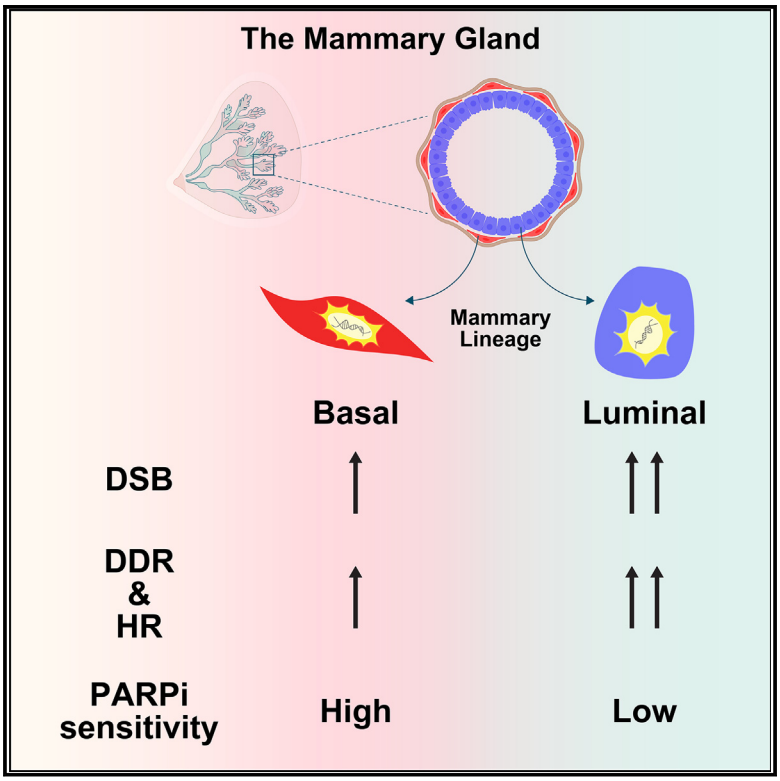
Cell Reports

Article

Differential DNA damage repair and PARP inhibitor vulnerability of the mammary epithelial lineages

Graphical abstract



Authors

Hyeyeon Kim, Kazeera Aliar, Pirashaanthy Tharmapalan, ..., Miquel A. Pujana, Hal K. Berman, Rama Khokha

Correspondence

rama.khokha@uhn.ca

In brief

Kim et al. study DNA damage response following γ-irradiation in mammary cells. They find basal cells less proficient in homologous recombination repair than luminal cells and vulnerable to PARP inhibitors. These results highlight that differences can exist in DDR and drug response based on lineage identity even within a single organ.

Highlights

- Mammary epithelial cells exhibit differential DDR protein expression based on cell lineage
- Independent of the cell cycle, the basal lineage exhibits lower HR repair capacity
- Basal progenitors are intrinsically sensitive to PARP inhibitors
- Breast tumors with a basal proteomic profile also show PARP inhibitor sensitivity

Kim et al., 2023, Cell Reports *42*, 113256
October 31, 2023
https://doi.org/10.1016/j.celrep.2023.113256

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Differential DNA damage repair and PARP inhibitor vulnerability of the mammary epithelial lineages

Hyeyeon Kim,[1,2] Kazeera Aliar,[1] Pirashaanthy Tharmapalan,[1,2] Curtis W. McCloskey,[1] Abhijith Kuttanamkuzhi,[1] Barbara T. Grünwald,[1] Luis Palomero,[3] Mathepan J. Mahendralingam,[1,2] Matthew Waas,[1] Arvind S. Mer,[4] Mitchell J. Elliott,[1] Bowen Zhang,[1,2] Khalid N. Al-Zahrani,[5] Ellen R. Langille,[5,6] Michael Parsons,[4] Swami Narala,[1] Stefan Hofer,[1] Paul D. Waterhouse,[2] Razqallah Hakem,[1,2,7] Benjamin Haibe-Kains,[1,2] Thomas Kislinger,[1,2] Daniel Schramek,[5,6] David W. Cescon,[1] Miquel A. Pujana,[3] Hal K. Berman,[1] and Rama Khokha[1,2,7,8,*]

