Identification of Cellular and Genetic Drivers of Breast Cancer Heterogeneity in Genetically Engineered Mouse Tumour Models

Lorenzo Melchor*1,2, Gemma Molyneux*1,3, Alan Mackay1, Fiona-Ann Magnay1, María Atienza4, Howard Kendrick1,5, Daniel Nava-Rodrigues1, María Ángeles López-García4, Fernanda Milanezi1,6, Kirsty Greenow5, David Robertson1, José Palacios7, Jorge S. Reis-Filho8 and Matthew J. Smalley1,5,9.

*These authors contributed equally to this study.

1Division of Breast Cancer Research, Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, United Kingdom

2Current address: Centre for Myeloma Research, Division of Molecular Pathology, The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK

3Current address: Centre for Molecular and Cellular Biology of Inflammation, Division of Immunology Infection & Inflammatory Diseases, King's College London, 1st Floor New Hunt's House, Great Maze Pond, London SE1 1UL, UK

4Department of Pathology, Instituto de Biomedicina de Sevilla, CSIC-Universidad de Sevilla, Hospital Universitario del Rocio, Seville, Spain
Current address: European Cancer Stem Cell Research Institute, Cardiff School of Biosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, CF24 4HQ, United Kingdom

Current address: Salomao & Zoppi Diagnósticos, 876 Divino Salvador Avenue, Sao Paulo, SP, Brazil

Department of Pathology, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigaciones Sanitarias, Madrid, Spain

Department of Pathology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY, USA, 10065

Corresponding author. Telephone: +44(0)29 208 75862 FAX: +44(0)29 208 74116 E-mail: SmalleyMJ@Cardiff.ac.uk

Running title: Breast cancer heterogeneity in mouse tumour models

Conflict of interest statement: The authors declare they have no conflicts of interest in the publication of this manuscript.

Word count (beginning of introduction to end of discussion): 3999
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 tumors 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 that 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 etiology.

Keywords: Brca2, Pten, p53, tumour heterogeneity, breast cancer molecular subtypes, basal-like.
Introduction

Breast cancer is a heterogeneous disease encompassing different histological and molecular subtypes with distinct clinical behaviours\cite{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\cite{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\cite{7-10}, but current models suggest that a substantial proportion, if not all, of these tumours derive from luminal-progenitor cells\cite{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 \textit{Brca1} and \textit{p53} loss were targeted to either basal or luminal ER negative mammary (lumER$^{\text{neg}}$) cells in mouse models, the balance of tumour phenotypes depended on the cell-of-origin. Though all tumours were molecularly classified as ‘basal-like’, histologically the basal-cell origin tumours were mostly adenomyoepitheliomas (AMEs) while the lumER$^{\text{neg}}$-cell origin tumours were high grade Invasive Ductal Carcinomas of No Special Type (IDC-NSTs)\cite{13}. It remains to be defined, however, whether the cell-of-origin is the prime determinant of tumour subtype or if 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 where 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^{+/+}\cdot p53^{+/+}, BlgCre:Brca2^{+/+}\cdot p53^{+/+} virgin and parous, Pten^{+/−}, K14Cre:Pten^{+/−}; BlgCre:Pten^{+/−} virgin and parous, BlgCre:Pten^{+/−}\cdot p53^{+/−}, BlgCre:Pten^{+/−}\cdot p53^{+/−}. Parous mice went through 2-3 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**

Haematoxilin 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 (alphaPR6; 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 hybridisation 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 normalised 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/). Sub-group 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 Supplementary Online Material. 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 lumERneg-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 Information 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 days; range 47-243 days) animals compared with either virgin (242d; 185-334d) or parous (275.5d; 133-445d) 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 (Figure S1). Droplet digital PCR (ddPCR) demonstrated that tumours consistently had fewer copies of floxed Brca2 and p53 exons compared to unfloxed exons (Figure S2). However, the presence of infiltrating immune cells (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%; Figure 1B; Figure 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 (MI; Figure 1C, Figure 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 (Figure 1D-F; Figure 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 (Figure 1B, Figure 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 (Figure 1D,E; Figure 20,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 $\leq$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 (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/f mice as well as germ-line Pten heterozygote mice (Pten^+/f/). Due to a strong skin phenotype, only fourteen K14cre:Pten^f/f mice could age older than four months old. From these, only 5 mammary tumours were obtained (from four mice). Mammary tumour
latencies in K14Cre:Ptenfloxed (141-386d), virgin BlgCre:Ptenfloxed (340d; 128-711d), parous BlgCre:Ptenfloxed (357d; 245-771d), and Pten+/− mice (368d; 100-434d), were not significantly different (Figure 1G). However, BlgCre:Ptenfloxed:p53f/+&f/f mice developed tumours (312d; 139-361d) significantly faster than parous BlgCre:Ptenfloxed mice (p<0.05, log rank test) and also faster, though not significantly potentially due to the small sample size, than virgin BlgCre:Ptenfloxed mice (p=0.078, log rank test; Figure 1G).

All tumours had lower expression for Pten floxed exon 4 compared with exon 6 (Figure S3A,B), confirming recombination of the conditional allele during tumourigenesis. Expression of p53 exon 4 was higher in Ptenfloxed tumours, similar or lower in Ptenfloxed:p53f/+ tumours, and always reduced in Ptenfloxed:p53f/f tumours relative to control spleen (Figure S3C). This is consistent with Pten loss causing p53 induction in p53f/+ 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:Ptenfloxed:p53f/+&f/f mice had fewer copies of floxed p53 exons compared to unfloxed exons (Figure S2). The same caveats regarding infiltrating immune cells apply (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 (Tables S2, S3). Benign tumours were classified as sclerosing adenosis and benign-AMEs (Figure 3A-F), with 5/15 (33%) displaying papillary architecture (Figure S4A-F). All showed strong K14, K18, p63, and ER staining (Figure 3C-F, 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 (Figure 1H, Figure 3G-R, Table S2). A subset displayed papillary architecture (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 (Figure 1J-L; Figure 3I-K, 3O-Q). Remarkably, strong ER expression was seen in 7/8 (87.5%) analysed malignant tumours (15-40% cells; Figure 3L, 3R; Figure S5). PRA expression was observed in 8/9 analysed tumours (1-30% cells; Figure S5). Notably, both malignant Pten^{+/-} and K14Cre:Pten tumours had similar phenotypes to lumER^{neg}-cell origin tumours (Table S2).

