![The Journal of Clinical Investigation](JCI logo)

# Breast cancer: origins and evolution

Kornelia Polyak

**Science in Medicine**

Breast cancer is not a single disease, but rather is composed of distinct subtypes associated with different clinical outcomes. Understanding this heterogeneity is key for the development of targeted cancer-preventative and -therapeutic interventions. Current models explaining inter- and intratumoral diversity are the cancer stem cell and the clonal evolution hypotheses. Although tumor initiation and progression are predominantly driven by acquired genetic alterations, recent data implicate a role for microenvironmental and epigenetic changes as well. Comprehensive unbiased studies of tumors and patient populations have significantly advanced our molecular understanding of breast cancer, but translating these findings into clinical practice remains a challenge.

**Find the latest version:**

https://jci.me/33295/pdf

# Breast cancer: origins and evolution

Kornelia Polyak

Department of Medical Oncology, Dana-Farber Cancer Institute, and Department of Medicine,
Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts, USA.

**Breast cancer is not a single disease, but rather is composed of distinct subtypes associated with different clinical outcomes. Understanding this heterogeneity is key for the development of targeted cancer-preventative and -therapeutic interventions. Current models explaining inter- and intratumoral diversity are the cancer stem cell and the clonal evolution hypotheses. Although tumor initiation and progression are predominantly driven by acquired genetic alterations, recent data implicate a role for microenvironmental and epigenetic changes as well. Comprehensive unbiased studies of tumors and patient populations have significantly advanced our molecular understanding of breast cancer, but translating these findings into clinical practice remains a challenge.**

## Major clinical and scientific problems and advances

Breast cancer is the leading cause of cancer-related death in women world-wide (1). Despite significant advances in diagnosing and treating breast cancer, several major unresolved clinical and scientific problems remain. These are related to (a) prevention (who needs it and when), (b) diagnosis (we need more specific and sensitive methods), (c) tumor progression and recurrence (what causes it and how to predict it), (d) treatment (who should be treated and how), and (e) therapeutic resistance (how to predict, prevent, and overcome it). Resolving all these problems is complicated by the fact that breast cancer is not a single disease but is highly heterogeneous at both the molecular and clinical level (2, 3). Comprehensive gene expression profiling of large sets of tumors by multiple independent groups and technologies have revealed five major molecular subtypes of breast cancer: basal-like, luminal A, luminal B, HER2$^+$/ER$^-$, and normal breast–like (2, 4, 5). The molecular differences result in distinct clinical outcomes and responses to treatment; in general, the basal-like tumors have the worst, and luminal A-type tumors the best, prognosis (3). These subtypes are conserved across ethnic groups and are already evident at the ductal carcinoma in situ (DCIS) stage (6), suggesting distinct tumor progression pathways for each tumor type. Distinct cell of origin (e.g., cancer stem cells) and tumor subtype–specific genetic and epigenetic events are two possible, not necessarily mutually exclusive explanations of this extensive intra- and intertumoral heterogeneity (Figure 1); investigating these is one of the most intensely studied areas of research.

While the exact etiology of breast cancer is unknown, family history is one of the strongest determinants of risk, implying hereditary factors. Because germline mutations in high-penetrance cancer susceptibility genes, including *BRCA1, BRCA2,* and *TP53,* account for less than 25% of excess risk, variation in moderate- and low-penetrance genes is likely to explain the majority of cases. In line with this, several large-scale studies recently executed have identified such allele variants and SNPs in *FGFR2* (fibroblast growth factor receptor 2), *TNRC9* (thymocyte selection–associated high mobility group box 9), *MAP3K1* (mitogen-activated kinase kinase kinase 1), *LSP1* (lymphocyte-specific protein), *CASP8* (caspase 8), and *TGFB1* as being associated with breast cancer risk (7–10). Although with a few exceptions, the mechanisms by which these variants may influence breast tumorigenesis is largely unknown; their assessment could be used for the identification of individuals with higher risk of developing breast cancer. These women could potentially be candidates for cancer-preventative therapies and increased surveillance, since despite all improvements in the treatment of advanced-stage tumors, early diagnosis and prevention have the most dramatic effect on overall disease-specific outcome.

