Title: Evolutionary pathways in BRCA1-associated breast tumors

Authors: Filipe C. Martins¹,²,³,⁴,⁵, Subhajyoti De⁶,⁷,⁸, Vanessa Almendro¹,²,³,⁹, Mithat Gönen¹⁰, So Yeon Park¹¹,¹², Joanne L. Blum¹³, William Herlihy¹⁴, Gabrielle Ethington¹³, Stuart J. Schnitt¹⁵,¹⁶, Nadine Tung³,¹⁷, Judy E. Garber²,³,¹⁸, Katharina Fetten¹⁷, Franziska Michor⁶,⁷, Kornelia Polyak¹,²,³,¹⁹

Authors’ Affiliations: Departments of ¹Medical Oncology, ⁶Biostatistics and Computational Biology, and ¹⁷Radiation Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA. ²Department of Medicine, Brigham and Women’s Hospital, Boston, MA 02115, USA. Departments of ³Medicine and ¹⁵Pathology, Harvard Medical School, Boston, MA 02115, USA. ⁴Obstetrics and Gynaecology Department, Hospitais da Universidade de Coimbra, Coimbra, Portugal. ⁵Gulbenkian Programme for Advanced Medical Education, Portugal. ⁷Department of Biostatistics, Harvard School of Public Health, Boston, MA 02215, USA. ⁸King’s College, University of Cambridge, Cambridge, CB2 1ST, United Kingdom. ⁹Department of Medical Oncology, Hospital Clinic, Institut d’Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain. ¹⁰Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. ¹¹Department of Pathology, Seoul National University College of Medicine, Seoul, South Korea. ¹²Department of Pathology, Seoul National University Bundang Hospital, Seongnam, Gyeonggi, South Korea. ¹³Baylor-Charles A. Sammons Cancer Center, Dallas, TX 75246, USA. Department of Pathology, ¹⁴Baylor University Medical Center, Dallas, TX 75246, USA. Departments of ¹⁶Pathology and ¹⁷Hematology-Oncology, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA. ¹⁹Harvard Stem Cell Institute, Cambridge, MA 02138, USA.

Running title: BRCA1-associated breast tumorigenesis

Key words: BRCA1, breast cancer, evolution

List of abbreviations: LOH – loss of heterozygosity, DCIS – ductal carcinoma in situ

Notes: Financial support: This work was supported by the National Cancer Institute P50 CA89383 (to K. Polyak) and U54CA143798 (to F. Michor and M. Gönen), the Cellex foundation (to V. Almendro), Human Frontier Science Program long-term fellowship (to S. De), the Susan G. Komen Foundation
Corresponding authors: Kornelia Polyak, Dana-Farber Cancer Institute, 450 Brookline Ave. D740C, Boston, MA 02215. Phone: 617-632-2106; Fax: 617-582-8490; E-mail: Kornelia_polyak@dfci.harvard.edu. Franziska Michor, Dana-Farber Cancer Institute, 450 Brookline Ave. CLS-11007, Boston, MA 02215. Phone: 617-632-5045; Fax: 617-632-2444; E-mail: michor@jimmy.harvard.edu.

Potential conflicts of interest: No potential conflicts of interest were disclosed.
ABSTRACT
BRCA1-associated breast tumors display loss of BRCA1 and frequent somatic mutations of PTEN and TP53. Here we describe the analysis of BRCA1, PTEN, and p53 at the single cell level in 55 BRCA1-associated breast tumors and computational methods to predict the relative temporal order of somatic events based on the frequency of cells with single or combined alterations. Although there is no obligatory order of events, we found that loss of PTEN is the most common first event and is associated with basal-like subtype, whereas in the majority of luminal tumors, mutation of TP53 occurs first and mutant PIK3CA is rarely detected. We also observed intra-tumor heterogeneity for the loss of wild-type BRCA1 and increased cell proliferation and centrosome amplification in the normal breast epithelium of BRCA1 mutation carriers. Our results have important implications for the design of chemopreventive and therapeutic interventions in this high-risk patient population.

SIGNIFICANCE: Defining the temporal order of tumor-driving somatic events is critical for early detection, risk stratification, and the design of chemopreventive therapies. Our combined experimental and computational approach reveals that the loss of wild-type BRCA1 may not be the first event in the majority of BRCA1-associated breast tumors and may not be present in all cancer cells within tumors.
INTRODUCTION

BRCA1 germline mutations confer a high risk of breast and ovarian cancer. Somatic loss of the wild-type BRCA1 allele is thought to be a rate-limiting initiating step of tumorigenesis (1). BRCA1-associated breast tumors also acquire additional somatic genetic events during their progression as mutations of PTEN and TP53 are frequently observed in these cases (2). The BRCA1 tumor suppressor gene is thought to be a prototypical cancer susceptibility gene insofar as the somatic loss of the wild-type allele, most commonly through loss of heterozygosity (LOH) is a required rate-limiting step of tumor initiation (1). However, multiple lines of evidence suggest that even normal cells of BRCA1 mutation carriers often display an altered phenotype, indicating haploinsufficiency (3-5). For example, the distribution and characteristics of breast epithelial progenitors are altered in BRCA1 mutation carriers, thereby potentially increasing the probability of neoplastic transformation (6, 7). The inability of normal cells to survive the acute loss of BRCA1 (8) also suggests that the loss of wild-type BRCA1 may not be the initiating step of tumorigenesis. Correlating with this, the loss of wild-type BRCA2 was shown to be a relatively late event in pancreatic tumorigenesis of BRCA2 mutation carriers (9). Furthermore, loss of wild-type BRCA1 may only occur in preexisting TP53 mutant foci in ovarian cancer (10) and preinvasive and invasive breast tumors in BRCA1/2 mutation carriers display a high degree of heterogeneity for BRCA1/2 LOH (11). Here we describe the analysis of BRCA1, p53, and PTEN at the single cell level and computational methods to identify the most likely evolutionary pathways in BRCA1-associated breast tumors.
RESULTS

