Journal of Pathology
J Pathol 2014; **233:** 124–137
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.4345

ORIGINAL PAPER

Identification of cellular and genetic drivers of breast cancer heterogeneity in genetically engineered mouse tumour models

Lorenzo Melchor,[1,#,##] Gemma Molyneux,[1†#] Alan Mackay,[1] Fiona-Ann Magnay,[1] María Atienza,[2] Howard Kendrick,[1§] Daniel Nava-Rodrigues,[1] María Ángeles López-García,[2] Fernanda Milanezi,[1,‡] Kirsty Greenow,[3] David Robertson,[1] José Palacios,[4] Jorge S Reis-Filho[5] and Matthew J Smalley[1,*,§]

[1] *Division of Breast Cancer Research, Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK*
[2] *Department of Pathology, Instituto de Biomedicina de Sevilla, CSIC–Universidad de Sevilla, Hospital Universitario del Rocío, Seville, Spain*
[3] *European Cancer Stem Cell Research Institute, Cardiff School of Biosciences, Cardiff University, Cardiff, UK*
[4] *Department of Pathology, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigaciones Sanitarias, Madrid, Spain*
[5] *Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, USA*

**Correspondence to: MJ Smalley, European Cancer Stem Cell Research Institute, Cardiff School of Biosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK. E-mail: SmalleyMJ@Cardiff.ac.uk*

[#]These authors contributed equally to this study.
[##]Current address: Centre for Myeloma Research, Division of Molecular Pathology, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK
[†]Current address: Centre for Molecular and Cellular Biology of Inflammation, Division of Immunology Infection and Inflammatory Diseases, King's College London, 1st Floor New Hunt's House, Great Maze Pond, London SE1 1UL, UK
[‡]Current address: Salomao and Zoppi Diagnósticos, 876 Divino Salvador Avenue, Sao Paulo, SP, Brazil
[§]Current address: European Cancer Stem Cell Research Institute, Cardiff School of Biosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK

Abstract

The heterogeneous nature of mammary tumours may arise from different initiating genetic lesions occurring in distinct cells of origin. Here, we generated mice in which *Brca2*, *Pten* and *p53* were depleted in either basal mammary epithelial cells or luminal oestrogen receptor (ER)-negative cells. Basal cell-origin tumours displayed similar histological phenotypes, regardless of the depleted gene. In contrast, luminal ER-negative cells gave rise to diverse phenotypes, depending on the initiating lesions, including both ER-negative and, strikingly, ER-positive invasive ductal carcinomas. Molecular profiling demonstrated that luminal ER-negative cell-origin tumours resembled a range of the molecular subtypes of human breast cancer, including basal-like, luminal B and 'normal-like'. Furthermore, a subset of these tumours resembled the 'claudin-low' tumour subtype. These findings demonstrate that not only do mammary tumour phenotypes depend on the interactions between cell of origin and driver genetic aberrations, but also multiple mammary tumour subtypes, including both ER-positive and -negative disease, can originate from a single epithelial cell type. This is a fundamental advance in our understanding of tumour aetiology.

Keywords: *Brca2*; *Pten*; *p53*; tumour heterogeneity; breast cancer molecular subtypes; basal-like

Received 11 June 2013; Revised 24 January 2014; Accepted 26 February 2014

No conflicts of interest were declared.

Introduction

Breast cancer is a heterogeneous disease encompassing different histological and molecular subtypes, with distinct clinical behaviours [1–4]. The biological basis of this heterogeneity remains poorly understood; improving this understanding is key to better patient stratification. Although distinct molecular events occurring in different target cells may explain the variety of breast cancer phenotypes [5,6], there is not necessarily a direct correlation between tumour phenotype and its cell of origin. For instance, breast cancers of 'basal-like' subtype were proposed to arise from basal stem cells [7–10], but current models suggest that a substantial proportion, if not all, of these tumours derive from luminal-progenitor cells [11–13]. Disentangling the complex relationship between tumour-initiating genetic events, target cells and tumour phenotypes is ideally suited to studies using genetically engineered mouse models.

We previously demonstrated that when *Brca1* and *p53* loss were targeted to either basal or luminal ER-negative mammary (lumER^{neg}) cells in mouse models, the balance of tumour phenotypes depended on the cell of origin. Although all tumours were molecularly classified as 'basal-like', histologically the basal cell-origin tumours were mostly adenomyoepitheliomas (AMEs), while the lumER^{neg} cell-origin tumours were high-grade invasive

ductal carcinomas of no special type (IDC-NSTs)[13]. It remains to be defined, however, whether the cell of origin is the prime determinant of tumour subtype, or whether initiating genetic hits also play a role in shaping phenotype, in addition to simply stimulating tumourigenesis.

To address this question, we generated conditional mouse models in which *Brca2*, *p53* and/or *Pten* were deleted in distinct cell populations of the mouse mammary gland. To fully describe the tumours these animals developed, detailed histopathological, immunohistochemical and gene expression analyses were performed. We demonstrate that the relative contributions of cell of origin and molecular lesion to determining mammary tumour heterogeneity are context-dependent. The final tumour phenotype is the result of both interactions between the cell of origin and genetic aberrations, and epistatic interactions between genetic aberrations within a cancer.

Materials and methods

Tumour cohorts

The following genotypes were established and maintained until tumours developed: *K14Cre:Brca2$^{f/f}$:p53$^{f/f}$*, *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* virgin and parous; *Pten$^{+/-}$*, *K14Cre:Pten$^{f/f}$*; *BlgCre:Pten$^{f/f}$* virgin and parous; *BlgCre:Pten$^{f/+}$:p53$^{f/+}$*, *BlgCre:Pten$^{f/f}$:p53$^{f/f}$*. Parous mice went through two or three pregnancy cycles. Tumours were excised from humanely killed mice and half was fixed in 4% phosphate-buffered formalin (BIOS Europe Ltd, Skelmesdale, UK) overnight for paraffin embedding. The remainder was snap-frozen on dry ice for nucleic acid isolation.

Histology and immunohistochemistry

Haematoxylin and eosin (H&E) staining was performed using standard methods. Immunohistochemistry for ERα, p63, K14 and K18 and double p63/ER immunofluorescence were carried out as described [13,14]. Immunohistochemistry for PRA (hPRa7; ThermoScientific, UK) and PRB (αPR6; Abcam, Cambridge, UK) were performed using the ER protocol. Immunohistochemistry for human CLDN3 (Z23.JM, Invitrogen-Life, Paisley, UK), CLDN4 (3E2C1, Invitrogen-Life), CDH1 (Zymed, CA, USA) and PTEN (6H2.1; Dako, Denmark) were performed as described [15,16].

Gene expression microarray analysis

Samples which underwent gene expression analysis were morphologically checked to be representative. Microarray hybridization was performed by UCL Genomics (UCL, London, UK), using the Affymetrix GeneChip Mouse Genome 430 2.0 Array (Affymetrix, Santa Clara, CA, USA). Data were read using the Affymetrix package in R (v 2.11.0) and annotated using Bioconductor 2.8. Arrays were normalized with the RMA method in Expression Console 1.1 and annotated with corresponding human orthologue annotation based upon the Mouse Genome Informatics database (http://www.informatics.jax.org/). Subgroup assignment was performed based upon nearest-centroid Spearman rank correlation over 0.1, as described [13,17], using published centroid data [18]. Meta-analysis of the mouse tumour signatures in human breast cancers is fully described in the online supplementary material (see Supplementary experimental procedures). MIAME-compliant data are available (ArrayExpress, E-MEXP-3663).

Results

To determine how different cells of origin interact with different initiating genetic lesions to drive tumour heterogeneity, we generated mouse cohorts carrying conditional alleles of *Brca2*, *p53* and *Pten* together with either *K14Cre* or *BlgCre*, which preferentially target tumour formation to basal or lumER^{neg} cells, respectively [13]. Cohorts of virgin/parous *BlgCre* animals were established. For additional information about mouse cohorts, cells of origin of the tumours and full tumour details, see supplementary material (Supplementary experimental procedures and Tables S2, S3).

Cell of origin drives tumour phenotype in *Brca2*-deleted mammary tumours

All three cohorts of mice carrying conditional *Brca2* and *p53* alleles (*K14Cre:Brca2$^{f/f}$:p53$^{f/f}$*, virgin *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* and parous *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$*) developed mammary tumours. Median latency was significantly shorter ($p < 0.0001$, log-rank test) in *K14Cre:Brca2$^{f/f}$:p53$^{f/f}$* [197 (range 47–243) days] animals compared with either virgin [242 (range 185–334) days] or parous [275.5 (range 133–445) days] *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* mice (Figure 1A). Significant reduction in conditional *p53* and *Brca2* expression was shown in all tumours relative to control spleens, concordant with deletion of *floxed* exons (see supplementary material, Figure S1). Droplet digital PCR (ddPCR) demonstrated that tumours consistently had fewer copies of floxed *Brca2* and *p53* exons compared to unfloxed exons (see supplementary material, Figure S2). However, the presence of infiltrating immune cells (see supplementary material, Table S2) and likely contamination of tumour samples by other wild-type host cells meant that tumours rarely showed a floxed allele number which approached zero.