[1]Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada
[2]Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada
[3]ProCURE, Catalan Institute of Oncology, Oncobell, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, 08908 Barcelona, Catalonia, Spain
[4]Department of Biochemistry, Microbiology & Immunology, University of Ottawa, Ottawa, ON K1H 8M5, Canada
[5]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada
[6]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[7]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON M5G 2N2, Canada
[8]Lead contact
*Correspondence: rama.khokha@uhn.ca
https://doi.org/10.1016/j.celrep.2023.113256

SUMMARY

It is widely assumed that all normal somatic cells can equally perform homologous recombination (HR) and non-homologous end joining in the DNA damage response (DDR). Here, we show that the DDR in normal mammary gland inherently depends on the epithelial cell lineage identity. Bioinformatics, post-irradiation DNA damage repair kinetics, and clonogenic assays demonstrated luminal lineage exhibiting a more pronounced DDR and HR repair compared to the basal lineage. Consequently, basal progenitors were far more sensitive to poly(ADP-ribose) polymerase inhibitors (PARPis) in both mouse and human mammary epithelium. Furthermore, PARPi sensitivity of murine and human breast cancer cell lines as well as patient-derived xenografts correlated with their molecular resemblance to the mammary progenitor lineages. Thus, mammary epithelial cells are intrinsically divergent in their DNA damage repair capacity and PARPi vulnerability, potentially influencing the clinical utility of this targeted therapy.

INTRODUCTION

The DNA damage response (DDR) plays a fundamental role in oncology, spanning mechanisms of tumor development and progression to targeted cancer therapies. In somatic cells, DNA double-strand breaks (DSBs) are resolved by either homologous recombination (HR), ensuring accurate repair with sister chromatids as templates, or non-homologous end joining (NHEJ), which directly ligates DSB ends but is more prone to errors. Mathematical projections indicate that two-thirds of mutations in human cancers arise from random errors during DNA replication, and the total number of stem cell divisions in tissues positively correlates with their cumulative lifetime cancer risk.[1,2] Thus, DDR is especially imperative for genomic integrity in tissues that undergo repeated cycles of stem/progenitor cell-driven tissue remodeling, such as the breast epithelium during the menstrual cycle and pregnancy.[3,4] As such, many hereditary breast cancer susceptibility genes are involved in DDR (*ATM*, *CHEK2*, *MRE11*, *BRCA1*, *BRCA2*, *RAD51*),[5,6] in which deleterious mutations in the HR repair genes, such as *BRCA1/2*, are prevalent in over 20% of all breast cancers.[7] A number of poly (ADP-ribose) polymerase inhibitors (PARPis) have been developed to exploit synthetic lethality in cells harboring deleterious *BRCA1/2* mutations.[8–10] Notably, two PARPis, olaparib[11,12] and talazoparib,[13,14] were approved as single-agent therapies for the treatment of advanced breast cancer with germline *BRCA1/2* mutation in 2018, and currently, combinational trials with chemo- or immunotherapy are being tested, and a new generation of PARPis is underway.[15]

Intriguingly, accumulating preclinical and clinical evidence shows that BRCA mutation status does not solely account for PARPi sensitivity. For instance, the inhibitory effects of PARPis on human breast cancer cell lines are observed independent of BRCA status,[16,17] and several clinical studies on ovarian and breast cancer highlighted the therapeutic benefits of PARPis even in individuals without BRCA mutations.[18–22] These findings collectively suggest that functional "BRCAness"[23] extends beyond mutations in DNA repair pathways. Thus, understanding

the DNA damage repair processes and various determinants of functional HR in breast epithelial cells will contribute to improving PARPi treatment efficacy.