Addition of p53 conditional alleles into the BlgCre:Pten cohort increased the ratio of malignant to benign tumours, with 20/21 (95%) tumours being malignant (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, 1H), as half of all BlgCre:Pten^{f/f}:p53^{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 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) (Figure 1L; Figure 4E,K,Q). Again, IDC-NSTs and MSCTs had few p63-positive cells and these were scattered through the tumour whereas AMEs had high levels of p63 staining organised 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; 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 (Figure S5 and Table S2). Double immunofluorescence staining of benign and malignant BlgCre:Ptentf/f AMEs, as well as malignant BlgCre:Ptentf/f:p53f/+&f/f AMEs, demonstrated that in benign tumours p63-positive and ER-positive cell populations were mutually exclusive but in malignant tumours most of p63-positive cells were also ER-positive (Figure S6). This double positivity for ER/p63 suggests an aberrant differentiation in these tumour cells.

Therefore, BlgCre:Brca2f/f:p53f/f and BlgCre:Ptentf/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 \textit{Pten} and \textit{p53} to lumER$^\text{neg}$-cells accelerated tumour formation and, notably, resulted in a range of phenotypes that once again included IDC-NSTs and MSCTs. \textit{BlgCre}:\textit{Pten}^{f/f}:\textit{p53}^{f/+}\&^{f/f} IDC-NSTs, however, were strongly ER-positive, unlike IDC-NSTs from other cohorts (\textit{BlgCre}:\textit{Brca2}^{f/f}:\textit{p53}^{f/f} and \textit{BlgCre}:\textit{Brca1}^{f/f}:\textit{p53}^{+/-}) \cite{13,20}.

\textbf{Luminal ER$^\text{neg}$-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 \textit{K14Cre}:\textit{Brca1}^{f/f}:\textit{p53}^{+/-} and \textit{BlgCre}:\textit{Brca1}^{f/f}:\textit{p53}^{+/-} tumours)\cite{13} using the Affymetrix MouseChip Genome platform. Unsupervised hierarchical clustering showed the tumours broadly clustered into three molecular groups (Figure 5). One included the \textit{Brca2}:\textit{p53} tumours and the \textit{Pten}:\textit{p53} tumours; the second group consisted of most of \textit{Pten}-only tumours; the third group included \textit{Brca1}:\textit{p53} tumours and some \textit{Pten} tumours. Pairwise SAM comparisons between groups delivered a list of significantly-associated genes, which were interrogated for GO terms and KEGG pathway analysis (Table S4). The \textit{Brca2}:\textit{p53}/\textit{Pten}:\textit{p53} group (group 1) and the \textit{Brca1}:\textit{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 (\textit{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 (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 nor 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, 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 categorised 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 ($p<0.0001$, $\chi^2$ test) in pairwise genotype comparisons (Table 1).

PAM50 analysis is sensitive to sample cohort normalisation 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 heatmaps for the mouse transcriptome signature in the human datasets (Figure 5F, 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 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$^{\text{neg}}$-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 characterised by upregulation of mesenchymal-associated genes and downregulation 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 categorised 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 downregulated in the claudin-low phenotype[22] (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 if 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 (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-cells carcinomas and metaplastic carcinomas with mesenchymal differentiation (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$^\text{neg}$-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 etiology of tumour formation is largely deterministic and depends on the earliest events in carcinogenesis (i.e. the founder genetic/ epigenetic events).
Germ-line mutations in human BRCA2 predispose to breast and ovarian cancers[27]. Although 66-93% 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 unlikely 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 (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<sup>neg</sup>-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 germ-line PTEN-loss-syndrome families are enriched for molecular apocrine differentiation, which is characterised by
elevated levels of androgen signalling[31], and our mouse Pten tumour cohort also expressed high levels of the Androgen Receptor (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\textsuperscript{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\textsuperscript{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 \textit{K14Cre}-driven models (\textit{K14Cre:Brca1}^{\text{\texttt{f/f}}}:p53^{\text{\texttt{f/f}}} \text{ and } \textit{K14Cre:Ecad}^{\text{\texttt{f/f}}}:p53^{\text{\texttt{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 \textit{Brca1} and \textit{Ecad} alleles used by Liu and colleagues and Derksen and colleagues are dominant over the \textit{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 \textit{BlgCre} transgene preferentially drives tumour formation in lumER\textsuperscript{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 Supplemental Online Material); whereas equivalent mouse and human mammary epithelial cell types can be inferred (i.e. 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 Supplemental Online Material) 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; finally, we have not yet observed tumours that resemble sporadic human ER-positive IDC-NSTs with a luminal A molecular profile. We hypothesise that either lumER\textsuperscript{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 by a Marie-Curie Fellowship to LM. LM was funded by Fundación Ramón Areces and Marie Curie Actions (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.

Statement of 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 optimised 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. MJS designed the study, assisted with analysis of tumour pathology and bioinformatic analysis and wrote the manuscript.