Targeted therapies for cancer provide the promise of high therapeutic efficacy with minimal side effects. Successful drug development requires the identification of tumor cell–specific molecular pathways amenable to drug development, the targeting of which leads to therapeutic response, as well as the choice of appropriate patient populations likely to realize clinical benefit from the therapeutic intervention. An elusive goal in oncology has been to improve the identification of such targets and patient populations prior to empirical clinical testing. One approach toward achieving this goal is the comprehensive sequencing of cancer genomes, which is now a real possibility following the availability of the human genome sequence and advances in sequencing technologies. Several such studies have recently been completed, some focusing only on kinases (11–13), others sequencing all characterized genes (14). Based on these results it is becoming apparent that the number of genes mutated in breast cancer is fairly high, and with a few exceptions such as components of the PI3KCA/AKT/PTEN, TP53, and NF-κB pathways, the frequency of mutations in any given gene in sporadic tumors is low. Thus the discovery of all mutated genes will require the sequencing of large sets of tumors of each breast cancer subtype (e.g., basal-like, luminal, etc.) in diverse ethnic populations. Furthermore, additional studies are necessary to address the functional relevance of the mutations detected and to differentiate random changes from casual mutations that provide growth or survival advantage for the tumor cells. The low frequency of mutations in a large set of genes seemingly presents a daunting task for the development of targeted therapies. However, the concentration of the mutated genes in a limited number of pathways (e.g., PI3KCA/AKT/PTEN and IKBKE/NF-κB) indicates that targeting pathways rather than individual genes may be a viable therapeutic approach.
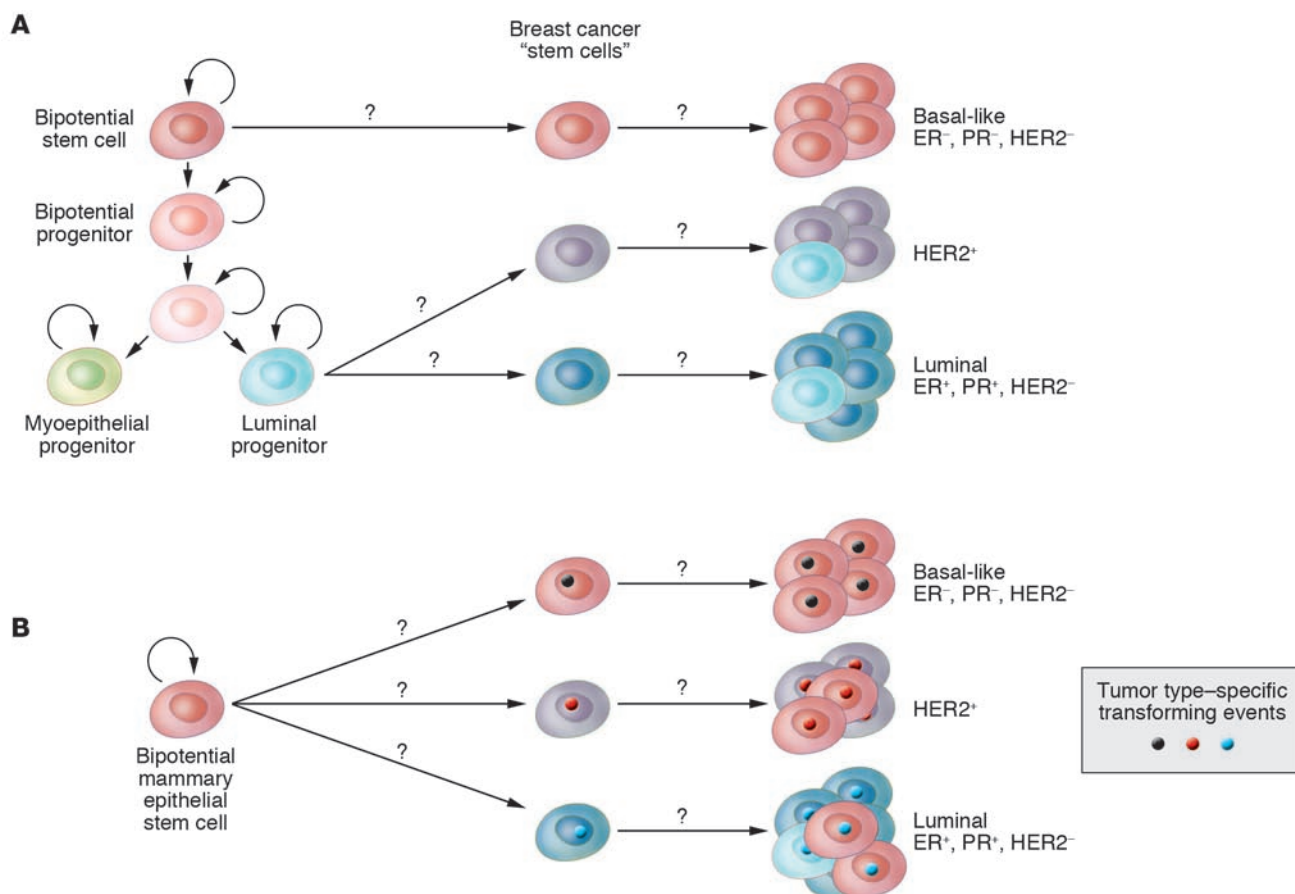
**Figure 1**
Hypothetical models explaining breast tumor subtypes. Cell of origin (**A**) and tumor subtype–specific transforming event (**B**) models. Based on the cell of origin hypothesis, each tumor subtype is initiated in a different cell type (presumably stem or progenitor cell), whereas according to the model depicted in **B**, the cell of origin can be the same for different tumor subtypes and the tumor phenotype is primarily determined by acquired genetic and epigenetic events.
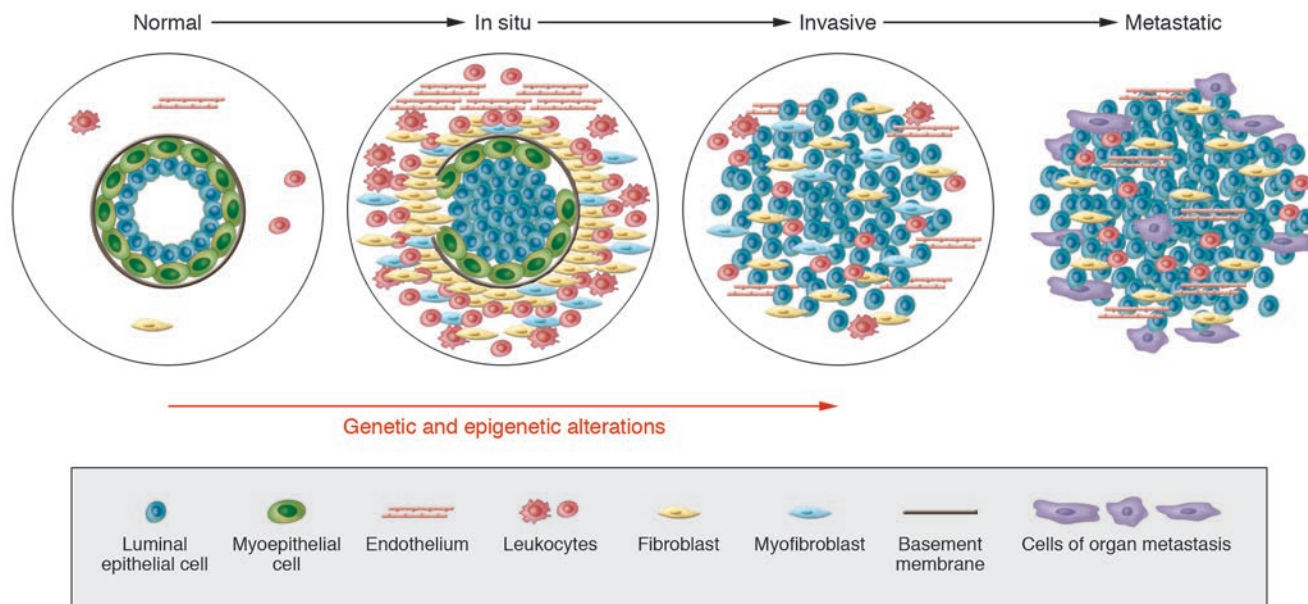
## Microenvironmental influences on breast cancer initiation and progression

The natural history of breast cancer involves progression through defined pathological and clinical stages, starting with ductal hyperproliferation, with subsequent evolution into in situ and invasive carcinomas, and finally into metastatic disease (15, 16) (Figure 2). DCIS is thought to be a precursor of invasive ductal carcinoma based on molecular, epidemiological, and pathological studies (16, 17). While there has been a dramatic improvement in our ability to detect early-stage disease, our understanding of the factors involved in tumor progression and our ability to selectively interfere with tumor progression lag far behind.

The initiation of breast cancer is due to transforming (genetic and epigenetic) events in a single cell. Subsequent tumor progression is driven by the accumulation of additional genetic changes combined with clonal expansion and selection. Based on these assumptions several groups have performed comprehensive gene expression and genetic profiling studies comparing in situ, invasive, and metastatic breast carcinomas but have failed to identify tumor stage–specific gene signatures (18–22). However, these studies have focused mainly on the tumor epithelial cells, while the potential involvement of other epithelial

and myoepithelial cells and the stroma in tumor progression have not been explored in sufficient depth.

