Cell Stem Cell
Article

BRCA1 Basal-like Breast Cancers Originate from Luminal Epithelial Progenitors and Not from Basal Stem Cells

Gemma Molyneux,[1] Felipe C. Geyer,[1] Fiona-Ann Magnay,[1] Afshan McCarthy,[1] Howard Kendrick,[1] Rachael Natrajan,[1] Alan MacKay,[1] Anita Grigoriadis,[2] Andrew Tutt,[2] Alan Ashworth,[1] Jorge S. Reis-Filho,[1] and Matthew J. Smalley[1,*]
[1]The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK
[2]Breakthrough Breast Cancer Research Unit, Guy's Hospital, King's Health Partners AHSC, London SE1 9RT, UK
*Correspondence: matthew.smalley@icr.ac.uk
DOI 10.1016/j.stem.2010.07.010

SUMMARY

Breast cancers in *BRCA1* mutation carriers frequently have a distinctive basal-like phenotype. It has been suggested that this results from an origin in basal breast epithelial stem cells. Here, we demonstrate that deleting *Brca1* in mouse mammary epithelial luminal progenitors produces tumors that phenocopy human *BRCA1* breast cancers. They also resemble the majority of sporadic basal-like breast tumors. However, directing *Brca1* deficiency to basal cells generates tumors that express molecular markers of basal breast cancers but do not histologically resemble either human *BRCA1* or the majority of sporadic basal-like breast tumors. These findings support a derivation of the majority of human *BRCA1*-associated and sporadic basal-like tumors from luminal progenitors rather than from basal stem cells. They also demonstrate that when target cells for transformation have the potential for phenotypic plasticity, tumor phenotypes may not directly reflect histogenesis. This has important implications for cancer prevention strategies.

INTRODUCTION

The epithelial glandular network of the breast comprises two main cell types, outer basal cells and inner luminal cells. The majority of the basal cells are differentiated myoepithelial cells (Sleeman et al., 2007), but this layer also includes the mammary stem cells (Shackleton et al., 2006; Sleeman et al., 2006; Stingl et al., 2006; Taddei et al., 2008). The luminal cells include differentiated milk-secreting cells (during lactation) but other functional luminal cell types can be found, such as hormone receptor-expressing cells (Sleeman et al., 2007). The luminal layer is also thought to contain a number of different classes of progenitor, including progenitors for the hormone receptor-expressing cells (Booth and Smith, 2006) and secretory cells (Oakes et al., 2008). These different epithelial lineages can be separated by flow cytometric analysis of expression levels of CD24, Sca-1, and CD49f antigens (Kendrick et al., 2008; Sleeman et al., 2007; Stingl et al., 2006). The myoepithelial cells of the basal layer are defined as $CD24^{+/low}$ Sca-1^{-} $CD49f^{low}$, whereas the basal stem cells are enriched in the $CD24^{+/low}$ Sca-1^{-} $CD49f^{high}$ cell fraction. Luminal estrogen receptor-positive (ER^{+}) cells have the phenotype $CD24^{+/high}$ Sca-1^{+} whereas luminal $CD24^{+/high}$ Sca-1^{-} cells are largely ER^{-} progenitors.

Tissue stem cells are good candidates for the cells of origin of tumors, because their long in vivo lifespan means they can accumulate the mutations required for tumor formation (Smalley and Ashworth, 2003). It is also possible that progenitors downstream of stem cells could accumulate enough genetic lesions to generate a tumor (Dontu et al., 2003), depending on their rate of turnover. In breast cancer, it is proposed that different tumor subtypes (Sorlie et al., 2003) may originate from different classes of stem/progenitor cells, with ER^{-} "basal-like" breast cancers originating from ER^{-} basal stem cells and ER^{+} luminal breast cancers from progenitors committed to luminal ER^{+} differentiation (Dontu et al., 2003; Melchor and Benítez, 2008). Strikingly, the majority of breast tumors arising in carriers of germline mutations in *BRCA1* have the basal-like phenotype (Lakhani et al., 2005; Palacios et al., 2005). It has therefore been proposed that loss of *BRCA1* function in basal stem cells results in tumor formation associated with a block in luminal differentiation (Foulkes, 2004; Liu et al., 2008; Vassilopoulos et al., 2008). In contrast, recent work showed an increase in luminal progenitor numbers in breast tissue of *BRCA1* mutation carriers and a correlation between the gene expression profile of normal human luminal progenitors and basal-like breast cancers (Lim et al., 2009). However, despite the association of aberrant luminal progenitors with *BRCA1* mutations, an origin of *BRCA1* breast cancers in luminal progenitors was not directly demonstrated. To achieve this, a direct comparison of the effects of creating identical tumor predisposing events in stem versus progenitor cells is required. Here we describe such a comparison. We have directed *Brca1* loss of function either predominantly to basal or luminal mammary epithelial cells, thereby biasing tumor origin to basal stem cells, or to progenitors of the luminal layer. In both systems, the molecular features of the tumors that arise are similar to those of human basal-like tumors. However, only tumors originating in luminal progenitors have the same histological appearance as human *BRCA1* tumors and indeed resemble the majority of human sporadic (nonfamilial) basal-like breast cancers. In contrast, *Brca1* loss in basal stem cells preferentially generated malignant adenomyoepitheliomas, which are rare

tumors in humans, with features consistent with a stem cell origin. Importantly, these results reconcile conflicting data from other investigators, showing that deletion of *Brca1* in stem cells causes accumulation of cells with a very strong basal phenotype (Liu et al., 2008), but only deletion of *Brca1* in luminal progenitors produces tumors resembling human *BRCA1* mutation-associated breast cancers (Lim et al., 2009). Our findings suggest that malignant adenomyoepitheliomas are true stem cell tumors but, given the comparative rarity of this tumor type in humans, also highlight luminal progenitors as the key to understanding much of the etiology of breast cancer. The finding that normal mammary gland stem cells may not be common targets for transformation in the breast has important implications for strategies to treat and prevent breast cancer.

RESULTS

The *Blg* Promoter Is Active in Progenitor Cells with a Luminal Sca-1$^-$ ER$^-$ Phenotype

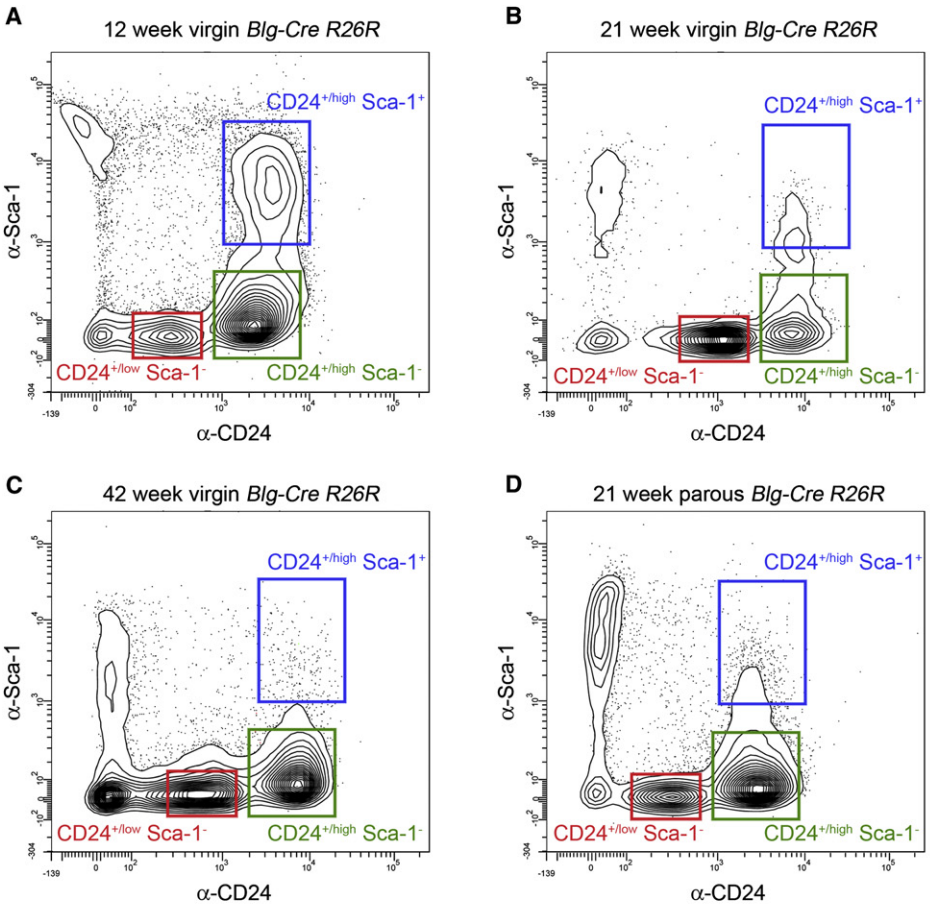
We have previously described a Cre-*loxP*-based mouse model of breast cancer in which Cre expression was under the control of the *Blg* promoter and animals were put through two rounds of pregnancy to drive maximum levels of conditional allele recombination (McCarthy et al., 2007). Mammary tumors developed in these mice but the cells of origin of these tumors were unknown, although the *Blg* promoter is thought to be mainly active in luminal cells (Naylor et al., 2005). We therefore set out to define in detail the cell lineages in which *Blg-Cre*-mediated recombination occurs in the mammary epithelium by using *Rosa26 Reporter* (*R26R*) mice, in which Cre activity permanently switches on β-galactosidase expression.