List of Online Supporting information

Supplementary Experimental Procedures

Human material
Origins of mouse lines and breeding strategy
Cell types targeted in BlgCre models
Cell types targeted in K14Cre models
Evaluation of histological phenotypes and immunohistochemistry results
Quantitative Real-Time rtPCR
Droplet Digital Polymerase Chain Reaction (ddPCR) for floxed/unfloxed allele detection
Meta-analysis of signatures developed in the mouse in human breast cancers
Assessment of microarray batch variation
Supplementary Figure and Table Legends

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-localisation of p63 and ER in malignant but not benign adenomyoepitheliomas.

Figure S7. Affymetric gene expression analysis is not confounded by batch variation.

Figure S8. Mouse mammary tumor molecular signatures are differentially correlated with human breast cancer molecular subtypes.

Figure S9. Mouse mammary metaplastic spindle cell tumours have molecular signatures similar to the human claudin-low breast cancer subtype.
Figure S10. Staining of human IDC-NSTs, AMEs and metaplastic tumours for CLDN4 and PTEN expression.

Table S1. Primers used for genotyping and TaqMan gene expression assays.

Table S2. Full malignant tumour histological features and immunohistochemical findings.

Table S3. Full benign tumour histological features and immunohistochemical findings.

Table S4. Genes upregulated in tumour molecular clusters determined by SAM pair-wise comparisons and analysed by Gene Ontology (GO) and KEGG pathway analysis.

Table S5. Detailed PAM analysis results.

Table S6. Results of staining of a pilot cohort of human tumours for PTEN, CLDN3, CLDN4 and CDH1 expression.
References


staining in normal and neoplastic endometrial, breast and prostatic tissues. *Human Pathology.*


<table>
<thead>
<tr>
<th></th>
<th>Basal</th>
<th>Her2</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Normal Breast-like</th>
<th>Undefined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brca1 tumors (n=21)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>K14Cre</em> (n=3)</td>
<td>14 (66.67)*</td>
<td>1 (4.76)</td>
<td>0 (0)</td>
<td>1 (4.76)</td>
<td>5 (23.81)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>BlgCre</em> (n=18)</td>
<td>2 (66.67)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (33.33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>Virgin</em> (n=12)</td>
<td>12 (66.67)</td>
<td>1 (5.56)</td>
<td>0 (0)</td>
<td>1 (5.56)</td>
<td>4 (22.22)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>Parous</em> (n=6)</td>
<td>7 (58.33)</td>
<td>1 (8.33)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>4 (33.33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Brca2 tumors (n=13)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>K14Cre</em> (n=5)</td>
<td>5 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>BlgCre</em> (n=8)</td>
<td>4 (50)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (12.5)</td>
<td>2 (25)</td>
<td>1 (12.5)</td>
</tr>
<tr>
<td><em>Virgin</em> (n=4)</td>
<td>2 (50)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (25)</td>
<td>1 (25)</td>
</tr>
<tr>
<td><em>Parous</em> (n=4)</td>
<td>2 (50)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (25)</td>
<td>1 (25)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Pten tumors (n=21)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pten</em>&lt;sup&gt;−/−&lt;/sup&gt; (n=6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>6 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>Pten</em>&lt;sup&gt;+/−&lt;/sup&gt; (n=15)</td>
<td>1 (6.67)</td>
<td>0 (0)</td>
<td>3 (20)</td>
<td>0 (0)</td>
<td>11 (73.33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>K14Cre</em> (n=2)</td>
<td>1 (50)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (50)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>BlgCre</em> (n=13)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>3 (23.08)</td>
<td>0 (0)</td>
<td>10 (76.92)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>Virgin</em> (n=7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (28.57)</td>
<td>0 (0)</td>
<td>5 (71.43)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><em>Parous</em> (n=6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (16.67)</td>
<td>0 (0)</td>
<td>5 (83.33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Pten p53 tumors (n=10)</strong></td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (20)</td>
<td>4 (40)**</td>
<td>3 (30)</td>
<td>1 (10)</td>
</tr>
<tr>
<td><em>p53&lt;sup&gt;−/−&lt;/sup&gt;</em> (n=6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (33.33)</td>
<td>1 (16.67)</td>
<td>2 (33.33)</td>
<td>1 (16.67)</td>
</tr>
<tr>
<td><em>p53&lt;sup&gt;−/−&lt;/sup&gt;</em> (n=4)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>3 (75)</td>
<td>1 (25)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Table 1. Correlation of tumor molecular profiles to human breast cancer molecular phenotypes. Number and percentage of tumors 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 tumors listed by genotype, Cre promoter and parity status. Shaded cells indicate the modal correlation. (*) Basal vs non-basal $\chi^2$ test $p<0.0001$. (**) Normal vs non-normal $\chi^2$ test $p<0.0001$. (***) Luminal B vs non-luminal B $\chi^2$ test $p<0.0001$. See Table S4.
Figure Legends

Figure 1. *Brca2*, *Pten*, and *Pten p53* derived tumours have distinct features. (A) Survival curve for \(K14\text{Cre}:\text{Brca2}^{f/f}\cdot\text{p53}^{f/f}\) \((n=29)\), virgin \((n=21)\), and parous \((n=18)\) \(\text{BlgCre}:\text{Brca2}^{f/f}\cdot\text{p53}^{f/f}\) mice. Only data from animals developing mammary tumours are shown. \(K14\text{Cre}\) tumours had a significantly shorter latency \((p<0.0001; \text{Log Rank Test})\). (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, D, E, and F) Mitotic index (number of mitotic figures in ten 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^{+/−}\) \((n=15)\), \(K14\text{cre}:\text{Pten}^{f/f}\) \((n=3)\), virgin \((n=43)\) and parous \((n=29)\) \(\text{BlgCre}:\text{Pten}^{f/f}\); and \(\text{Blgcre}:\text{Pten}^{f/f}\cdot\text{p53}^{f/+\cdot f/f}\) \((n=12)\) 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, J, K, and 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. Asterisks indicate statistically significant differences in t-tests: (*) \(P<0.05\), (**) \(P<0.01\), (***) \(P<0.001\).
**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*°°:p53°° IDC-NST. (A) Low power H&E showing tumour with mixed borders. (B) High power H&E showing mitotic figures and nuclear pleomorphism. (C-F) Staining for K14 (C), K18 (D), p63 (E) and ER (F). (G-L) *BlgCre:Brca2*°°:p53°° metaplastic spindle cell carcinoma. (G) Low power H&E showing tumour with central necrosis. (H) High power H&E 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*°°:p53°° malignant AME. (M, N) Low power (M) and high power (N) H&Es showing 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 in (A, G, M) = 1.5 mm. Bars in (B-F), (H-L), and (N-R) = 100 μm. Inset boxes are magnified x3. See also Tables S2 and S3 and Figure S1.