Epithelial-mesenchymal interactions are known to be important for the normal development of the mammary gland and for breast tumorigenesis (23). In vivo and in vitro studies have demonstrated that cells composing the microenvironment (myoepithelial and endothelial cells, fibroblasts, myofibroblasts, leukocytes, and other cell types) and the ECM molecules modulate tissue specificity of the normal breast as well as the growth, survival, polarity, and invasive behavior of breast cancer cells (24, 25). The observations that deletion of the type II TGF-β receptor in fibroblasts in mice (26, 27) and carcinogen treatment of mammary fat pad stroma in rats (28) promote tumor initiation and progression highlight the importance of stromal influences on tumorigenesis. Human epidemiological data linking chronic inflammation to increased and prolonged use of NSAIDs to decreased cancer risk (29) also implicate microenvironmental changes in tumorigenesis as inflammation is primarily a stromal reaction. Among all leukocytes, macrophages have been the most extensively analyzed in breast tumor progression. Using various model systems, macrophages have been shown to play a role in promoting angiogenesis, invasion, and metastatic spread (30, 31). Most importantly, elimination of

**Figure 2**
Hypothetical model of breast tumor progression. Schematic view of normal, in situ, invasive, and metastatic carcinoma progression. Normal breast ducts are composed of the basement membrane and a layer of luminal epithelial and myoepithelial cells. Cells composing the stroma include various leukocytes, fibroblasts, myofibroblasts, and endothelial cells. In in situ carcinomas the myoepithelial cells are epigenetically and phenotypically altered and their number decreases, potentially due to degradation of the basement membrane. At the same time, the number of stromal fibroblasts, myofibroblasts, lymphocytes, and endothelial cells increases. Loss of myoepithelial cells and basement membrane results in invasive carcinomas, in which tumor cells can invade surrounding tissues and can migrate to distant organs, eventually leading to metastases.

tumor-associated macrophages using a vaccine-based approach significantly decreased tumor growth and progression of breast and other carcinomas, demonstrating that targeting the tumor microenvironment may be a feasible therapeutic strategy (32, 33).

To characterize the molecular changes that occur in the microenvironment and to explore their potential role in tumor progression, a recent study described the purification of all major cell types from normal human breast tissue, DCIS, and invasive carcinomas and the analysis of their comprehensive gene expression and genetic profiles using serial analysis of gene expression (SAGE) and SNP arrays, respectively (34). Using these approaches, dramatic gene expression changes were found in all cell types during breast tumor progression, whereas clonally selected genetic alterations were limited to tumor epithelial cells. This latter result contradicts findings of several other labs describing frequent genetic alterations in breast tumor stroma (35–41). The discrepancy is likely to be due to the use of different experimental approaches, and resolving it would be important for clarifying the involvement of stromal cells in tumorigenesis. A significant fraction of genes identified by Allinen and colleagues (34) as abnormally expressed in tumor epithelial and stromal cells encode secreted proteins and receptors, implicating a role for abnormal autocrine/paracrine signaling in breast tumor progression. Correlating with this, several of the chemokines (e.g. CXC motif chemokine ligand 12 [CXCL12] and CXCL14) overexpressed in tumor myoepithelial cells and myofibroblasts, respectively, enhance tumor cell proliferation, migration, and invasion and promote angiogenesis and metastatic spread (34, 42, 43). Several other chemokines including CXCL1, IL-8, CCL5/RANTES, and MCP1 have also been implicated in breast tumorigenesis (44, 45). Because chemokine receptors are G protein–coupled receptors that are fairly good drug targets, efforts are ongoing to test whether the inhibition of chemokine receptors could potentially be exploited for the treatment of breast and other cancer types (46).

Some of the microenvironmental changes in tumors appear to be permanent, since normal and tumor-associated stromal cells maintain their differences even after their removal from the patients and following in vitro culture (24, 42, 47, 48), raising the possibility of underlying heritable alterations such as epigenetic modifications (e.g., DNA methylation and chromatin changes). To explore this possibility Hu et al. analyzed the comprehensive DNA methylation profiles of epithelial and myoepithelial cells and fibroblasts isolated from normal and neoplastic breast tissue using a new method — methylation-specific digital karyotyping (49). Numerous genes differentially methylated between normal and cancer tissue were identified in each cell type, demonstrating that epigenetic alterations may be in part responsible for changes in the tumor microenvironment. Interestingly, transcription factors with known developmental function, including *Homeobox* genes, appeared to be preferentially affected by DNA methylation in each cell type, suggesting that differences in methylation between the normal and cancer-associated microenvironment may reflect abnormal differentiation of the cells in tumors. Similar to the changes observed in tumor epithelial cells, epigenetic changes observed in the stroma may also be different depending on the tumor subtype, suggesting co-evolution of tumor epithelial and stromal compartments. For example the DNA methylation status of progesterone receptor (PR), 17-β-hydroxysteroid dehydrogenase IV (HSD17B4), and cadherin 13 (CDH13) in epithelial and stromal cells is very distinct between HER2+- and luminal A–type tumors (50).

### Key tumor progression steps: invasion and metastasis
Critical and poorly understood events in breast tumor progression that have dramatic effects on clinical management and outcome
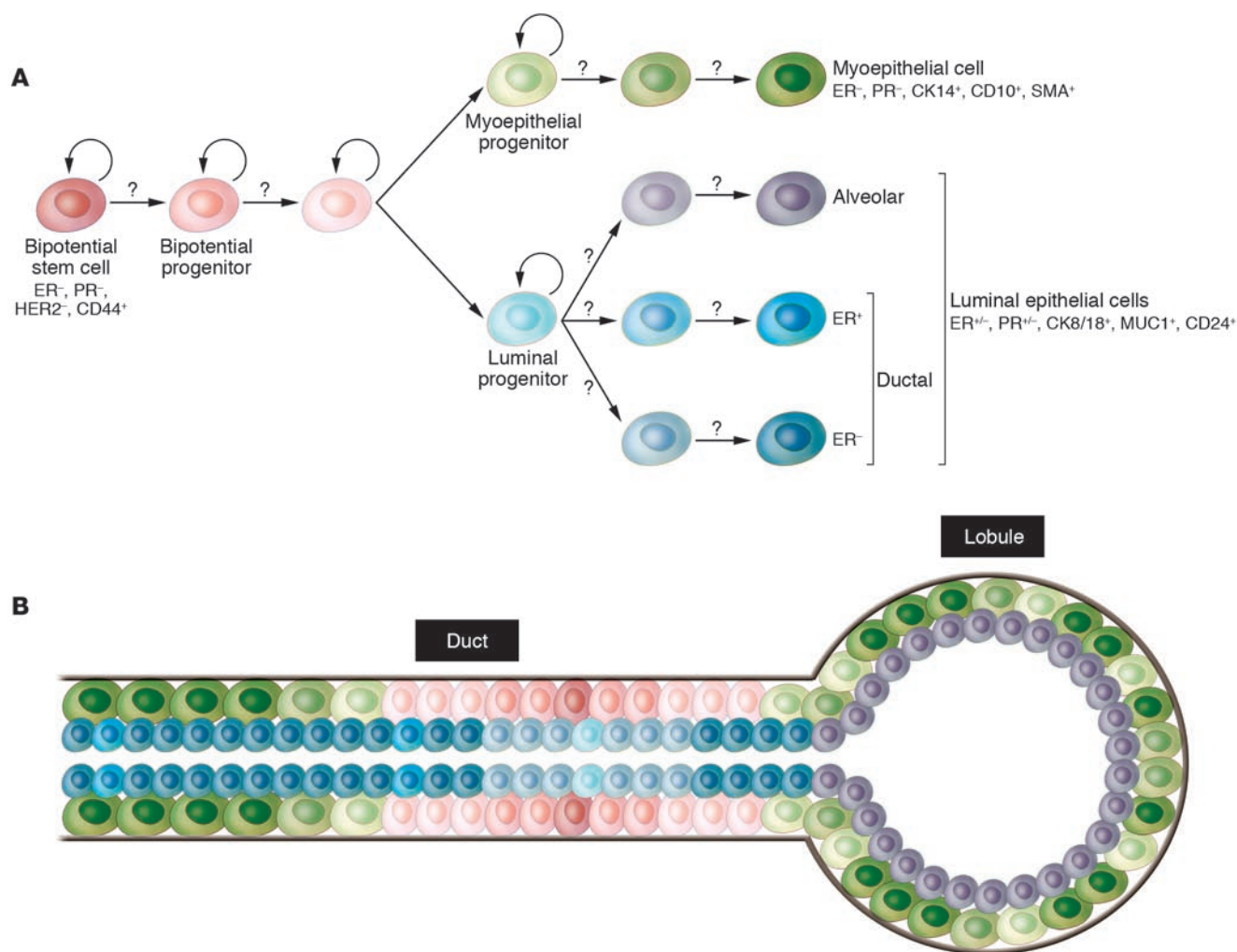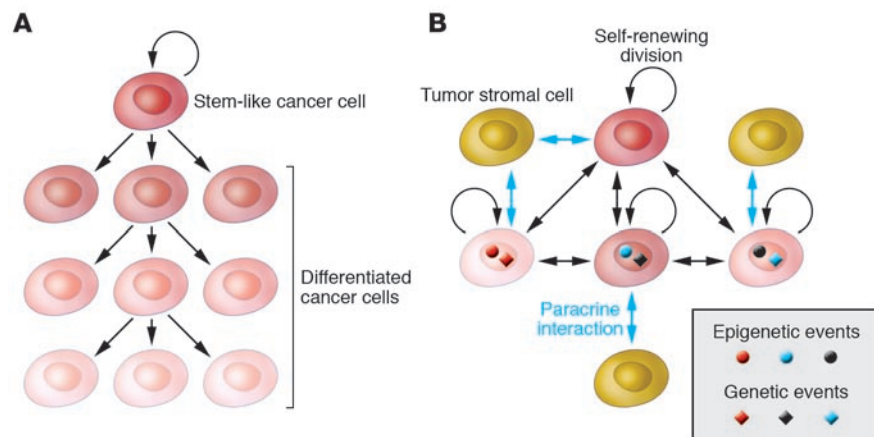
**Figure 3**
Hypothetical model of human mammary epithelial stem cell hierarchy and differentiation. (**A**) Hypothetical depiction of mammary epithelial stem cells and their various progeny. A bipotential stem cell gives rise to luminal epithelial and myoepithelial cells, but the intermediary steps and their regulation are largely unknown (question marks). The model is likely to oversimplify the real situation, since there are many different types of luminal epithelial cells and both the myoepithelial and luminal cells are likely different in the ducts and alveoli. (**B**) Schematic picture of a normal terminal duct lobular unit with the putative location of the various stem and differentiated cells indicated. Gray line denotes the basement membrane; color of cell types correlates with that in **A**. CK14, cytokeratin 14; MUC1, mucin 1.