BlgCre:Brca2$^{f/f}$:p53$^{f/f}$ tumours were classified mainly as either IDC-NSTs (13/29, 44.8%; Figures 1B, 2A–F) or metaplastic spindle cell tumours (MSCTs; 14/29, 48.3%; Figure 2G–L). Tumours were high-grade, with pushing/mixed borders, a high nuclear pleomorphism, little or no tubule formation and a high mitotic index

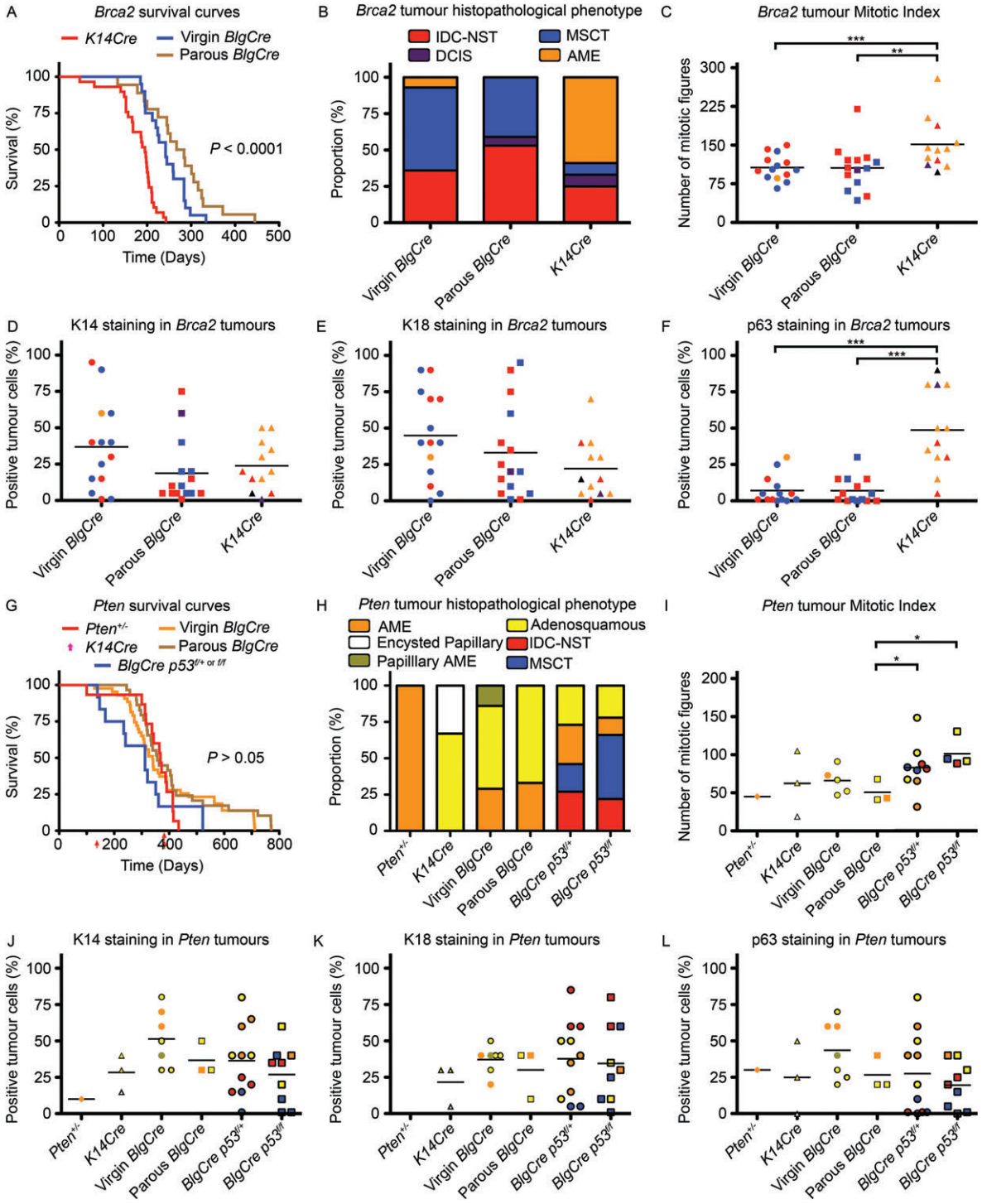

Figure 1. *Brca2*, *Pten* and *Pten p53*-derived tumours have distinct features. (A) Survival curve for *K14Cre:Brca2*$^{f/f}$*:p53*$^{f/f}$ ($n = 29$), virgin ($n = 21$) and parous ($n = 18$) *BlgCre:Brca2*$^{f/f}$*:p53*$^{f/f}$ mice. Only data from animals developing mammary tumours are shown. *K14Cre* tumours had a significantly shorter latency ($p < 0.0001$; log rank test). (B) Relative abundance of histological phenotypes in *Brca2* tumours arising in distinct cells of origin. IDC-NST, invasive ductal carcinoma of no special type; AME, adenomyoepithelioma; DCIS, ductal carcinoma *in situ*; MSCT, metaplastic spindle cell tumour. (C–F) Mitotic index (number of mitotic figures in 10 high-power fields) and percentages of K14-, K18- and p63-positive cells within *Brca2* tumours. Bars indicate median values; data points represent individual tumours and are coloured according to histological phenotype. (G) Survival curve for *Pten*$^{+/-}$ ($n = 15$), *K14cre:Pten*$^{f/f}$ ($n = 3$), virgin ($n = 43$) and parous ($n = 29$) *BlgCre:Pten*$^{f/f}$; and *Blgcre:Pten*$^{f/f}$*:p53*$^{f/+\text{&}f/f}$ ($n = 12$) mice. Only data from animals developing malignant mammary tumours are shown. (H) Relative abundance of histological phenotypes in malignant *Pten* and *Pten p53* tumours. AME, adenomyoepithelioma; Encysted Papillary, encysted papillary carcinoma; Adenosquamous, metaplastic adenosquamous carcinoma; IDC-NST, invasive ductal carcinoma of no special type; MSCT, metaplastic spindle cell tumour. (I–L) Mitotic index and distribution of K14-, K18- and p63-positive cells within *Pten* and *Pten:p53* tumours. Bars indicate median values; data points represent individual tumours and are coloured according to histological phenotype. Statistically significant differences in *t*-tests: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

(MI; Figures 1C, 2B). Most tumours (25/29; 86%) were positive for keratin 14 and keratin 18 (K14 and K18, expressed in basal or luminal cells, respectively, in the normal mammary epithelium) and were weakly positive for p63. The majority (24/29) were also ER-negative (Figures 1D–F, 2C–F, 2I–L). In contrast, histological analysis of *K14Cre:Brca2$^{f/f}$:p53$^{f/f}$* tumours demonstrated that most tumours (7/11; 64%) were malignant AMEs (Figures 1B, 2M–R). All tumours were of high histological grade, with significantly higher MI (Figure 1C; $p < 0.01$, *K14Cre*-vs-parous-*BlgCre*; $p < 0.001$, *K14Cre*-vs-virgin-*BlgCre*, unpaired two-tailed *t*-test), and had multifocal necrosis. Tumours were K14/K18-positive, with a distinguishable distribution of K14- and K18-positive cells in abluminal and luminal cell layers, respectively, consistent with the AME diagnosis (Figures 1D, E, 2O, P). Compared with *BlgCre* tumours, *K14Cre* tumours had stronger p63 staining in significantly more (Figure 1F, $p < 0.001$, unpaired two-tailed *t*-test) cells in each tumour (range 5–90%, predominantly in the abluminal cell layer; Figure 2Q). Of *K14Cre* tumours in which ER staining was determined, half (4/8; 50%) were ER-negative (Figure 2R), three contained ≤ 5% ER-positive cells, but one had around 40% ER-positive cells. In both *BlgCre* and *K14Cre Brca2* tumours, PR staining was concordant with ER staining, although typically fewer cells were PRA-positive than ER-positive, and fewer still were PRB-positive than PRA-positive. Thus, in some cases, weakly ER-positive tumours were PR-negative (see supplementary material, Table S2). Therefore, targeted deletion of *Brca2* and *p53* in basal or lumER^{neg} cells resulted in tumours with different latencies and histopathological features.

The tumour-initiating lesion determines the phenotype of luminal ER^{neg}-origin tumours

Next, we examined *K14Cre:Pten$^{f/f}$*, *BlgCre:Pten$^{f/f}$* and *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* mice as well as germline *Pten* heterozygote mice (*Pten$^{+/-}$*). Due to a strong skin phenotype, only 14 *K14cre:Pten$^{f/f}$* mice could age older than 4 months old. From these, only five mammary tumours were obtained (from four mice). Mammary tumour latencies in *K14Cre:Pten$^{f/f}$* (141–386 days), virgin *BlgCre:Pten$^{f/f}$* [340 (range 128–711) days], parous *BlgCre:Pten$^{f/f}$* [357 (range 245–771) days] and *Pten$^{+/-}$* mice [368 (range 100–434) days] were not significantly different (Figure 1G). However, *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* mice developed tumours [312 (range 139–361) days] significantly faster than parous *BlgCre:Pten$^{f/f}$* mice ($p < 0.05$, log rank test) and also faster, although not significantly, potentially due to the small sample size, than virgin *BlgCre:Pten$^{f/f}$* mice ($p = 0.078$, log rank test; Figure 1G).