CD24$^{+/low}$ Sca-1$^-$ basal, CD24$^{+/high}$ Sca-1$^-$ luminal ER$^-$, and CD24$^{+/high}$ Sca-1$^+$ luminal ER$^+$ mammary epithelial cells were isolated from four groups of mice: virgin *Blg-Cre R26R* mice at 12 weeks, 21 weeks, and 42 weeks of age and also mice that had undergone two full-term pregnancies, each followed by 3 weeks of lactation and normal weaning. These animals (hereafter termed "parous" mice) had weaned their last litter by 18 weeks of age and cells were harvested from them at 21 weeks, to allow for postweaning involution and remodeling. In each group, separated cells were cultured as single-cell clones for 10 days in vitro and then stained for β-galactosidase (R26R reporter) activity (Figure 1; Figure S1 available online). The proportion of cells in each population with colony-forming potential (progenitors) that were β-galactosidase$^+$ ("recombined" progenitors) was calculated (see Experimental Procedures). In the 12-week-old virgin mice, the majority of colonies derived from luminal ER$^-$ cells were β-galactosidase positive, whereas very few luminal ER$^+$-derived colonies were positive and no detectable β-galactosidase activity was found in basal-derived colonies. By 21 weeks, almost all virgin luminal ER$^-$ cell-derived colonies were β-galactosidase positive. However, approximately half of colonies derived from luminal ER$^+$ cells were also positive and β-galactosidase activity could be detected in occasional basal cell-derived colonies. By 42 weeks, many positive colonies were seen in the virgin basal cell cultures, although this varied considerably from preparation to preparation. Interestingly, the distribution of β-galactosidase-positive colonies derived from the 21-week parous mice was almost identical to that of the 42-week virgin mice. These results demonstrated that Cre-dependent recombination saturated the luminal ER$^-$ population relatively quickly in virgin animals but also suggested that Cre-expressing progenitors from this population were contributing to estrous cycle-dependent remodeling of the tissue and regeneration of the basal and luminal ER$^+$ cell layers. Pregnancy accelerated the accumulation of recombined cells in the basal population. In all groups, however, the *Blg-Cre* transgene was clearly most strongly active in the luminal ER$^-$ population. Furthermore, it was clear that at all developmental stages, the luminal ER$^-$ population contained a far higher proportion of recombined progenitors than any other population.

Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$ Mice Develop Mammary Tumors that Phenocopy Human *BRCA1* Loss-of-Function Breast Cancers

Although *Blg-Cre* transgene activity resulted in conditional allele recombination in almost all progenitors in the luminal ER$^-$ population, in parous and older virgin animals recombinant alleles could be detected in cells from the basal layer, which contains the mammary stem cells. Therefore, it was still possible that basal stem cells could be the primary targets for *Blg-Cre*-driven tumor formation. To address this, we established cohorts of virgin and parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* mice and monitored them for tumor formation. The results were compared with samples from a cohort of virgin *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* mice, in which Cre expression was driven by the *Keratin 14* promoter. Keratin 14 (K14) is expressed in basal cells in the mouse mammary epithelium (Sleeman et al., 2006), although in humans, K14 can also be expressed in luminal cells (Gusterson et al., 2005). Importantly, K14 is expressed by human (Pece et al., 2010) and mouse (Jonkers et al., 2001; Sleeman et al., 2007) mammary stem cell populations. We confirmed this by quantitative real-time PCR (qPCR) analysis of *Krt14* gene expression in luminal ER$^-$, luminal ER$^+$, myoepithelial (CD24$^{+/low}$ Sca-1$^-$ CD49f^{low}), and mammary stem cell-enriched (CD24$^{+/low}$ Sca-1$^-$ CD49f^{high}) populations (Sleeman et al., 2007). Interestingly, this demonstrated that *Krt14* gene expression in mammary stem cell-enriched cells was significantly higher (1.6-fold) even than that in myoepithelial cells (Figure S2A).

If recombined cells in the basal population made a significant contribution to tumor formation in the *Blg-Cre* animals, then parity would be expected to accelerate the rate of tumor onset as it accelerates the accumulation of these cells. However, there was no significant difference in age or rate of mammary tumor onset between virgin and parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* mice, with a median latency of 336 and 319 days, respectively. In contrast, onset of *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors was significantly faster than both *Blg-Cre* cohorts (log-rank test, $p < 0.001$), with a median latency of 249 days (Figure 2A). All tumors except three (from the virgin *Blg-Cre* cohort) showed a significant ($p < 0.05$) reduction in *p53* expression compared to spleens from wild-type animals of a similar mixed genetic background (Figure S2B). For *Brca1*, exon 22 expression was decreased in the majority of tumors to a greater extent than exon 19 expression, consistent with deletion of *floxed* exon 22 from the conditional *Brca1* allele. In one tumor (*K14-Cre* 3265), *Brca1* expression levels were increased compared to spleen. The reasons for this are unclear (Figures S2C and S2D).



E

Population	12 week virgin			21 week virgin			42 week virgin			21 week parous		
	Colonies counted (plates)	X-Gal$^+$ colonies % ± SD	% Recombined progenitors	Colonies counted (plates)	X-Gal$^+$ colonies % ± SD	% Recombined progenitors	Colonies counted (plates)	X-Gal$^+$ colonies % ± SD	% Recombined progenitors	Colonies counted (plates)	X-Gal$^+$ colonies % ± SD	% Recombined progenitors
CD24$^{+/low}$ Sca-1$^-$	482 (11)	0.0 ± 0.0	0.0	141 (14)	2.0 ± 3.5	0.2	344 (46)	21.7 ± 27.9	1.7	81 (19)	28.7 ± 29.5	1.3
CD24$^{+/high}$ Sca-1$^-$	709 (11)	76.8 ± 19.1	51.6	497 (11)	98.8 ± 1.0	46.5	1189 (23)	97.0 ± 1.1	52.2	646 (19)	95.1 ± 1.0	33.7
CD24$^{+/high}$ Sca-1$^+$	530 (11)	1.7 ± 0.9	0.8	143 (13)	56.1 ± 34.6	6.4	191 (23)	34.2 ± 8.8	2.9	231 (19)	46.1 ± 6.5	5.8

Figure 1. The *Blg-Cre* Transgene Is Active Primarily in Luminal ER$^-$ Cells

(A–D) Flow cytometry of cells from virgin *Blg-Cre R26R* mice at 12 weeks (A), 21 weeks (B), and 42 weeks (C) and from 21-week parous mice (D) indicating the CD24$^{+/low}$ Sca-1$^-$ basal (red), CD24$^{+/high}$ Sca-1$^-$ luminal ER$^-$ (green), and CD24$^{+/high}$ Sca-1$^+$ luminal ER$^+$ (blue) regions used to isolate cells for in vitro colony-forming assays.

(E) Results of the colony-forming assays. Each data set is from three independent isolations of each population. The total number of colonies counted, the total number of 96-well plates analyzed, the mean percentage (±standard deviation; SD) of β-galactosidase$^+$ colonies, and the proportion of recombined progenitors within each population at each developmental stage are indicated.

See also Figure S1.

The phenotype of both the virgin and parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors was similar (Figures 2B–2I; the properties of all the tumors are described in Table S1). They were classified mainly as invasive ductal carcinomas of no special type (IDC-NST), typically with central necrosis (Figure 2B; asterisk), and pushing (Figures 2B and 2C; arrows) or mixed borders.

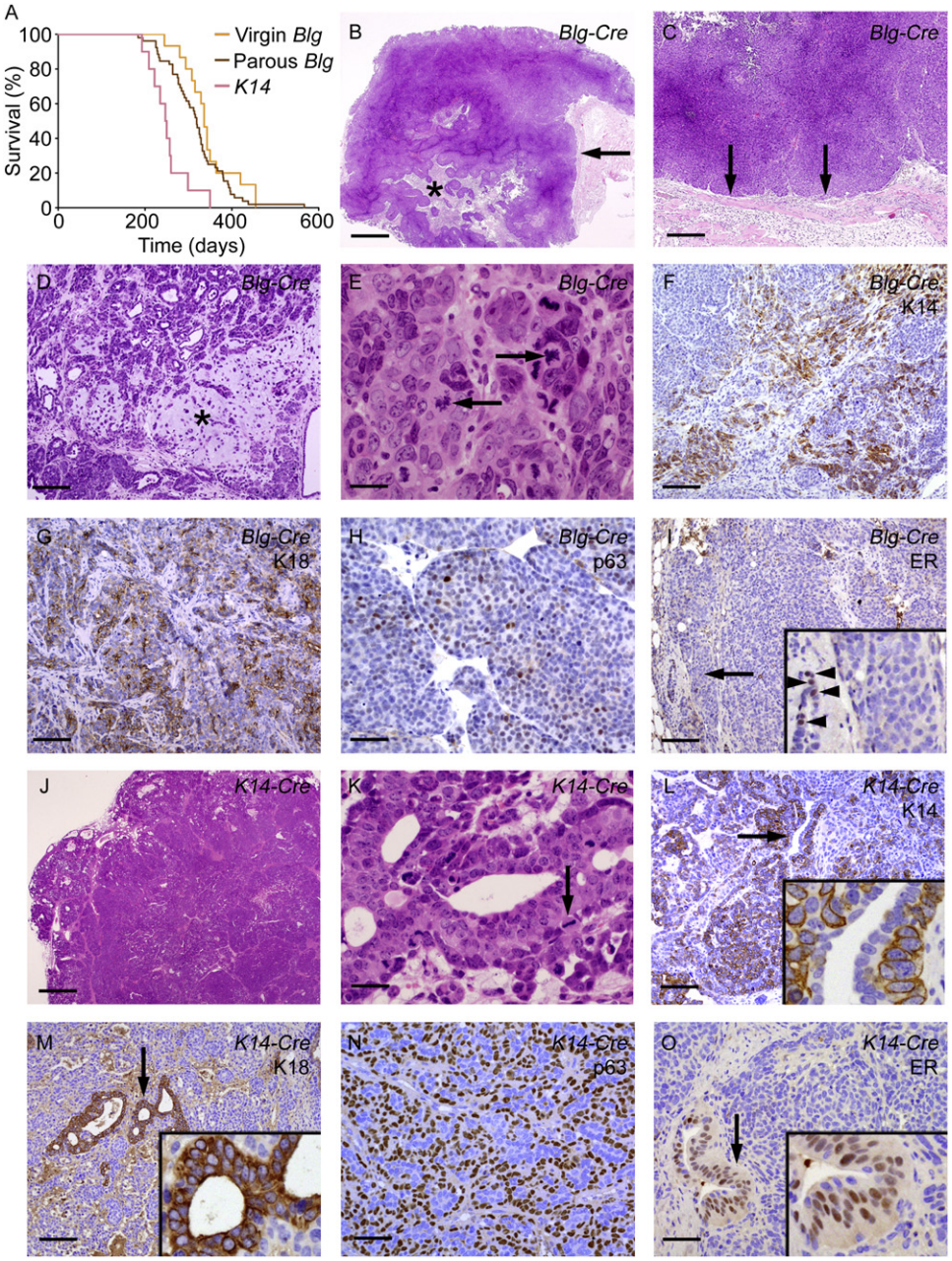


Figure 2. Histological Features of Tumors from Virgin *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* Mice but Not *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* Mice Resemble Those of the Majority of Human *BRCA1* Tumors

(A) Kaplan-Meier tumor latency curve of *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$*, virgin *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$*, and parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* animals. Only data from animals developing mammary tumors is shown. The latency of *K14-Cre* tumors was significantly shorter ($p < 0.001$; Log rank test) than that of either virgin or parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors.

(B–I) Histological features of *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors.