are the transition of DCIS to invasive carcinoma and the metastatic spread of primary tumors to distant organs. Thus far no genetic events have been identified that underlie these transitions. However, the diagnostic criterion that distinguishes invasive from in situ carcinomas is the disappearance of the myoepithelial cell layer as an organized entity (51). Molecular studies revealed that myoepithelial cells associated with DCIS are not phenotypically normal; they have lost some of their differentiation markers and have upregulated genes promoting angiogenesis and invasion (34). While the physiological relevance of these molecular changes is unknown, these data lead to the hypothesis that the in situ to invasive carcinoma progression may be regulated by myoepithelial cells (52–54). Specifically, abnormal DCIS-associated myoepithelial cells together with various stromal cells may degrade the basement membrane, resulting in the progression of in situ carcinomas to invasive tumors. Correlating with this hypothesis, DCIS-associated myoepithelial cells have increased levels

of ECM-degrading enzymes, such as several MMPs, compared with their normal counterparts (34). An alternative explanation of the histopathologic observations is that the differentiation of mammary epithelial stem cells to myoepithelial cells may be progressively lost during tumorigenesis, resulting in the disappearance of the myoepithelial cell layer and progression to invasive cancer.

Similarly, microenvironmental factors may also play a key role in metastatic progression. Observations obtained in several experimental systems support this hypothesis. In mice and rats, lactation- and involution-induced changes in the ECM of the mammary fat pad promote breast cancer invasion and metastatic progression (55–57). The mammary gland tissue architecture undergoes dramatic remodeling during pregnancy and subsequent involution, during which numerous genes associated with invasion and ECM degradation are significantly upregulated, potentially promoting the spread of cancer cells. Thus pregnancy- and lactation-associat-

**A**

Stem-like cancer cell

Differentiated cancer cells

**B**

Self-renewing division

Tumor stromal cell

Paracrine interaction

Epigenetic events

Genetic events

**Figure 4**

Hypothetical models explaining intratumoral heterogeneity and tumor evolution. (**A**) Based on the cancer stem cell hypothesis, differentiated cancer cells are progeny of cancer stem cells and they are not able to undergo self-renewing cell division. Thus, only the cancer stem cell can accumulate additional genetic changes that can drive tumor progression and drug resistance. (**B**) Based on the clonal evolution model, tumor cell phenotypes are determined based on the combination of cell type of origin of the tumor-initiating cell, acquired genetic and epigenetic alterations, and paracrine signals from surrounding cells. Cellular phenotypes are not stable and can change as the tumor evolves. All tumor cells have the capacity to undergo self-renewing division; thus they all have the potential to contribute to tumor progression and drug resistance. The two models do not have to be mutually exclusive, and their combination (e.g., clonal evolution of cancer stem cells) is also plausible.

protein acidic cysteine–rich (SPARC) (65). Importantly, the expression of these genes was also increased in human primary invasive breast carcinomas that developed distant metastases to the same organs as predicted by the xenograft model (64–66). In addition, in human breast tumors a "wound-response" gene expression signature characteristic of activated fibroblasts is predictive of the risk of distant metastasis and shorter survival (67), further emphasizing the role of the microenvironment in tumor progression.

Further supporting evidence for the role of stromal cells in promoting metastatic progression comes from the finding that the genotype of the host influences the risk of metastasis (68), although this may not be entirely independent of the influence of the host on the tumor phenotype itself. Specifically, African-American women, and even more noticeably, women in Africa, more frequently develop basal-like breast cancer than women of European decent, which is associated with higher risk of brain and visceral metastases than other subtypes (69–72).

ed changes in the ECM may potentially explain the worse prognosis and increased frequency of distant metastases of breast tumors diagnosed during or shortly after pregnancy in women (58).

In xenograft models of breast cancer, co-injection of various fibroblasts or mesenchymal stem cells with cancer cells promotes tumor growth (59) and metastasis (60). This tumor-promoting effect of coinjected stromal cells can be explained by several potential mechanisms. For example, the fibroblasts can become activated by the tumor cells and may secrete angiogenic and invasion-promoting factors. In support of this hypothesis are the findings that irradiated, senescent, cancer-associated, or inflammatory fibroblasts promote tumor growth more effectively than normal fibroblasts (61–63). Alternatively, fibroblasts can modify tumor epithelial cells, for example by possibly upregulating the expression of chemokines and thus creating a paracrine positive-feedback loop. Most likely, each of these possible interactions occurs in all models, but their relative importance and the specific genes involved could depend on the particular cancer and stromal cell type used for the experiments.