All tumours had lower expression for *Pten floxed* exon 4 compared with exon 6 (see supplementary material, Figure S3A, B), confirming recombination of the conditional allele during tumourigenesis. Expression of *p53* exon 4 was higher in *Pten$^{f/f}$* tumours, similar or lower in *Pten$^{f/f}$:p53$^{f/+}$* tumours and always reduced in *Pten$^{f/f}$:p53$^{f/f}$* tumours relative to control spleen (see supplementary material, Figure S3C). This is consistent with *Pten* loss causing p53 induction in *p53$^{+/+}$* mice and a dose-dependent reduction in this response following loss of one or two *p53* alleles [19]. Again, ddPCR demonstrated that tumours from *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* mice had fewer copies of floxed *p53* exons compared to unfloxed exons (see supplementary material, Figure S2). The same caveats regarding infiltrating immune cells apply (see supplementary material, Table S2). For technical reasons, the ddPCR assay could not be performed on the floxed *Pten* allele.

Pten depletion generated both malignant and benign neoplasms regardless of the origin cell type (see supplementary material, Tables S2, S3). Benign tumours were classified as sclerosing adenosis and benign AMEs (Figure 3A–F), with 5/15 (33%) displaying papillary architecture (see supplementary material, Figure S4A–F). All showed strong K14, K18, p63 and ER staining (Figure 3C–F; see also supplementary material, Figure S4C–J).

In contrast to IDC-NSTs/MSCTs developing from lumER^{neg} cells in *Brca2:p53* mice, malignant lumER^{neg}-origin *Pten$^{f/f}$* tumours were AMEs (4/10; 40%) or metaplastic adenosquamous carcinomas (ASQCs; 4/10; 40%), with two additional tumours showing both elements (Figures 1H, 3G–R; see also supplementary material, Table S2). A subset displayed papillary architecture (see supplementary material, Figure S4L–Q). Tumours had pushing/mixed borders, central/multifocal necrosis and low/intermediate histological grades with intermediate nuclear pleomorphism, tubule formation and MI (Figure 1I). Metaplastic squamous cells were found in 8/10 (80%) tumours (5–75% cells; Figure 3N). All tumours were positive for K14, K18 and p63 (Figures 1J–L, 3I–K, 3O–Q). Remarkably, strong ER expression was seen in 7/8 (87.5%) of analysed malignant tumours (15–40% cells, Figure 3L, R; see also supplementary material, Figure S5). PRA expression was observed in 8/9 analysed tumours (1–30% cells; see supplementary material, Figure S5). Notably, both malignant *Pten$^{+/-}$* and *K14Cre:Pten* tumours had similar phenotypes to lumER^{neg} cell-origin tumours (see supplementary material, Table S2).

Addition of *p53* conditional alleles into the *BlgCre:Pten* cohort increased the ratio of malignant:benign tumours, with 20/21 (95%) tumours being malignant (see supplementary material, Table S2). It also shifted the spectrum of histopathological phenotypes closer to that seen with *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* cohorts (Figure 1B, H), as half of all *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* tumours were classed as either IDC-NSTs (5/20; Figure 4A–F) or MSCTs (6/20; Figure 4G–L). The remainder were diagnosed as malignant-AMEs (2/20, 10%; Figure 4M–R), ASQCs (4/20, 20%) or mixed tumours (3/20, 15%).

As opposed to the low/intermediate histological grades of *Pten* tumours, *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* carcinomas showed high histological grade, with high

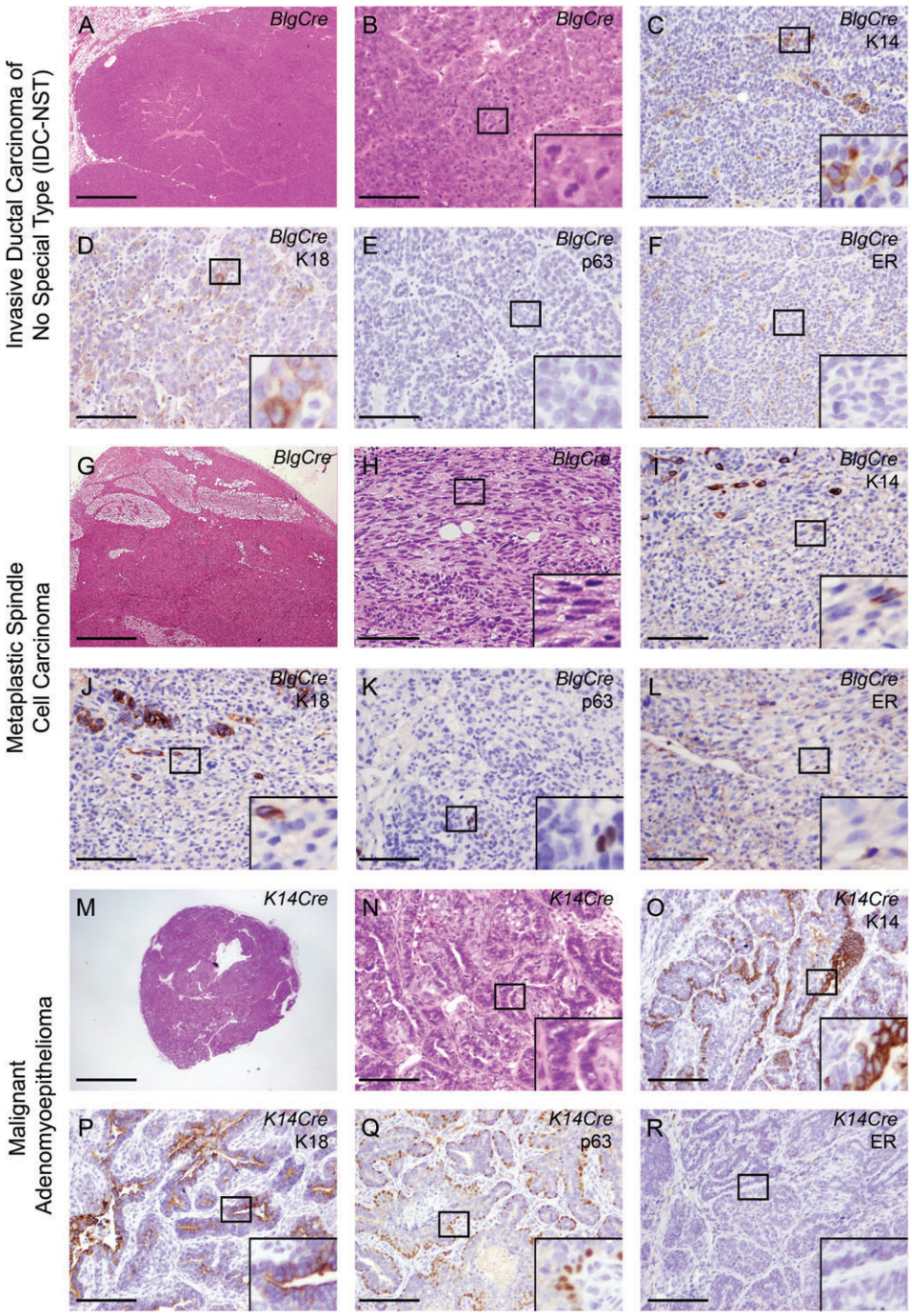

Figure 2. *Brca2* depletion in the mammary epithelium results in a spectrum of tumour phenotypes dependent on the cell of origin. (A–F) *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* IDC-NST. (A) Low-power H&E-stained section showing tumour with mixed borders. (B) High-power H&E-stained section showing mitotic figures and nuclear pleomorphism. (C–F) Staining for K14 (C), K18 (D), p63 (E) and ER (F). (G–L) *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* metaplastic spindle cell carcinoma: (G) low-power H&E-stained section showing tumour with central necrosis; (H) high-power H&E-stained section showing neoplastic epithelioid and spindloid cells; (I, J) K14 (I) and K18 (J) expression in epithelioid cells, but lack of expression in spindle cells; (K) p63 staining; (L) ER staining. (M–R) *K14Cre:Brca2$^{f/f}$:p53$^{f/f}$* malignant AME: (M, N) low-power (M) and high-power (N) H&E-stained sections illustrating the pushing margins of a tumour; (O) K14 staining in basal-like neoplastic population; (P) K18 staining in luminal-like neoplastic population; (Q) strong p63 staining in basal-like neoplastic cells; (R) lack of ER staining in neoplastic cells. Bars = (A, G, M) 1.5 mm; (B–F, H–L and N–R) 100 μm. Inset boxes are magnified ×3. (See also supplementary material, Tables S2, S3, Figure S1)

nuclear pleomorphism, lack of tubule formation and high MI (Figure 1I). Tumours had mixed borders, with central/multifocal necrosis. Spindle (5–100%) and squamous (1–50%) metaplastic cells were seen in all tumours. All were K14/K18-positive (Figure 1J, K), but staining tended to be at low levels in IDC-NSTs (5–25%; Figure 4C, D), at very low levels in MSCTs (1–10%; Figure 4I, J) and at the highest levels in AMEs (40–60%; Figure 4O, P). p63 staining was positive in 18/20 (90%) tumours (1–80% cells) (Figures 1L, 4E, K, Q). Again, IDC-NSTs and MSCTs had few p63-positive cells and these were scattered through

