the cell type used to initiate them (Nguyen *et al,* 2015). A similar result was found for tumors derived by transducing primary cells from the same normal donors with SV40 T antigen + KRAS^{G12V} (Keller *et al,* 2012) or cells from donors with a different *BRCA1* mutation status using vectors encoding TP53^{R175H} + CCND1 + myristoylated PIK3CA + KRAS^{G12V} (Proia *et al,* 2011). Thus, the initiating cell type may not necessarily make a major contribution to the transcriptional profile of the cells constituting the bulk of any breast cancer. Such a concept is of interest as it challenges the idea that globally acquired molecular profiles of breast cancers will provide informative indications of the cell of origin or the cells from which relapses are most likely to emerge.

Conclusions

Heterogeneity is a pronounced feature of human breast cancer genomes and epigenomes. These variable features likely explain the corresponding heterogeneity evident in the transcriptomes of these malignant populations. The multitude of these alterations, plus the still partial elucidation of the molecular networks governing the properties of normal human mammary cells, still obscures identification of critical initial transforming events. Nevertheless, early changes that lead to human breast cancer development remain important potential targets for more effective strategies. Expansion of *de novo* models now appears possible with established robust transduction protocols and new screening approaches on the horizon. The coupling of these strategies with clonal analyses, highly multiplexed gene manipulations, and exposure to small molecules thus holds new promise for the future more rapid identification of targetable mechanisms critical to breast cancer development.

Acknowledgements

This review was prepared with support from the BC Cancer Foundation Strategic Priorities fund, grant #702851 from the Canadian Cancer Society Research Institute and grant #22416 jointly funded by The Cancer Research Society and the Canadian Institutes of Health Research (CIHR). S Tan held a CIHR Banting and Best Research Studentship.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aboussekhra A (2011) Role of cancer-associated fibroblasts in breast cancer development and prognosis. *Int J Dev Biol* 55: 841 – 849

Adams JR, Xu K, Liu JC, Agamez NMR, Loch AJ, Wong RG, Wang W, Wright KL, Lane TF, Zacksenhaus E, Egan SE (2011) Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. *Cancer Res* 71: 2706 – 2717

Addadi Y, Moskovits N, Granot D, Lozano G, Carmi Y, Apte RN, Neeman M, Oren M (2010) p53 status in stromal fibroblasts modulates tumor growth in an SDF1-dependent manner. *Cancer Res* 70: 9650 – 9658

Allinen M, Beroukhim R, Cai L, Brennan C, Lahti-Domenici J, Huang H, Porter D, Hu M, Chin L, Richardson A, Schnitt S, Sellers WR, Polyak K (2004) Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* 6: 17 – 32

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

Annunziato S, Kas SM, Nethe M, Yücel H, Del Bravo J, Pritchard C, Bin Ali R, van Gerwen B, Siteur B, Drenth AP, Schut E, van de Ven M, Boelens MC, Klarenbeek S, Huijbers IJ, van Miltenburg MH, Jonkers J (2016) Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev* 30: 1470 – 1480

Arendt LM, St Laurent J, Wronski A, Caballero S, Lyle SR, Naber SP, Kuperwasser C (2014) Human breast progenitor cell numbers are regulated by WNT and TBX3. *PLoS One* 9: e111442

Arendt LM, Kuperwasser C (2015) Form and function: how estrogen and progesterone regulate the mammary epithelial hierarchy. *J Mammary Gland Biol Neoplasia* 20: 9 – 25

Asselin-Labat M-L, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, Hartley L, Robb L, Grosveld FG, van der Wees J, Lindeman GJ, Visvader JE (2007) Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat Cell Biol* 9: 201 – 209

Asselin-Labat M-L, Vaillant F, Sheridan JM, Pal B, Wu D, Simpson ER, Yasuda H, Smyth GK, Martin TJ, Lindeman GJ, Visvader JE (2010) Control of mammary stem cell function by steroid hormone signalling. *Nature* 465: 798 – 802

Asselin-Labat M-L, Sutherland KD, Vaillant F, Gyorki DE, Wu D, Holroyd S, Breslin K, Ward T, Shi W, Bath ML, Deb S, Fox SB, Smyth GK, Lindeman GJ, Visvader JE (2011) Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Mol Cell Biol* 31: 4609 – 4622

Augello MA, Hickey TE, Knudsen KE (2011) FOXA1: master of steroid receptor function in cancer. *EMBO J* 30: 3885 – 3894

Bachelard-Cascales E, Chapellier M, Delay E, Pochon G, Voeltzel T, Puisieux A, Caron de Fromentel C, Maguer-Satta V (2010) The CD10 enzyme is a key player to identify and regulate human mammary stem cells. *Stem Cells* 28: 1081 – 1088

Bailey SD, Desai K, Kron KJ, Mazrooei P, Sinnott-Armstrong NA, Treloar AE, Dowar M, Thu KL, Cescon DW, Silvester J, Yang SYC, Wu X, Pezo RC, Haibe-Kains B, Mak TW, Bedard PL, Pugh TJ, Sallari RC, Lupien M (2016) Noncoding somatic and inherited single-nucleotide variants converge to promote ESR1 expression in breast cancer. *Nat Genet* 48: 1260 – 1266

Balani S, Nguyen LV, Eaves CJ (2017) Modeling the process of human tumorigenesis. *Nat Commun* 8: 15422

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

Behbod F, Kittrell FS, LaMarca H, Edwards D, Kerbawy S, Heestand JC, Young E, Mukhopadhyay P, Yeh H-W, Allred DC, Hu M, Polyak K, Rosen JM, Medina D (2009) An intraductal human-in-mouse transplantation model mimics the subtypes of ductal carcinoma *in situ*. *Breast Cancer Res* 11: R66

Bernardo GM, Lozada KL, Miedler JD, Harburg G, Hewitt SC, Mosley JD, Godwin AK, Korach KS, Visvader JE, Kaestner KH, Abdul-Karim FW, Montano MM, Keri RA (2010) FOXA1 is an essential determinant of ERalpha expression and mammary ductal morphogenesis. *Development* 137: 2045 – 2054