**Figure 3.** *Pten* depletion in luminal progenitors generates benign and malignant tumours. (A-F) Features of *BlgCre:Pten*°° benign AME. (A, B) Low power (A) and high power (B) magnification H&Es. (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*°° malignant AME. (G) Low power H&E showing multifocal necrosis and pushing borders. (H) High power H&E 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$^{ff}$ tumour showing an adenosquamous carcinoma clone originating from an AME. (M) Low power H&E showing multifocal necrosis and mixed borders. (N) High power H&E 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 in (A, G, M) = 1.5 mm. Bars in (B-F), (H-L), and (N-R) = 100 μm. Inset boxes are magnified x3. See also Tables S2 and S3 and Figures S3 and S4.

**Figure 4.** *p53* loss alters tumour phenotypes in *Pten* knockout mice. (A-F) Features of a BlgCre:Pten$^{ff}$:p53$^{ff}$ IDC-NST. (A) Low power H&E showing central necrosis and mixed borders. (B) High-power H&E 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$^{ff}$:p53$^{ff}$ metaplastic spindle cell carcinoma. (G) Low power H&E showing multifocal necrosis and mixed borders. (H) High-power H&E 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$^{ff}$:p53$^{ff}$ malignant AME. (M) Low power H&E showing multifocal necrosis and pushing borders. (N) High-power H&E 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 in (A, G, M) = 1.5 mm. Bars in (B-F), (H-L), and (N-R) = 100 μm. Inset boxes are magnified x3. See also Table S2 and S3 and Figures S5 and S6.

**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 4,657 Affymetrix probes. (B) Zoomed image of the gene expression heatmap showing a selection of significantly upregulated 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 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 Figure S7 and Supplementary Online Material 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 tumors), 2 (enriched in Pten tumors), and 3 (enriched in Brca1 tumors) in human breast cancer datasets[3, 22]. Spearman rank correlation values for each signature were plotted against dataset molecular phenotypes as correlation heatmaps. Note Group 2 and Group 3 signatures correlated with the human luminal A / normal breast-like and the basal-like subtypes, respectively (both heatmaps),
whereas the Group 1 signature was highly correlated with the human claudin-low subtype (right heatmap). See also Figure S8.

Figure 6. Low expression of claudin-related genes in metaplastic spindle cell carcinomas. Boxplots showing expression levels for mouse orthologues of human genes (Cldn3, Cldn4, Cldn7, and Cdh1) characteristically downregulated 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 MSCTs overall show lower expression for each of the genes as compared with the other phenotypes. These differences are statistically significant (one way ANOVA test, p<0.05, indicated by *) for all genes except Cldn4 (p=0.056). See also Figure S9.
Supplementary Online Information

Supplementary Experimental Procedures

*Origins of mouse lines and breeding strategy*

All animal work was carried out following local ethical review and under the authority of UK Home Office Project and Personal Licences.

*K14Cre:Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f}* mice (mixed FVB/129 Ola background) were obtained from Dr Jos Jonkers (Netherlands Cancer Institute, Amsterdam, The Netherlands) [1]. *Pten\textsuperscript{+/-}* mice (FVB background) and *Pten\textsuperscript{f/f}* mice (C57Bl6 background) were obtained from JAX (The Jackson Laboratory, Bar Harbor, Maine, USA). *BlgCre:Brca1\textsuperscript{f/+}:p53\textsuperscript{+/-}* mice (C57Bl6 background) have been previously described [2, 3].

To obtain *BlgCre:Brca2\textsuperscript{f/+}:p53\textsuperscript{f/+}* mice, the *K14Cre:Brca2\textsuperscript{f/+}:Pten\textsuperscript{f/f}* strain was crossed with wild type C57Bl6 mice to generate *Brca2\textsuperscript{f/+}:p53\textsuperscript{+/-}* genotypes. These were crossed with the *BlgCre:Brca1\textsuperscript{f/+}:Brca2\textsuperscript{f/+}:p53\textsuperscript{+/-}* mice to generate *BlgCre Brca1\textsuperscript{f/+}:Brca2\textsuperscript{f/+}:p53\textsuperscript{+/-}* mice. These were crossed with each other until the *BlgCre:Brca2\textsuperscript{f/+}:p53\textsuperscript{f/+}* genotype was established.

The *K14Cre:Pten\textsuperscript{f/+}* strain was obtained by crossing *K14Cre:Brca2\textsuperscript{f/+}:p53\textsuperscript{f/+}* mice with *Pten\textsuperscript{f/+}* mice and then mating the offspring until the desired genotype was generated.