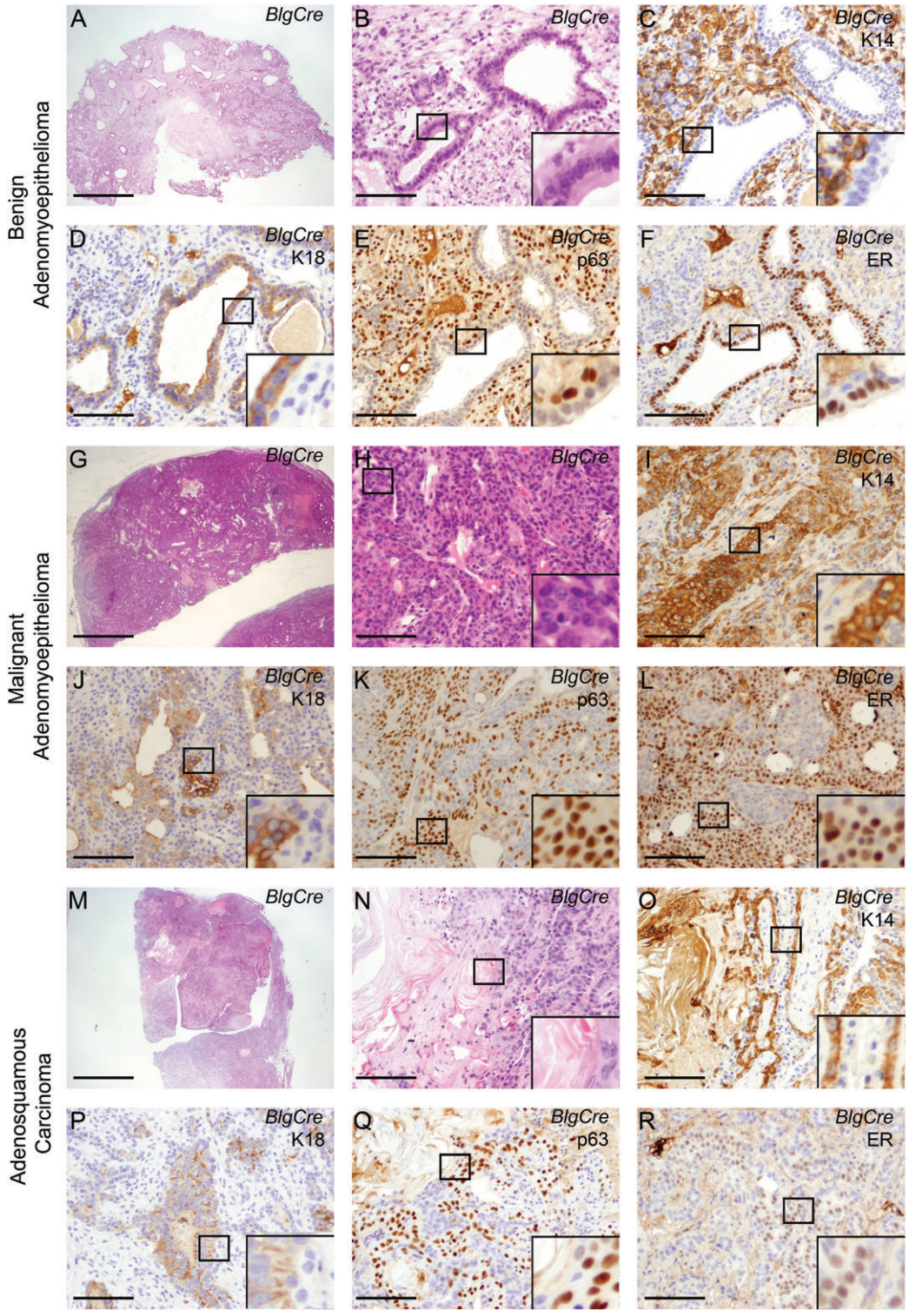

Figure 3. *Pten* depletion in luminal progenitors generates benign and malignant tumours. (A–F) Features of *BlgCre:Pten*$^{f/f}$ benign AME: (A, B) low-power (A) and high-power (B) H&E-stained sections; (C) strong K14 staining in proliferating basal cells encasing the glandular structures and in spindle cells; (D) strong K18 staining in luminal cells; (E) strong p63 staining in basal cells; (F) strong ER staining in luminal cells. (G–L) Features of *BlgCre:Pten*$^{f/f}$ malignant AME: (G) low-power H&E-stained section showing multifocal necrosis and pushing borders; (H) high-power H&E-stained section showing gland-like structures and expansion of the abluminally located cells – note increased nuclear pleomorphism; (I) strong K14 staining in proliferating basal cells and in spindle cells; (J) strong K18 staining in luminal cells; (K) p63 and (L) ER staining; note the similarity of expression pattern of p63 and ER. (M–R) Features of a *BlgCre:Pten*$^{f/f}$ tumour, showing an adenosquamous carcinoma clone originating from an AME: (M) low-power H&E-stained section showing multifocal necrosis and mixed borders; (N) high-power H&E-stained section showing metaplastic squamous elements emerging from AME area; (O) K14 staining in neoplastic basal-like cells; (P) K18 staining in neoplastic luminal cells; (Q) strong p63 expression in epithelioid cells; (R) ER staining in neoplastic epithelioid cells. Bars = (A, G, M) 1.5 mm; (B–F, H–L and N–R) 100 μm. Inset boxes are magnified ×3. (See also supplementary material, Tables S2, S3, Figures S3, S4)