Bloushtain-Qimron N, Yao J, Snyder EL, Shipitsin M, Campbell LL, Mani SA, Hu M, Chen H, Ustyansky V, Antosiewicz JE, Argani P, Halushka MK, Thomson JA, Pharoah P, Porgador A, Sukumar S, Parsons R, Richardson AL, Stampfer MR, Gelman RS *et al* (2008) Cell type-specific DNA methylation patterns in the human breast. *Proc Natl Acad Sci USA* 105: 14076 – 14081

Böcker W, Bier B, Freytag G, Brömmelkamp B, Jarasch ED, Edel G, Dockhorn-Dworniczak B, Schmid KW (1992) An immunohistochemical study of the breast using antibodies to basal and luminal keratins, alpha-smooth muscle actin, vimentin, collagen IV and laminin. Part I: normal breast and benign proliferative lesions. *Virchows Arch A Pathol Anat Histopathol* 421: 315 – 322

Bouchard L, Lamarre L, Tremblay PJ, Jolicoeur P (1989) Stochastic appearance of mammary tumors in transgenic mice carrying the MMTV/c-neu oncogene. *Cell* 57: 931 – 936

van Bragt MPA, Hu X, Xie Y, Li Z (2014) RUNX1, a transcription factor mutated in breast cancer, controls the fate of ER-positive mammary luminal cells. *Elife* 4: e03881

Brisken C, O'Malley B (2010) Hormone action in the mammary gland. *Cold Spring Harb Perspect Biol* 2: a003178

Britschgi A, Duss S, Kim S, Couto JP, Brinkhaus H, Koren S, De Silva D, Mertz KD, Kaup D, Varga Z, Voshol H, Vissieres A, Leroy C, Roloff T, Stadler MB, Scheel CH, Miraglia LJ, Orth AP, Bonamy GMC, Reddy VA *et al* (2017) The Hippo kinases LATS1 and 2 control human breast cell fate via crosstalk with ERα. *Nature* 541: 541 – 545

Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490: 61 – 70

Cancer Genome Atlas Research Network (2015) The molecular taxonomy of primary prostate cancer. *Cell* 163: 1011 – 1025

Cardiff RD, Anver MR, Gusterson BA, Hennighausen L, Jensen RA, Merino MJ, Rehm S, Russo J, Tavassoli FA, Wakefield LM, Ward JM, Green JE (2000) The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene* 19: 968 – 988

Carr JR, Kiefer MM, Park HJ, Li J, Wang Z, Fontanarosa J, DeWaal D, Kopanja D, Benevolenskaya EV, Guzman G, Raychaudhuri P (2012) FoxM1 regulates mammary luminal cell fate. *Cell Rep* 1: 715 – 729

Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* 122: 33 – 43

Chakrabarti R, Hwang J, Andres Blanco M, Wei Y, Lukačišin M, Romano R-A, Smalley K, Liu S, Yang Q, Ibrahim T, Mercatali L, Amadori D, Haffty BG, Sinha S, Kang Y (2012a) Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat Cell Biol* 14: 1212 – 1222

Chakrabarti R, Wei Y, Romano R-A, DeCoste C, Kang Y, Sinha S (2012b) Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells* 30: 1496 – 1508

Chakrabarti R, Wei Y, Hwang J, Hang X, Andres Blanco M, Choudhury A, Tiede B, Romano R-A, DeCoste C, Mercatali L, Ibrahim T, Amadori D, Kannan N, Eaves CJ, Sinha S, Kang Y (2014) ΔNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and Wnt signalling. *Nat Cell Biol* 16: 1004 – 1015, 1–13

Chakrabarti R, Celià-Terrassa T, Kumar S, Hang X, Wei Y, Choudhury A, Hwang J, Peng J, Nixon B, Grady JJ, DeCoste C, Gao J, van Es JH, Li MO,

Aifantis I, Clevers HC, Kang Y (2018) Notch ligand Dll1 mediates cross-talk between mammary stem cells and the macrophageal niche. *Science* 360: eaan4153

Chatterjee S, Basak P, Buchel E, Safneck J, Murphy LC, Mowat M, Kung SK, Eirew P, Eaves CJ, Raouf A (2018) Breast cancers activate stromal fibroblast-induced suppression of progenitors in adjacent normal tissue. *Stem Cell Reports* 10: 196 – 211

Chen Q, Zhang N, Gray RS, Li H, Ewald AJ, Zahnow CA, Pan D (2014) A temporal requirement for Hippo signaling in mammary gland differentiation, growth, and tumorigenesis. *Genes Dev* 28: 432 – 437

Cheuk IWY, Shin VY, Kwong A (2017) Detection of methylated circulating DNA as noninvasive biomarkers for breast cancer diagnosis. *J Breast Cancer* 20: 12 – 19

Chia SK, Bramwell VH, Tu D, Shepherd LE, Jiang S, Vickery T, Mardis E, Leung S, Ung K, Pritchard KI, Parker JS, Bernard PS, Perou CM, Ellis MJ, Nielsen TO (2012) A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. *Clin Cancer Res* 18: 4465 – 4472

Choi YS, Chakrabarti R, Escamilla-Hernandez R, Sinha S (2009) Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of Elf5. *Dev Biol* 329: 227 – 241

Chou J, Provot S, Werb Z (2010) GATA3 in development and cancer differentiation: cells GATA have it!. *J Cell Physiol* 222: 42 – 49

Chou J, Lin JH, Brenot A, Kim J-W, Provot S, Werb Z (2013) GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. *Nat Cell Biol* 15: 201 – 213

Choudhury S, Almendro V, Merino VF, Wu Z, Maruyama R, Su Y, Martins FC, Fackler MJ, Bessarabova M, Kowalczyk A, Conway T, Beresford-Smith B, Macintyre G, Cheng Y-K, Lopez-Bujanda Z, Kaspi A, Hu R, Robens J, Nikolskaya T, Haakensen VD *et al* (2013) Molecular profiling of human mammary gland links breast cancer risk to a p27(+) cell population with progenitor characteristics. *Cell Stem Cell* 13: 117 – 130