(B) IDC-NST showing central necrosis (asterisk) and pushing margins (arrow). Scale bar represents 0.4 mm.

(C) Tumor with pushing margins (arrows). Scale bar represents 270 μm.

(D) Tumor showing chondroid metaplasia (asterisk). Scale bar represents 110 μm.

(E) High-power image showing aberrant mitotic figures (arrows) and conspicuous nuclear pleomorphism. Scale bar represents 27 μm.

(F) K14 expression in epithelioid and spindloid neoplastic cells.

(G) K18 expression in epithelioid neoplastic cells.

Scale bars in (F) and (G) represent 110 μm.

(H) p63 expression in neoplastic cells. Scale bar represents 55 μm.

(I) ER staining in an ER$^-$ tumor. The arrow indicates the region enlarged in the inset. Note that this region includes a normal duct entrapped at the edge of the tumor containing ER$^+$ luminal cells (arrowheads). Scale bar represents 110 μm.

Metaplastic elements in the form of spindle and/or squamous cell metaplasia were found in 12 out of 14 cases from virgin animals and 9 out of 10 from parous animals. In one virgin tumor, chondroid metaplasia was also observed (Figure 2D; asterisk). All *Blg-Cre* tumors were classified as histological grade 3, displaying a high degree of nuclear pleomorphism, little or no tubule formation, and a high mitotic index, often with aberrant mitotic figures (Figure 2E, arrows). All were positive for basal K14 (Figure 2F). Ten cases from virgin animals and 10 from parous animals were positive for luminal Keratin 18 (K18) (Figure 2G), with staining for both K14 and K18 seen in both epithelioid and metaplastic spindle cells. Eight cases from virgin animals and five from parous animals were positive for the basal marker p63 (Figure 2H). Only three tumors from virgin animals and two from parous animals were ER^+ (Figure 2I). In summary, these tumors displayed histological and immunohistochemical features that closely recapitulated the cardinal features of human *BRCA1* breast tumors. In particular, most displayed features characteristic of the majority of human basal-like breast cancers (Rakha et al., 2008).

K14-Cre Brca1$^{f/f}$ p53$^{+/-}$ Mice Develop Mammary Tumors that Do Not Resemble Human *BRCA1* Loss-of-Function Breast Cancers

Compared to the *Blg-Cre* cohorts, the balance of tumor phenotypes in the *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors was reversed (Figures 2J–2O and 3A). Only one tumor was classified as an IDC-NST but seven of the eight available for analysis were classed as adenosquamous (metaplastic) carcinoma or malignant adenomyoepithelioma (Figure 2J). In the latter, distinct malignant luminal-like and basal-like populations could be observed in an arrangement reminiscent of that of the normal mammary ductal morphology (Figure 2K), features consistent with a stem cell origin. All tumors were strongly K14 and p63 positive (Figures 2L and 2N). Seven tumors were also K18 positive (Figure 2M) but five were ER^- (Figure 2O). The percentage of K14-positive cells within the tumors in the three cohorts fell into a similar range (although it tended to be slightly higher in the *K14-Cre* tumors) (Figure 3B). However, the percentage of K18-positive cells was higher in the parous *Blg-Cre* tumors compared to the *K14-Cre* tumors, whereas the virgin *Blg-Cre* tumors had a bimodal distribution of K18 staining (Figure 3B). In contrast, the percentage of p63-positive cells was higher in the *K14-Cre* tumors compared to the parous *Blg-Cre* tumors, whereas the virgin *Blg-Cre* tumors again had a bimodal distribution (Figure 3B). This suggested that the *K14-Cre* tumors had a more overt basal-like differentiation, the parous *Blg-Cre* tumors were more luminal-like, and the virgin *Blg-Cre* tumors were, as a group, intermediate between the two.

Tumors from all three cohorts had a high mitotic index, which contributed to them being classified as grade 3. However, when mitotic counts were considered, parous *Blg-Cre* tumors were found to be slightly, but significantly ($p < 0.05$, unpaired two-tailed t test), more proliferative than virgin *Blg-Cre* tumors, with on average 145 mitotic figures (95% confidence limits, 102 to 188) and 95 mitotic figures (95% confidence limits, 72 to 118) per 10 high-power fields, respectively. Remarkably, the *K14-Cre* tumors were even more proliferative, with 296 mitotic figures (95% confidence limits, 218 to 373) in 10 fields ($p < 0.001$, unpaired two-tailed t test against parous *Blg-Cre* tumor data) (Figure 3B). The features of these *K14-Cre* tumors were, therefore, consistent with an origin in stem cells with a high proliferative potential.

The profound differences in the histopathological features between tumor cohorts that vary only in the promoter used to drive gene expression are highly likely to be the result of different cells of origin. The patterns of expression of the *K14* and *Blg* promoters within the mammary gland (Figure 1; Figure S2A; Jonkers et al., 2001) and the comparative phenotypes of the tumors support a model in which the *K14-Cre* promoter biases tumor formation to basal stem cells whereas the *Blg-Cre* promoter biases tumor formation to progenitors in the luminal layer.

Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$ Tumors Have a Basal Flow Cytometric Profile

To further investigate the appearance of cells with a basal-like phenotype in luminal progenitor-origin *Brca1* tumors, we used flow cytometry to compare normal epithelial cells isolated from young (12-week) and aged (42-week) virgin mice with tumor cells stained for CD24, Sca-1, and CD49f expression. Changes in the proportion of $CD24^-$ nonepithelial, $CD24^{+/low}$ basal, and $CD24^{+/high}$ luminal cells (Figure 3C) as well as in numbers of cells with the basal stem cell-enriched phenotype ($CD24^{+/low}$ $Sca\text{-}1^-$ $CD49f^{high}$) were quantified (Figure 3D). Both the tumors and aged virgin normal cells showed an increase in the proportion of $CD24^{+/low}$ cells and a decrease in $CD24^{+/high}$ cells compared to young normal mice. However, the tumors also showed significant enrichment for cells with a stem cell-like phenotype ($CD24^{+/low}$ $Sca\text{-}1^-$ $CD49f^{high}$) as well as the appearance of a population with the phenotype $CD24^{+/high}$ $Sca\text{-}1^-$ $CD49f^{high}$ (Figures 3C and 3D; Figure S3A).

Enrichment within the tumors for cells with a stem cell-like flow cytometry phenotype led us to speculate that they may contain

(J–O) Histological features of *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors.
(J) Low-magnification micrograph of malignant adenomyoepithelioma. Scale bar represents 0.4 mm.
(K) High-power image showing gland-like structures lined by two neoplastic cell populations. Note the mitotic figure with plane of division along axis of pseudobasal cells (arrow), suggesting that the pseudobasal cells are contributing to the luminally located cells. Scale bar represents 27 μm.
(L) K14 expression in neoplastic cells (arrow indicates area enlarged in inset). Note that the luminally located cells often lack K14 expression (inset).
(M) K18 was strongly expressed in luminally located cells (inset; arrow indicates enlarged area).
Scale bars in (L) and (M) represent 110 μm.
(N) p63 expression in pseudobasal neoplastic cells. Scale bar represents 55 μm.
(O) ER expression in an ER^- K14-Cre tumor. The arrow indicates an area of hyperplasia entrapped within the tumor that acts as an internal positive control for the staining (enlarged in the inset). Scale bar represents 110 μm.
See also Figure S2 and Table S1.

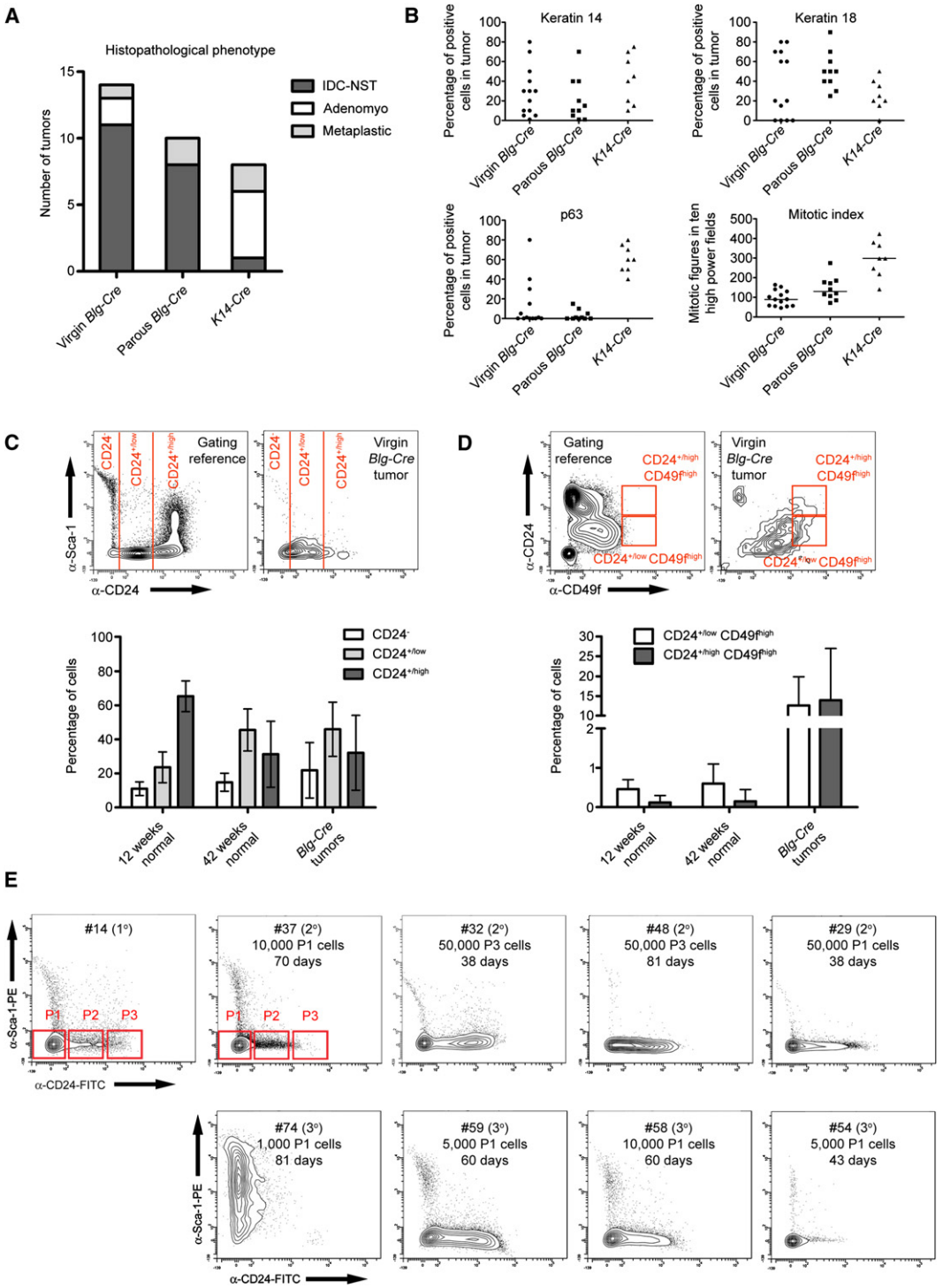


Figure 3. *Blg-Cre* Tumors Have Enhanced Luminal Histological Features Compared to *K14-Cre* Tumors but Are Enriched for CD49f Expression

(A) Prevalence of histopathological phenotypes of *Blg-Cre* and *K14-Cre* tumors. Abbreviations: IDC-NST, invasive ductal carcinoma, no special type; adenomyo, malignant adenomyoepithelioma; metaplastic, metaplastic carcinoma (includes metaplastic spindle cell carcinoma and adenosquamous carcinoma).