To investigate the relative order of somatic loss of PTEN, BRCA1 LOH, and mutation in TP53 in BRCA1-associated breast tumors, we used a combined computational and experimental approach based on the following assumptions: (1) invasive tumors still contain cancer cells from earlier progression steps, (2) if all tumor cells have mutation X but only a subset of them harbor mutation Y, then mutation X must have occurred before Y, (3) the rate of cell proliferation and death are not significantly different between cells with single and combined alterations, (4) by quantifying the number of tumor cells with single mutations and combinations thereof, the probable evolutionary path of a tumor can be identified, and (5) the analysis of a part of a tumor provides information about the whole. Thus, we analyzed the frequency of individual tumor cells with single and combined alterations in 55 malignant breast tumors from BRCA1 germline mutation carriers (Supplementary Table S1).

Because mutation detection by sequencing from single cells in situ in tissue slices is currently technically not feasible and silencing by DNA methylation or other epigenetic mechanisms may be an alternative mode of gene inactivation, we used a combination of immunofluorescence, immuno-FISH (fluorescence in situ hybridization combined with immunofluorescence), and dual immunohistochemistry for assessing the status of these three proteins at the single cell level in archived tissue samples. Specifically, the expression of PTEN was evaluated by immunohistochemistry, BRCA1 LOH by FISH (defining LOH when both BAC and CEP signal counts $\neq$ 2), whereas the mutational status of p53 was evaluated by immunohistochemistry (when combined with immunostaining for PTEN) or by immunofluorescence (when combined with BRCA1 FISH).

Assays were optimized using xenografts and formalin-fixed paraffin-embedded cell blocks derived from breast cancer cell lines with known PTEN, TP53, and BRCA1 status, and also sections from mice with conditional deletion of PTEN in the prostate epithelium (12) as the antibody used also detects mouse PTEN (Supplementary Fig. S1A-C and Supplementary Table S2) followed by testing of sporadic breast tumor samples. The PTEN antibody used has been previously validated in BRCA1-
associated breast tumors for specificity for wild type PTEN (13). Staining in genetically normal stromal cells was used as internal control on each tumor section. Each of these methods has limitations such as the inability to detect all TP53 mutations by immunohistochemistry and the limited accuracy to predict loss of wild-type BRCA1 allele by FISH. On the other hand, by assessing proteins instead of mutations in DNA, non-genetic causes of PTEN inactivation or activation of p53 signaling could also be detected. To assess our accuracy of predicting BRCA1 LOH based on FISH, we analyzed a set of slides derived from cell lines with wild type or mutant BRCA1 mixed at different ratios (10%, 25%, and 40%) and found good agreement (D=0.85, Somer’s D) between the predicted and known BRCA1 LOH percentage (Supplementary Fig. S1D). To estimate the concordance between results obtained using methods we employed for single cell analysis and those used for bulk cell populations, we compared BRCA1 LOH frequencies estimated based on FISH to that based on PCR analysis of laser capture microdissection-purified tumor cells (14). By employing certain assumptions and converting PCR values to the number of mutant alleles (details in Methods and in Supplementary Table S3), using Dixon’s Q test (15) we only detected four outliers (out of 36 tumors analyzed by both methods), where the two techniques led to different results; this finding was potentially due to uniparental isodisomy for the BRCA1 locus in these cases. Excluding these four cases, the percentage of cells with BRCA1 LOH inferred by FISH was on average 12.7% lower than that inferred by PCR, and the average difference between the two techniques was only 20.5% (±13.3%) (Supplementary Fig. S1E). Therefore, despite the technical limitations of methods applicable for the analysis of single cells in situ in intact tissues, the obtained results are in good agreement with those using other current technologies. Nevertheless, mutational analysis of single cells in situ, which may be feasible in the future, would likely give the most accurate results.