The *BlgCre:Pten\textsuperscript{f/+}, BlgCre:Pten\textsuperscript{f/+}:p53\textsuperscript{f/+}* and *BlgCre:Pten\textsuperscript{f/+}:p53\textsuperscript{f/+}* strains were generated by first crossing *BlgCre:Brca1\textsuperscript{f/+}:p53\textsuperscript{+/-}* mice with *Pten\textsuperscript{f/+}* mice to
obtain BlgCre:Brca1\textsuperscript{f/+}:Pten\textsuperscript{f/+} and then crossing these animal to obtain BlgCre:Pten\textsuperscript{f/f}. Then, the BlgCre:Pten\textsuperscript{f/f} mice were crossed with BlgCre:Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f} mice to obtain BlgCre:Pten\textsuperscript{f/+}:Brca2\textsuperscript{f/+}:p53\textsuperscript{f/+}. The remaining Brca2 floxed allele was bred out, the Pten floxed allele was bred to homogeneity and then the lines maintained by crossing BlgCre:Pten\textsuperscript{f/+}:p53\textsuperscript{f/+} with BlgCre:Pten\textsuperscript{f/f}:p53\textsuperscript{f/f} animals.

Therefore, the majority of the mouse tumours, with the exception of the K14Cre:Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f} line, have come from a mainly C57Bl6 background, although this is still not pure. Furthermore, to generate new combinations of alleles we have intercrossed between the strains in order to increase the homogeneity of their backgrounds. The distinctly different genetic background of the K14Cre:Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f} mice is unlikely to be the reason they developed mainly AMEs as this tumour phenotype is seen in other lines in other genetic backgrounds, and in particular in the BlgCre:Brca1\textsuperscript{f/f}:p53\textsuperscript{+/-} mouse, which is C57Bl6. It is also unlikely that it is genes linked to the p53 floxed allele carried over from the K14Cre:Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f} background, rather than the p53 itself, that are responsible for the rescue of the MSCT and IDC-NST phenotypes in the BlgCre:Pten\textsuperscript{f/f}:p53\textsuperscript{f/+} animals. This is because IDC-NSTs and MSCTs are seen in the BlgCre:Brca1\textsuperscript{f/f}:p53\textsuperscript{+/-} mouse, which is a germline p53 heterozygote from a completely different origin [4].

**Cell types targeted in BlgCre models**

A key consideration in the interpretation of these data is the cell of origin of the different tumours, particularly in the BlgCre models. Data from our
previous studies [3, 5] support a lumER\textsuperscript{neg} origin of BlgCre tumours. In 12 week old virgin female mice, BlgCre drives recombination in 51\% of the total lumER\textsuperscript{neg} population and in 77\% of cells with \textit{in vitro} colony forming potential (colony forming cells; CFCs) within that population. The ability to form colonies \textit{in vitro} is the standard operational definition of a progenitor cell [6]. In contrast, the transgene drives recombination in 0.6\% of the total lumER\textsuperscript{pos} population and in 1\% of CFCs within that population. By 43 weeks of age, the numbers are 97\% of the CFCs (52\% of the total) for the lumER\textsuperscript{neg} population and 34\% of the CFCs (3\% of the total) for the lumER\textsuperscript{pos} population [3]. At 10-12 weeks of age, the ratio of lumER\textsuperscript{neg} progenitors to lumER\textsuperscript{pos} progenitors (as determined by c-Kit expression, which marks luminal progenitors) is 17:1 [5] and hence the ratio of recombined lumER\textsuperscript{neg} to lumER\textsuperscript{pos} progenitors in the resting mammary epithelium is >1000:1 at 12 weeks. Moreover, the overall proportion of lumER\textsuperscript{pos} cells declines with age [3]. The most likely cell of origin for all BlgCre model tumours, therefore, is lumER\textsuperscript{neg} progenitors.

However, while an origin of ER- tumours in lumER\textsuperscript{neg} progenitors was unremarkable, the consistent appearance of ER+ tumours in the BlgCre:Pten:p53 models was surprising. The possibility that ER- tumours were generated from lumER\textsuperscript{neg} progenitors in BlgCre:Brca1:p53 and BlgCre:Brca2:p53 mice and ER+ tumours were generated from lumER\textsuperscript{pos} progenitors in BlgCre:Pten:p53 mice must be considered. However, it is highly unlikely that recombination of the conditional alleles only occurs in lumER\textsuperscript{pos} progenitors in the BlgCre:Pten:p53 mice and in lumER\textsuperscript{neg} progenitors in the BlgCre:Brca1/2:p53 mice, as the only difference between these mice is the
floxed alleles and any downstream effects of loss of Pten or Brca1/2 must happen after switching has occurred, although one cannot definitively exclude the possibility that chromatin structure around the Brca1, Brca2 and Pten loci is different in these two cell types, altering the likelihood of recombination. If one accepts, rather, that BlgCre drives conditional allele recombination equally efficiently, on a cell to cell basis, in both lumER<sup>neg</sup> and lumER<sup>pos</sup> progenitors, and that the ratio of recombined lumER<sup>neg</sup> to lumER<sup>pos</sup> progenitors is >1000:1, then there must be substantially more lumER<sup>neg</sup> than lumER<sup>pos</sup> progenitors with deleted conditional alleles in both BlgCre:Pten:p53 and the BlgCre:Brca1/2:p53 mice. The lumER<sup>neg</sup> progenitors are, thus, the most likely origin of both ER+ and ER- tumours in these models.

A final caveat is that while deletion of Pten and p53 alleles in the BlgCre:Pten:p53 mice may be occurring in the same cell type (lumER<sup>neg</sup> progenitors) as deletion of Brca1/2 and p53 alleles in the BlgCre:Brca1/2:p53 mice, this does not necessarily mean that the cells that actually ‘transform’ into a tumour are the same in both cases. Loss of Brca1/2 or Pten alleles in the same cell may initially drive that cell along different differentiation pathways before the cell actually becomes neoplastic. Thus, it is formally possible that loss of Pten could cause ER negative cells to become ER positive prior to transformation and thus generate ER positive tumours. Without extensive formal lineage tracing analysing the very early stages of tumour formation in these models, it is not possible to determine whether or not this is occurring. However, the ultimate origin of these tumours, the cell type in which recombination occurs, is still the same. The only difference in
these scenarios is that in one alternative, recombination is followed by transformation and then deleted-allele-specific tumour differentiation, whereas in the other recombination is followed by deleted-allele-specific cellular differentiation and then transformation. The outcomes of both these scenarios are essentially identical.