(B) Distribution of K14-, K18-, and p63-positive cells within the tumor cohorts and comparative mitotic indices (mitotic figures counted in 10 high-power fields in each tumor of each cohort; bars indicate median values). Each data point represents an individual tumor.

(C and D) CD24 and Sca-1 (C) or CD24 and CD49f (α_6 integrin) (D) flow cytometry staining of normal mouse mammary epithelium (left plots; used as gating reference) and cells isolated from virgin *Blg-Cre* tumors (right plots). Below the plots comparisons are shown of the proportions (±SD) of $CD24^-$ (unshaded bars),

a distinct population of "tumor-initiating cells" or "cancer stem cells." Therefore, we carried out limiting dilution transplant assays on cells from virgin *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors separated on the basis of levels of CD24 expression. These experiments proved technically challenging, because very few viable cells could be recovered from tumors after dissociation and flow sorting (Figure S3A). Nevertheless, we were able to carry out limiting dilution transplant series of dissociated cells from two tumors (three further samples were transplanted at limiting dilutions but failed to form any secondary tumors and so are excluded from the analysis). The results showed that tumor cells with an undifferentiated CD24^{-} Sca-1^{-} phenotype were most highly enriched for tumor-initiating cells, with 1 in 1949 cells having tumor-initiating potential compared to 1 in 7061 cells in the CD24$^{+/low}$ Sca-1^{-} population (Table S2). It is the latter phenotype which is associated with the stem cell population in normal mammary epithelium, so this demonstrated that in tumors the stem cell surface marker phenotype does not necessarily correlate with functional potential.

During this analysis, we collected extensive data on the flow cytometric profiles of primary, secondary, and tertiary transplanted tumors. Remarkably, they were very different, with at least four different types of profile observed (Figure 3E). There was no correlation between the type of profile and the tumor subpopulation from which the new generation of tumor had been grown. Rather, the profiles suggested that clonal selection was occurring, possibly within a pool of cancer stem cell-like cells (Marusyk and Polyak, 2010).

We examined whether the balance of mammary epithelial populations was altered in *Blg-Cre* mice carrying floxed *Brca1* alleles at an age where recombination has occurred only in the luminal ER^{-} progenitor population but prior to tumor formation (10–12 weeks). There were no significant differences (two-tailed t test on Log10 transformed data) between the proportions of basal and luminal ER^{+} cells in mice with various combinations of *Brca1* and *p53* alleles compared to each other or control *Blg-Cre R26R* mice. However, *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* animals had a significantly reduced (two-tailed t test on Log10 transformed data, $p < 0.05$) proportion of luminal ER^{-} cells compared to *Blg-Cre Brca1$^{f/+}$ p53$^{+/-}$* animals (Figure S3C). The significance of this is unclear, however, because in neither genotype was the proportion of luminal ER^{-} cells significantly different to wild-type. Nevertheless, the results do support the luminal ER^{-} progenitor population as the primary target of *Blg-Cre* activity.

Gene Expression Profiles of *Brca1$^{f/f}$ p53$^{+/-}$* Tumors Resemble Those of Luminal Sca-1^{-} ER^{-} Cells

To determine whether the basal-like differentiation observed by histopathology and flow cytometry in luminal origin *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors was reflected at the molecular level, the gene expression profiles of 12 virgin and 6 parous *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors and 3 *K14-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors were analyzed. A set of 2184 Affymetrix probes known to robustly distinguish between normal basal, luminal ER^{-}, and luminal ER^{+} cells (Kendrick et al., 2008) was used for unsupervised hierarchical clustering of the tumors (Figure 4A). Further, both the 2184 subset and the entire probe data set were used for a prediction analysis of microarrays (PAM) approach in which the expression pattern of the normal populations was used as a training set to build a classifier. This was applied to the expression data from the tumors to identify which normal cell population the tumors most closely resembled (Figure 4A and Table S3; similar results were obtained for both the 2184 subset and the entire data set). These data showed that 16 out of 21 tumors, including 2 of the 3 *K14-Cre* tumors, most closely resembled luminal ER^{-} cells in their gene expression profiles. Four out of five tumors classified as basal-like by PAM clustered with each other and with the normal basal cells. However, the virgin *Blg-Cre*, parous *Blg-Cre*, and *K14-Cre* tumors did not segregate into separate groups.

The gene expression microarray data suggested substantial similarity in the molecular pathological phenotype between the three tumor cohorts. To analyze this further, we used qPCR-based gene expression analysis on a subset of genes whose expression was associated with the normal epithelial subtypes (Kendrick et al., 2008). We were also able to use qPCR to address gene expression profiles in a larger number of *K14-Cre* tumors than were available for analysis by microarray. The results confirmed the microarray data, showing that all cohorts expressed genes associated with luminal ER^{-} and basal cell populations but not with luminal ER^{+} cells. Unsupervised hierarchical clustering of the qPCR data showed that, when the additional *K14-Cre* tumors were included, they tended to cluster with the normal basal cell populations. Three *Blg-Cre* tumors classified by PAM as most similar to normal basal cells (9, 36.1, and 7165) also clustered with this group (Figure 4B). Although this clustering did separate tumors more obviously according to their cells of origin, it was clear that overall the tumor cohorts had similar gene expression patterns (Figures 4B and 5). There were some exceptions consistent with the histopathology of the tumors. For instance, *Trp63* expression was higher in the *K14-Cre* cohort than the parous *Blg-Cre* cohort, but in the virgin *Blg-Cre* cohort, *Trp63* expression ranged from very low to high expression levels. The basal marker *Id4* had a similar pattern of expression (Figure 5). In conclusion, despite the different histological features of the *K14-Cre* and *Blg-Cre* tumors, they had a similar (but not identical) gene expression profile that most closely resembled that of the normal luminal ER^{-} cells.

CD24$^{+/low}$ (light gray bars), and CD24$^{+/high}$ cells (dark gray bars) (C) or of CD24$^{+/low}$ CD49f^{high} (unshaded bars) and CD24$^{+/high}$ CD49f^{high} cells (dark gray bars) (D) in three independent preparations of 12-week and 42-week normal virgin mice and nine virgin *Blg-Cre* tumors.

(E) Flow cytometric analysis of transplanted tumors. Top row, CD24 and Sca-1 staining profile of a primary tumor, #14, and four secondary tumors that originated from transplantation of tumor #14. Bottom row, CD24 and Sca-1 profiles of four tertiary tumors from transplantation of tumor #37. The populations P1, P2, and P3 isolated for transplant analysis are indicated. The population of origin, number of cells transplanted, and tumor latency are indicated on each secondary and tertiary tumor.

See also Figure S3 and Table S2.

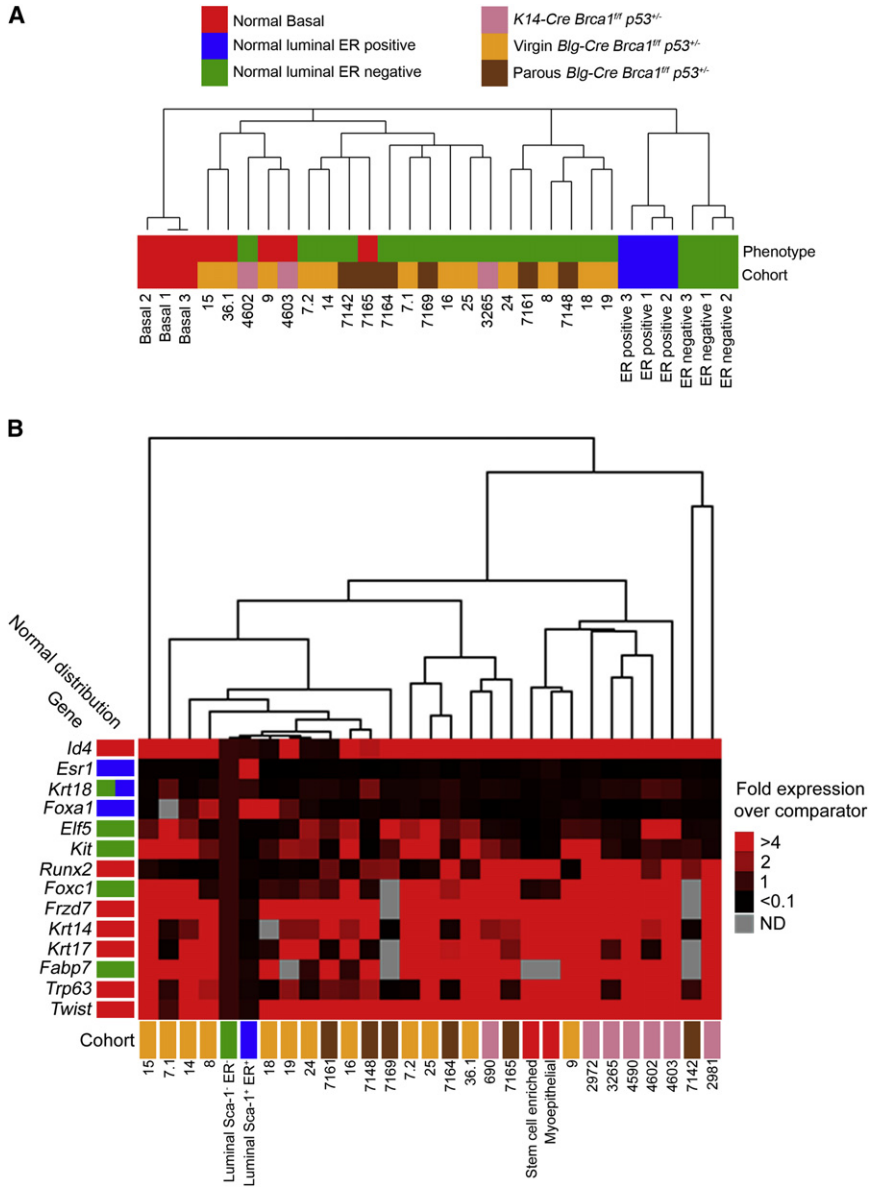


Figure 4. Most *Brca1$^{f/f}$ p53$^{+/-}$* Tumors Have a Luminal ER$^-$ Gene Expression Profile