Clarke RB, Howell A, Potten CS, Anderson E (1997) Dissociation between steroid receptor expression and cell proliferation in the human breast. *Cancer Res* 57: 4987 – 4991

Cowper-Sal Lari R, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoute J, Moore JH, Lupien M (2012) Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat Genet* 44: 1191 – 1198

Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Gräf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, METABRIC Group, Langerød A, Green A, Provenzano E *et al* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486: 346 – 352

Cyr AR, Kulak MV, Park JM, Bogachek MV, Spanheimer PM, Woodfield GW, White-Baer LS, O'Malley YQ, Sugg SL, Olivier AK, Zhang W, Domann FE, Weigel RJ (2015) TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis. *Oncogene* 34: 436 – 444

Davalos V, Martinez-Cardus A, Esteller M (2017) The epigenomic revolution in breast cancer: from single-gene to genome-wide next-generation approaches. *Am J Pathol* 187: 2163 – 2174

Debnath J, Muthuswamy SK, Brugge JS (2003) Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* 30: 256 – 268

Del Barco Barrantes I, Stephan-Otto Attolini C, Slobodnyuk K, Igea A, Gregorio S, Gawrzak S, Gomis RR, Nebreda AR (2018) Regulation of mammary luminal cell fate and tumorigenesis by p38α. *Stem Cell Reports* 10: 257 – 271

Deming SL, Nass SJ, Dickson RB, Trock BJ (2000) C-myc amplification in breast cancer: a meta-analysis of its occurrence and prognostic relevance. *Br J Cancer* 83: 1688 – 1695

Deplancke B, Alpern D, Gardeux V (2016) The genetics of transcription factor DNA binding variation. *Cell* 166: 538 – 554

Dontu G, Jackson KW, McNicholas E, Kawamura MJ, Abdallah WM, Wicha MS (2004) Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Res* 6: R605 – R615

Dos Santos CO, Dolzhenko E, Hodges E, Smith AD, Hannon GJ (2015) An epigenetic memory of pregnancy in the mouse mammary gland. *Cell Rep* 11: 1102 – 1109

Dravis C, Spike BT, Harrell JC, Johns C, Trejo CL, Southard-Smith EM, Perou CM, Wahl GM (2015) Sox10 regulates stem/progenitor and mesenchymal cell states in mammary epithelial cells. *Cell Rep* 12: 2035 – 2048

Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M (2007) Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res* 67: 6477 – 6483

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

Eirew P, Stingl J, Eaves CJ (2010) Quantitation of human mammary epithelial stem cells with *in vivo* regenerative properties using a subrenal capsule xenotransplantation assay. *Nat Protoc* 5: 1945 – 1956

Eirew P, Kannan N, Knapp DJHF, Vaillant F, Emerman JT, Lindeman GJ, Visvader JE, Eaves CJ (2012) Aldehyde dehydrogenase activity is a biomarker of primitive normal human mammary luminal cells. *Stem Cells* 30: 344 – 348

Eirew P, Steif A, Khattra J, Ha G, Yap D, Farahani H, Gelmon K, Chia S, Mar C, Wan A, Laks E, Biele J, Shumansky K, Rosner J, McPherson A, Nielsen C, Roth AJL, Lefebvre C, Bashashati A, de Souza C *et al* (2014) Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* 518: 422 – 426

Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, Popescu NC, Hahn WC, Weinberg RA (2001) Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev* 15: 50 – 65

Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, Van Tine BA, Hoog J, Goiffon RJ, Goldstein TC, Ng S, Lin L, Crowder R, Snider J, Ballman K, Weber J, Chen K, Koboldt DC, Kandoth C, Schierding WS *et al* (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 486: 353 – 360

Esquivel-Velázquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J (2015) The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res* 35: 1 – 16

Fallah Y, Brundage J, Allegakoen P, Shajahan-Haq AN (2017) MYC-driven pathways in breast cancer subtypes. *Biomolecules* 7: 53

Fang L, Hodge J, Saaoud F, Wang J, Iwanowycz S, Wang Y, Hui Y, Evans TD, Razani B, Fan D (2017) Transcriptional factor EB regulates macrophage polarization in the tumor microenvironment. *Oncoimmunology* 6: e1312042

Ferrari N, Riggio AI, Mason S, McDonald L, King A, Higgins T, Rosewell I, Neil JC, Smalley MJ, Sansom OJ, Morris JR, Cameron ER, Blyth K (2015) Runx2 contributes to the regenerative potential of the mammary epithelium. *Sci Rep* 5: 15658

Forster N, Saladi SV, van Bragt M, Sfondouris ME, Jones FE, Li Z, Ellisen LW (2014) Basal cell signaling by p63 controls luminal progenitor function and lactation via NRG1. *Dev Cell* 28: 147 – 160

Fridriksdottir AJ, Kim J, Villadsen R, Klitgaard MC, Hopkinson BM, Petersen OW, Rønnov-Jessen L (2015) Propagation of oestrogen receptor-positive and oestrogen-responsive normal human breast cells in culture. *Nat Commun* 6: 8786

Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR (2004) A census of human cancer genes. *Nat Rev Cancer* 4: 177 – 183

Garbe JC, Pepin F, Pelissier FA, Sputova K, Fridriksdottir AJ, Guo DE, Villadsen R, Park M, Petersen OW, Borowsky AD, Stampfer MR, LaBarge MA (2012) Accumulation of multipotent progenitors with a basal differentiation bias during aging of human mammary epithelia. *Cancer Res* 72: 3687 – 3701

Gascard P, Bilenky M, Sigaroudinia M, Zhao J, Li L, Carles A, Delaney A, Tam A, Kamoh B, Cho S, Griffith M, Chu A, Robertson G, Cheung D, Li I, Heravi-Moussavi A, Moksa M, Mingay M, Hussainkhel A, Davis B *et al* (2015) Epigenetic and transcriptional determinants of the human breast. *Nat Commun* 6: 6351