**Cell types targeted in K14Cre models**

The *Krt14* gene is most highly expressed in the basal mammary epithelium in the adult. However, it is also expressed from early on in development during the formation of the embryonic mammary gland. Therefore, the *Krt14Cre* has the potential to be driving recombination at any point from very early to very late during mammary development. One might expect, therefore, that the *Krt14Cre* would drive recombination, and thus tumour formation, throughout the mammary gland in all animals. However, this is not the case. It has been demonstrated that adult mice carrying the allele we have used have recombined flox alleles in the majority of basal cells but only in occasional luminal cells [1]. This would be consistent with the occasional appearance of IDC-NSTs in the *K14Cre:Brca1* and *K14Cre:Brca2* flox models, but the majority of tumours being a different phenotype. Furthermore, the *BlgCre* and *Krt14Cre:Brca1* and *Brca2* models clearly develop different tumour types. As the only difference between these models is the Cre driver (and two different *KrtCre* alleles were used in the *Brca1* and *Brca2* models) the most likely explanation for the different tumour types is different cells of origin. As the lineage data supports the lumER<sup>neg</sup> population as the origin for BlgCre tumours, the basal cells must be the origin of the *K14Cre* tumours.
**Human material**

This study was performed following standard ethical procedures of the Spanish regulation (Ley de Investigación Orgánica Biomédica, 14 July 2007) and was approved by the ethics committee of the Hospital Virgen del Rocío de Sevilla and the Fundación Pública Andaluza para la Gestión de la Investigación en Salud de Sevilla (FISEVI), Spain. Written informed consent was obtained and all clinical investigation was been conducted according to the principles expressed in the Declaration of Helsinki.

**Evaluation of histological phenotypes and immunohistochemistry results**

Mouse tumour sections were scored with observers blinded to the genotypes of animals on a multiheaded microscope (by LM, MJS and JSR-F). Consensus scores were calculated for each marker on each tumour. The final histological subtypes were further discussed with two pathologists with an interest in breast pathology (DNR and/or FM). Markers were semiquantitatively assessed by estimating percentages of morphologically unequivocal neoplastic cells displaying either nuclear (ER/p63/PRA/PRB) or cytoplasmic (K14/K18) staining. Human tumour samples were assessed by MA, MAL-G and JP essentially according to previous criteria [7, 8] and as described in the legend to Table S6.

**Quantitative Real-Time rtPCR**

Total RNA was isolated from up to 30 mg of frozen tumour material using RLT Buffer and RNeasy Mini Kit (Qiagen, Crawley, UK) according to
manufacturer’s guidelines. Isolated RNA was subjected to cDNA synthesis using QuantiTect Reverse Transcription Kit (Qiagen). qRT-PCR was performed using TaqMan Assays-on-Demand probes (Applied Biosystems, Paisley, UK; Table S1). Assays were tested on three independent cDNA preparations from each tumour. Results were analysed with the Δ-ΔC_t method normalised to β-Actin and relative to a comparator sample (normal lumER$^{\text{neg}}$ cells).

**Droplet Digital Polymerase Chain Reaction (ddPCR) for floxed/unfloxed allele detection**

To measure the tumor depletion of the gene floxed alleles as compared with the unfloxed alleles, 50 ng of genomic tumor DNA were also used in a ddPCR reaction following manufacturer’s guidelines (ddPCR Supermix for Probes #186-3010, Bio-Rad, Hemel Hempstead, Hertfordshire, UK). Taqman made-to-order mouse copy number assays for Brca2, Pten, and Trp53, as well as Tfrc mouse copy number reference assays were used (Life Technologies, Paisley, UK; Table S1). Reactions were performed in duplicates; reference (DNA from normal mammary gland cells) and non-template controls were added to the study.

**Meta-analysis of signatures developed in the mouse in human breast cancers**

Based upon SAM pairwise comparison analyses of mouse mammary tumours, three different molecular signatures were obtained after applying filtering thresholds: Group 1 (120 probes 15-fold up regulated and 91 probes
10-fold down regulated in Brca2 and Pten:p53 tumours), Group 2 (92 probes 10-fold up regulated and 115 probes 10-fold down regulated in Pten tumours), and Group 3 (42 probes 5-fold up regulated and 91 probes 5-fold down regulated in Brca1:p53 tumours). These gene lists were then matched to each external dataset [9-18]. Datasets were annotated for subtype and gene symbol. Probes with more than 20% missing values were removed. Missing values were imputed using KNN imputation. Replicate probes were removed retaining the one with the greatest variance for each gene.

For each dataset, recovered expression values for each gene were tested for significant differences across subtypes using a two-tailed ANOVA. Average expression values for both up- and down-regulated gene lists across the subtypes were bar plotted and pair wise compared using t-test. Additionally, for each gene list combining both up and down regulated genes, a Spearman rank correlation was performed for each sample. Correlation values were reported for each mouse signature and plotted against dataset ordered based upon molecular subtype or mutation status as correlation heatmaps.