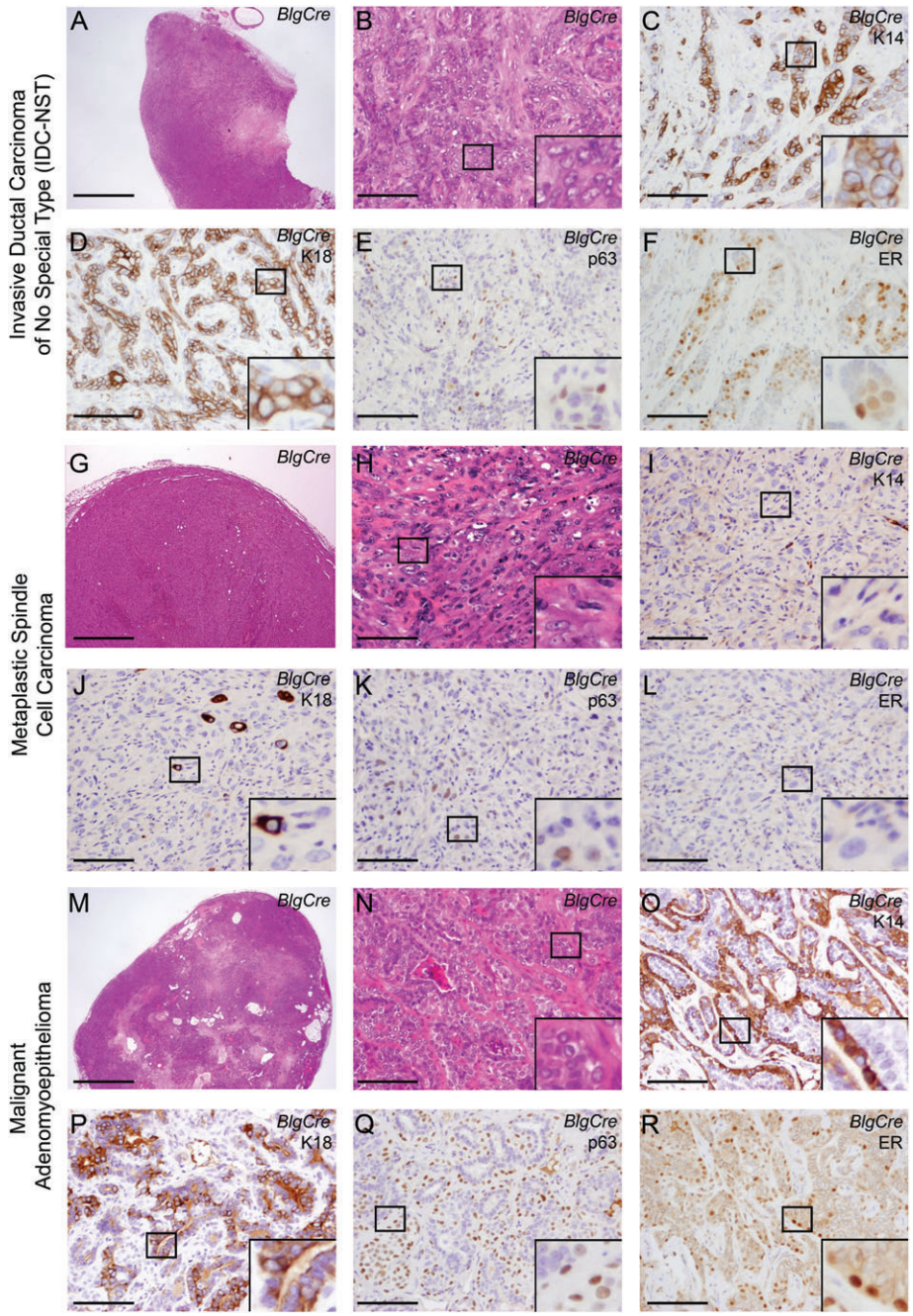

Figure 4. *p53* loss alters tumour phenotypes in *Pten* knockout mice. (A–F) Features of a *BlgCre:Pten*$^{f/f}$*:p53*$^{f/f}$ IDC-NST: (A) low-power H&E-stained section showing central necrosis and mixed borders; (B) high-power H&E-stained section showing aberrant proliferation of highly pleomorphic neoplastic epithelioid cells; (C) K14 staining in neoplastic epithelioid cells; (D) K18 staining in neoplastic epithelioid cells; (E) p63 and (F) ER; note staining of epithelioid neoplastic cells. (G–L) Features of a *BlgCre:Pten*$^{f/f}$*:p53*$^{f/+}$ metaplastic spindle cell carcinoma: (G) low-power H&E-stained section showing multifocal necrosis and mixed borders; (H) high-power H&E-stained section showing abundant spindle cells; (I) lack of K14 expression in spindle cells; (J) K18 expression in small nests of epithelioid cells, no K18 staining in spindle cells; (K) p63; (L) tumour cells are negative for ER. (M–R) Features of a *BlgCre:Pten*$^{f/f}$*:p53*$^{f/+}$ malignant AME: (M) low-power H&E-stained section showing multifocal necrosis and pushing borders; (N) high-power H&E-stained section illustrating heterogeneous neoplastic populations; (O) K14 staining in abluminal cells; (P) K18 expression in pseudo-luminal cells; (Q) strong p63 expression in proliferating abluminal cells; (R) ER expression in cells located in the abluminal layer. Bars = (A, G, M) 1.5 mm; (B–F, H–L and N–R) 100 µm. Inset boxes are magnified ×3. (See also supplementary material, Tables S2, S3, Figures S5, S6)

the tumour, whereas AMEs had high levels of p63 staining organized into distinct abluminal epithelial layers and 'nests' of p63-positive neoplastic cells (Figure 4E, K, Q). ER staining was intense and frequent in tumour cells in 17/20 (85%) carcinomas (1–30% cells; see supplementary material, Figure S5) including IDC-NSTs (Figure 4F, R), but was either absent or expressed at low levels in MSCTs. Like the *Brca2* cohorts, PR staining was concordant with ER staining, with fewer cells PRA-positive than ER-positive, and fewer still PRB-positive than PRA-positive (see supplementary material, Figure

S5, Table S2). Double immunofluorescence staining of benign and malignant *BlgCre:Pten$^{f/f}$* AMEs, as well as malignant *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* AMEs, demonstrated that in benign tumours p63-positive and ER-positive cell populations were mutually exclusive, but in malignant tumours most p63-positive cells were also ER-positive (see supplementary material, Figure S6). This double positivity for ER/p63 suggests an aberrant differentiation in these tumour cells.

Therefore, *BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* and *BlgCre:Pten$^{f/f}$* mouse models showed distinct differences in tumour latency and phenotype, despite the initiating genetic lesions being targeted to the same cell population. This demonstrated that in these cases the cell of origin was not the sole determinant of tumour phenotype. Rather, the initiating genetic hits underpinned tumour behaviour and phenotype. Targeted deletion of both *Pten* and *p53* to lumER^{neg} cells accelerated tumour formation and, notably, resulted in a range of phenotypes that once again included IDC-NSTs and MSCTs. *BlgCre:Pten$^{f/f}$:p53$^{f/+\&f/f}$* IDC-NSTs, however, were strongly ER-positive, unlike IDC-NSTs from other cohorts (*BlgCre:Brca2$^{f/f}$:p53$^{f/f}$* and *BlgCre:Brca1$^{f/f}$:p53$^{+/-}$*) [13,20].

Luminal ER^{neg}-origin tumours display diverse molecular profiles determined by the initiating genetic lesion

We performed whole-transcriptome analysis of a subset of tumours from each genotype (including a previous collection of *K14Cre:Brca1$^{f/f}$:p53$^{+/-}$* and *BlgCre:Brca1$^{f/f}$:p53$^{+/-}$* tumours) [13], using the Affymetrix MouseChip Genome platform. Unsupervised hierarchical clustering showed that the tumours broadly clustered into three molecular groups (Figure 5). One included the *Brca2:p53* and *Pten:p53* tumours; the second group consisted of most of *Pten*-only tumours; the third group included *Brca1*:*p53* and some *Pten* tumours. Pairwise SAM comparisons between groups delivered a list of significantly-associated genes, which were interrogated for GO terms and KEGG pathway analysis (see supplementary material, Table S4). The *Brca2:p53*/*Pten:p53* group (group 1) and the *Brca1*:*p53* group (group 3) genes were highly enriched for GO Bioprocess annotations associated with transcription, metabolism, biosynthesis and regulation of cell death. In contrast, the group 2 (*Pten*) genes were enriched for development, homeostasis, signalling and regulation of cell death bioprocesses and expressed genes involved in 'response to hormone stimulus' and 'steroid metabolic process'. Pathway analysis showed a great similarity between all tumour groups (see supplementary material, Table S4), although with some differences. For instance, group 1 was enriched for genes associated with adhesion, junctional complexes and JAK–STAT signalling pathways, group 2 with genes associated with calcium signalling and vascular smooth muscle pathways, and group 3 with genes associated with the cell cycle and DNA replication pathways. Interestingly, genes for cysteine and methionine metabolism pathways were enriched in groups 1 and 3, while genes for glycine, serine, threonine and tyrosine metabolism pathways were enriched in group 2, suggesting fundamental differences in the metabolism of these tumour groups.

Importantly, these molecular clusters were determined by the initiating genetic lesion (Figure 5C), with expression profiles being consistent across tumours carrying the same initiating lesion. Tumours with different lesions were not randomly interspersed, neither did tumours cluster by *Cre* promoter. Thus, the tumour molecular profile was governed by its initiating genetic lesion, not by the cell to which those lesions were targeted.

Luminal ER^{neg} cells generate basal-like, 'normal breast-like', luminal A and luminal B tumours

We next asked which human breast cancer molecular subtypes the mouse tumours of this analysis most closely resembled, using a single sample predictor gene set (SSP) [18] (Table 1, Figure 5E; see also supplementary material, Table S5). Consistent with their lack of ER expression, 9/13 (70%) *Brca2:p53* mouse tumours classed as basal-like using the PAM50 gene set, irrespective of whether they were from the *K14Cre* or *BlgCre* cohorts. Of the *Pten* tumours, 17/21 (81%) tumours were categorized as 'normal breast-like', three tumours classed as luminal A and one as basal-like. Conversely, *Pten:p53* tumours were classified as luminal B (4/10), 'normal breast-like' (3/10), luminal A (2/10) and one could not be assigned to any subtype. Differences in the proportions of the predominant subtypes within each genotype were highly significant ($p < 0.0001$, χ^2 test) in pairwise genotype comparisons (Table 1).

PAM50 analysis is sensitive to sample cohort normalization issues [17]. We therefore interrogated different human breast tumour transcriptome datasets [1,3,18,21–23], including three enriched for *BRCA1/2* mutation carriers [24–26], using mouse tumour transcriptome signatures. We built three mouse molecular signatures based upon the top probes up- and down-regulated within each mouse group (*Pten* only; *Brca1:p53* only; combined *Brca2:p53*/*Pten*$^{-}$*p53*), identified by SAM pairwise comparisons. Signatures were applied to each sample from each dataset. Correlation heat maps for the mouse transcriptome signature in the human datasets (Figure 5F; see also supplementary material, Figure S8) confirmed, first, that the *Brca1:p53* mouse signature was associated with the human basal-like subtype and with human *BRCA1* breast cancers; second, that luminal A, normal breast-like and non-*BRCA1/2* cancers were enriched in breast cancer samples with a gene signature similar to the *Pten* mouse tumours; and third, that the *Brca2:p53*/*Pten:p53* signature was observed across the range of human breast cancer molecular subtypes. Notably, when testing human breast cancer datasets that included the claudin-low subtype, a particular enrichment for the *Brca2:p53*/*Pten:p53* signature was