Giraddi RR, Shehata M, Gallardo M, Blasco MA, Simons BD, Stingl J (2015) Stem and progenitor cell division kinetics during postnatal mouse mammary gland development. *Nat Commun* 6: 8487

Gu B, Watanabe K, Sun P, Fallahi M, Dai X (2013) Chromatin effector Pygo2 mediates Wnt-notch crosstalk to suppress luminal/alveolar potential of mammary stem and basal cells. *Cell Stem Cell* 13: 48 – 61

Guen VJ, Chavarria TE, Kröger C, Ye X, Weinberg RA, Lees JA (2017) EMT programs promote basal mammary stem cell and tumor-initiating cell stemness by inducing primary ciliogenesis and Hedgehog signaling. *Proc Natl Acad Sci USA* 114: E10532 – E10539

Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, Itzkovitz S, Noske A, Zürrer-Härdi U, Bell G, Tam W-L, Mani SA, van Oudenaarden A, Weinberg RA (2012) Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* 148: 1015 – 1028

Guy CT, Cardiff RD, Muller WJ (1992) Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* 12: 954 – 961

Gyorffy B, Pongor L, Bottai G, Li X, Budczies J, Szabó A, Hatzis C, Pusztai L, Santarpia L (2018) An integrative bioinformatics approach reveals coding and non-coding gene variants associated with gene expression profiles and outcome in breast cancer molecular subtypes. *Br J Cancer* 118: 1107 – 1114

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144: 646 – 674

Herz H-M, Hu D, Shilatifard A (2014) Enhancer malfunction in cancer. *Mol Cell* 53: 859 – 866

Hollern DP, Swiatnicki MR, Andrechek ER (2018) Histological subtypes of mouse mammary tumors reveal conserved relationships to human cancers. *PLoS Genet* 14: e1007135

Holm K, Hegardt C, Staaf J, Vallon-Christersson J, Jönsson G, Olsson H, Borg Å, Ringnér M (2010) Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. *Breast Cancer Res* 12: R36

Holm K, Staaf J, Lauss M, Aine M, Lindgren D, Bendahl P-O, Vallon-Christersson J, Barkardottir RB, Höglund M, Borg Å, Jönsson G, Ringnér M (2016) An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells. *Breast Cancer Res* 18: 27

Howard B, Ashworth A (2006) Signalling pathways implicated in early mammary gland morphogenesis and breast cancer. *PLoS Genet* 2: e112

Huh SJ, Clement K, Jee D, Merlini A, Choudhury S, Maruyama R, Yoo R, Chytil A, Boyle P, Ran FA, Moses HL, Barcellos-Hoff MH, Jackson-Grusby L, Meissner A, Polyak K (2015) Age- and pregnancy-associated DNA methylation changes in mammary epithelial cells. *Stem Cell Reports* 4: 297 – 311

Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS (2011) FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet* 43: 27 – 33

Hynes NE, Stoelzle T (2009) Key signalling nodes in mammary gland development and cancer: Myc. *Breast Cancer Res* 11: 210

Hynes NE, Watson CJ (2010) Mammary gland growth factors: roles in normal development and in cancer. *Cold Spring Harb Perspect Biol* 2: a003186

Joshi PA, Jackson HW, Beristain AG, Di Grappa MA, Mote PA, Clarke CL, Stingl J, Waterhouse PD, Khokha R (2010) Progesterone induces adult mammary stem cell expansion. *Nature* 465: 803 – 807

Jozwik KM, Chernukhin I, Serandour AA, Nagarajan S, Carroll JS (2016) FOXA1 directs H3K4 monomethylation at enhancers via recruitment of the methyltransferase MLL3. *Cell Rep* 17: 2715 – 2723

Kahata K, Maturi V, Moustakas A (2017) TGF-β family signaling in ductal differentiation and branching morphogenesis. *Cold Spring Harb Perspect Biol* 10: a031997

Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, Riggs M, Banerjee N, Synnestvedt M, Schlichting E, Kåresen R, Shama Prasada K, Rotti H, Rao R, Rao L, Eric Tang M-H, Satyamoorthy K, Lucito R, Wigler M, Dimitrova N *et al* (2011) DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. *Mol Oncol* 5: 77 – 92

Kannan N, Huda N, Tu L, Droumeva R, Aubert G, Chavez E, Brinkman RR, Lansdorp P, Emerman J, Abe S, Eaves CJ, Gilley D (2013) The luminal progenitor compartment of the normal human mammary gland constitutes a unique site of telomere dysfunction. *Stem Cell Reports* 1: 28 – 37

Kannan N, Nguyen LV, Makarem M, Dong Y, Shih K, Eirew P, Raouf A, Emerman JT, Eaves CJ (2014) Glutathione-dependent and -independent oxidative stress-control mechanisms distinguish normal human mammary epithelial cell subsets. *Proc Natl Acad Sci USA* 111: 7789 – 7794

Kaur G, Dufour JM (2012) Cell lines: valuable tools or useless artifacts. *Spermatogenesis* 2: 1 – 5

Keller PJ, Arendt LM, Skibinski A, Logvinenko T, Klebba I, Dong S, Smith AE, Prat A, Perou CM, Gilmore H, Schnitt S, Naber SP, Garlick JA, Kuperwasser C (2012) Defining the cellular precursors to human breast cancer. *Proc Natl Acad Sci USA* 109: 2772 – 2777

Khaled WT, Choon Lee S, Stingl J, Chen X, Raza Ali H, Rueda OM, Hadi F, Wang J, Yu Y, Chin S-F, Stratton M, Futreal A, Jenkins NA, Aparicio S, Copeland NG, Watson CJ, Caldas C, Liu P (2015) BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat Commun* 6: 5987

Knapp DJHF, Kannan N, Pellacani D, Eaves CJ (2017) Mass cytometric analysis reveals viable activated caspase-3(+) luminal progenitors in the normal adult human mammary gland. *Cell Rep* 21: 1116 – 1126

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

Kuperwasser C, Chavarria T, Wu M, Magrane G, Gray JW, Carey L, Richardson A, Weinberg RA (2004) Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc Natl Acad Sci USA* 101: 4966 – 4971