Next, we analyzed 55 BRCA1-associated and 20 sporadic breast tumors from women with no family history of breast cancer (Fig. 1A and Supplementary Tables S1 and S4). We counted 100-200 individual cancer cells in each tumor, and for each cell, we recorded the status of all three genes (i.e.,
wild-type=wt and mutant=mut; details of counting is described in Methods section). Thus, each cell belonged to one of 8 possible states (i.e., PTEN_{wt}TP53_{wt}BRCA1_{wt}, PTEN_{wt}TP53_{wt}BRCA1_{mut}, PTEN_{wt}TP53_{mut}BRCA1_{wt}, PTEN_{wt}TP53_{mut}BRCA1_{mut}, PTEN_{mut}TP53_{wt}BRCA1_{wt}, PTEN_{mut}TP53_{wt}BRCA1_{mut}, PTEN_{mut}TP53_{mut}BRCA1_{wt}, and PTEN_{mut}TP53_{mut}BRCA1_{mut}). For each tumor, we recorded the number of cells in each state (Supplementary Table S1). To determine the most probable first mutation event, we then compared the numbers of cells belonging to the three single-mutation states – for instance, PTEN_{wt}TP53_{wt}BRCA1_{mut}, PTEN_{wt}TP53_{mut}BRCA1_{wt}, and PTEN_{mut}TP53_{wt}BRCA1_{wt}, and determined whether any state had a significantly higher number of cells compared to the two other states (Supplementary Table S3, see Methods for details). The state with the larger number of cells was then designated as indicating the first event. Once the first event was determined, we compared the cell counts of the remaining two mutational states in a similar manner to determine the most probable second event. If a sample contained any cells with all three mutations, then we identified the mutation of the third gene as the last event. For most samples, the order of events was determined unambiguously, while for a few samples, the cell counts of the 8 states suggested complex evolutionary trajectories.

Using this approach, we found that there were two main evolutionary pathways in BRCA1 tumors defined by the presence or absence of PTEN. In the majority (28/55) of tumors, loss of PTEN was the most probable first event followed by mutation in TP53 or BRCA1 LOH with about equal probability (Fig. 1B). Mutation in TP53 was the second most common first event detected in 17/55 cases, and was almost always followed by BRCA1 LOH. BRCA1 LOH was the least common first event, observed in only 10/55 tumors and the majority of these cases had mutant p53 as the only other alteration. The relative order of events and thus evolutionary paths to tumorigenesis were strongly associated with the tumor subtype, as triple negative (i.e., negative for estrogen and progesterone receptors and HER2; ER-/PR-/HER2-) tumors almost always had PTEN loss as the first event (Path 1), whereas luminal tumors showed mutant TP53 or BRCA1 LOH as first event (Path 2).
We categorized all tumors with PTEN loss as the first event into Path 1, regardless of their TP53 status based on immunohistochemistry, because some TP53 mutations such as protein truncation cannot be detected by the antibody-based technology we employed (16). In three tumors, we found evidence for both evolutionary paths; thus, these samples could not be unambiguously assigned to either trajectory (Fig 1B and Supplementary Fig. S2A). In addition, we cannot exclude the possibility of other mutational events taking place that could define different evolutionary paths. Interestingly, all commonly used BRCA1 mutant breast cancer cell lines (HCC-1937, MDA-MB-436, SUM-1315, and SUM-149) are triple negative with loss of PTEN and mutant TP53 and all other PTEN-null cell-lines also have TP53 mutations (13) implying a selective advantage of clones with a combination of these changes both in vivo in the tumors from which the cell lines were derived from and in cell culture.

As opposed to the BRCA1-associated hereditary cases, loss of PTEN was detected at lower frequency in sporadic triple negative breast tumors (Supplementary Table S3). Allelic imbalance of the BRCA1 locus was also rarely (<5%) observed in sporadic cases (Supplementary Tables S3 and S4) and due to the lack of germline mutation, it is not equivalent to BRCA1 LOH in BRCA1-linked tumors. In sporadic breast tumors, PTEN loss and PIK3CA mutations are associated with the basal-like and luminal subtypes, respectively (17). In our BRCA1-associated patient cohort, PTEN loss was strongly associated with ER status (p=5.36x10⁻⁷, Fisher’s Exact test; Supplementary Table S5) and even in the luminal tumors, the percentage of ER+ tumor cells was lower than in sporadic cases (p=0.01, Mann Whitney test; Supplementary Tables S1 and S4). All but one ER+ tumor (a DCIS) expressed wild-type PTEN, whereas most ER- tumors were PTEN negative. To determine if PIK3CA mutation may be an alternative mechanism to PTEN loss for the activation of the PIK3CA/AKT pathway in BRCA1 luminal tumors (though loss of PTEN and mutation in PIK3CA are not functionally equivalent), we assessed the most common mutational hotspots in PIK3CA (E542K, E545A, E545K, H1047L, and H1047R) by mass-spectrometry (18). Only two out of 55 BRCA1 tumors (one ER+ and
one ER-, but 6 out 10 luminal and 0/10 triple negative sporadic tumors, had mutant PIK3CA (Supplementary Tables S1 and S4), suggesting that even luminal BRCA1 tumors display tumorigenic paths distinct from those of sporadic cases. These results also imply that these luminal breast tumors observed in BRCA1 mutation carriers are not likely to be "sporadic" cases.