Human breast cancers can originate from different "cells of origin" within the stem/progenitor and differentiated cell populations of the mammary epithelium.[24,25] The breast epithelium is a bilayer with luminal and basal lineages that exhibit salient differences, such as in metabolism,[26,27] epigenetics,[28,29] and telomere biology.[30] Such unique lineage traits are thought to be retained within the resulting cancers and can thus provide fundamental insights into the biological identity and specific vulnerabilities of major tumor subsets.[31] However, given that the essential capacity to engage DNA repair is assumed to be uniform across non-neoplastic somatic cell types, this key process in the mammary epithelium has not been explored in a lineage-specific manner. We report here that, in an unsupervised analysis across humans and mice, DDR differences dominated the proteomic landscape of luminal vs. basal cells in the mammary epithelium. This led us to discover that mammary epithelial cell lineages differ in their ability to engage HR repair. Furthermore, this differential DSB repair capacity at the mammary progenitor level correlated with selective PARPi sensitivities, revealing cell lineage as a novel determinant of PARPi response. Correspondingly, human and mouse cancer cells reminiscent of a basal lineage origin displayed higher PARPi sensitivity in cell lines or patient-derived xenografts compared with luminal-like cancer models. Altogether, this study uncovers baseline and post-irradiation repair preferences in different breast cancer precursor cell populations within a single tissue.

RESULTS

Breast epithelial cells differ in DDR protein expression

We aimed to identify key biological characteristics that set apart the basal and luminal mammary epithelial lineages. Established cell-surface markers allow the specific isolation of three main mammary epithelial cell populations[32–34]: basal cells (Lin$^-$CD24$^+$CD49f^{hi}), luminal progenitor cells (Lin$^-$CD24$^+$CD49f^{lo}Sca1$^-$CD49b$^+$), and mature luminal cells (Lin$^-$CD24$^+$CD49f^{lo}Sca1$^+$CD49b$^{+/-}$). We previously generated a multimodal profiling resource for these three cell populations, sorted by fluorescence-activated cell sorting (FACS) directly from unirradiated, virgin female wild-type mice.[28] To define potential key differences between these epithelial lineages in the normal mammary gland, we interrogated the mouse mammary proteomic dataset (Figure 1A). Gene set enrichment analysis[35] revealed 57, 63, and 8 upregulated pathways (false discovery rate [FDR] <0.01) in basal, luminal progenitor, and mature luminal populations, respectively (Figure 1B; Table S1). Unexpectedly, luminal progenitors exhibited a remarkable enrichment in pathways related to "DNA replication and repair," with nearly one-third (23/63) of all enriched pathways related to DDR function (Figure 1C), suggesting the possibility of differences in DDR between the mammary epithelial lineages.

Indeed, of the 156 proteins known to participate in the DSB repair pathway (Reactome: R-HSA-5693532; Reactome database ID release 63), 32 were detected in the mouse mammary proteome dataset, and approximately two-thirds of these were enriched in luminal progenitors (Figure S1A), leading to spontaneous grouping of the samples by cell lineage. Likewise, in our analogous human breast proteomes (Figure 1D),[26] unsupervised hierarchical clustering of DDR proteins resulted in the grouping by breast epithelial lineage with 127 of the 276 human DDR genes[6] detected in this dataset (Figure S1B). In addition, the most differentially and consistently upregulated genes in luminal lineages (luminal progenitors, Figure 1E, left; mature luminal, Figure 1E, center) compared with basal cells comprised master regulators of DDR (including ATM and PRKDC/DNA-PK), core components of the MRN complex responsible for sensing of DSBs (MRE11A and RAD50),[36] and several positive regulators of homology-directed repair (such as PARP1, RPA2, FEN1). On the other hand, basal cells showed an upregulation of MDC1, another key mediator of DDR, along with XPC, DDB1, and DDB2, which are involved in nucleotide excision repair of ultraviolet light-damaged DNA. Altogether, the proteomic DDR landscape differed significantly between mammary epithelial lineages in both human and mouse.