LaMarca HL, Visbal AP, Creighton CJ, Liu H, Zhang Y, Behbod F, Rosen JM (2010) CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland. *Stem Cells* 28: 535 – 544

Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, Zhou A, Eyob H, Balakrishnan S, Wang C-Y, Yaswen P, Goga A, Werb Z (2015) Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* 526: 131 – 135

Lee HK, Willi M, Wang C, Yang CM, Smith HE, Liu C, Hennighausen L (2017) Functional assessment of CTCF sites at cytokine-sensing mammary enhancers using CRISPR/Cas9 gene editing in mice. *Nucleic Acids Res* 45: 4606 – 4618

Leung CT, Brugge JS (2012) Outgrowth of single oncogene-expressing cells from suppressive epithelial environments. *Nature* 482: 410 – 413

Li Y, Hively WP, Varmus HE (2000) Use of MMTV-Wnt-1 transgenic mice for studying the genetic basis of breast cancer. *Oncogene* 19: 1002 – 1009

Li N, Singh S, Cherukuri P, Li H, Yuan Z, Ellisen LW, Wang B, Robbins D, DiRenzo J (2008) Reciprocal intraepithelial interactions between TP63 and hedgehog signaling regulate quiescence and activation of progenitor elaboration by mammary stem cells. *Stem Cells* 26: 1253 – 1264

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, Asselin-Labat M-L, Gyorki DE, Ward T, Partanen A, Feleppa F, Huschtscha LI, Thorne HJ, kConFab, Fox SB, Yan M, French JD, Brown MA, Smyth GK, Visvader JE *et al* (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 15: 907 – 913

Lim E, Wu D, Pal B, Bouras T, Asselin-Labat M-L, Vaillant F, Yagita H, Lindeman GJ, Smyth GK, Visvader JE (2010) Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res* 12: R21

Liu X, Robinson GW, Wagner K-U, Garrett L, Wynshaw-Boris A, Hennighausen L (1997) Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev* 11: 179 – 186

Liu S, Dontu G, Wicha MS (2005) Mammary stem cells, self-renewal pathways, and carcinogenesis. *Breast Cancer Res* 7: 86 – 95

Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M (2008) FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* 132: 958 – 970

Maguire SL, Peck B, Wai PT, Campbell J, Barker H, Gulati A, Daley F, Vyse S, Huang P, Lord CJ, Farnie G, Brennan K, Natrajan R (2016) Three-dimensional modelling identifies novel genetic dependencies associated with breast cancer progression in the isogenic MCF10 model. *J Pathol* 240: 315 – 328

Makarem M, Kannan N, Nguyen LV, Knapp DJHF, Balani S, Prater MD, Stingl J, Raouf A, Nemirovsky O, Eirew P, Eaves CJ (2013a) Developmental changes in the *in vitro* activated regenerative activity of primitive mammary epithelial cells. *PLoS Biol* 11: e1001630

Makarem M, Spike BT, Dravis C, Kannan N, Wahl GM, Eaves CJ (2013b) Stem cells and the developing mammary gland. *J Mammary Gland Biol Neoplasia* 18: 209 – 219

Malhotra GK, Zhao X, Edwards E, Kopp JL, Naramura M, Sander M, Band H, Band V (2014) The role of Sox9 in mouse mammary gland development and maintenance of mammary stem and luminal progenitor cells. *BMC Dev Biol* 14: 47

Mani SA, Guo W, Liao M-J, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Brisken C, Yang J, Weinberg RA (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133: 704 – 715

Maruyama R, Choudhury S, Kowalczyk A, Bessarabova M, Beresford-Smith B, Conway T, Kaspi A, Wu Z, Nikolskaya T, Merino VF, Lo P-K, Liu XS, Nikolsky Y, Sukumar S, Haviv I, Polyak K (2011) Epigenetic regulation of cell type-specific expression patterns in the human mammary epithelium. *PLoS Genet* 7: e1001369

Memmi EM, Sanarico AG, Giacobbe A, Peschiaroli A, Frezza V, Cicalese A, Pisati F, Tosoni D, Zhou H, Tonon G, Antonov A, Melino G, Pelicci PG, Bernassola F (2015) p63 Sustains self-renewal of mammary cancer stem cells through regulation of Sonic Hedgehog signaling. *Proc Natl Acad Sci USA* 112: 3499 – 3504

Menezes ME, Das SK, Emdad L, Windle JJ, Wang X-Y, Sarkar D, Fisher PB (2014) Genetically engineered mice as experimental tools to dissect the critical events in breast cancer. *Adv Cancer Res* 121: 331 – 382

Michalak EM, Visvader JE (2016) Dysregulation of histone methyltransferases in breast cancer – opportunities for new targeted therapies? *Mol Oncol* 10: 1497 – 1515

Miller DH, Jin DX, Sokol ES, Cabrera JR, Superville DA, Gorelov RA, Kuperwasser C, Gupta PB (2018) BCL11B drives human mammary stem cell self-renewal *in vitro* by inhibiting basal differentiation. *Stem Cell Reports* 10: 1131 – 1145

Miyano M, Sayaman RW, Stoiber MH, Lin C-H, Stampfer MR, Brown JB, LaBarge MA (2017) Age-related gene expression in luminal epithelial cells is driven by a microenvironment made from myoepithelial cells. *Aging (Albany NY)* 9: 2026 – 2051

Mohibi S, Mirza S, Band H, Band V (2011) Mouse models of estrogen receptor-positive breast cancer. *J Carcinog* 10: 35

Moody SE, Sarkisian CJ, Hahn KT, Gunther EJ, Pickup S, Dugan KD, Innocent N, Cardiff RD, Schnall MD, Chodosh LA (2002) Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. *Cancer Cell* 2: 451 – 461

Morel A-P, Ginestier C, Pommier RM, Cabaud O, Ruiz E, Wicinski J, Devouassoux-Shisheboran M, Combaret V, Finetti P, Chassot C, Pinatel C, Fauvet F, Saintigny P, Thomas E, Moyret-Lalle C, Lachuer J, Despras E, Jauffret J-L, Bertucci F, Guitton J *et al* (2017) A stemness-related ZEB1-MSRB3 axis governs cellular pliancy and breast cancer genome stability. *Nat Med* 23: 568 – 578