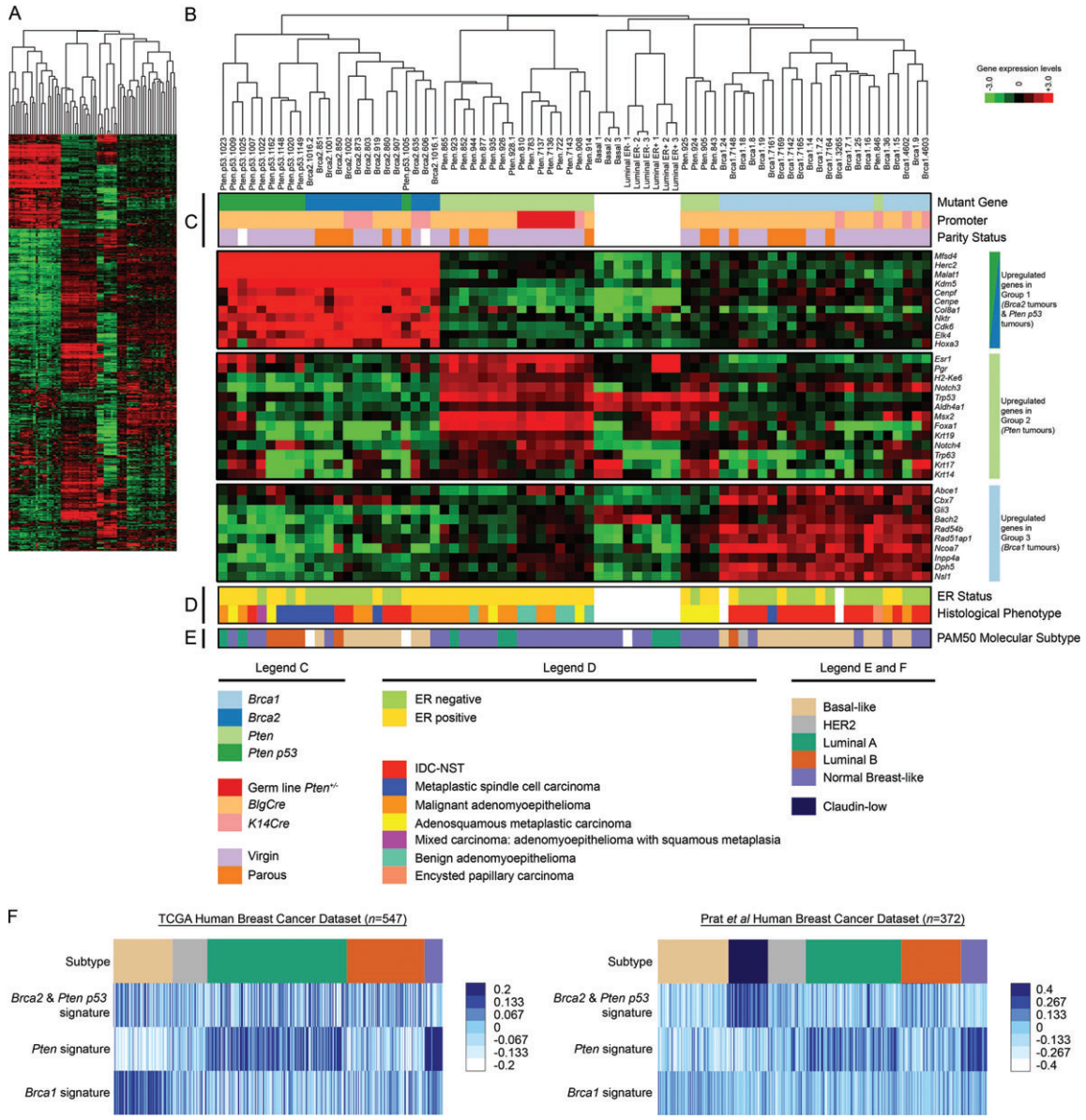

Figure 5. The initiating genetic lesion is the primary determinant of the molecular expression pattern of the resulting tumour. (A) Thumbnail image of the unsupervised hierarchical clustering of 4657 Affymetrix probes. (B) Zoomed image of the gene expression heat map, showing a selection of significantly up-regulated genes which define each tumour cluster. (C) Initiating genetic lesion, *Cre* promoter and virgin/parity status information is colour-coded. (D) Further dendrogram branches can be defined by ER status and histological phenotype. (E) Summary of results of SSP analysis using the PAM50 dataset [18] to identify the human breast cancer subtypes the mouse tumours most closely resemble; white squares indicate no association. Expression data of normal mouse populations and *Brca1* tumours was taken from previously published work [13,34]. Pvclust analysis confirmed that the stability of the three main tumour molecular clusters was > 90% (see supplementary material, Figure S7 and Supplementary experimental procedures, for confirmatory analysis that microarray batch variation did not affect clustering). (F) Analysis for enrichment of up- and down-regulated gene sets in mouse signatures of groups 1 (enriched in *Brca2* and *Pten:p53* tumours), 2 (enriched in *Pten* tumours) and 3 (enriched in *Brca1* tumours) in human breast cancer datasets [3,22]. Spearman rank correlation values for each signature were plotted against dataset molecular phenotypes as correlation heat maps. Note that group 2 and 3 signatures correlated with the human luminal A/normal breast-like and the basal-like subtypes, respectively (both heat maps), whereas the group 1 signature was highly correlated with the human claudin-low subtype (right heat map). (See also supplementary material, Figure S8)

obtained in this group. The *Brca2:p53/Pten:p53* signature was not enriched in human *BRCA2* tumours; indeed, in one study [25] the association was with *BRCA1* tumours.

These results showed that tumours deriving from the same cell of origin, lumER^{neg} cells, not only had very different molecular features, depending on the initiating genetic lesions, but also spanned a broad range of human-equivalent molecular signatures. Hence, the 'intrinsic subtype' classification of a tumour does not necessarily reflect its cell of origin.

Luminal ER^{neg} cells generate 'claudin-low' tumours

The claudin-low subtype is not distinguished by the PAM50 gene set. This subtype is characterized by

Table 1. Correlation of tumour molecular profiles to human breast cancer molecular phenotypes

	PAM50 correlation to human molecular subtypes *n* (%)					
	Basal	Her2	Luminal A	Luminal B	Normal breast-like	Undefined
Brca1 tumours (*n* = 21)	14 (66.67)*	1 (4.76)	0 (0)	1 (4.76)	5 (23.81)	0 (0)
K14Cre (*n* = 3)	2 (66.67)	0 (0)	0 (0)	0 (0)	1 (33.33)	0 (0)
BlgCre (*n* = 18)	12 (66.67)	1 (5.56)	0 (0)	1 (5.56)	4 (22.22)	0 (0)
Virgin (*n* = 12)	7 (58.33)	1 (8.33)	0 (0)	0 (0)	4 (33.33)	0 (0)
Parous (*n* = 6)	5 (83.33)	0 (0)	0 (0)	1 (16.67)	0 (0)	0 (0)
Brca2 tumours (*n* = 13)	9 (69.23)*	0 (0)	0 (0)	1 (7.69)	2 (15.38)	1 (7.69)
K14Cre (*n* = 5)	5 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
BlgCre (*n* = 8)	4 (50)	0 (0)	0 (0)	1 (12.5)	2 (25)	1 (12.5)
Virgin (*n* = 4)	2 (50)	0 (0)	0 (0)	0 (0)	1 (25)	1 (25)
Parous (*n* = 4)	2 (50)	0 (0)	0 (0)	1 (25)	1 (25)	0 (0)
Pten tumours (*n* = 21)	1 (4.76)	0 (0)	3 (14.29)	0 (0)	17 (80.95)**	0 (0)
$Pten^{+/-}$ (*n* = 6)	0 (0)	0 (0)	0 (0)	0 (0)	6 (100)	0 (0)
$Pten^{f/f}$ (*n* = 15)	1 (6.67)	0 (0)	3 (20)	0 (0)	11 (73.33)	0 (0)
K14Cre (*n* = 2)	1 (50)	0 (0)	0 (0)	0 (0)	1 (50)	0 (0)
BlgCre (*n* = 13)	0 (0)	0 (0)	3 (23.08)	0 (0)	10 (76.92)	0 (0)
Virgin (*n* = 7)	0 (0)	0 (0)	2 (28.57)	0 (0)	5 (71.43)	0 (0)
Parous (*n* = 6)	0 (0)	0 (0)	1 (16.67)	0 (0)	5 (83.33)	0 (0)
Pten p53 tumours (*n* = 10)	0 (0)	0 (0)	2 (20)	4 (40)***	3 (30)	1 (10)
$p53^{f/+}$ (*n* = 6)	0 (0)	0 (0)	2 (33.33)	1 (16.67)	2 (33.33)	1 (16.67)
$p53^{f/f}$ (*n* = 4)	0 (0)	0 (0)	0 (0)	3 (75)	1 (25)	0 (0)

Numbers and percentages of tumours most closely correlating to each category are shown. Undefined indicates samples not classified into any centroid group due to a Spearman correlation rank < 0.1. Mouse tumours are listed by genotype, *Cre* promoter and parity status. Shaded cells indicate the modal correlation.
*Basal versus non-basal; χ^2 test, $p < 0.0001$.
**Normal versus non-normal; χ^2 test, $p < 0.0001$.
***Luminal B versus non-luminal B; χ^2 test, $p < 0.0001$ (see supplementary material, Table S4).

up-regulation of mesenchymal-associated genes and down-regulation of genes related to epithelial cell–cell junctions, particularly claudins *CLDN3*, *−4* and *−7* and *CDH1* [22]. As a 'mesenchymal-like' appearance was typical of the MSCTs from our tumour cohorts, and enrichment for the *Brca2:p53*/*Pten:p53* signature, both tumour genotypes with high numbers of MSCTs, was observed in the claudin-low subtype in breast cancer datasets which included that group, we analysed expression of *Cldn3*, *4* and *7* and *Cdh1* across the tumour panel categorized by histological phenotype. The results confirmed that MSCTs had significantly lower expression levels of these four genes compared to other tumour types (Figure 6), and indeed of the whole geneset reported as down-regulated in the claudin-low phenotype [22] (see supplementary material, Figure S9). This demonstrates that the transcriptomic signature of MSCTs recapitulates that of claudin-low tumours, suggesting that this tumour type can also originate from $lumER^{neg}$ cells.