Normal mammary epithelial populations have variable ability to respond to DNA damage

We next investigated whether this differential DDR protein abundance translated to functional differences in DDR capacity between mammary lineages upon genotoxic insult. First, we FACS purified primary mammary epithelial cell populations from unirradiated wild-type female mice and performed neutral comet assay to measure tail moment as a measure of DSBs. Notably, at baseline, without DNA-damaging insult, luminal progenitors exhibited a longer tail moment than both mature luminal and basal cells, indicating a higher homeostatic level of DSBs in this cell population (Figure 2A). Then, we mapped DDR kinetics in the murine mammary gland *in situ* by examining the induction and resolution of the DSB markers, γ-H2AX and 53BP1, following *in vivo* irradiation. We initially optimized the *in vivo* irradiation by testing modest, sublethal radiation doses (2–6 Gy), which led to a dose-dependent increase in punctate nuclear γ-H2AX foci in the mammary epithelium (Figure S2A). At 6 Gy, γ-H2AX foci were readily detected in proliferating luminal cells (Ki67$^+$ K14$^-$; Figure S2B), while 53BP1 foci were abundant in Ki67$^-$ cells at early time-points (Figure S2C). Most γ-H2AX and 53BP1 foci were resolved within 24 h (Figures S2B and S2C), demonstrating an overall robust DDR in the mammary epithelium *in vivo*. Then, we established an intracellular flow cytometry assay (Figure S2D) to quantitatively assess γ-H2AX kinetics in the three mammary cell types. γ-H2AX peaked at 0.5–1 h post-irradiation and resolved within 24–48 h in all cell types, yet each lineage displayed unique kinetics of γ-H2AX foci formation and resolution (Figure 2B). The highest proportions of γ-H2AX$^+$ cells were observed across all cell types at early time points following irradiation (Figure 2C, top). Basal cells resolved DSBs and returned to baseline median fluorescence intensity levels of γ-H2AX in the shortest time, while both luminal progenitors and mature luminal cells had more unresolved γ-H2AX at 24 and 48 h (Figure 2C, top). Moreover, luminal progenitors consistently displayed the highest absolute number of γ-H2AX$^+$ cells, even in the unirradiated state, indicating persistent damage (Figure S2E). This was furthermore accompanied by higher median