Our findings were intriguing: the loss of wild-type BRCA1 may not be the first event in most BRCA1-associated breast tumors, and even in tumors that display apparent loss of the wild-type BRCA1 allele, not all tumor cells show this change (Fig. 2A, B). We thus investigated these findings in further detail. First, we analyzed if the intra-tumor frequency of BRCA1 LOH and the intra-tumor diversity in cell types were associated with tumor subtype and evolutionary paths. Interestingly, the percentage of tumor cells with BRCA1 LOH was significantly higher (p=0.047, Mann Whitney test) in basal-like tumors for which loss of PTEN was the first event (Fig. 2B). To assess associations between evolutionary paths and intra-tumor diversity for BRCA1 LOH, we grouped the tumor samples according to the probable first event (i.e., PTEN, BRCA1 or TP53 mutation) and plotted the distribution of Shannon’s and Simpson’s indices for each group (Supplementary Table S1); these indices are routinely used in evolutionary biology and ecology to determine species diversity (19). In general, the samples in which BRCA1 LOH was the most probable first event had a greater extent of diversity, whereas the opposite (lower diversity) was observed in tumors with TP53 mutation as first event, although neither of these was statistically significant (Fig. 2C).

Next, we sought to confirm the presence of functional wild-type BRCA1 protein in tumors with heterogeneous BRCA1 LOH based on FISH. Thus, we performed immunofluorescence analysis of BRCA1 in all tumors and analyzed foci formation in S-phase cells, which is regarded as definitive evidence for wild-type BRCA1 (20). We found a good agreement between the FISH and immunofluorescence data (D=0.98, Somer’s D) (Supplementary Table S1 and Supplementary Fig. S2B), and BRCA1 foci were readily observed in cells of tumors with heterogeneous but not with complete loss of BRCA1 (Fig. 2D and Supplementary Fig. S2C). We validated the functional
relevance of the BRCA1 foci we observed by confirming their co-localization with Rad51 (Supplementary Fig. S2D).

Our finding that loss of BRCA1 is rarely the first event in BRCA1-associated breast tumors suggests a haplo-insufficient phenotype in the mammary epithelium that may explain the increased risk for breast cancer in mutation carriers. For example, the number of cells-of-origin for breast cancer might be higher in BRCA1 mutation carriers, contributing to their higher risk of breast cancer (21). Correlating with this, we (Su et al., unpublished data) and others (6, 7) have observed that the relative fraction of breast epithelial progenitors is higher in BRCA1 mutation carriers compared to control women, potentially implying a higher rate of cell proliferation in mutation carriers. Estrogen and progesterone are potent mitogens for normal breast epithelial cells and prior studies in Brca1 -/- mice demonstrated that blocking progesterone signaling inhibits mammary tumorigenesis (22). Thus, we analyzed the number of breast epithelial cells positive for Ki67, a proliferation marker, and for progesterone receptor (PR) by multi-color immunofluorescence in control and BRCA1 tissues. We detected significantly more Ki67+, PR+, and Ki67+PR+ cells in contra-lateral normal breast tissue of BRCA1 mutation carriers diagnosed with breast cancer compared to that observed in controls, whereas normal prophylactic mastectomy tissues of BRCA1 mutation carriers without breast cancer and reduction mammoplasty tissue from controls were not significantly different (Fig. 3A,B and Supplementary Table S6). The % of PR+ and Ki67+ cells fluctuate during the menstrual cycle with higher fraction of cells being positive in the luteal phase of the menstrual cycle. However, it is unlikely that all control and all BRCA1 mutation carriers would be in the same phase of their cycle leading to the observed differences.

During the analysis of Ki67+ cells, we noticed occasional multi-polar mitoses in normal breast tissues from BRCA1 mutation carriers, suggesting aberrant centrosome function (Fig. 3C). Thus, we analyzed the number of centrosomes in the normal breast epithelium of BRCA1 mutation carriers with and without breast cancer and that of matched controls by immunofluorescence for polyglutamylated
tubulin, a centrosome marker (23). We found significantly higher (p=<0.01) numbers of cells with more than two centrosomes in BRCA1 mutation carriers compared to controls (Fig. 3D,E).

**DISCUSSION**

The development of most human tumors is predicted to take many years and require the progressive accumulation of tumor-driving somatic alterations (24). Identification of genes and pathways that play key roles in tumor initiation and progression is the necessary first step towards designing therapies that may interfere with them, which would be especially important in germline mutation carriers of high risk cancer susceptibility genes. Unfortunately most human tumors are diagnosed at relatively late stage, when they already accumulated numerous genetic and epigenetic alterations making it difficult to decipher which one(s) are functionally relevant for tumorigenesis and at what progression stage. To address this issue, several mathematical models have been developed for predicting the relative order of somatic genetic alterations during tumor progression including profiling tumors at different stages (25) and inferring the order of events based on cross-sectional genomic data of late stage tumors (26, 27). We have investigated putative evolutionary pathways in BRCA1-associated breast tumors by assessing the status of BRCA1, PTEN, and p53 at the single cell level and predicting the probable order of events using a mathematical model based on the frequency of cells with single and combined alterations.