The importance of paracrine and autocrine factors in promoting metastasis is also supported by results obtained in a xenograft model of breast cancer progression. Massague and colleagues have been characterizing derivatives of the MDA-MB-231 breast cancer cell line that develop distant metastases at high frequency in particular organs such as bone, adrenal gland, or lung (64–66). Many of the genes overexpressed in these highly metastatic variants encode for secreted or cell surface proteins. For example, the expression of IL-11, CXC chemokine receptor 4, and connective tissue growth (CTGF) is highly elevated in cells metastatic to the bone (64), whereas MDA-MB-231 derivatives that preferentially metastasize to the lungs have increased levels of CXCL1, MMP1, MM2, IL-13Ra2, VCAM1, and secreted

## Mammary epithelial stem cells

The mammary gland is a unique organ that undergoes extensive remodeling and differentiation even in adults. Hormonal changes in each menstrual cycle induce waves of proliferation in the mammary epithelium, whereas pregnancy leads to extensive ductal branching and alveogenesis (73). Based on these observations the existence of normal human adult mammary epithelial stem cells has been proposed, but the cellular identity and molecular characteristics of these cells have not yet been defined. Stem cells are defined as cells with the capacity for self renewal and the ability to give rise to multiple differentiated cell types. Recent in vitro and in vivo clonality and implantation studies have demonstrated the existence of cells with stem cell–like properties in normal human and mouse breast tissue that are required for normal mammary gland development (74–84). However, the molecular mechanisms underlying the maintenance and differentiation of normal mammary stem cells are largely unknown. Normal mammary epithelial stem cells are thought to give rise to luminal epithelial (both estrogen receptor–positive and –negative) and myoepithelial cells, but the regulators of these different differentiation pathways remain elusive (81, 85) (Figure 3A).

In the mouse, numerous experiments demonstrated the existence of progenitor cells in all stages of mammary gland development that can regenerate a functional mammary gland when transplanted into the cleared fat pad (i.e., endogenous mammary epithelium is removed) of a recipient mouse (86–88). Using retrovirus-infected cells and, more recently, cells purified using specific cell surface markers (Lin$^-$CD29$^{hi}$CD24$^+$), it was also shown that a single cell is able to regenerate the entire mammary gland (83, 84, 89). Bromodeoxyuridine and tritiated thymidine pulse-chase experiments have also demonstrated that a subset of mouse mammary epithelial cells retain DNA label (90, 91),

a property of stem cells in various tissues, indicating that they undergo rare and asymmetric cell divisions.

In the human breast, putative mammary epithelial progenitors have been identified using in vitro clonogenicity assays as a measure of "stemness". Specifically, several studies have shown that a subset of human mammary epithelial cells are able to form colonies in vitro and give rise to both luminal epithelial and myoepithelial cells; thus they appear to represent bipotential mammary epithelial progenitors (80, 92–101). Similar to other tissue types, culturing mammary epithelial cells in mammospheres enriched for progenitors, and the gene expression profiles of the cells, also correlated with stem cell–like function (80). Most recently Ole Petersen and colleagues described evidence that the most primitive mammary epithelial stem cells are located in the ducts and not in the end buds as suggested in mice (102) (Figure 3B).

Only a few studies report injecting these cells into mammary fat pads (humanized or just cleared) of immunodeficient mice in an effort to generate ductal outgrowth, but even in these cases the resulting branching structures were not well characterized and the assay was performed using fairly large numbers of cells (75). The culture conditions and the markers the authors found to be expressed in the various progenitors varied in the different reports. Most studies agree that cytokeratin 8 (CK8) and CK18 are luminal markers, while CK5 is expressed in progenitors. However, mucin 1 (MUC1) was described as present on both luminal and progenitor cells and CD10 was described as present on progenitor and myoepithelial cells. Similarly contradictory results have been reported for epithelial cellular adhesion molecule (ESA), SMA, CK14, and integrin α6 (ITGA6). All these studies were performed in vitro, and the properties and gene expression patterns of progenitor-like cells can be very distinct in vitro and in vivo. Thus experiments performed in vitro have to be interpreted with caution and have to be validated in vivo using physiological conditions. Furthermore, additional markers are necessary for the more accurate definition of stem cells, lineage-committed progenitors, and terminally differentiated luminal epithelial and myoepithelial cells.

The development and function of the mammary gland is under hormonal control, and a subset of mammary epithelial cells express ERs and PRs both in the mouse mammary gland and in the human breast (103, 104). Contrary to ER+ breast cancer cells, these normal ER+ cells do not proliferate in response to estrogen, but cells surrounding normal ER+ cells are frequently proliferating, indicating that paracrine factors may mediate the mitogenic effects of estrogen in the normal mammary epithelium (104). Lifetime accumulated hormonal exposure is one of the most important risk factors for breast cancer. Since stem cells are thought to be the cellular targets of transforming events, the hormone responsiveness of putative mammary epithelial stem cells has been extensively investigated. However, in both human and mouse there have been conflicting reports regarding the hormone receptor status of the mammary epithelial stem cells. In the mouse, based on pulse-chase experiments, the putative mouse mammary epithelial stem cells were predicted to be ER+ (105), while stem cells purified using specific cell surface markers were hormone receptor–negative (both ER and PR) and HER2− (82). One potential explanation of this differing result could be that for the pulse-chase experiment mice were injected daily with fairly large amounts of exogenous estrogen that might have led to the upregulation of ER or induced proliferation of ER+ cells. Similarly conflicting results have been described for the human mammary epithelial progenitors — some

studies reported these cells as hormone receptor negative, while others reported them as ER+ cells (93, 97). Thus resolving this issue requires further studies in both species.

In addition to using cells from freshly purified human or mouse breast epithelium, several cell lines with progenitor properties have also been described. The COMMA-Dβ mouse mammary epithelial cell line has properties of stem cells and is able to generate outgrowths that appear to differentiate into functional mammary ducts and alveoli (106). Similarly, the MCF10ADCIS.com cells appear to have properties of bipotential progenitors, since they are able to generate both luminal and myoepithelial cell lineages when injected into immunodeficient mice (107). The recently developed MCF-15, HMT348, and EM-G3 breast epithelial cell lines also possess characteristics of bipotential progenitor cells, although this was defined only on the basis of in vitro expression of certain markers (108–110). Although the use of cell lines has its limitations, they could be useful for various functional studies that are difficult to perform on primary cells as long as findings are validated by subsequent in vivo studies using primary cells or human tissue samples.

## Tumor heterogeneity and evolution

There is a current controversy regarding tumor heterogeneity, progression, and therapeutic resistance. The two competing, although not necessarily mutually exclusive, models are the cancer stem cell and the clonal evolution and selection hypotheses (Figure 4). Although both models agree that tumors originate from a single cell that acquired multiple mutations and unlimited proliferative potential, important differences exist between the models. Specifically, the two models explain tumor heterogeneity with different mechanisms: either as a program of aberrant differentiation or as a competition among tumor cells with different phenotypes. They define the target cells of the tumor-initiating transforming event differently: either as normal stem cells or not a specific normal cell in particular. They identify different subpopulations of tumor cells as drivers of tumor progression: either a small pool of cells, the "cancer stem cells," or any tumor cell with selective advantage. Finally, they explain therapeutic resistance differently: either cancer stem cells are inherently drug resistant or therapy selects for resistant clones. The most important issues related to the two models are the need to (a) determine which model correlates with clinical behavior in patients, (b) identify new approaches for the therapeutic targeting of distinct cell populations, and (c) identify and characterize regulators of the stem cell–like and more differentiated tumor cell phenotypes and (d) determine whether these represent stable or plastic characteristics.