CellPress
OPEN ACCESS

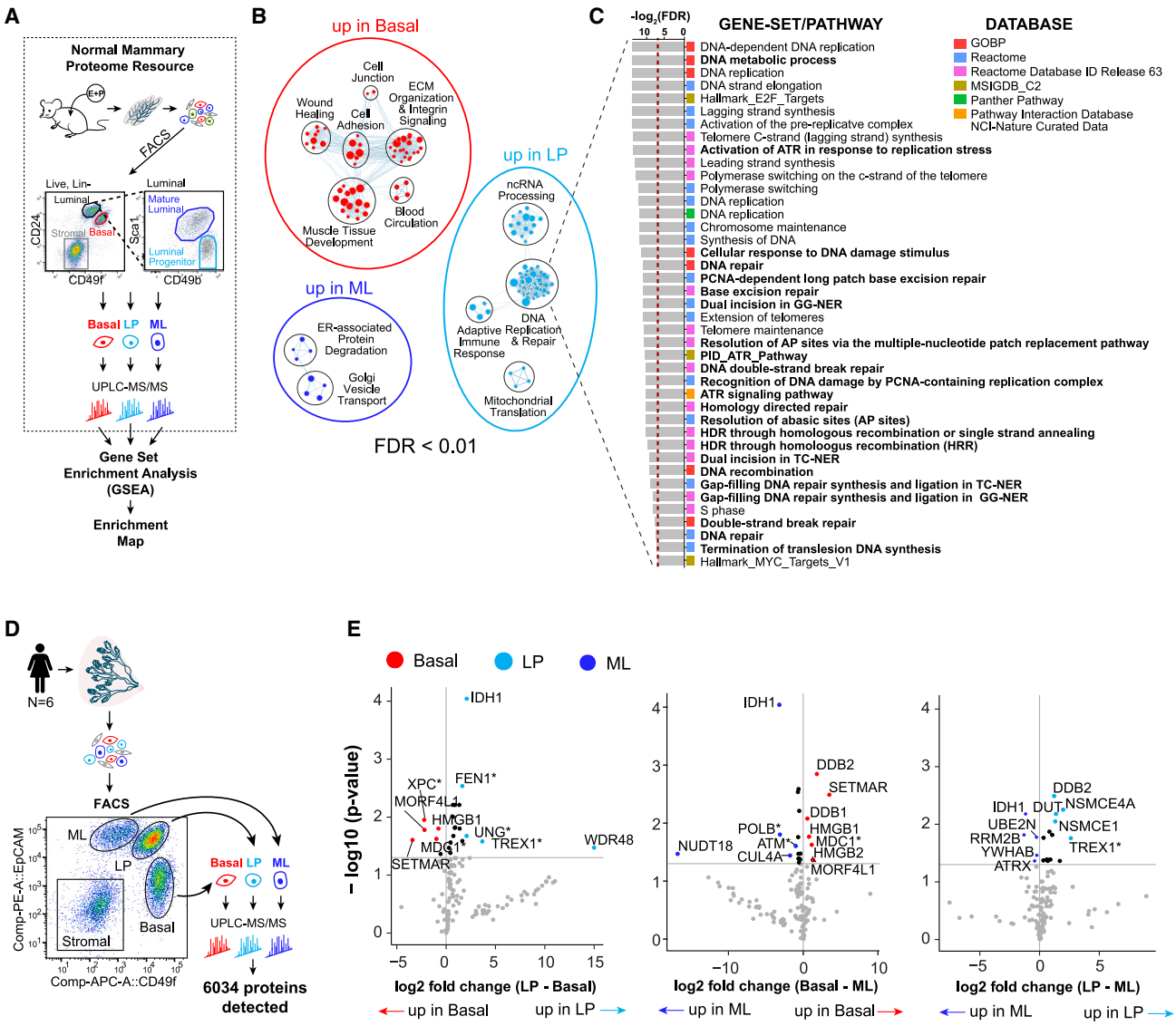


Figure 1. Global proteomics of normal mouse and human mammary epithelial populations reveal lineage differences in DDR

(A) A schematic diagram illustrating interrogation of our previously published normal mouse mammary proteomic dataset. Global proteomic profiles of FACS-purified basal (red; Lin$^-$CD24$^+$CD49^{hi}), luminal progenitor (LP; light blue; Lin$^-$CD24$^+$CD49^{lo}Sca1$^-$CD49$^+$), or mature luminal (ML; blue; Lin$^-$CD24$^+$CD49^{lo}Sca1$^+$CD49$^{+/-}$) epithelial populations from the mammary glands of wild-type female mice under ovarian hormone stimulation (estrogen plus progesterone; E + P) were subjected to pathway enrichment analysis. The resulting enriched pathways for each cell population were visualized into a network by using Enrichment Map.

(B) Enrichment map illustrating clusters of gene sets/pathways (nodes) significantly upregulated in the basal, LP, or ML population compared with the other two as determined by GSEA (false discovery rate [FDR] < 0.01). Each cluster was manually annotated with a common biological theme.

(C) Each gene set/pathway is color-coded by corresponding curated database, and pathways related to DNA damage and repair are in bold. The red dashed line indicates the false discovery rate (FDR) <0.01 cutoff.

(D) Schematic diagram illustrating global proteomic profiling of FACS-purified human breast epithelial population cells, basal (Lin$^-$EpCAM^{lo}CD49^{hi}), LP (Lin$^-$EpCAM^{hi}CD49^{med}), and ML (Lin$^-$EpCAM^{lo}CD49^{lo}), freshly dissociated from six premenopausal reduction mammoplasty breast tissue specimens.

(E) Volcano plots showing differential expression of 127 DDR proteins (defined by Knijnenburg et al.[6]) from the human mammary proteomic dataset generated from FACS-purified basal, LP (light blue) and ML (blue) cell populations. The x axis indicates the relative difference in protein abundance between the two selected populations (log2 fold change); the y axis indicates p values adjusted for multiple correction testing. The gray horizontal line depicts a q value of 0.05. The top 5 significantly upregulated proteins in each pairwise comparison were colored according to cell type. Asterisks identify "core members" of one of the 10 DDR pathways.[6]