Following Knudson’s two hit model for familial cancer syndromes (28), loss of wild type BRCA1 allele is presumed to be an essential rate-limiting step of BRCA1-associated tumorigenesis. However, several lines of evidence suggest that even the normal breast tissue of BRCA1 germline mutation carriers display an abnormal phenotype including altered frequency, gene expression profiles, and functional properties of breast epithelial progenitors (6, 7) that may increase the risk of breast cancer. Our result that loss of PTEN and TP53 mutation occur before BRCA1 LOH in most cases, support the idea of a haploinsufficient phenotype in the normal breast epithelium of BRCA1 mutation carriers.
Even in tumors where BRCA1 LOH is the probable first event, not all tumor cells appear to have this alteration, potentially suggesting the presence of not-yet-identified somatic genetic (or epigenetic) alterations that may precede the loss of wild type BRCA1 in these tumors. Our finding of increased cellular proliferation and abnormal mitoses coupled with centrosome amplification in normal breast epithelium of BRCA1 mutation carriers further support a haploinsufficient phenotype that may increase breast cancer risk.

The first hint suggesting that BRCA1 may regulate genomic instability by influencing centrosome function was its centrosomal localization in mitosis (29), which is maintained throughout the cell cycle, albeit at lower concentrations (30). Mammary tumors in mice with conditional deletion of Brca1 exhibit gross genomic instability and centrosome amplification with recurrent genomic imbalances resembling those in human BRCA1-associated breast cancer (31). Studies in human breast cancer cell lines also showed that centrosome numbers are regulated by a BRCA1-dependent ubiquitination (32) and that expression of an enzymatically inactive BRCA1 mutant leads to supernumerary and hyperactive centrosomes (30). However, as all these prior studies used cells completely devoid of wild type BRCA1, our results is the first demonstration of centrosome abnormalities in the normal breast epithelium of BRCA1 mutation carriers implying haploinsufficient phenotype for this function. The high frequency of PTEN loss by gross genomic rearrangements (13) and truncating mutations in TP53 in BRCA1-associated breast tumors (16) could potentially be due to genomic instability induced by abnormal centromeres.

Our analysis of the probable order of loss of wild type BRCA1, PTEN, and TP53 also allows predictions regarding possible interactions among these tumor suppressor pathways. For example, in BRCA1 tumors with both PTEN loss and p53 mutation the former was always predicted to precede the later. Indeed, prior studies have described complex cooperative interactions between PTEN and TP53 (33) that could lead to selection against loss of wild type PTEN in cells with mutant TP53. The near complete lack of PI3KCA mutations in BRCA1-associated tumors also implies that some genetic
changes that commonly occur in sporadic breast tumors, may not confer advantage during BRCA1-driven tumorigenesis possibly due to differences in cell-of-origin or presence of other genetic or epigenetic alterations.

Our results have several potentially important clinical implications. First, PARP inhibitors are promising new agents for the treatment and prevention of tumors in *BRCA1* and *BRCA2* mutation carriers due to their synthetic lethal interaction in cells lacking wild type BRCA-associated DNA repair function (34, 35). However, our data showing that loss of wild-type BRCA1 may not be the first event in most BRCA1-associated breast tumors and that this loss may not occur in all tumor cells raises concerns about the efficacy of such approaches. In contrast, PARP inhibitors were also shown to have synthetic lethal interaction with loss of PTEN in cell culture (36, 37). If this interaction also occurs in breast tumors *in vivo*, then PARP inhibitors may still be effective in the majority of BRCA1-linked breast tumors characterized by PTEN loss and triple-negative subtype (Evolutionary Path 1) but not in the luminal subset (Evolutionary Path 2). Furthermore, as the frequency of *BRCA1* LOH is also higher in PTEN null cases, the probability of efficient therapeutic response targeting BRCA1-null tumor cells (e.g., PARP inhibitors) is also expected to be higher in these cases than in the luminal subset. Second, our results also imply that AKT pathway inhibitors and agents designed for targeting p53 mutant tumor cells may show promise for the prevention and treatment of breast tumors in a subset of *BRCA1* mutation carriers. Lastly, our methodology of single cell profiling and computational identification of the evolutionary paths to tumorigenesis can also be applied to other tumor types and promises to provide information about the natural history of a range of human tumors.
METHODS

Tissue samples and cell lines

Breast tissue samples were collected at Harvard-affiliated hospitals (Dana-Farber Cancer Institute, Brigham and Women’s Hospital, and Beth-Israel Deaconess Medical Center, Boston, MA), Seoul National University Bundang Hospital (Seoul, Korea), Texas Oncology-Baylor Charles A. Sammons Cancer Center (Dallas, TX), University of California San Francisco (San Francisco, CA), and Johns Hopkins University (Baltimore, MD) using protocols approved by the institutional review boards. Breast cancer cell lines used in the study were obtained from ATCC (MDA-MB-468, MDA-MB-436, and HCC-1937), Marc Lippman (MCF7), and Steve Ethier (SUM149PT) and their identity confirmed by SNP6 array analysis. None of the cell lines were maintained as continuous cultures in the lab and were used only at early passage.