**Assessment of microarray batch variation**

The unexpected clustering of the Brca2:p53 tumours with the Pten:p53 tumours, rather than the Brca1:p53 tumours may have been an artefact of microarray analyses (e.g. batch variation). To test this, we carried out quantitative real-time rtPCR analysis of the expression of 22 lineage-associated genes in all tumour cohorts and in normal mammary epithelial subpopulations. We then extracted the gene expression data for the same 22
genes from the whole transcriptome arrays. Unsupervised hierarchical clustering analyses were performed on both sets of data. The results were in strong agreement, this time separating the tumours into two groups (Figure S7). One consisted of the *Pten* cohort together with normal lumER\textsuperscript{neg} and lumER\textsuperscript{pos} populations. The other consisted of the *Brca1* and *Brca2* cohorts together with normal basal cell populations. The *Pten:p53* tumours were scattered amongst the two groups (Figure S7). These results obtained with an orthogonal RNA expression analysis method argue against a batch effect in the whole transcriptome clustering.

**Supplementary Figure and Table Legends**

**Figure S1.** Expression of *Trp53* and *Brca2* in *Brca2\textsuperscript{f/f}:p53\textsuperscript{f/f}* mouse tumours. (A-B) *Trp53* expression in tumour versus matched (A) or average (B) spleens determined with a probe against (floxed) exon 4. (C) *Brca2* expression levels in tumour versus matched spleens determined with probes against exon 6 and (floxed) exon 11. Gene expression levels were determined by qPCR on triplicate samples from each tumour and are shown as mean gene expression levels ±95% confidence limits relative to the comparator sample. *No detectable expression.*

**Figure S2.** Copy number analysis for floxed versus unfloxed alleles determined by droplet digital-PCR results. (A) Detection of *Brca2* alleles in *Brca2 p53* tumors. (B) Detection of *Trp53* alleles in *Brca2:p53* tumors. (C) Detection of the *Trp53* alleles in *Pten:p53* tumors. Reference copy number value is 2 copies as predicted by gene reference (*Tfrc*) and further confirmed
by normal mammary gland (MG) cells. Error bars indicate the Poisson 95% confidence intervals for each copy number determination. X axis crosses Y at reference copy number value 2. Note all tumors consistently show lower number of copies of the floxed alleles (in asterisks) as compared with the unfloxed alleles. *Pten* analysis could not be performed due to technical limitations to assess copy number data in the floxed allele (see Supplementary Methods).

**Figure S3. Expression of *Pten* and *Trp53* in *Pten*\(^{ff}\) and *Pten*\(^{ff}:p53^{ff}\) mouse tumours.** (A-B) *Pten* expression levels in tumour versus matched spleens determined with probes against (floxed) exon 4 and exon 11 in *Pten*\(^{ff}\) (A) and *Pten*\(^{ff}:p53^{ff}\) (B) tumours. (C) *Trp53* expression in tumour versus matched spleens determined with a probe against (floxed) exon 4. Gene expression levels were determined by qPCR on triplicate samples from each tumour and are shown as mean gene expression levels ±95% confidence limits relative to the comparator sample.

**Figure S4. Papillary features in benign and malignant AMEs of *Pten* tumours.** (A-F) Histological features of a benign *BlgCre:Pten*\(^{ff}\) papillary adenomyoepithelioma. (A) Low power H&E of benign tumour. (B) High power H&E showing tumour cells with palisade-like nuclei tumour cells, typical of papillary tumours. (C) K14 staining in basal cells. (D) K18 staining in luminal cells. (E) p63 expression in basal cells. (F) ER expression in neoplastic cells located in the luminal cells of gland-like structures. (G-L) Histological features of a malignant *BlgCre:Pten*\(^{ff}\) papillary adenomyoepithelioma. (G) Low power
H&E of tumour multifocal necrosis and mixed borders. (H) High power H&E showing tumour cells with pleomorphic palisade-like nuclei tumour cells. (I) K14 staining in basal cells. (J) K18 staining in luminal cells. (K) p63 expression in basal cells. (L) Neoplastic cells in both luminal and basal regions express ER. Bars in (A, B) = 1.5 mm. Bars in (B-F, H-L) = 100 μm. Inset boxes are magnified x3.

Figure S5. Features of benign Pten and Pten:p53 tumours and comparison of ER staining in malignant Pten and Pten:p53 tumours. (A-E) Percentage of tumour cells positive for K14 (A), K18 (B), p63 (C), and ER (D) in benign Pten and Pten:p53 tumours. (E) Percentage of tumour cells positive for ER in malignant Pten and Pten p53 tumours. (F) Percentage of tumour cells positive for PRA in malignant Pten and Pten:p53 tumours. (G – I) Staining of a BlgCre:Ptenf/f:p53f/f tumour for ER (G), PRA (H) and PRB (I); bars in (G, H, I) = 50 μm. Insets are magnified 2.5 x.

Figure S6. Co-localisation of p63 and ER in malignant but not benign adenomyoepitheliomas. (A) Dual immunofluorescence staining of benign (top, A) and malignant (bottom, B) AME for p63 (green) and ER (red). Sections are counterstained with DAPI. Note co-localisation of p63 and ER in the malignant tumour (e.g. white arrows) while occasional p63+ ER- nuclei (green arrows) and p63- ER+ nuclei (red arrows) demonstrate that the double staining was not due to cross reactivity of antibodies. Bar = 70 μm. Inset panels magnified x3.
Figure S7. Affymetric gene expression analysis is not confounded by batch variation. Unsupervised hierarchical clustering of gene expression data from qPCR analysis of 22 lineage-associated genes (A) and from Affymetrix gene expression data (see Figure 5) of the same 22 genes.

Figure S8. Mouse mammary tumor molecular signatures are differentially correlated with human breast cancer molecular subtypes. (A) Bar plots of the mean expression values for both up- (top) and down-regulated (bottom) mouse gene lists across the human subtypes in the TCGA dataset [10, 15] Error bars indicate the Poisson 95% confidence intervals (standard t-test in pairwise comparisons). (B) Spearman rank correlation values for each signature plotted against dataset molecular phenotypes from additional human breast cancer datasets [13, 14]. (C) As in (A) but tested in a dataset in which claudin-low tumours have been annotated [12]. (D) As in (B) but again for additional human breast cancer datasets in which claudin-low tumours have been annotated [9, 11]. (E) As in (B) but considering human familial breast cancers datasets [16-18]. Human samples were BRCA1-mutated, BRCA2-mutated, non-BRCA1/2 mutated, and sporadic breast cancer cases.