Human metaplastic tumours have variable PTEN expression but express low claudin levels

Using a pilot cohort of human breast cancers, including some very rare human AMEs, we examined whether there was an association between PTEN expression and the human histological phenotypes equivalent to those in our mouse cohorts. We found that staining of metaplastic tumours and IDC-NSTs was variable, but the few human AMEs we examined were very strongly PTEN-positive (see supplementary material, Table S6, Figure S10).

We also examined the same tumour group for expression of CLDN3, CLDN4 and CDH1. Unlike PTEN staining, these results were concordant with the mouse data, as human IDC-NSTs expressed high protein levels, but there was absence of expression in non-epithelial areas of spindle-cell carcinomas and metaplastic carcinomas with mesenchymal differentiation (see supplementary material, Table S6, Figure S10).

Discussion

Inter-tumour heterogeneity must arise from different (epi)genetic lesions occurring in different cells of origin. Here, we have applied histopathological and molecular pathology approaches to analyse tumours arising in genetically engineered mouse models from different initiating lesions in distinct cells of origin. We show that, in our model system, targeting tumour-initiating lesions to basal cells results primarily in adenomyoepitheliomas, whereas targeting $lumER^{neg}$ cells results in tumours with a range of histopathological features including metaplastic tumours and invasive ductal carcinomas [4]. Importantly, we have generated both ER-positive luminal-like and ER-negative basal-like tumours from this target population. We have also shown that the initial genetic lesion is the prime determinant of the molecular profile of the subsequent tumours arising from these cells. This suggests that, rather than being a truly stochastic process, the aetiology of tumour formation is largely deterministic and depends on the earliest events in carcinogenesis (ie the founder genetic/epigenetic events).

Germline mutations in human *BRCA2* predispose to breast and ovarian cancers [27]. Although 66–93%

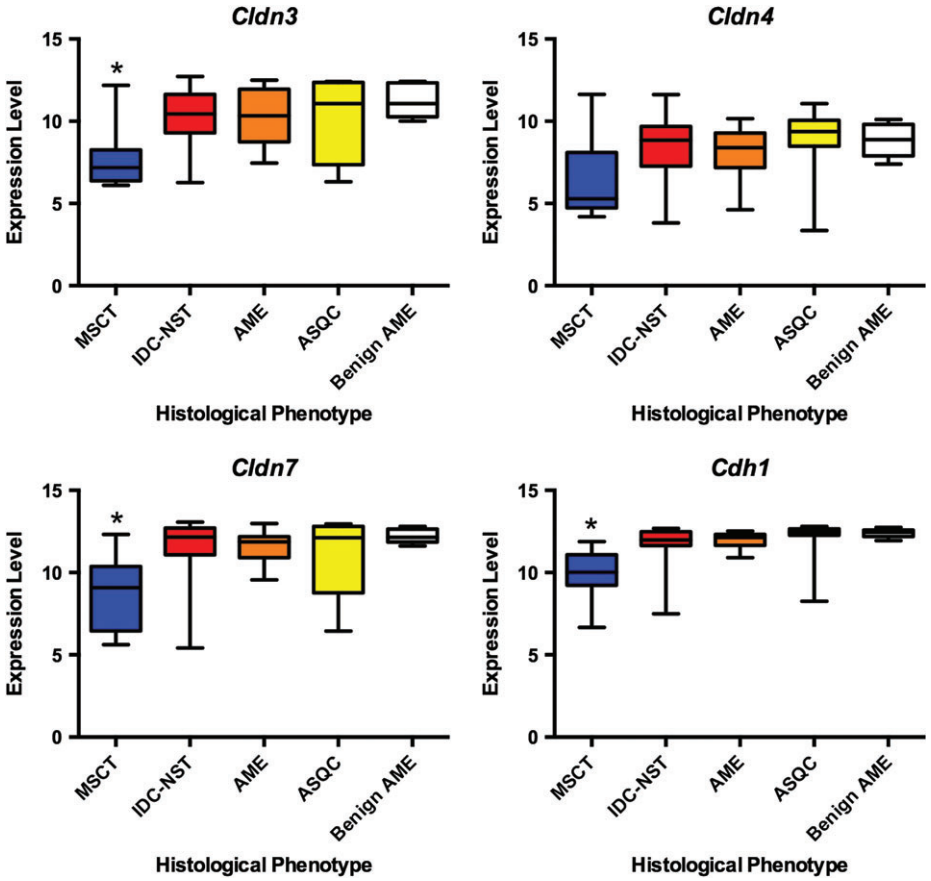

Figure 6. Low expression of claudin-related genes in metaplastic spindle cell carcinomas. Box plots showing expression levels for mouse orthologues of human genes (*Cldn3*, *Cldn4*, *Cldn7* and *Cdh1*) characteristically down-regulated in the human claudin-low molecular subtype. Mouse tumours were grouped in five categories, based solely on histological phenotype: MSCT, metaplastic spindle cell tumours; IDC-NST, invasive ductal carcinoma of no special type; AME, malignant adenomyoepithelioma; ASQC, adenosquamous carcinoma; Benign AME, benign adenomyoepithelioma. Note that MSCTs overall show lower expression for each of the genes as compared with the other phenotypes. These differences were statistically significant (one-way ANOVA, $^{*}p < 0.05$) for all genes except *Cldn4* ($p = 0.056$). (See also supplementary material, Figure S9)

of *BRCA2*-associated human cancers are ER-positive in > 10% tumour cells [28,29], only 3/15 (20%) of the *Brca2* IDC-NSTs described here were ER-positive (1–10% tumour cells). As the same cells of origin could generate ER-positive IDC-NSTs in the *Pten*:*p53* model, this is not likely explained by a lack of potential to differentiate along this lineage. Moreover, most of the *Brca2* mouse tumours had a molecular profile similar to human basal-like breast cancers (Table 1, Figure 5F) and did not resemble a typical human BRCA2 tumour profile (see supplementary material, Figure S8). It should be noted, however, that a subset (13–19%) of human *BRCA2*-mutated breast cancers have a basal-like molecular profile and are also ER-negative [30], which would be consistent with the similarity of the *Brca2:p53*/*Pten:p53* signature to human BRCA1 tumours in data from one study [25]. It is possible that *BlgCre:Brca2:p53* tumours model the basal-like subset of human *BRCA2* breast cancers.

Loss of PTEN expression is recurrent in human breast cancers, in both basal and luminal subtypes [3]. In our study, targeting conditional depletion of *Pten* alone to mouse lumER^{neg}-cells resulted in a different effect to *Brca1/2*:*p53* loss, leading to the development of benign and malignant AMEs and ASQCs. Notably, AMEs were highly differentiated and ER-positive. In contrast, analysis of a pilot cohort of human breast cancers, including very rare human AMEs, found strong PTEN expression in these tumours. Larger numbers are required to confirm these findings, but they suggest that human AMEs are not associated with somatic PTEN loss, unlike in the mouse. Notably, however, breast tumours from germline PTEN loss-syndrome families are enriched for molecular apocrine differentiation, which is characterized by elevated levels of androgen signalling [31], and our mouse Pten tumour cohort also expressed high levels of the androgen receptor (see supplementary material, Table S4). The mouse tumour phenotypes were altered when conditional *Pten*:*p53* alleles were combined, resulting in the development of MSCTs and ER-positive IDC-NSTs. In these tumours, the molecular changes observed (loss of claudins in MSCTs) were reflected in the equivalent human tumours.

Our findings show that a broad spectrum of tumour phenotypes can emerge from the lumER^{neg} cell population, they suggest that p53 loss-of-function is a

prime driver of histopathological phenotype, and they demonstrate that, in contrast, cell of origin is not a strict driver of tumour phenotype. Our results are consistent with the notion that mammary tumour heterogeneity is a result of context-dependent interactions between cell of origin and early genetic hits. In the K14Cre basal-origin tumours we describe, the AME phenotype is the default tumour type, irrespective of the driving genetic lesion. Conversely, lumER^{neg} cells are able to generate a broad spectrum of tumour histological and molecular phenotypes, including highly aggressive ER/PR-negative and ER/PR-positive neoplasms. Since tumours had long latency periods, and additional genetic mutations must have arisen in all genetic backgrounds to permit tumour formation, the stability of histological phenotypes within each genetic background was notable. Either any additional genetic hits were stochastic and had little effect on overall tumour phenotype, or each cell of origin/genetic background combination developed a set of stereotypical lesions that contributed to the tumour phenotype. Future massively-parallel sequencing studies may lead to a deeper understanding of the mutational changes in these genetic backgrounds.

Interestingly, other groups described *K14Cre*-driven models (*K14Cre:Brca1$^{f/f}$:p53$^{f/f}$* and *K14Cre:Ecad$^{f/f}$:p53$^{f/f}$*) [32,33], in which the predominant tumour phenotype was not an AME but rather a more typical luminal-like tumour. We have discussed this issue previously [13] but our current results support a model in which the *Brca1* and *Ecad* alleles used by Liu and colleagues and Derksen and colleagues are dominant over the *K14Cre* cell of origin in driving tumour phenotype, in a way in which the alleles we have used are not.