Moses H, Barcellos-Hoff MH (2011) TGF-beta biology in mammary development and breast cancer. *Cold Spring Harb Perspect Biol* 3: a003277

Moumen M, Chiche A, Deugnier M-A, Petit V, Gandarillas A, Glukhova MA, Faraldo MM (2012) The proto-oncogene Myc is essential for mammary stem cell function. *Stem Cells* 30: 1246 – 1254

Muller WJ, Sinn E, Pattengale PK, Wallace R, Leder P (1988) Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. *Cell* 54: 105 – 115

Nguyen LV, Makarem M, Carles A, Moksa M, Kannan N, Pandoh P, Eirew P, Osako T, Kardel M, Cheung AMS, Kennedy W, Tse K, Zeng T, Zhao Y, Humphries RK, Aparicio S, Eaves CJ, Hirst M (2014) Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells. *Cell Stem Cell* 14: 253 – 263

Nguyen LV, Pellacani D, Lefort S, Kannan N, Osako T, Makarem M, Cox CL, Kennedy W, Beer PA, Carles A, Moksa M, Bilenky M, Balani S, Babovic S, Sun I, Rosin M, Aparicio S, Hirst M, Eaves CJ (2015) Barcoding reveals complex clonal dynamics of *de novo* transformed human mammary cells. *Nature* 528: 267 – 271

Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, Phung AT, Willey E, Kumar R, Jabart E, Driver I, Rock J, Goga A, Khan SA, Lawson DA, Werb Z, Kessenbrock K (2018) Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nat Commun* 9: 2028

Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MCU, Mardis ER, Perou CM, Bernard PS, Ellis MJ (2010) A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* 16: 5222 – 5232

Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, Raine K, Jones D, Marshall J, Ramakrishna M, Shlien A, Cooke SL, Hinton J, Menzies A, Stebbings LA, Leroy C, Jia M, Rance R, Mudie LJ, Gamble SJ *et al* (2012) The life history of 21 breast cancers. *Cell* 149: 994 – 1007

Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjærde OC, Langerød A, Ringnér M, Ahn S-M *et al* (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534: 47 – 54

Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 27: 1160 – 1167

Pasculli B, Barbano R, Parrella P (2018) Epigenetics of breast cancer: biology and clinical implication in the era of precision medicine. *Semin Cancer Biol* 51: 22 – 35

Pelissier FA, Garbe JC, Ananthanarayanan B, Miyano M, Lin C, Jokela T, Kumar S, Stampfer MR, Lorens JB, LaBarge MA (2014) Age-related dysfunction in mechanotransduction impairs differentiation of human mammary epithelial progenitors. *Cell Rep* 7: 1926 – 1939

Pellacani D, Bilenky M, Kannan N, Heravi-Moussavi A, Knapp DJHF, Gakkhar S, Moksa M, Carles A, Moore R, Mungall AJ, Marra MA, Jones SJM, Aparicio S, Hirst M, Eaves CJ (2016) Analysis of normal human mammary epigenomes reveals cell-specific active enhancer states and associated transcription factor networks. *Cell Rep* 17: 2060 – 2074

Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. *Nature* 406: 747 – 752

Pfefferle AD, Herschkowitz JI, Usary J, Harrell JC, Spike BT, Adams JR, Torres-Arzayus MI, Brown M, Egan SE, Wahl GM, Rosen JM, Perou CM (2013) Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome Biol* 14: R125

Phillips S, Prat A, Sedic M, Proia TA, Wronski A, Mazumdar S, Skibinski A, Shirley SH, Perou CM, Gill G, Gupta PB, Kuperwasser C (2014) Cell-state transitions regulated by SLUG are critical for tissue regeneration and tumor initiation. *Stem Cell Reports* 2: 633 – 647

Pierce DF, Gorska AE, Chytil A, Meise KS, Page DL, Coffey RJ, Moses HL (1995) Mammary tumor suppression by transforming growth factor beta 1 transgene expression. *Proc Natl Acad Sci USA* 92: 4254 – 4258

Place AE, Jin Huh S, Polyak K (2011) The microenvironment in breast cancer progression: biology and implications for treatment. *Breast Cancer Res* 13: 227

Podsypanina K, Politi K, Beverly LJ, Varmus HE (2008) Oncogene cooperation in tumor maintenance and tumor recurrence in mouse mammary tumors induced by Myc and mutant Kras. *Proc Natl Acad Sci USA* 105: 5242 – 5247

Poli V, Fagnocchi L, Fasciani A, Cherubini A, Mazzoleni S, Ferrillo S, Miluzio A, Gaudioso G, Vaira V, Turdo A, Giaggianesi M, Chinnici A, Lipari E, Bicciato S, Bosari S, Todaro M, Zippo A (2018) MYC-driven epigenetic reprogramming favors the onset of tumorigenesis by inducing a stem cell-like state. *Nat Commun* 9: 1024

Pouliot M-C, Labrie Y, Diorio C, Durocher F (2015) The role of methylation in breast cancer susceptibility and treatment. *Anticancer Res* 35: 4569 – 4574

Proia DA, Kuperwasser C (2006) Reconstruction of human mammary tissues in a mouse model. *Nat Protoc* 1: 206 – 214

Proia TA, Keller PJ, Gupta PB, Klebba I, Jones AD, Sedic M, Gilmore H, Tung N, Naber SP, Schnitt S, Lander ES, Kuperwasser C (2011) Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell* 8: 149 – 163

Purkis PE, Steel JB, Mackenzie IC, Nathrath WB, Leigh IM, Lane EB (1990) Antibody markers of basal cells in complex epithelia. *J Cell Sci* 97(Pt 1): 39 – 50

Qiao Y, He H, Jonsson P, Sinha I, Zhao C, Dahlman-Wright K (2016) AP-1 is a key regulator of proinflammatory cytokine TNFα-mediated triple-negative breast cancer progression. *J Biol Chem* 291: 5068 – 5079