Figure S9. Mouse mammary metaplastic spindle cell tumours have molecular signatures similar to the human claudin-low breast cancer subtype. From low (left) to high expression (right), barplots indicate the average expression of genes in the downregulated geneset of the human ‘claudin-low’ breast cancer subtype [12]. Bars are color coded by genotype
and histological phenotypes are indicated below the X-axis. Note that all metaplastic spindle cell tumours in this series (blue boxes) have low expression for genes characteristically downregulated in human claudin-low tumours.

**Figure S10. Immunostaining of human tumours.** (A – C) CLDN4 staining of (A) a human triple-negative IDC-NST showing strong staining, (B) a human triple-negative IDC-NST showing reduced staining and (C) a human metaplastic spindle cell tumour showing absence of CLDN4 staining in spindle cells. (D – L) Analysis of PTEN expression in human histological subtypes. (D) PTEN staining of normal human breast. (E) H&E of benign AME and (F) PTEN staining of benign AME. (G) H&E of malignant AME and (H) PTEN staining of malignant AME. (I) Triple-negative IDC-NST showing PTEN staining. (J) Triple-negative IDC-NST showing absence of PTEN staining in tumour cells. (K) PTEN positive metaplastic spindle cell tumour. (L) PTEN negative metaplastic spindle cell tumour. Bars A – C = 50 μm; bars D – L = 25 μm.

**Table S1.** (A) Primers used for genotyping. (B) TaqMan gene expression assays.

**Table S2.** Full malignant tumour histological features and immunohistochemical findings. Histological type was defined based on the World Health Organisation classification [19, 20]; metaplastic carcinomas were subtyped according to previously published criteria [21]. Tumours were
considered positive for a marker by immunohistochemistry if they contained ≥1% morphologically unequivocal neoplastic cells with discrete staining clearly above background levels. ¹Nuclear pleomorphism graded according to Elston and Ellis [22]. ²Tubule formation graded according to Elston and Ellis [22]. ³Number of mitotic figures in ten high power (x63) fields. ⁴Borders classified as pushing if in >75% of the circumference of the tumour, neoplastic cells arranged in sheets and nests compressed without encasing the adjacent mammary tissues, as infiltrative if in >75% of the circumference of the tumour, neoplastic cells arranged in cords, sheets and nests invaded the adjacent mammary tissues encasing normal structures and/or eliciting stromal reaction, and as mixed if they contained both elements [23]. ⁵DCIS was graded based on nuclear morphology according to the UK guidelines. ⁶RNA could be extracted from these tumours for qPCR but not either the amount obtained was not sufficient for gene expression microarray or the RNA quality failed QC. ⁷Tumours were small and processed for histology only – no frozen material was available for molecular analysis. Browse tabs for Brca2, Pten, and Pten:p53 tumour data. For Brca1 tumour information, see [3].

Table S3. Full benign tumour histological features and immunohistochemical findings. For details see legend to Table S1. Benign tumours were not scored for tumour formation, nuclear pleomorphism or mitotic indices. Browse tabs for Pten, and Pten:p53 benign tumour data.

Table S4. Genes upregulated in tumour molecular clusters determined by SAM pair-wise comparisons and analysed by Gene Ontology (GO)
and KEGG Pathway Analysis. For the SAM analysis, only genes with a fold difference in expression of >2 and a P-value and local FDR <0.05 are shown. For GO and KEGG analysis, gene lists were uploaded to DAVID [24]. For the Group 1 genes, only the most highly expressed 3000 genes were uploaded. GO terms associated with a less than 10 genes or a P value > 0.05 were excluded from the lists. All KEGG terms are included with no filters.

Table S5. Detailed PAM analysis results. Correlations for each tumour comparing tumour molecular profiles human molecular subtypes [11].

Table S6. Staining of human tumour samples for PTEN, CLDN3, CLDN4 and CDH1. For claudins and CDH1 sample evaluation, separate assessments were performed for the epithelial and mesenchymal components in metaplastic carcinomas [25]. For claudins, two parameters were evaluated: first, cell membrane staining intensity: 0 (absent), 1 (low), 2 (moderate), and 3 (high); and second, the percentage of cells with positive staining: 0, 1 (0-1%), 2 (2-10%), 3 (10-33%), 4 (34-66%), 5 (>66%). The addition of cell membrane staining plus percentage of cells with positive staining gave a score from 0 to 8 which was further categorized. Scores 0, 1 and 2 were considered as absent or negligible expression (indicated by a 0 on the table); 3 and 4, as low expression (1); 5 and 6, as moderate expression (2); and, lastly, 7 and 8 as high expression (3). CDH1 (E-cadherin) was scored as 0, absent; 1, reduced expression, less than 50% of cells with positive cell membrane staining; 2, conserved expression, when >50% of cells showing strong membrane staining. For PTEN, sample evaluation was carried out as described [7] to
generate a histoscore. The maximum PTEN histoscore per case is 600. This is the result of the addition of the nuclear and cytoplasmic PTEN histoscores. IDC-NST-TN, invasive ductal carcinoma of no special type but with triple negative phenotype; MC_M, mesenchymal-type metaplastic cercinoma carcinoma, which can have chondrosarcomatous differentiation (MC_M_CHR), or osteosarcomatous differentiation (MC_M_OST); MC_SC, spindle cell-type metaplastic carcinoma equivalent to mouse MSCT; AME, adenomyoepithelioma; NA, not analyzed because component was not observed or no immunostaining performed.

Supplementary References