Immunohistochemistry

Immunohistochemistry was performed using whole sections of formalin-fixed paraffin-embedded (FFPE) tissues and antibodies for PTEN (Cell Signaling, Clone 138G6, rabbit monoclonal antibody - mAb), p53 (Dako, Clone DO-7, mouse mAb), and estrogen receptor (ER, Neomarkers, Clone 6F11, mouse IgG1 mAb). Heat induced antigen retrieval was performed in 10 mM citric acid (pH=6.0) in a steamer at 95°C for 40 minutes. Sections were incubated with primary antibodies diluted in 5% goat serum as follows: PTEN – 1:100 dilution, overnight at 4°C, P53 – 1:100 dilution, 2h at RT, and ER – 1:50 dilution, 3h at RT, followed by incubation with appropriate biotinylated secondary antibodies and peroxidase-conjugated avidin-biotin complexes (Elite ABC, Vector Laboratories). Formed immunocomplexes were visualized using diaminobenzidine (DAB; Sigma) or ImmPACT-VIP (Vector Laboratories). Sections were rinsed in phosphate-buffered saline (PBS) between each step. Double-immunohistochemistry was performed by sequentially incubating the sections with PTEN and p53 or PTEN and ER antibodies. Peroxidase activity and non-specific biotin binding were blocked by
incubation with 3% hydrogen peroxide and 10% goat serum, respectively, between the sequential staining with the two different antibodies. Slides were counterstained with methyl-green to visualize nuclei.

**Immunofluorescence**

Immunofluorescence was performed using antibodies for BRCA1 (Calbiochem, Clone MS110, mouse IgG1 mAb), p53 (Dako, Clone DO-7, mouse mAb), Ki67 (DAKO, clone MIB-1, mouse IgG1 mAb), progesterone receptor (PR) (Neomarkers, RB9017, rabbit polyclonal antibody), polyglutamylated tubulin (Enzo Life Sciences, clone GT335, mouse IgG1 mAb), and RAD51 (Santa Cruz, H-92, rabbit polyclonal antibody). Antigen retrieval, blocking, and primary antibody dilutions were the same as for immunohistochemistry described above. Tissue sections were incubated with primary (BRCA1 – 1:100, 3h; RAD51 – 1:100, 3h), and secondary antibodies (Alexa Fluor 555-conjugated goat anti-mouse IgG1 and Alexa Fluor 488-conjugated goat anti-rabbit; Invitrogen, both 1:100, 45 min). Sections of normal breast tissue from 12 BRCA1 carriers and 9 non-carrier controls (with and without contra-lateral breast cancer) were stained using antibodies for polyglutamylated tubulin at 1:800 dilution for 1h at RT, or a combination of Ki67 and PR, at 1:100 dilution for 2 hr at RT. Alexa Fluor 555-conjugated goat anti-mouse IgG1 and Alexa Fluor 488-conjugated goat anti-rabbit secondary antibodies were used at 1:100 dilution and incubated for 45 min. Samples were washed twice with PBS-Tween 0.05% between incubations and protected for long-term storage with VECTASHIELD HardSet Mounting Medium with DAPI (Vector Laboratories, cat #H-1500).

**ImmunoFISH (Immunofluorescence combined with FISH)**

A combination of immunofluorescence for p53 and FISH for BRCA1 was performed essentially as previously described (38) but using probes specific for BRCA1 (BAC Clone ID: 831F13, Invitrogen) and chromosome 17 (CEP 17 Spectrum Green Probe, Abbott), and antibodies for p53 (Dako, Clone
DO-7, mouse mAb). BRCA1 BAC probe was labeled by nick translation and Alexa Fluor 647-conjugated nucleotides (Abbott).

**PIK3CA mutation analysis**

DNA preparation and mutation analyses were performed essentially as previously described (18); detailed procedures are available upon request.

**Confocal microscopy analysis**

Samples were stored at -20°C for at least 48 hrs before image analysis. For combined PR and Ki67 immunofluorescence, different images from multiple areas of each sample were acquired with a Nikon Ti microscope attached to a Yokogawa spinning-disk confocal unit, 60x plan apo objective, and OrcaER camera controlled by Andor iQ software. For BRCA1 immunofluorescence and immunoFISH, images were acquired with a SP5 Leica Confocal Microscope, 60x plan objective, and analyzed by Leica software (Leica Application Suite – Advanced Fluorescence 2.2.0). For immunoFISH, several images from different regions of each tumor were acquired by overlapping plans (Z-stack with 0.4 μm intervals) in order to capture all FISH signals in the section.

**Scoring for the expression of markers in individual tumor cells**

For immunoFISH, we analyzed 100-200 individual cancer cells from different areas of each tumor, and scored the p53 status and the number of CEP and BAC probes in each cell. Because tumor cells that were positive or lacked PTEN were very well demarcated within tumors, PTEN status was defined by double-immunohistochemical staining of adjacent sections for PTEN and p53. Because double-immunohistochemistry provides a more complete overview of the section than immunoFISH, a larger area was assessed by double-immunohistochemistry than by immunoFISH. Thus, in small regions we were able to assess PTEN and p53, but not BRCA1 status. These technical issues were
taken into consideration during our statistical analyses by performing all possible permutations and selecting the weakest p-value (see details below). Similarly, because immunofluorescence is less sensitive than immunohistochemistry, in a few cases we could not detect staining for mutant p53 by immunofluorescence, but still got good signal by immunohistochemistry. In these cases, we counted the proportion of p53 mutant cells based on immunohistochemistry and the BRCA1 status based on immunoFISH slides and performed all possible permutations and selected the weakest p-value (see details below). For immunofluorescence for BRCA1 and polyglutamylated tubulin, we analyzed ~200 cells, whereas for Ki67 and PR double- immunofluorescence, 1000-2000 cells were evaluated in each slide.