(A) Unsupervised hierarchical clustering and PAM analysis of tumor gene expression data with data from normal mouse mammary epithelial subpopulations. Clustering is based on the expression of the 2184 cell-specific probes previously identified (Kendrick et al., 2008). PAM results are indicated by color coding below the tumors samples, with green indicating closest similarity to normal luminal ER$^-$ cells and red to normal basal cells. No tumors were similar to luminal ER$^+$ cells (blue). (B) Heat map of gene expression patterns from qPCR analysis of tumor samples based on relative expression levels of 14 genes compared to a comparator population (normal CD24$^{+/high}$ Sca-1$^-$ luminal ER$^-$ cells). Genes analyzed were a selection known to be associated with normal basal (*Frd7*, *Id4*, *Krt14*, *Krt17*, *Runx2*, *Trp63*, and *Twist*), luminal ER$^-$ (*Elf5*, *Foxc1*, *Kit*, *Fabp7*), luminal ER$^+$ (*Esr1*, *Foxa1*) mammary cells, and cells from both luminal populations (*Krt18*) (Kendrick et al., 2008). Tumor samples and genes were clustered in an unsupervised manner. The data from the normal populations are from three separate isolates of the normal cell populations from mice with a mixed genetic background, each measured in triplicate. Data from individual tumor samples were also measured in triplicate. See also Table S3.

Mouse *Brca1$^{f/f}$ p53$^{+/-}$* Tumors Are Classified as Basal-like via Human Class Predictor Gene Sets

Only tumors from the luminal *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* cohorts resembled human *BRCA1* tumors with respect to histopathological phenotype (Figure 2). However, the majority of human *BRCA1* tumors are classified by gene expression analysis as most similar to "basal-like" nonfamilial (sporadic) breast cancers (Sorlie et al., 2003), whereas the mouse tumors tended to have gene expression patterns that most closely resembled normal mouse luminal ER$^-$ cells (Figure 4A). We therefore asked which human breast tumor subtype the mouse tumors most closely resembled. Two single sample predictor gene sets (SSPs) that have been used previously to predict human breast cancer molecular subtypes (Hu et al., 2006; Parker et al., 2009) were used to predict, according to the human classification, the phenotype of the mouse tumors. Although doubt has recently been cast on the ability of SSPs to reliably identify luminal, HER2, and normal breast-like breast cancer subtypes, it is clear that they do robustly distinguish basal-like from non-basal-like breast cancers (Weigelt et al., 2010). Notably, all tumors were classified as basal-like, according to the Hu306 centroids (Hu et al., 2006). Furthermore, 18 out of 21 tumors were classified as basal-like according to the more recently published PAM50 gene set (Figure 6; Table S3; Parker et al., 2009).

We next considered whether sporadic human basal-like breast cancers expressed genes associated with luminal ER$^-$ cells. First, we used self-organizing mapping analysis of the patterns of expression of the 2184 probe subset in the virgin *Blg-Cre* tumors to identify groups of genes expressed at high levels across all tumors (Table S4). Then we identified genes expressed at high levels in the tumors whose expression was associated with the normal luminal ER$^-$ cell population. We examined the distribution of these genes in a panel of 53 grade 3 human breast cancers (Natrajan et al., 2010) (including both luminal and basal subtypes defined according to Nielsen's criteria) (Nielsen et al., 2004) and in the NKI295 tumor gene expression data set (Table 1; Table S5; van de Vijver et al., 2002). In the former set, 69 of the mouse genes could be robustly mapped to human orthologs. 21 had a significant subtype association and of these, 18 (86%) were associated with the

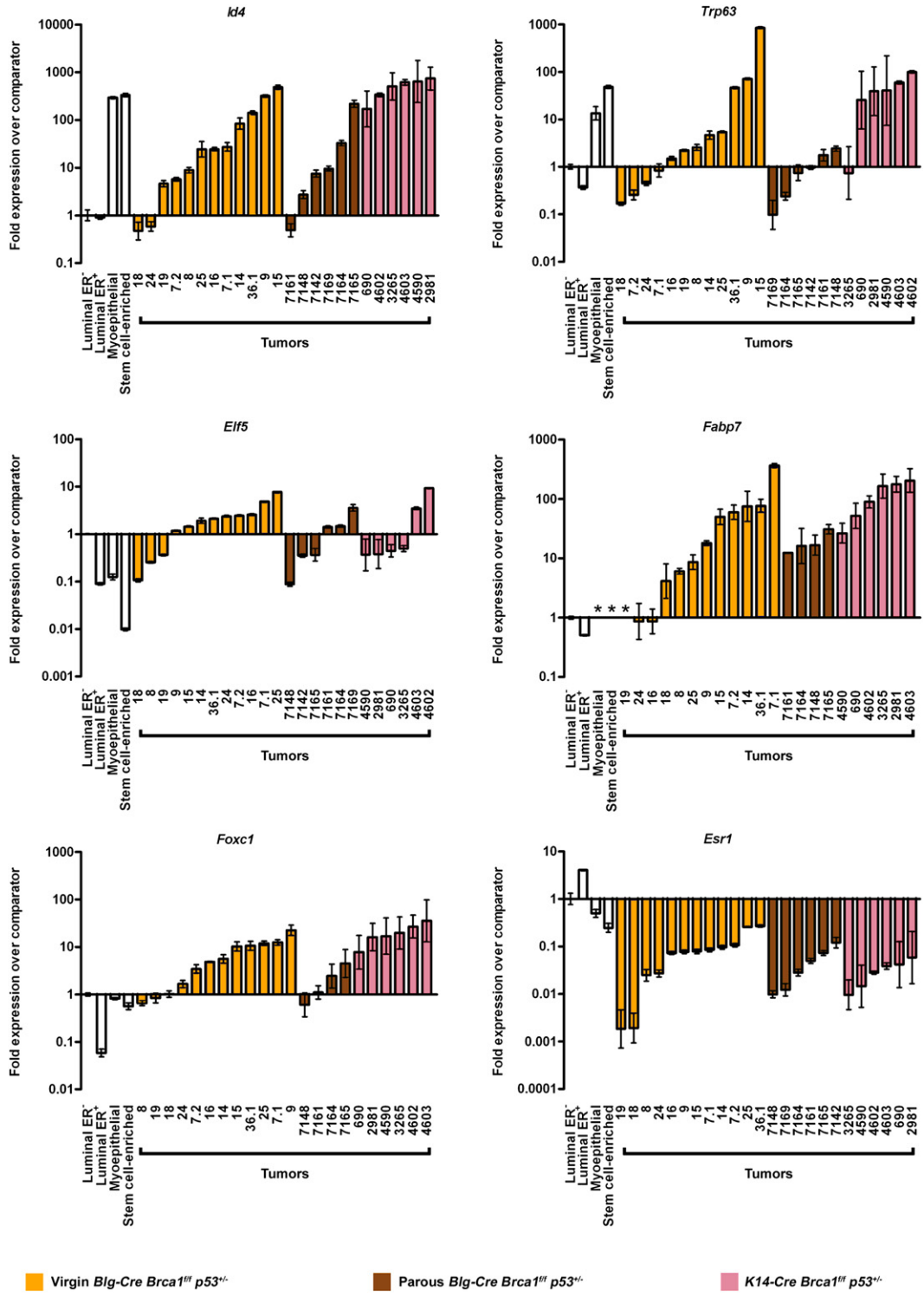


Figure 5. qPCR Analysis of Gene Expression in Tumor Cohorts

Mean gene expression levels (±95% confidence limits) for a subset of the genes shown in Figure 4B (*Id4*, *Trp63*, *Elf5*, *Fabp7*, *Foxc1*, and *Esr1*) in normal luminal ER$^-$, luminal ER$^+$, myoepithelial (CD24$^{+/low}$ Sca-1$^-$ CD49f^{low}), and basal stem cell-enriched (CD24$^{+/low}$ Sca-1$^-$ CD49f^{high}) cells and in the individual tumors, relative to the comparator population (which were the luminal ER$^-$ cells). The data from the normal populations are from three separate isolates of the normal cell populations from mice with a mixed genetic background similar to the *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* mice, each measured in triplicate. Data from individual tumor samples were also measured in triplicate. The normal populations and the three tumor cohorts are grouped together. Within each cohort data are ordered left to right from lowest to highest gene expression levels. If a tumor sample is not present on the graph, then RNA was not available for analysis. Asterisk indicates that expression of a particular gene was undetectable in that sample.

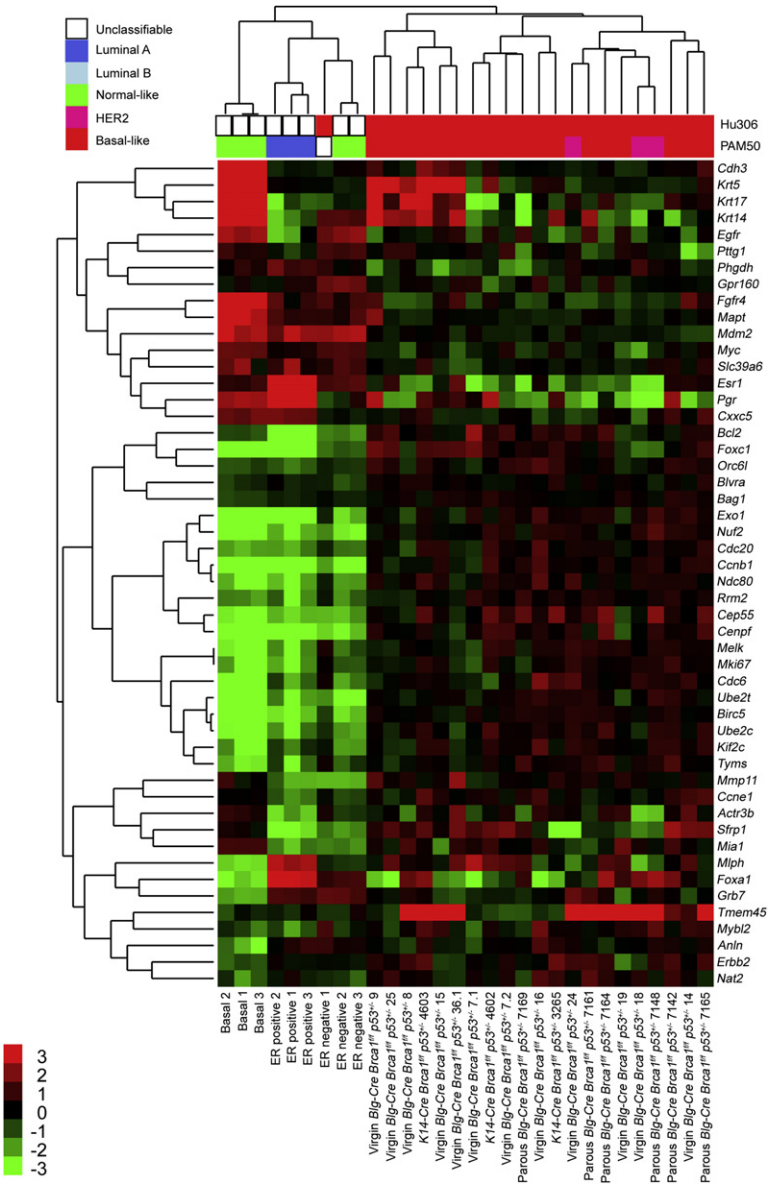


Figure 6. *Brca1*$^{f/f}$ *p53*$^{+/-}$ Tumors Resemble Human Sporadic Basal-like Breast Cancers