Putative stem and differentiated epithelial cells have been purified from cancerous and normal breast tissue using cell surface markers CD24 and CD44 associated with stem cell–like properties (111) and their gene expression and genetic profiles analyzed (112, 113). The gene expression profiling data indicated that CD44+ cells are indeed more stem cell like and less differentiated than CD24+ cells based on their expression of various stem and differentiated cell markers. An apparent decrease in the number of cells expressing CD44+ cell markers was detected in normal breast tissue from women at late pregnancy stage, correlating with the predicted decline of progenitors due to pregnancy-induced differentiation (114). In addition, comparison of the gene expression profiles of CD44+ and CD24+ cells isolated from normal or neoplastic breast tissue revealed that cells of the same type (e.g.,

normal CD44$^+$ cells and cancer CD44$^+$ cells) are more similar to each other than to cells of the same tissue (e.g., normal CD44$^+$ cells and normal CD24$^+$ cells).

Several signaling pathways were identified to be specifically active in CD44$^+$ stem cell–like breast cancer cells, including TGF-β, WNT, hedgehog (Hh), integrin, urinary plasminogen activator (uPA), and ephrin signaling, and it was demonstrated that the inhibition TGF-β signaling induced a more epithelial phenotype (112). CD44$^+$ cells were also found to be more invasive, proliferative, and angiogenic, predicting aggressive tumor cell behavior (112). Correlating with this, analysis of a gene signature characteristic of CD44$^+$ cells in microarray data of primary lymph node–negative invasive tumors identified groups of patients with shorter distant metastasis-free survival (112, 113). However, the frequency of cells expressing CD24$^+$ cell markers increased in distant metastases compared with matched primary tumors. Most importantly, SNP array and FISH analyses determined that CD24$^+$ (more differentiated) and CD44$^+$ (putative stem) cells from individual tumors were clonally related but not always identical, a result that contradicts the strictest interpretation of the cancer stem cell hypothesis and supports the clonal evolution model or the combination of the two models instead. Specifically, it is possible that the cancer stem cell itself changes and undergoes clonal evolution during tumor progression or following therapeutic interventions. In this respect the stem cell–like phenotype can be one of the characteristics driving selection or associated with other genetic or epigenetic alterations that confer growth and survival advantage for the tumor cell.

Tumor progression can be viewed as an evolution of a large population of genetically and epigenetically distinct and unstable individuals in an ever-changing environment (114). From this point of view, it is statistically more likely that all or the majority of tumor cells have the potential to evolve and the dominant clone can be different at different stages of tumor progression as well as before and after treatment. This hypothesis seemingly contradicts findings that the gene expression profiles of primary invasive tumors and their matched metastases (lymph node or distant) are more similar to each other than to other tumors of the same clinical and pathologic stage (3–6). However, these studies were conducted using bulk tumor samples and array-based platforms; both of which have several drawbacks that could confound the results. This model could also explain why gene expression signatures in primary tumors can predict the risk of distant metastases and why we have not been able to define tumor stage–specific signatures and genetic events underlying invasion and metastatic spread. The development and application of technologies that allow the comprehensive molecular profiling of individual cells in situ would be necessary to answer many of the lingering questions in breast tumor biology. Based on the advances made in the past decade, it will not be too long before this happens.

## Acknowledgments

Address correspondence to: Kornelia Polyak, Department of Medical Oncology, Dana-Farber Cancer Institute, and Department of Medicine, Harvard Medical School and Brigham and Women's Hospital, 44 Binney Street, Boston, Massachusetts 02115, USA. Phone: (617) 632-2106; Fax: (617) 582-8490; E-mail: kornelia_polyak@dfci.harvard.edu.

1. Kamangar, F., Dores, G.M., and Anderson, W.F. 2006. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J. Clin. Oncol.* **24**:2137–2150.
2. Perou, C.M., et al. 2000. Molecular portraits of human breast tumours. *Nature.* **406**:747–752.
3. Sorlie, T., et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**:10869–10874.
4. Hu, Z., et al. 2006. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics.* **7**:96.
5. Sorlie, T., et al. 2006. Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. *BMC Genomics.* **7**:127.
6. Yu, K., Lee, C.H., Tan, P.H., and Tan, P. 2004. Conservation of breast cancer molecular subtypes and transcriptional patterns of tumor progression across distinct ethnic populations. *Clin. Cancer Res.* **10**:5508–5517.
7. Cox, A., et al. 2007. A common coding variant in CASP8 is associated with breast cancer risk. *Nat. Genet.* **39**:352–358.
8. Easton, D.F., et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature.* **447**:1087–1093.
9. Hunter, D.J., et al. 2007. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.* **39**:870–874.
10. Stacey, S.N., et al. 2007. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* **39**:865–869.
11. Futreal, P.A., et al. 2004. A census of human cancer genes. *Nat. Rev. Cancer.* **4**:177–183.
12. Bardelli, A., and Velculescu, V.E. 2005. Mutational analysis of gene families in human cancer. *Curr. Opin. Genet. Dev.* **15**:5–12.
13. Greenman, C., et al. 2007. Patterns of somatic mutation in human cancer genomes. *Nature.* **446**:153–158.
14. Sjoblom, T., et al. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science.* **314**:268–274.
15. Allred, D.C., Mohsin, S.K., and Fuqua, S.A. 2001. Histological and biological evolution of human premalignant breast disease. *Endocr. Relat. Cancer.* **8**:47–61.
16. Burstein, H.J., Polyak, K., Wong, J.S., Lester, S.C., and Kaelin, C.M. 2004. Ductal carcinoma in situ of the breast. *N. Engl. J. Med.* **350**:1430–1441.
17. Simpson, P.T., Reis-Filho, J.S., Gale, T., and Lakhani, S.R. 2005. Molecular evolution of breast cancer. *J. Pathol.* **205**:248–254.
18. Chin, K., et al. 2004. In situ analyses of genome instability in breast cancer. *Nat. Genet.* **36**:984–988.
19. Porter, D., et al. 2003. Molecular markers in ductal carcinoma in situ of the breast. *Mol. Cancer Res.* **1**:362–375.
20. Porter, D.A., et al. 2001. A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res.* **61**:5697–5702.
21. Ma, X.J., et al. 2003. Gene expression profiles of human breast cancer progression. *Proc. Natl. Acad. Sci. U. S. A.* **100**:5974–5979.
22. Yao, J., et al. 2006. Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. *Cancer Res.* **66**:4065–4078.
23. Howlett, A.R., and Bissell, M.J. 1993. The influence of tissue microenvironment (stroma and extracellular matrix) on the development and function of mammary epithelium. *Epithelial Cell Biol.* **2**:79–89.
24. Bissell, M.J., and Radisky, D. 2001. Putting tumours in context. *Nat. Rev. Cancer.* **1**:46–54.
25. Elenbaas, B., and Weinberg, R.A. 2001. Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. *Exp. Cell Res.* **264**:169–184.
26. Cheng, N., et al. 2005. Loss of TGF-beta type II receptor in fibroblasts promotes mammary carcinoma growth and invasion through upregulation of TGF-alpha-, MSP-and HGF-mediated signaling networks. *Oncogene.* **24**:5053–5068.
27. Bhowmick, N.A., et al. 2004. TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science.* **303**:848–851.
28. Maffini, M.V., et al. 2004. The stroma as a crucial target in rat mammary gland carcinogenesis. *J. Cell Sci.* **117**:1495–1502.
29. Ulrich, C.M., Bigler, J., and Potter, J.D. 2006. Non-steroidal anti-inflammatory drugs for cancer prevention: promise, perils and pharmacogenetics. *Nat. Rev. Cancer.* **6**:130–140.
30. Lin, E.Y., and Pollard, J.W. 2007. Tumor-associated macrophages press the angiogenic switch in breast cancer. *Cancer Res.* **67**:5064–5066.
31. Lewis, C.E., and Hughes, R. 2007. Inflammation and breast cancer. Microenvironmental factors reg-