Ramakrishnan R, Khan SA, Badve S (2002) Morphological changes in breast tissue with menstrual cycle. *Mod Pathol* 15: 1348 – 1356

Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, Iscove N, Jones S, McKinney S, Emerman J, Aparicio S, Marra MA, Eaves CJ (2008) Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell* 3: 109 – 118

Ren Y, Jia H-H, Xu Y-Q, Zhou X, Zhao X-H, Wang Y-F, Song X, Zhu Z-Y, Sun T, Dou Y, Tian W-P, Zhao X-L, Kang C-S, Mei M (2018) Paracrine and epigenetic control of CAF-induced metastasis: the role of HOTAIR stimulated by TGF-ß1 secretion. *Mol Cancer* 17: 5

Rheinbay E, Parasuraman P, Grimsby J, Tiao G, Engreitz JM, Kim J, Lawrence MS, Taylor-Weiner A, Rodriguez-Cuevas S, Rosenberg M, Hess J, Stewart C, Maruvka YE, Stojanov P, Cortés ML, Seepo S, Cibulskis C, Tracy A, Pugh TJ, Lee J *et al* (2017) Recurrent and functional regulatory mutations in breast cancer. *Nature* 547: 55 – 60

Richter T, von Zglinicki T (2007) A continuous correlation between oxidative stress and telomere shortening in fibroblasts. *Exp Gerontol* 42: 1039 – 1042

Rios AC, Fu NY, Lindeman GJ, Visvader JE (2014) *In situ* identification of bipotent stem cells in the mammary gland. *Nature* 506: 322 – 327

Roarty K, Rosen JM (2010) Wnt and mammary stem cells: hormones cannot fly wingless. *Curr Opin Pharmacol* 10: 643 – 649

Robinson JLL, Holmes KA, Carroll JS (2013) FOXA1 mutations in hormone-dependent cancers. *Front Oncol* 3: 20

Roswall P, Bocci M, Bartoschek M, Li H, Kristiansen G, Jansson S, Lehn S, Sjölund J, Reid S, Larsson C, Eriksson P, Anderberg C, Cortez E, Saal LH, Orsmark-Pietras C, Cordero E, Haller BK, Häkkinen J, Burvenich IJG, Lim E *et al* (2018) Microenvironmental control of breast cancer subtype elicited through paracrine platelet-derived growth factor-CC signaling. *Nat Med* 24: 463 – 473

Rutkowski MR, Allegrezza MJ, Svoronos N, Tesone AJ, Stephen TL, Perales-Puchalt A, Nguyen J, Zhang PJ, Fiering SN, Tchou J, Conejo-Garcia JR (2014) Initiation of metastatic breast carcinoma by targeting of the ductal epithelium with adenovirus-cre: a novel transgenic mouse model of breast cancer. *J Vis Exp* 85: e51171

Sandgren EP, Schroeder JA, Qui TH, Palmiter RD, Brinster RL, Lee DC (1995) Inhibition of mammary gland involution is associated with transforming growth factor alpha but not c-myc-induced tumorigenesis in transgenic mice. *Cancer Res* 55: 3915 – 3927

Saunderson EA, Stepper P, Gomm JJ, Hoa L, Morgan A, Allen MD, Jones JL, Gribben JG, Jurkowski TP, Ficz G (2017) Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat Commun* 8: 1450

Schulze-Garg C, Löhler J, Gocht A, Deppert W (2000) A transgenic mouse model for the ductal carcinoma *in situ* (DCIS) of the mammary gland. *Oncogene* 19: 1028 – 1037

Senoo M, Pinto F, Crum CP, McKeon F (2007) p63 Is essential for the proliferative potential of stem cells in stratified epithelia. *Cell* 129: 523 – 536

Shah SH, Miller P, Garcia-Contreras M, Ao Z, Machlin L, Issa E, El-Ashry D (2015) Hierarchical paracrine interaction of breast cancer associated fibroblasts with cancer cells via hMAPK-microRNAs to drive ER-negative breast cancer phenotype. *Cancer Biol Ther* 16: 1671 – 1681

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

Shi P, Feng J, Chen C (2015) Hippo pathway in mammary gland development and breast cancer. *Acta Biochim Biophys Sin (Shanghai)* 47: 53 – 59

Shi W, Fornes O, Mathelier A, Wasserman WW (2016) Evaluating the impact of single nucleotide variants on transcription factor binding. *Nucleic Acids Res* 44: 10106 – 10116

Shin HY, Willi M, Yoo KH, Zeng X, Wang C, Metser G, Hennighausen L (2016) Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat Genet* 48: 904 – 911

Shou J, Lai Y, Xu J, Huang J (2016) Prognostic value of FOXA1 in breast cancer: a systematic review and meta-analysis. *Breast* 27: 35 – 43

Sikandar SS, Kuo AH, Kalisky T, Cai S, Zabala M, Hsieh RW, Lobo NA, Scheeren FA, Sim S, Qian D, Dirbas FM, Somlo G, Quake SR, Clarke MF (2017) Role of epithelial to mesenchymal transition associated genes in mammary gland regeneration and breast tumorigenesis. *Nat Commun* 8: 1669

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

Skibinski A, Breindel JL, Prat A, Galván P, Smith E, Rolfs A, Gupta PB, LaBaer J, Kuperwasser C (2014) The Hippo transducer TAZ interacts with the SWI/SNF complex to regulate breast epithelial lineage commitment. *Cell Rep* 6: 1 – 14

Soule HD, Maloney TM, Wolman SR, Peterson WD, Brenz R, McGrath CM, Russo J, Pauley RJ, Jones RF, Brooks SC (1990) Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res* 50: 6075 – 6086

Spike BT, Engle DD, Lin JC, Cheung SK, La J, Wahl GM (2012) A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* 10: 183 – 197

Sproul D, Nestor C, Culley J, Dickson JH, Dixon JM, Harrison DJ, Meehan RR, Sims AH, Ramsahoye BH (2011) Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. *Proc Natl Acad Sci USA* 108: 4364 – 4369