Our study has important limitations. While the *BlgCre* transgene preferentially drives tumour formation in lumER^{neg} cells, we cannot definitively exclude that promoter 'leakiness' may, in a modest number of cases, result in tumours originating from other cell types, or that initial gene deletions may affect cell differentiation and thus alter the phenotype of the cell that finally transforms (see supplementary material, Supplementary experimental procedures); whereas equivalent mouse and human mammary epithelial cell types can be inferred (ie cells which are luminal or basal, ER-positive or ER-negative), the cell types in which allele recombination occurs in the mouse have not been directly mapped to human cell types; the mouse strains we have used, while mainly on a C57Bl6 background, are not pure-bred (see supplementary material, Supplementary experimental procedures) and there may be background strain-specific alleles linked to the conditional alleles which could affect tumour phenotypes; in our models, and in all current mouse models involving more than one conditional allele, it is not possible to control the order in which allele recombination occurs; and finally, we have not yet observed tumours that resemble sporadic human ER-positive IDC-NSTs with a luminal A molecular profile. We hypothesize that either lumER^{pos} progenitors will need to be targeted as the cell of origin for this tumour type, or that these tumours are simply too indolent to be modelled within the mouse lifespan. In general we note that, while mouse models are important as models of breast cancer, mice are not humans and caution must be exercised in extrapolating results between species, as is illustrated by the case of PTEN expression in AMEs.

Despite these limitations, this study does provide a fundamental advance in our understanding of the origins of mammary tumour heterogeneity. We provide multiple lines of evidence to demonstrate that the phenotype of a cancer is not a mere reflection of its cell of origin, calling into question conclusions about the histogenesis of malignancies derived from histopathological, immunophenotypical and transcriptomic analyses of fully developed tumours.

Acknowledgements

This study was supported by Breakthrough Breast Cancer, Cancer Research UK and a Marie-Curie Fellowship (to LM). LM was funded by Fundación Ramón Areces and Marie Curie Actions (Grant No. IEF-236788). MJS is supported by Cancer Research UK, Breast Cancer Campaign, the European Cancer Stem Cell Research Institute and Cardiff University. The authors would like to thank the Breakthrough Breast Cancer Histopathology Unit for their work, as well as Suzanne Carreiras for her technical assistance.

Author contributions

LM and GM bred mouse lines, carried out post-mortem analyses, collected tumours, prepared RNA, carried out qPCR and bioinformatics analyses and assisted with experimental design, analysis of pathology and writing of manuscript; FAM carried out histopathological staining and assisted with its interpretation; HK genotyped mice, assisted with qPCR gene expression analysis and interpretation of qPCR data; DN-R and FM reviewed histopathological material; DR optimized and carried out double immunofluorescence staining for p63 and ER and assisted with its interpretation; KG optimized and carried out PRA and PRB staining and assisted with its interpretation; MA, MAL-G and JP carried out staining and analysis of human tissue samples; AM carried out bioinformatic analysis of gene expression array data and assisted with its interpretation; JSR-F analysed mouse tumour pathology, advised on its interpretation and assisted with writing the manuscript; and MJS designed the study, assisted with analysis of tumour pathology and bioinformatic analysis and wrote the manuscript.

References

1. Sorlie T, Tibshirani R, Parker J, *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003; **100**: 8418–8423.

2. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet* 2011; **378**: 1812–1823.
3. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012; **490**: 61–70.
4. Lakhani SR, Ellis IO, Schnitt SJ, *et al. WHO Classification of Tumors of the Breast*, 4th edn. IARC Press: Lyon, France, 2012.
5. Visvader JE. Cells of origin in cancer. *Nature* 2011; **469**: 314–322.
6. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* 2011; **5**: 5–23.
7. Melchor L, Benitez J. An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. *Carcinogenesis* 2008; **29**: 1475–1482.
8. Smalley M, Ashworth A. Stem cells and breast cancer: a field in transit. *Nat Rev Cancer* 2003; **3**: 832–844.
9. Foulkes WD. BRCA1 functions as a breast stem cell regulator. *J Med Genet* 2004; **41**: 1–5.
10. Dontu G, Al-Hajj M, Abdallah WM, *et al.* Stem cells in normal breast development and breast cancer. *Cell Prolif* 2003; **36**(suppl 1): 59–72.
11. Lim E, Vaillant F, Wu D, *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in *BRCA1* mutation carriers. *Nat Med* 2009; **15**: 907–913.
12. Proia TA, Keller PJ, Gupta PB, *et al.* Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell* 2011; **8**: 149–163.
13. Molyneux G, Geyer FC, Magnay FA, *et al.* BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* 2010; **7**: 403–417.
14. Robertson D, Savage K, Reis-Filho JS, *et al.* Multiple immunofluorescence labelling of formalin-fixed paraffin-embedded (FFPE) tissue. *BMC Cell Biol* 2008; **9**: 13.
15. Maiques O, Santacana M, Valls J, *et al.* Optimal protocol for PTEN immunostaining: role of analytical and pre-analytical variables in PTEN staining in normal and neoplastic endometrial, breast and prostatic tissues. *Human Pathology* 2014; **45**: 522–532.
16. Castilla MÁ, Díaz-Martín J, Sarrió D, *et al.* MicroRNA-200 Family modulation in distinct breast cancer phenotypes. *PLoS One* 2012; **7**: e47709.
17. Weigelt B, Mackay A, A'Hern R, *et al.* Breast cancer molecular profiling with single sample predictors: a retrospective analysis. *Lancet Oncol* 2010; **11**: 339–349.
18. Parker JS, Mullins M, Cheang MC, *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009; **27**: 1160–1167.
19. Berger AH, Pandolfi PP. Haplo-insufficiency: a driving force in cancer. *J Pathol* 2011; **223**: 137–146.
20. McCarthy A, Savage K, Gabriel A, *et al.* A mouse model of basal-like breast carcinoma with metaplastic elements. *J Pathol* 2007; **211**: 389–398.
21. Hu Z, Fan C, Oh DS, *et al.* The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genom* 2006; **7**: 96.
22. Prat A, Parker JS, Karginova O, *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 2010; **12**: R68.
23. van de Vijver MJ, He YD, van't Veer LJ, *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; **347**: 1999–2009.
24. Waddell N, Arnold J, Cocciardi S, et al. Subtypes of familial breast tumours revealed by expression and copy number profiling. *Breast Cancer Res Treat* 2010; **123**: 661–677.
25. Jonsson G, Staaf J, Vallon-Christersson J, *et al.* The retinoblastoma gene undergoes rearrangements in *BRCA1*-deficient basal-like breast cancer. *Cancer Res* 2012; **72**: 4028–4036.
26. Larsen MJ, Kruse TA, Tan Q, *et al.* Classifications within molecular subtypes enables identification of *BRCA1*/*BRCA2* mutation carriers by RNA tumor profiling. *PLoS One* 2013; **8**: e64268.
27. Peto J, Collins N, Barfoot R, *et al.* Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst* 1999; **91**: 943–949.
28. Lakhani SR, Van De Vijver MJ, Jacquemier J, *et al.* The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in *BRCA1* and *BRCA2*. *J Clin Oncol* 2002; **20**: 2310–2318.
29. Palacios J, Honrado E, Osorio A, *et al.* Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to *BRCA1* or *BRCA2* mutations: differences from breast carcinomas arising in *BRCA1* and *BRCA2* mutation carriers. *Clin Cancer Res* 2003; **9**: 3606–3614.
30. Melchor L, Benitez J. The complex genetic landscape of familial breast cancer. *Hum Genet* 2013; **132**: 845–863.
31. Banneau G, Guedj M, MacGrogan G, *et al.* Molecular apocrine differentiation is a common feature of breast cancer in patients with germline PTEN mutations. *Breast Cancer Res* 2010; **12**: R63.
32. Derksen PW, Liu X, Saridin F, *et al.* Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* 2006; **10**: 437–449.
33. Liu X, Holstege H, van der Gulden H, *et al.* 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* 2007; **104**: 12111–12116.
34. Kendrick H, Regan JL, Magnay FA, *et al.* Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. *BMC Genom* 2008; **9**: 591.

SUPPLEMENTARY MATERIAL ON THE INTERNET

The following supplementary material may be found in the online version of this article:

Supplementary experimental procedures

Figure S1. Expression of *Trp53* and *Brca2* in *Brca2*$^{f/f}$ *p53*$^{f/f}$ mouse tumours

Figure S2. ddPCR analysis of conditional allele recombination in tumours

Figure S3. Expression of *Pten* and *Trp53* in *Pten*$^{f/f}$ *and Pten*$^{f/f}$ *p53*$^{f/f}$ mouse tumours

Figure S4. Papillary features in benign and malignant AMEs of *Pten* tumours

Figure S5. Features of benign *Pten* and *Pten p53* tumours and comparison of ER staining in malignant *Pten* and *Pten p53* tumours. Comparison of ER, PR-A and PR-B staining

Figure S6. Co-localization of p63 and ER in malignant but not benign adenomyoepitheliomas