ulating macrophage function in breast tumours: hypoxia and angiopoietin-2. *Breast Cancer Res.* **9**:209.

32. Luo, Y., et al. 2006. Targeting tumor-associated macrophages as a novel strategy against breast cancer. *J. Clin. Invest.* **116**:2132–2141. doi:10.1172/JCI27648.

33. Wu, W., et al. 2006. Targeting cell-impermeable prodrug activation to tumor microenvironment eradicates multiple drug-resistant neoplasms. *Cancer Res.* **66**:970–980.

34. Allinen, M., et al. 2004. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell.* **6**:17–32.

35. Moinfar, F., et al. 2000. Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. *Cancer Res.* **60**:2562–2566.

36. Fukino, K., et al. 2004. Combined total genome loss of heterozygosity scan of breast cancer stroma and epithelium reveals multiplicity of stromal targets. *Cancer Res.* **64**:7231–7236.

37. Kurose, K., et al. 2002. Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. *Nat. Genet.* **32**:355–357.

38. Kurose, K., et al. 2001. Genetic model of multi-step breast carcinogenesis involving the epithelium and stroma: clues to tumour-microenvironment interactions. *Hum. Mol. Genet.* **10**:1907–1913.

39. Weber, F., et al. 2005. Variability in organ-specific EGFR mutational spectra in tumour epithelium and stroma may be the biological basis for differential responses to tyrosine kinase inhibitors. *Br. J. Cancer.* **92**:1922–1926.

40. Weber, F., et al. 2006. Total-genome analysis of BRCA1/2-related invasive carcinomas of the breast identifies tumor stroma as potential landscaper for neoplastic initiation. *Am. J. Hum. Genet.* **78**:961–972.

41. Fukino, K., Shen, L., Patocs, A., Mutter, G.L., and Eng, C. 2007. Genomic instability within tumor stroma and clinicopathological characteristics of sporadic primary invasive breast carcinoma. *JAMA.* **297**:2103–2111.

42. Orimo, A., et al. 2005. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell.* **121**:335–348.

43. Muller, A., et al. 2001. Involvement of chemokine receptors in breast cancer metastasis. *Nature.* **410**:50–56.

44. Ben-Baruch, A. 2003. Host microenvironment in breast cancer development: inflammatory cells, cytokines and chemokines in breast cancer progression: reciprocal tumor-microenvironment interactions. *Breast Cancer Res.* **5**:31–36.

45. Zlotnik, A. 2006. Chemokines and cancer. *Int. J. Cancer.* **119**:2026–2029.

46. Taborga, M., Corcoran, K.E., Fernandes, N., Ramkissoon, S.H., and Rameshwar, P. 2007. G-coupled protein receptors and breast cancer progression: potential drug targets. *Mini Rev. Med. Chem.* **7**:245–251.

47. Bissell, M.J., Radisky, D.C., Rizki, A., Weaver, V.M., and Petersen, O.W. 2002. The organizing principle: microenvironmental influences in the normal and malignant breast. *Differentiation.* **70**:537–546.

48. Tlsty, T.D., and Hein, P.W. 2001. Know thy neighbor: stromal cells can contribute oncogenic signals. *Curr. Opin. Genet. Dev.* **11**:54–59.

49. Hu, M., et al. 2005. Distinct epigenetic changes in the stromal cells of breast cancers. *Nat. Genet.* **37**:899–905.

50. Fiegl, H., et al. 2006. Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res.* **66**:29–33.

51. Lerwill, M.F. 2004. Current practical applications of diagnostic immunohistochemistry in breast pathology. *Am. J. Surg. Pathol.* **28**:1076–1091.

52. Barsky, S.H., and Karlin, N.J. 2005. Myoepithelial cells: autocrine and paracrine suppressors of breast cancer progression. *J. Mammary Gland Biol. Neoplasia.* **10**:249–260.

53. Barsky, S.H. 2003. Myoepithelial mRNA expression profiling reveals a common tumor-suppressor phenotype. *Exp. Mol. Pathol.* **74**:113–122.

54. Polyak, K., and Hu, M. 2005. Do myoepithelial cells hold the key for breast tumor progression? *J. Mammary Gland Biol. Neoplasia.* **10**:231–247.

55. McDaniel, S.M., et al. 2006. Remodeling of the mammary microenvironment after lactation promotes breast tumor cell metastasis. *Am. J. Pathol.* **168**:608–620.

56. Bemis, L.T., and Schedin, P. 2000. Reproductive state of rat mammary gland stroma modulates human breast cancer cell migration and invasion. *Cancer Res.* **60**:3414–3418.

57. Schedin, P. 2006. Pregnancy-associated breast cancer and metastasis. *Nat. Rev. Cancer.* **6**:281–291.

58. Polyak, K. 2006. Pregnancy and breast cancer: the other side of the coin. *Cancer Cell.* **9**:151–153.

59. Zhu, W., et al. 2006. Mesenchymal stem cells derived from bone marrow favor tumor cell growth in vivo. *Exp. Mol. Pathol.* **80**:267–274.

60. Karnoub, A.E., et al. 2007. Bone marrow-derived mesenchymal stem cells promote breast cancer metastasis. *Nature.* **449**:557–563.

61. Barcellos-Hoff, M.H. 1998. The potential influence of radiation-induced microenvironments in neoplastic progression. *J. Mammary Gland Biol. Neoplasia.* **3**:165–175.

62. Krtolica, A., and Campisi, J. 2002. Cancer and aging: a model for the cancer promoting effects of the aging stroma. *Int. J. Biochem. Cell Biol.* **34**:1401–1414.

63. Liu, D., and Hornsby, P.J. 2007. Senescent human fibroblasts increase the early growth of xenograft tumors via matrix metalloproteinase secretion. *Cancer Res.* **67**:3117–3126.

64. Kang, Y., et al. 2003. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell.* **3**:537–549.

65. Minn, A.J., et al. 2005. Genes that mediate breast cancer metastasis to lung. *Nature.* **436**:518–524.

66. Nguyen, D.X., and Massague, J. 2007. Genetic determinants of cancer metastasis. *Nat. Rev. Genet.* **8**:341–352.

67. Chang, H.Y., et al. 2005. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl. Acad. Sci. U. S. A.* **102**:3738–3743.