Comparison of BRCA1 LOH based on FISH and PCR and correlation between FISH and immunofluorescence data

In order to compare BRCA1 LOH data based on PCR (14) to % LOH based on FISH we made the following assumptions: 1. there was no contamination from normal tissue/stroma; 2. there was no copy gain or loss; 3. both original and mutated tumor cell populations were homogeneous with respect to BRCA1 locus. If these assumptions hold, then the BRCA1 LOH PCR data can be expressed as \(|(p-50)*2|\), where p is the percentage of mutant allele. We used Somer’s D to analyze the concordance between FISH and immunofluorescence data.

Prediction of the order of events: BRCA1 LOH status

BRCA1 LOH was determined using the ratio of copy number ratio of the BAC probe versus the CEP (centromeric) probe in each cell. Ideally a BRCA1 wild-type cell should have a BAC to CEP ratio of 2:2 (referred to as BAC/CEP:2/2), whereas copy number alteration leads to a different ratio. Incidentally, the sectioning of tissue samples can affect nuclei of a subset of cells in a slide (e.g., those residing in the section plane) such that those cells have an apparently different BAC to CEP...
ratio even though they are genetically wild type. For example, during sectioning a wild-type cell may lose part of its nucleus, including one copy of chr17, leading to the BAC to CEP ratio of 1:1, which can be misinterpreted as loss of BRCA1 due to whole chromosome loss during tumorigenesis. On the other hand, cells with BRCA1 copy number alterations may have apparently normal BAC to CEP ratio for similar reasons. To estimate the bias introduced by the above phenomenon, and to make necessary corrections in our analyses, we performed a set of control analyses. We analyzed 15 independent slides prepared using BRCA1 wild-type cells (MDA-MB-468 cell line), used as negative controls, and 15 independent slides prepared from cells with known BRCA1 LOH (SUM-149, HCC-1937, and MDA-MB-436 cell lines), used as positive controls. The negative and positive control sets had approximately 35%, and 3%, respectively, cells with BAC/CEP:2/2 ratios (Supplementary Table S3), as opposed to 100% and 0%, respectively, in the ideal case (Supplementary Table S3). Therefore, even a sample with only BRCA1 wild-type cells would have about 65% cells with apparent signatures of BRCA1 LOH (i.e., BAC/CEP ≠ 2/2). The thickness of the slides had minimal effects on these proportions for both control sets. From the distribution of the proportions of BAC/CEP:2/2 cells in positive and negative control sets, we calculated the most probable number of cells present with genuine BRCA1 LOH, after adjusting for the variability arising from the technical issue described above. We used a residual bootstrapping technique with 15 iterations. At each iteration, we drew a value (v_i; i=1;15) from the distribution of BAC/CEP:2/2 ratios in the negative control, and again one value (v_j; j=1;15) from the positive control. If, in a given slide, the proportion of cells with BRCA1 LOH (with a certain status for PTEN and P53 genes) (i) was lower than v_i (the value expected to arise due to technical reasons), as observed in the negative control sets, we considered all those cells as wild-type, and (ii) was higher than that v_i, we subtracted the contribution of variability due to technical reasons, and adjusted the number of BRCA1 wild-type cells accordingly. We made similar adjustments based on the positive control set and recalibrated the number of cells with wild-type BRCA1 and BRCA1 LOH in the dataset. We then used the recalibrated dataset for further analysis.
We also determined the order of events after adjustments using the median value of BAC/CEP:2/2, instead of the distribution of the proportion of BAC/CEP:2/2, in the positive and negative control sets. The order of events was highly similar to that depicted in Fig. 1 (Supplementary Table S4). To further validate this approach, we applied it to a set of slides derived from tissue blocks of different mixtures of cell lines with wild type and mutant BRCA1 (10%, 25% and 40% of wt and mutant mix). We then assessed correlations between estimated and known % using Somer’s D test.

**Combinatorial mutation status**

For a subset of the samples, it was difficult to determine the mutation status for multiple genes on a cell-by-cell basis, as we described above in the scoring section, because we used two sequential tissue sections to analyze the status of the three markers. To avoid any bias arising due to this ambiguity, for each of those samples, we generated an ensemble of cases with cell counts for different combinations of the mutational status for the three genes, which satisfied the cell counts observed for individual mutations and their unambiguous combinations. Using a missing value imputation strategy, we then calculated a single p-value for the order of events in that sample from that ensemble of cases.

**Order of events**

To determine the first event in a given sample, we compared the number of cells in the three groups – (i) PTEN loss, BRCA1 and P53 wild-type (N100), (ii) P53 loss, BRCA1 and PTEN wild-type (N010) and (iii) BRCA1 gain (or loss), PTEN and P53 wild-type (N001). We accepted the most abundant group to indicate the first mutation event and compared the number of cells in that group with each of the two others using a binomial test, and then used the union-intersection test to obtain a conservative estimate:

\[
P(PTEN) = \max\left\{ F(N010,N100 + N010,0.5), F(N001,N100 + N001,0.5) \right\}
\]
\[ P(P53) = \max \left\{ F(N100, N100 + N010, 0.5), F(N001, N010 + N001, 0.5) \right\} \]

\[ P(BRCA1) = \max \left\{ F(N100, N100 + N001, 0.5), F(N010, N001 + N010, 0.5) \right\} \]

where

\[ F(x, y, p) = \sum_{i=1}^{n} \frac{y}{x} i \cdot (1 - p)^{y-x} \]

If two of the most abundant groups had the same number of cells, we flagged the first event as ‘unresolved’. Once the first mutation event was flagged, we compared the number of cells in the two groups, which (i) were mutated in the first and second gene and wild-type for the third and (ii) wild-type for the second gene and mutated in the first and third genes, using the binomial test as above.