Heat map of expression of genes (red, high expression; green, low expression) from the PAM50 breast subtype classifier gene sets in mouse tumors. Samples are clustered on the basis of expression of the PAM50 genes and the single sample predictor correlations with human tumor subtypes for both the PAM50 and Hu306 gene sets are shown below the clustering tree (green, luminal A; blue, luminal B; gray, HER2; black, normal breast-like; red, basal-like; open, unclassifiable). See also Table S3.

basal-like subtype ($p < 0.0001$, two-tailed Fisher exact test). In the NKI data set, 61 orthologs of the mouse genes could be identified. 39 of these had a significant subtype association and 26 (67%) were associated with the basal-like subtype ($p < 0.01$, two-tailed Fisher exact test). These analyses demonstrate that genes typically expressed by ER$^-$ cells of the luminal mammary epithelium are also expressed by sporadic human basal-like breast cancers, consistent with the SSP classification of the mouse tumors.

Finally, gene set enrichment analysis (GSEA) (Subramanian et al., 2005) was used on the expression profiles to identify functionally or phenotypically relevant sets of genes enriched in the tumors compared to the normal cell populations. Notably, this analysis demonstrated that the tumors were enriched for genes expressed at high levels in human ER$^-$ *BRCA1* loss-of-function tumors ($p < 0.01$; FDR q value 0.393) (Table S6).

DISCUSSION

We have directly compared tumor formation in mammary epithelial basal stem and luminal progenitor cells to address not only the roles of stem and progenitor cells in tumorigenesis but also the relationship between the cell of tumor origin and tumor phenotype. We have shown that targeting deletion of *Brca1* to the basal cell layer, and thus biasing tumor formation to basal stem cells, preferentially resulted in the development of aggressive malignant adenomyoepitheliomas. In humans these are rare tumors consisting of two distinct neoplastic populations with basal-like and luminal-like features, consistent with a stem cell origin. Importantly, they are distinct from benign proliferations of myoepithelial cells. In contrast, deleting *Brca1* in luminal progenitors preferentially generated IDC-NSTs that not only phenocopied human *BRCA1* basal-like breast cancers but also had a strong resemblance to the majority of basal-like tumors not associated with a family history of breast cancer (“sporadic” basal-like tumors). These results suggest that malignant adenomyoepitheliomas are true stem cell tumors of the breast but, given the rarity of this tumor type and its lack of association with *BRCA1* disease, that normal mammary gland stem cells are not common targets for transformation in the breast and certainly make little contribution to the origins of *BRCA1*-associated breast cancer. Rather, luminal progenitors are probably the cell type most commonly associated with the initiation of breast cancer.

Although tumors originating in different cell types had distinct histological appearances, their gene expression patterns had similarities. Tumors from all cohorts tended to express genes associated with both normal basal and luminal ER$^-$ cells and to have low levels of expression of luminal ER$^+$ genes (Figures 4B and 5). All tumors were classified as “basal-like” by the Hu306 and PAM50 single sample predictor (SSP) analysis (with the exception of tumors 18, 24, and 7148, which were classed as HER2-like by the PAM50 SSP). However, the *K14-Cre* cohort had higher overall levels of expression of basal-associated genes, consistent with their origin, and clustered with the normal basal populations (Figure 4B). These results suggest that the cell of origin of a tumor is one of the key determinants of the tumor

histological features but can also influence its gene expression profile. In contrast, *BRCA1* loss primarily determines the molecular landscape of the tumor but also has some influence on its histological appearance.

In contrast to our *K14-Cre* model, Liu and colleagues reported a *K14-Cre Brca1$^{f/f}$ p53$^{f/f}$* mouse model in which the tumors obtained were mainly IDC-NSTs, with only one adenomyoepithelioma (Liu et al., 2007). Remarkably, the tumor spectrum they reported is similar to that seen in our *Blg-Cre* model. There are a number of differences between the two *K14-Cre* models, which may explain these discrepancies. These include different background strains, different *K14-Cre* transgenes, and different floxed *Brca1* alleles. Liu and colleagues also used a floxed *p53* allele, rather than the *p53* heterozygote background. Any of these factors could alter the relative likelihood of targeting *Brca1* loss to stem versus progenitor cells, or their relative susceptibility to transformation, resulting in enhanced tumor formation in luminal progenitor cells in the *K14* model of Liu and colleagues and hence a tumor spectrum similar to our *Blg-Cre* model. The importance of the interaction between *p53* and *Brca1* in transforming basal stem cells is suggested by data from Smart and colleagues, who reported no mammary phenotype in a *K14-Cre* conditional *Brca1* mouse line on a wild-type p53 background, unlike in *WAP-Cre* and *MMTV-Cre*-driven *Brca1* knockout models (Smart et al., 2008).

It is formally possible that use of the *p53* heterozygote background in the tumor models may alter the predicted cell type specificity of the promoters used to drive Cre expression, given the role of *p53* in controlling polarity in mammary stem cell division (Cicalese et al., 2009). However, *p53* heterozygosity has not been investigated in this context. In our model, *p53* gene expression levels in *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* spleens and three tumors (14, 15, and 25) were not significantly different to those of wild-type animals. Thus, although the mice were *p53* heterozygous, there was little evidence of *p53* haploinsufficiency, at least at the ages at which tumors were developing (Figure S2). *p53*-dependent changes in stem cell dynamics prior to deletion of the *Brca1* floxed allele are therefore unlikely to alter the targeting of Cre activity. We also cannot definitively exclude the possibility that modifier genes linked to the *Blg-Cre* or *K14-Cre* loci influenced the difference in tumor phenotypes. However, because the differences were a change in the balance of the tumor phenotypes rather than absolute changes in tumor type, this hypothesis is unlikely.

Blg-Cre activity resulted in a small number of recombined progenitors (i.e., cells with proliferative potential) in the luminal ER$^+$ population, suggesting that ER$^+$ progenitors could be the origin of some tumors. However, the *Blg-Cre* transgene primarily drove recombination in luminal ER$^-$ progenitors (Figure 1), supporting these cells as the main target for *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumor formation. When purified and transplanted into cleared fat pads, the luminal ER$^-$ population produced multilineage epithelial outgrowths containing myoepithelial, luminal ER$^-$, and luminal ER$^+$ cells, although at lower frequency and with less extensive outgrowths than are generated after transplantation of basal stem cells (Sleeman et al., 2007). This demonstrates that the luminal ER$^-$ population includes a proportion of cells with context-dependent multilineage differentiation capacity. Transformation of mouse luminal ER$^-$ cells and the generation of tumors containing cells with both luminal ER$^-$ and basal phenotypes could therefore be a result of destabilization of the progenitor phenotype and aberrant activation of normally tightly controlled differentiation programs. It has been reported that human luminal progenitors are lineage restricted and do not normally show basal differentiation potential (Lim et al., 2009). However, it is well established that some luminal epithelial cells within the human TDLU express "basal" keratins (Gusterson, 2009; Gusterson et al., 2005), suggesting that expression of some basal markers is an aspect of normal differentiation of a subset of human luminal cells. Furthermore, *BRCA1* loss is itself likely to result in upregulation of expression of basal-associated genes (Gorski et al., 2010; Hosey et al., 2007). This could contribute to the shift toward basal-like differentiation patterns, although clearly many cells in "basal-like" *BRCA1* tumors retain a fundamentally luminal identity.

Besides the inherent differentiation potential of the cells of origin of the *Blg-Cre* tumors and the effects of *BRCA1* loss on transcription, phenotypic plasticity in the *Blg-Cre* cohort was added to by metaplasia and, potentially, epithelial-mesenchymal transition (EMT). Metaplasia is widely seen in preneoplastic conditions such a Barrett's esophagus (Barbera and Fitzgerald, 2009) and is considered to facilitate tumor development (Mesquita et al., 2006) as well as contributing to tumor cellular heterogeneity (Geyer et al., 2010). EMT in epithelial tumors is associated with increased invasiveness, therapeutic resistance, and potentially, the generation of cells with cancer stem cell-like properties (Polyak and Weinberg, 2009). The *Blg-Cre* tumors expressed EMT-associated genes such as *Vim* and *Twist1* (Table S4); however, expression of these genes is a feature of normal myoepithelial cells (Kendrick et al., 2008). Furthermore, many (but not all) spindle cells observed in these tumors express keratins (Table S1). It is therefore unclear to what extent expression of EMT markers in these tumors was related to a basal differentiation program rather than to a full EMT.

Paradoxically, given the ER$^-$ phenotype of most *BRCA1* tumors, prophylactic oophorectomy provides protection against breast cancer in *BRCA1* mutation carriers (Rebbeck et al., 2002). Because estrogen stimulates proliferation of mammary epithelial cell growth in a paracrine manner (Mallepell et al., 2006), oophorectomy may block paracrine stimulation of a preneoplastic clone originating in an ER$^-$ stem or progenitor cell. Indeed, ovariectomy in mice resulted in a decrease in *cyclin D1* gene expression in basal, luminal progenitor, and mature luminal cells (Asselin-Labat et al., 2010). Alternatively, if there is a role for ER$^+$ progenitors in the origin of at least a subset of *BRCA1* breast cancers, as noted above, then oophorectomy might also act directly to block the early neoplastic progression ER$^+$ *BRCA1* tumors. Finally, it is formally possible that all *BRCA1* tumors, no matter their ER status, originate in ER$^+$ progenitors and that oophorectomy therefore directly blocks their early growth. Although we cannot definitively exclude a luminal ER$^+$ progenitor origin for some *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors, both our data and recent work from Lim and colleagues (Lim et al., 2009) strongly favor an origin of the majority of *BRCA1* tumors in luminal ER$^-$ progenitors and thus support a model of oophorectomy interfering with paracrine signals in the majority of cases.