68. Hunter, K. 2006. Host genetics influence tumour metastasis. *Nat. Rev. Cancer.* **6**:141–146.

69. Millikan, R.C., et al. 2007. Epidemiology of basal-like breast cancer. *Breast Cancer Res. Treat.* In press.

70. Carey, L.A., et al. 2007. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin. Cancer Res.* **13**:2329–2334.

71. Fulford, L.G., et al. 2007. Basal-like grade III invasive ductal carcinoma of the breast: patterns of metastasis and long-term survival. *Breast Cancer Res.* **9**:R4.

72. Minn, A.J., et al. 2007. Lung metastasis genes couple breast tumor size and metastatic spread. *Proc. Natl. Acad. Sci. U. S. A.* **104**:6740–6745.

73. Navarrete, M.A., et al. 2005. Assessment of the proliferative, apoptotic and cellular renovation indices of the human mammary epithelium during the follicular and luteal phases of the menstrual cycle. *Breast Cancer Res.* **7**:R306–R313.

74. Wicha, M.S., Liu, S., and Dontu, G. 2006. Cancer stem cells: an old idea — a paradigm shift. *Cancer Res.* **66**:1883–1890; discussion 1895–1886.

75. Liu, S., et al. 2006. Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malig-

nant human mammary stem cells. *Cancer Res.* **66**:6063–6071.

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

77. Dontu, G., and Wicha, M.S. 2005. Survival of mammary stem cells in suspension culture: implications for stem cell biology and neoplasia. *J. Mammary Gland Biol. Neoplasia.* **10**:75–86.

78. Dontu, G., El-Ashry, D., and Wicha, M.S. 2004. Breast cancer, stem/progenitor cells and the estrogen receptor. *Trends Endocrinol. Metab.* **15**:193–197.

79. Dontu, G., et al. 2004. Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Res* **6**:R605–R615.

80. Dontu, G., et al. 2003. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* **17**:1253–1270.

81. Dontu, G., Al-Hajj, M., Abdallah, W.M., Clarke, M.F., and Wicha, M.S. 2003. Stem cells in normal breast development and breast cancer. *Cell Prolif.* **36**(Suppl. 1):59–72.

82. Asselin-Labat, M.L., et al. 2006. Steroid hormone receptor status of mouse mammary stem cells. *J. Natl. Cancer Inst.* **98**:1011–1014.

83. Stingl, J., et al. 2006. Purification and unique properties of mammary epithelial stem cells. *Nature.* **439**:993–997.

84. Shackleton, M., et al. 2006. Generation of a functional mammary gland from a single stem cell. *Nature.* **439**:84–88.

85. Allred, D.C., Brown, P., and Medina, D. 2004. The origins of estrogen receptor alpha-positive and estrogen receptor alpha-negative human breast cancer. *Breast Cancer Res.* **6**:240–245.

86. Young, L.J., Medina, D., DeOme, K.B., and Daniel, C.W. 1971. The influence of host and tissue age on life span and growth rate of serially transplanted mouse mammary gland. *Exp. Gerontol.* **6**:49–56.

87. Smith, G.H., and Medina, D. 1988. A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. *J. Cell Sci.* **90**:173–183.

88. Wagner, K.U., et al. 2002. An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. *Development.* **129**:1377–1386.

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

90. Smith, G.H. 2005. Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. *Development.* **132**:681–687.

91. Zeps, N., Bentel, J.M., Papadimitriou, J.M., D'Antuono, M.F., and Dawkins, H.J. 1998. Estrogen receptor-negative epithelial cells in mouse mammary gland development and growth. *Differentiation.* **62**:221–226.

92. Stingl, J., Eaves, C.J., Kuusk, U., and Emerman, J.T. 1998. Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. *Differentiation.* **63**:201–213.

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

94. Bocker, W., et al. 2002. Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept. *Lab. Invest.* **82**:737–746.

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

96. Clarke, R.B., Anderson, E., Howell, A., and Potten, C.S. 2003. Regulation of human breast epithelial stem cells. *Cell Prolif.* **36**(Suppl. 1):45–58.

97. Clarke, R.B., et al. 2005. A putative human breast stem cell population is enriched for steroid receptor-positive cells. *Dev. Biol.* **277**:443–456.

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

99. Alvi, A.J., et al. 2003. Functional and molecular characterisation of mammary side population cells. *Breast Cancer Res.* **5**:R1–R8.

100. Boecker, W., and Buerger, H. 2003. Evidence of progenitor cells of glandular and myoepithelial cell lineages in the human adult female breast epithelium: a new progenitor (adult stem) cell concept. *Cell Prolif.* **36**(Suppl. 1):73–84.

101. Gudjonsson, T., et al. 2002. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. *Genes Dev.* **16**:693–706.

102. Villadsen, R., et al. 2007. Evidence for a stem cell hierarchy in the adult human breast. *J. Cell Biol.* **177**:87–101.

103. Petersen, O.W., Hoyer, P.E., and van Deurs, B. 1987. Frequency and distribution of estrogen receptor-positive cells in normal, nonlactating human breast tissue. *Cancer Res.* **47**:5748–5751.

104. Clarke, R.B., Howell, A., Potten, C.S., and Anderson, E. 1997. Dissociation between steroid receptor expression and cell proliferation in the human breast. *Cancer Res.* **57**:4987–4991.

105. Booth, B.W., and Smith, G.H. 2006. Estrogen receptor-alpha and progesterone receptor are expressed in label-retaining mammary epithelial cells that divide asymmetrically and retain their template DNA strands. *Breast Cancer Res.* **8**:R49.

106. Deugnier, M.A., et al. 2006. Isolation of mouse mammary epithelial progenitor cells with basal characteristics from the Comma-Dbeta cell line. *Dev. Biol.* **293**:414–425.

107. Tait, L.R., et al. 2007. Dynamic stromal-epithelial interactions during progression of MCF10DCIS.com xenografts. *Int. J. Cancer.* **120**:2127–2134.

108. Shen, K.C., et al. 2006. Isolation and characterization of a breast progenitor epithelial cell line with robust DNA damage responses. *Breast Cancer Res. Treat.* **98**:357–364.

109. Lewis, C.M., et al. 2006. Telomerase immortalization of human mammary epithelial cells derived from a BRCA2 mutation carrier. *Breast Cancer Res. Treat.* **99**:103–115.

110. Brozova, M., et al. 2007. Establishment, growth and in vivo differentiation of a new clonal human cell line, EM-G3, derived from breast cancer progenitors. *Breast Cancer Res. Treat.* **103**:247–257.

111. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. 2003. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**:3983–3988.

112. Shipitsin, M., et al. 2007. Molecular definition of breast tumor heterogeneity. *Cancer Cell.* **11**:259–273.

113. Liu, R., et al. 2007. The prognostic role of a gene signature from tumorigenic breast cancer cells. *N. Engl. J. Med.* **356**:217–226.

114. Merlo, L.M., Pepper, J.W., Reid, B.J., and Maley, C.C. 2006. Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer.* **6**:924–935.