Sreekumar A, Toneff MJ, Toh E, Roarty K, Creighton CJ, Belka GK, Lee D-K, Xu J, Chodosh LA, Richards JS, Rosen JM (2017) WNT-mediated regulation of FOXO1 constitutes a critical axis maintaining pubertal mammary stem cell homeostasis. *Dev Cell* 43: 436 – 448.e6

Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, Hilmarsdottir H, Olafsdottir E, Tryggvadottir L, Jonasson JG, Eyfjord J, Esteller M (2015) A DNA methylation-based definition of biologically distinct breast cancer subtypes. *Mol Oncol* 9: 555 – 568

Stewart TA, Pattengale PK, Leder P (1984) Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell* 38: 627 – 637

Stingl J, Eaves CJ, Zandieh I, Emerman JT (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 JT, Eaves CJ (2005) Epithelial progenitors in the normal human mammary gland. *J Mammary Gland Biol Neoplasia* 10: 49 – 59

Stingl J (2011) Estrogen and progesterone in normal mammary gland development and in cancer. *Horm Cancer* 2: 85 – 90

Takaku M, Grimm SA, Roberts JD, Chrysovergis K, Bennett BD, Myers P, Perera L, Tucker CJ, Perou CM, Wade PA (2018) GATA3 zinc finger 2 mutations reprogram the breast cancer transcriptional network. *Nat Commun* 9: 1059

Terry MB, McDonald JA, Wu HC, Eng S, Santella RM (2016) Epigenetic biomarkers of breast cancer risk: across the breast cancer prevention continuum. *Adv Exp Med Biol* 882: 33 – 68

Teulière J, Faraldo MM, Deugnier M-A, Shtutman M, Ben-Ze'ev A, Thiery JP, Glukhova MA (2005) Targeted activation of beta-catenin signaling in basal mammary epithelial cells affects mammary development and leads to hyperplasia. *Development* 132: 267 – 277

Valdez KE, Fan F, Smith W, Allred DC, Medina D, Behbod F (2011) Human primary ductal carcinoma *in situ* (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model. *J Pathol* 225: 565 – 573

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

Veltmaat JM, Mailleux AA, Thiery JP, Bellusci S (2003) Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation. *Differentiation* 71: 1 – 17

Verghese ET, Drury R, Green CA, Holliday DL, Lu X, Nash C, Speirs V, Thorne JL, Thygesen HH, Zougman A, Hull MA, Hanby AM, Hughes TA (2013) MiR-26b is down-regulated in carcinoma-associated fibroblasts from ER-positive breast cancers leading to enhanced cell migration and invasion. *J Pathol* 231: 388 – 399

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

Wade MA, Sunter NJ, Fordham SE, Long A, Masic D, Russell LJ, Harrison CJ, Rand V, Elstob C, Bown N, Rowe D, Lowe C, Cuthbert G, Bennett S, Crosier S, Bacon CM, Onel K, Scott K, Scott D, Travis LB *et al* (2015) c-MYC is a radiosensitive locus in human breast cells. *Oncogene* 34: 4985 – 4994

Wang X-X, Fu L, Li X, Wu X, Zhu Z, Fu L, Dong J-T (2011) Somatic mutations of the mixed-lineage leukemia 3 (MLL3) gene in primary breast cancers. *Pathol Oncol Res* 17: 429 – 433

Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, Liu M (2013) Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. *Stem Cells* 31: 1921 – 1931

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

Wang Z, Yang B, Zhang M, Guo W, Wu Z, Wang Y, Jia L, Li S, Cancer Genome Atlas Research Network, Xie W, Yang D (2018) lncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell-cycle progression in cancer. *Cancer Cell* 33: 706 – 720.e9

Watanabe K, Villarreal-Ponce A, Sun P, Salmans ML, Fallahi M, Andersen B, Dai X (2014) Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor. *Dev Cell* 29: 59 – 74

Wilson CL, Sims AH, Howell A, Miller CJ, Clarke RB (2006) Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue. *Endocr Relat Cancer* 13: 617 – 628

Wilson NK, Kent DG, Buettner F, Shehata M, Macaulay IC, Calero-Nieto FJ, Sánchez Castillo M, Oedekoven CA, Diamanti E, Schulte R, Ponting CP, Voet T, Caldas C, Stingl J, Green AR, Theis FJ, Göttgens B (2015) Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* 16: 712 – 724

Yalcin-Ozuysal O, Fiche M, Guitierrez M, Wagner K-U, Raffoul W, Brisken C (2010) Antagonistic roles of Notch and p63 in controlling mammary epithelial cell fates. *Cell Death Differ* 17: 1600 – 1612

Yamaji D, Na R, Feuermann Y, Pechhold S, Chen W, Robinson GW, Hennighausen L (2009) Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. *Genes Dev* 23: 2382 – 2387

Yeong J, Thike AA, Tan PH, Iqbal J (2017) Identifying progression predictors of breast ductal carcinoma *in situ*. *J Clin Pathol* 70: 102 – 108

Yoo KH, Oh S, Kang K, Wang C, Robinson GW, Ge K, Hennighausen L (2016) Histone demethylase KDM6A controls the mammary luminal lineage through enzyme-independent mechanisms. *Mol Cell Biol* 36: 2108 – 2120

Zacksenhaus E, Liu JC, Jiang Z, Yao Y, Xia L, Shrestha M, Ben-David Y (2017) Transcription factors in breast cancer-lessons from recent genomic analyses and therapeutic implications. *Adv Protein Chem Struct Biol* 107: 223 – 273

Zeng YA, 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

von Zglinicki T (2002) Oxidative stress shortens telomeres. *Trends Biochem Sci* 27: 339 – 344

Zhang Z, Christin JR, Wang C, Ge K, Oktay MH, Guo W (2016) Mammary-stem-cell-based somatic mouse models reveal breast cancer drivers causing cell fate dysregulation. *Cell Rep* 16: 3146 – 3156

Zhou S, Treloar AE, Lupien M (2016) Emergence of the noncoding cancer genome: a target of genetic and epigenetic alterations. *Cancer Discov* 6: 1215 – 1229