Our findings support a model in which *BRCA1* mutation-associated basal-like tumors develop in luminal progenitors.

Table 1. Mouse Luminal ER$^-$ Genes Are Expressed in Sporadic Human Basal-like Breast Cancers in a Panel of 53 Grade-Matched Tumors and in the NKI295 Tumor Set

		53 Grade-Matched Tumor Set		NKI295 Tumor Set	
Symbol	Unigene ID	Subtype Association	FDR Corrected p Value	Subtype Association	FDR Corrected p Value
ELF5	Hs.11713	basal-like	0	basal-like	0.00791176
FOXC1	Hs.348883	basal-like	0	basal-like	9.74E-16
RERG	Hs.199487	luminal	0	luminal	6.92E-13
GABRP	Hs.26225	basal-like	0.00138	basal-like	1.51E-14
SLC6A14	Hs.522109	basal-like	0.00345	basal-like	5.06E-07
KCNN4	Hs.10082	basal-like	0.003942857	basal-like	1.82E-15
CDK6	Hs.119882	basal-like	0.005366667	basal-like	3.85E-11
LYN	Hs.699154	basal-like	0.005366667	basal-like	5.25E-15
CMAS	Hs.311346	basal-like	0.019638462	basal-like	0.000281051
ZBED4	Hs.475208	basal-like	0.019638462	basal-like	1.88E-06
CA9	Hs.63287	basal-like	0.025875	basal-like	4.46E-10
NCALD	Hs.492427	basal-like	0.033282353	basal-like	0.000825839
PIK3AP1	Hs.310456	basal-like	0.03335	basal-like	1.73E-09
DCLK1	Hs.507755	luminal	0.037785714	luminal	2.03E-06
SMOX	Hs.433337	basal-like	0	NS	NS
TSPAN33	Hs.27267	basal-like	0.00759	NS	NS
NPAS2	Hs.156832	basal-like	0.019638462	NS	NS
ITGB8	Hs.592171	basal-like	0.019714286	NS	NS
SHROOM2	Hs.567236	luminal	0.02484	NS	NS
ATP10B	Hs.109358	basal-like	0.037785714	NS	NS
TRIB2	Hs.467751	basal-like	0.037785714	NS	NS
SLC1A1	Hs.444915	NS	NS	luminal	5.25E-15
SH3KBP1	Hs.719268	NS	NS	basal-like	1.28E-09
EPSTI1	Hs.546467	NS	NS	basal-like	1.52E-08
NAV3	Hs.655301	NS	NS	luminal	5.50E-08
CCL5	Hs.514821	NS	NS	basal-like	5.95E-08
MUC1	Hs.89603	NS	NS	luminal	1.48E-07
SOX6	Hs.368226	NS	NS	basal-like	1.48E-07
ALDOC	Hs.155247	NS	NS	basal-like	3.18E-07
TNFAIP2	Hs.525607	NS	NS	basal-like	8.38E-07
FUT9	Hs.49117	NS	NS	basal-like	2.03E-06
IKZF1	Hs.435949	NS	NS	basal-like	5.56E-06
ATP6V1B1	Hs.64173	NS	NS	luminal	9.87E-06
BEX4	Hs.184736	NS	NS	luminal	4.72E-05
BDH1	Hs.274539	NS	NS	luminal	8.43E-05
CLMN	Hs.301478	NS	NS	luminal	0.000276215
PIK3CB	Hs.239818	NS	NS	basal-like	0.000281051
CHMP1B	Hs.656244	NS	NS	basal-like	0.000326272
CGNL1	Hs.148989	NS	NS	luminal	0.000542776
RPS6KA5	Hs.510225	NS	NS	luminal	0.000738469
CSTF2T	Hs.591358	NS	NS	luminal	0.000825839
RAB32	Hs.287714	NS	NS	basal-like	0.000883334
ADCY5	Hs.593293	NS	NS	luminal	0.009537701
KIT	Hs.479754	NS	NS	basal-like	0.014232944

Table 1. *Continued*

		53 Grade-Matched Tumor Set		NKI295 Tumor Set	
Symbol	Unigene ID	Subtype Association	FDR Corrected p Value	Subtype Association	FDR Corrected p Value
ING3	Hs.489811	NS	NS	basal-like	0.020900677
DOCK8	Hs.132599	NS	NS	basal-like	0.027787215

Results of analysis on human breast cancers of expression of human orthologs of genes overexpressed in mouse *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors that in the normal mouse mammary tissue are associated with luminal ER$^-$ cells (Kendrick et al., 2008). The distribution of the gene expression patterns was compared in both luminal versus basal tumors (designated according to the Nielsen criteria) (Nielsen et al., 2004) and in all ER-positive versus ER-negative tumors in a panel of 53 grade-matched human breast tumors. The gene expression distribution was also analyzed in the NKI295 data set (van de Vijver et al., 2002). Only genes significantly associated with a basal-like versus luminal distribution in either of the two tumor sets are shown above. NS, no significant association in that tumor set. See also Tables S4, S5, and S6.

Furthermore, the majority of nonfamilial basal-like tumors have histopathological features consistent with a luminal progenitor origin. In contrast, true breast stem cells probably are the cell of origin of rare malignant adenomyoepithelioma. These results demonstrate that the inherent plasticity of some types of stem and progenitor cell, and the potential that this can be activated in the context of tumorigenesis, means that tumor phenotype may not simply reflect cell of origin (Gusterson, 2009). Accurate identification of the stem or progenitor cell of origin of a particular tumor type, however, offers the prospect of novel prophylactic and therapeutic approaches.

EXPERIMENTAL PROCEDURES

Harvest and Processing of Tumors

Tumors of 1–1.2 cm in diameter were excised from humanely killed mice. Part of the tumor was fixed in 4% phosphate-buffered formalin (BIOS Europe Ltd, Skelmersdale, UK) overnight for routine histological analysis and immunohistochemistry and part was snap frozen on dry ice for isolation of nucleic acids. Where tumors were of sufficient size, a portion was also kept for flow cytometric analysis.

Flow Cytometry and Clonal Analysis

Tumor and normal mammary cell suspensions were prepared by mechanical digestion followed by a short (30–60 min) enzymatic treatment. In the case of normal cells, this was followed by a second round of trypsinization to liberate single cells (Sleeman et al., 2007). Normal and tumor cell suspensions were stained with anti-CD45-PE-Cy7, anti-CD24-FITC, and anti-Sca-1-PE. Cells were then resuspended in medium containing 0.01% 4′,6-diamidino-2-phenylindole (DAPI) for analysis and cell sorting.

For clonal analysis, CD45$^-$ CD24$^{+/low}$ Sca-1$^-$ (basal), CD45$^-$ CD24$^{+/high}$ Sca-1$^-$ (luminal ER$^-$), and CD45$^-$ CD24$^{+/high}$ Sca-1$^+$ (luminal ER$^+$) cells were sorted into collection tubes and then resorted for extra purity. On the second round of sorting, the cells were single cell sorted into individual wells of 96-well plates (Nunc, Fisher Scientific UK Ltd., Leicestershire, UK) and cultured for 10 days (Sleeman et al., 2007). The plates were then fixed in 4% paraformaldehyde in PBS for 5 min, and stained with X-gal solution overnight at 37°C. β-galactosidase-positive colonies were scored and plates were then stained with crystal violet overnight to assess the total number of colonies. The proportion of recombined progenitors in each cell population at each developmental stage (Figure 1E) was calculated by multiplying the colony-forming efficiency of each population (the number of colonies formed as a proportion of the number of single wells seeded) by the percentage of colonies that formed which were β-galactosidase positive. Thus, for the virgin 12-week CD24$^{+/high}$ Sca-1$^-$ population, the proportion of recombined progenitors in this population was (709/(96 × 11))% × 76.8% = 51.6%.

Histology and Immunohistochemistry

Immunohistochemistry was carried out via standard techniques with ERα, p63, keratin 18, and keratin 14 antibodies. Sections were analyzed by three of the authors on a multiheaded microscope blinded to the genotype of the tumors and a consensus score was rendered for each marker in each case. All markers were semiquantitatively assessed by estimating the percentage of morphologically unequivocal neoplastic cells displaying either nuclear (ER and p63) or cytoplasmic (K14 and K18) staining.

Cleared Fat Pad Transplantation

Sorted mammary tumor subpopulations were transplanted at varying dilutions into cleared fat pads of 8–10 g *Rag*$^{-/-}$ immune-deficient mice. Once tumors appeared, they were processed as above. If tumors had not appeared after 6 months, mice were humanely killed and fat pads examined by wholemounting for any evidence of outgrowths.

Microarray Analysis of Tumor Gene Expression Patterns

Transcriptome analysis on tumor RNA isolated via the RNeasy Mini Kit (QIAGEN, West Sussex, UK) was carried out by Almac Diagnostics (Craigavon, UK) with Affymetrix 430_0.2 mouse gene expression chips. The raw data from normal cell populations was reanalyzed together with the raw data from the tumors to ensure that the analyses were comparable. Raw Affymetrix .CEL files were normalized via robust multiarray analysis (Irizarry et al., 2003) in the Affymetrix package from BioConductor (http://bioconductor.org/). With the previously established 2184 probe sets shown to robustly distinguish between normal mouse mammary basal, luminal ER$^-$, and luminal ER$^+$ mammary cells (Kendrick et al., 2008), all normal and tumor samples were clustered based on Ward clustering via an uncentered Pearson's correlation. The relationship of each individual tumor to the normal cell populations was determined by PAM (prediction analysis of microarrays) (Tibshirani et al., 2002) with either all probes or the 2184 cell-type-specific probe subset. SOM (self-organizing map) analysis was carried out on the ROCK platform (http://rock.icr.ac.uk), also using the 2184 probe subset, on a 3 × 3 grid using a Pearson's Correlation similarity metric.