We generated an ensemble of instances for each sample to account for the technical variability described in the BRCA1 LOH and combinatorial mutation status sections. Finally, using a missing value imputation strategy, we calculated a single p-value for the order of events in that sample from the ensemble of cases. If some cells had only one mutation and some other cells had only a different mutation, but no cells had both mutations, then the order of events was unclear. This scenario could potentially indicate two independent mutation events. Such cases are discussed as special cases.

**Shannon and Simpson indices of diversity**

Shannon and Simpson indices of diversity of BRCA1 LOH were calculated following standard procedures (19) and as described previously (38). The Pearson correlation coefficient between Shannon’s and Simpson’s indices for our dataset (i.e., Supplementary Table S1) is -0.9186. The anti-correlation is not surprising, because a population with high Shannon's index would usually have low Simpson's index and vice versa.
ACKNOWLEDGEMENTS

We thank Drs. Piere Paolo Pandolfi for providing sections from Pten^{loxP/loxP};Pb-Cre4 mice, Stanisława Werimowicz (Brigham and Women's Hospital, Boston, MA) for her help with evaluation of metaphase FISH images, Lisa Cameron in the DFCI Confocal and Light Microscopy Core Facility for technical assistance, and David Pellman and Susana Godinho for their recommendation of antibodies to assess centrosomes. We thank members of our laboratories for their critical reading of this manuscript and useful discussions.
REFERENCES


FIGURE LEGENDS

Figure 1. Evolutionary pathways in BRCA1-associated breast tumors. A, Examples of immuno-FISH and immunohistochemical analyses. B, Predicted order of somatic events and tumor characteristics. Bars mark loss of PTEN (green), BRCA1 LOH (brown), and mutation of p53 (orange). Hormone receptor (ER and PR) and HER2 status, presence of PIK3CA mutation (*), and probably evolutionary pathways (Path 1 and Path2) are indicated. In a few tumors both evolutionary paths were observed (T5, T24, T44) whereas a few others could not be assigned to either paths at high confidence. T30# is a DCIS (ductal carcinoma in situ). C, Summary of evolutionary paths in BRCA1-associated breast tumors. The thickness of the arrows and size of the circles are proportional to the number of tumors (indicated within circles and next to arrows) following the depicted paths. Blue and red indicate luminal and basal-like tumors, respectively. The three main paths are depicted by dashed red (basal-like tumors) and blue (mainly luminal tumors) arrows.

Figure 2. Heterogeneity for loss of wild-type BRCA1 allele in BRCA1-associated breast tumors. A, Representative immunoFISH analysis for mutant p53 (blue), chromosome 17 CEP (green), and BRCA1 BAC (red) probes depict heterogeneity for BRCA1 copy number among individual cancer cells. Insets highlight three cells with different BAC to CEP signal ratios corresponding to normal, and copy number gain and loss. B, Intra-tumor diversity for BRCA1 LOH. Frequency of tumor cells with BRCA1 LOH and Shannon and Simpson’s indices of diversity in each of the 55 BRCA1 tumors analyzed. The Pearson correlation coefficient between Shannon's and Simpson's indices is -0.9186. The anti-correlation is not surprising, because a population with high Shannon's index would usually have low Simpson's index and vice versa. C, Associations between first event and frequency of BRCA1 LOH and diversity indices. D, Representative immunofluorescence analyses of BRCA1 in tumors where BRCA1 foci are observed (top) or completely absent (bottom) indicating wild type and loss of function, respectively. Arrows mark
leukocytes that serve as internal positive control.

**Figure 3. Abnormalities in the normal breast epithelium of BRCA1 mutation carriers.** A, Representative examples of Ki67 and PR immunofluorescence in normal breast tissue from BRCA1 mutation carriers and non-carrier controls. B, Frequency of Ki67 and PR positive cells in the same tissues. C, Representative examples of multipolar mitoses in normal breast tissue from two distinct BRCA1 mutation carriers. D, Representative examples of polyglutamylated tubulin immunofluorescence in normal breast tissue from BRCA1 mutation carriers and non-carrier controls. E, Number of centrosomes/cell in the same tissues.
Figure 1
Figure 2
Figure 3
Evolutionary pathways in BRCA1-associated breast tumors
Filipe C. Martins, Subhajyoti De, Vanessa Almendro, et al.

Cancer Discovery  Published OnlineFirst April 10, 2012.

Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-11-0325

Supplementary Material
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2012/04/10/2159-8290.CD-11-0325.DC1
http://cancerdiscovery.aacrjournals.org/content/suppl/2012/06/04/2159-8290.CD-11-0325.DC2

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.