Quantitative Real-Time rtPCR

qPCR was carried out with TAQMAN (Applied Biosystems, Warrington, UK) Assays-on-Demand probes on freshly isolated RNA (Kendrick et al., 2008). Results were analyzed with the Δ-ΔC_t method normalized to β-actin and compared to a comparator sample (normal luminal ER$^-$ cells).

Correlation of *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* Tumors with Human Gene Expression Data Sets

Spearman's rank correlations were performed for each mouse tumor sample between the published human centroids for luminal A, luminal B, HER2, normal breast-like, and basal-like breast cancers as described by Hu and colleagues and Parker and colleagues (Hu et al., 2006; Parker et al., 2009). Tumors were assigned to one of the molecular subtypes based on the highest correlation coefficient between their expression profile and the five individual centroids of the molecular subtypes (i.e., basal-like, HER2, normal breast-like, luminal A, and luminal B).

Analysis of Gene Expression Data in Human Breast Cancers

A two-tailed t test was carried out on the expression levels within breast cancer subtypes of the human orthologs of genes expressed at high levels in both normal mouse luminal ER$^-$ cells and the *Blg-Cre Brca1$^{f/f}$ p53$^{+/-}$* tumors. Two tumor data sets were used: a cohort of 53 consecutive grade 3 IDC-NSTs

(Natrajan et al., 2010) and the NKI295 data set (van de Vijver et al., 2002). Results were corrected with false discovery rate (FDR) correction. Genes were considered to be significantly associated with a breast cancer subtype if the FDR-corrected p value was <0.05.

ACCESSION NUMBERS

Raw data files and methodological details fully compliant with MIAME guidelines have been submitted to ArrayExpress with the accession numbers E-TABM-683, E-TABM-684, and E-TABM-997.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and six tables and can be found with this article online at doi:10.1016/j.stem.2010.07.010.

ACKNOWLEDGMENTS

We thank the Breakthrough Histopathology Facility for their assistance with processing tumor tissues, Fredrik Wallberg for help with flow cytometry analysis, and Britta Weigelt for providing the centroids for human breast cancer molecular subtype assignment. This work was funded by Breakthrough Breast Cancer. We acknowledge NHS funding to the NIHR Biomedical Research Centre.

Received: August 25, 2009
Revised: June 4, 2010
Accepted: July 1, 2010
Published: September 2, 2010

REFERENCES

Asselin-Labat, M.L., Vaillant, F., Sheridan, J.M., Pal, B., Wu, D., Simpson, E.R., Yasuda, H., Smyth, G.K., Martin, T.J., Lindeman, G.J., and Visvader, J.E. (2010). Control of mammary stem cell function by steroid hormone signalling. Nature *465*, 798–802.

Barbera, M., and Fitzgerald, R.C. (2009). Cellular mechanisms of Barrett's esophagus development. Surg. Oncol. Clin. N. Am. *18*, 393–410.

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.

Cicalese, A., Bonizzi, G., Pasi, C.E., Faretta, M., Ronzoni, S., Giulini, B., Brisken, C., Minucci, S., Di Fiore, P.P., and Pelicci, P.G. (2009). The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. Cell *138*, 1083–1095.

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.

Foulkes, W.D. (2004). BRCA1 functions as a breast stem cell regulator. J. Med. Genet. *41*, 1–5.

Geyer, F.C., Weigelt, B., Natrajan, R., Lambros, M.B., de Biase, D., Vatcheva, R., Savage, K., Mackay, A., Ashworth, A., and Reis-Filho, J.S. (2010). Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. J. Pathol. *220*, 562–573.

Gorski, J.J., James, C.R., Quinn, J.E., Stewart, G.E., Staunton, K.C., Buckley, N.E., McDyer, F.A., Kennedy, R.D., Wilson, R.H., Mullan, P.B., and Harkin, D.P. (2010). BRCA1 transcriptionally regulates genes associated with the basal-like phenotype in breast cancer. Breast Cancer Res. Treat. *122*, 721–731.

Gusterson, B. (2009). Do 'basal-like' breast cancers really exist? Nat. Rev. Cancer *9*, 128–134.

Gusterson, B.A., Ross, D.T., Heath, V.J., and Stein, T. (2005). Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer. Breast Cancer Res. *7*, 143–148.

Hosey, A.M., Gorski, J.J., Murray, M.M., Quinn, J.E., Chung, W.Y., Stewart, G.E., James, C.R., Farragher, S.M., Mulligan, J.M., Scott, A.N., et al. (2007). Molecular basis for estrogen receptor alpha deficiency in BRCA1-linked breast cancer. J. Natl. Cancer Inst. *99*, 1683–1694.

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

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Jonkers, J., Meuwissen, R., van der Gulden, H., Peterse, H., van der Valk, M., and Berns, A. (2001). Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nat. Genet. *29*, 418–425.

Kendrick, H., Regan, J.L., Magnay, F.A., Grigoriadis, A., Mitsopoulos, C., Zvelebil, M., and Smalley, M.J. (2008). Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC Genomics *9*, 591.

Lakhani, S.R., Reis-Filho, J.S., Fulford, L., Penault-Llorca, F., van der Vijver, M., Parry, S., Bishop, T., Benitez, J., Rivas, C., Bignon, Y.J., et al; Breast Cancer Linkage Consortium. (2005). Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin. Cancer Res. *11*, 5175–5180.

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

Liu, X., Holstege, H., van der Gulden, H., Treur-Mulder, M., Zevenhoven, J., Velds, A., Kerkhoven, R.M., van Vliet, M.H., Wessels, L.F., Peterse, J.L., et al. (2007). Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. Proc. Natl. Acad. Sci. USA *104*, 12111–12116.

Liu, S., Ginestier, C., Charafe-Jauffret, E., Foco, H., Kleer, C.G., Merajver, S.D., Dontu, G., and Wicha, M.S. (2008). BRCA1 regulates human mammary stem/progenitor cell fate. Proc. Natl. Acad. Sci. USA *105*, 1680–1685.

Mallepell, S., Krust, A., Chambon, P., and Brisken, C. (2006). Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. Proc. Natl. Acad. Sci. USA *103*, 2196–2201.

Marusyk, A., and Polyak, K. (2010). Tumor heterogeneity: Causes and consequences. Biochim. Biophys. Acta *1805*, 105–117.

McCarthy, A., Savage, K., Gabriel, A., Naceur, C., Reis-Filho, J.S., and Ashworth, A. (2007). A mouse model of basal-like breast carcinoma with metaplastic elements. J. Pathol. *211*, 389–398.

Melchor, L., and Benítez, J. (2008). An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. Carcinogenesis *29*, 1475–1482.

Mesquita, P., Raquel, A., Nuno, L., Reis, C.A., Silva, L.F., Serpa, J., Van Seuningen, I., Barros, H., and David, L. (2006). Metaplasia—A transdifferentiation process that facilitates cancer development: The model of gastric intestinal metaplasia. Crit. Rev. Oncog. *12*, 3–26.

Natrajan, R., Weigelt, B., Mackay, A., Geyer, F.C., Grigoriadis, A., Tan, D.S., Jones, C., Lord, C.J., Vatcheva, R., Rodriguez-Pinilla, S.M., et al. (2010). An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers. Breast Cancer Res. Treat. *121*, 575–589.

Naylor, M.J., Li, N., Cheung, J., Lowe, E.T., Lambert, E., Marlow, R., Wang, P., Schatzmann, F., Wintermantel, T., Schüetz, G., et al. (2005). Ablation of beta1 integrin in mammary epithelium reveals a key role for integrin in glandular morphogenesis and differentiation. J. Cell Biol. *171*, 717–728.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., et al. (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin. Cancer Res. *10*, 5367–5374.

Oakes, S.R., Naylor, M.J., Asselin-Labat, M.L., Blazek, K.D., Gardiner-Garden, M., Hilton, H.N., Kazlauskas, M., Pritchard, M.A., Chodosh, L.A., Pfeffer, P.L., et al. (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev. *22*, 581–586.

Palacios, J., Honrado, E., Osorio, A., Cazorla, A., Sarrió, D., Barroso, A., Rodríguez, S., Cigudosa, J.C., Diez, O., Alonso, C., et al. (2005). Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. Breast Cancer Res. Treat. *90*, 5–14.

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

Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P.G., and Di Fiore, P.P. (2010). Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. Cell *140*, 62–73.

Polyak, K., and Weinberg, R.A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat. Rev. Cancer *9*, 265–273.

Rakha, E.A., Reis-Filho, J.S., and Ellis, I.O. (2008). Basal-like breast cancer: A critical review. J. Clin. Oncol. *26*, 2568–2581.

Rebbeck, T.R., Lynch, H.T., Neuhausen, S.L., Narod, S.A., Van't Veer, L., Garber, J.E., Evans, G., Isaacs, C., Daly, M.B., Matloff, E., et al; Prevention and Observation of Surgical End Points Study Group. (2002). Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. N. Engl. J. Med. *346*, 1616–1622.

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

Sleeman, K.E., Kendrick, H., Ashworth, A., Isacke, C.M., and Smalley, M.J. (2006). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast Cancer Res. *8*, R7.

Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J. Cell Biol. *176*, 19–26.

Smalley, M., and Ashworth, A. (2003). Stem cells and breast cancer: A field in transit. Nat. Rev. Cancer *3*, 832–844.

Smart, C.E., Clarke, C., Brooks, K.M., Raghavendra, A., Brewster, B.L., French, J.D., Hetherington, R., Fleming, J.S., Rothnagel, J.A., Wainwright, B., et al. (2008). Targeted disruption of Brca1 in restricted compartments of the mouse mammary epithelia. Breast Cancer Res. Treat. *112*, 237–241.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl. Acad. Sci. USA *100*, 8418–8423.

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

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Taddei, I., Deugnier, M.A., Faraldo, M.M., Petit, V., Bouvard, D., Medina, D., Fässler, R., Thiery, J.P., and Glukhova, M.A. (2008). Beta1 integrin deletion from the basal compartment of the mammary epithelium affects stem cells. Nat. Cell Biol. *10*, 716–722.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc. Natl. Acad. Sci. USA *99*, 6567–6572.

van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., et al. (2002). A gene-expression signature as a predictor of survival in breast cancer. N. Engl. J. Med. *347*, 1999–2009.

Vassilopoulos, A., Wang, R.H., Petrovas, C., Ambrozak, D., Koup, R., and Deng, C.X. (2008). Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. Int. J. Biol. Sci. *4*, 133–142.

Weigelt, B., Mackay, A., A'hern, R., Natrajan, R., Tan, D.S., Dowsett, M., Ashworth, A., and Reis-Filho, J.S. (2010). Breast cancer molecular profiling with single sample predictors: A retrospective analysis. Lancet Oncol. *11*, 339–349.