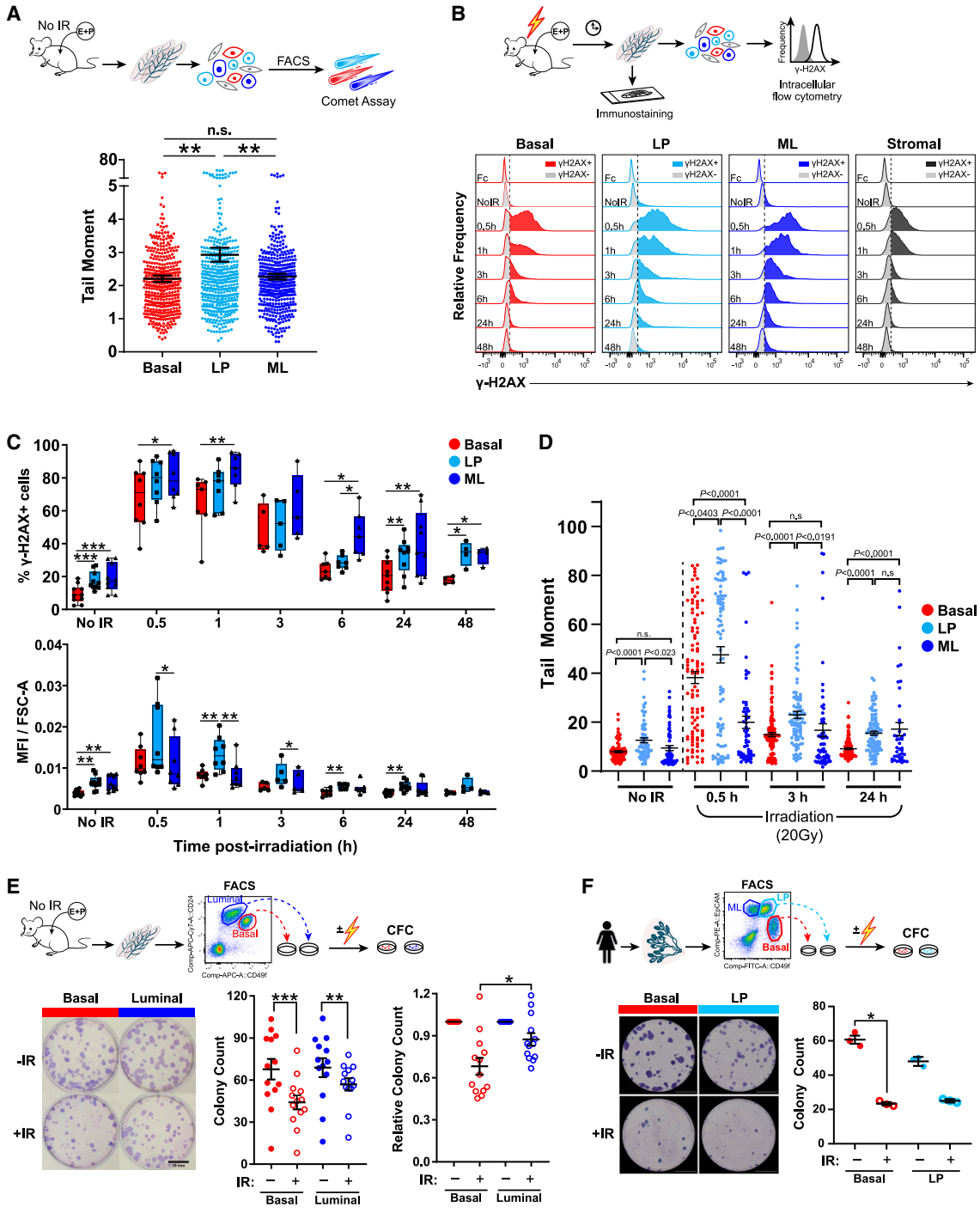


Figure 2. Differential DNA damage response across mammary epithelial populations after genotoxic insult

(A) Workflow of neutral comet assay. Scatterplot showing tail moment on sorted cells from basal, luminal progenitor (LP), and mature luminal (ML) populations from wild-type female mice under ovarian hormone stimulation (estrogen plus progesterone; E + P). A total of ~60–150 comets were scored per cell population, per mouse (n = 4). Statistical significance was calculated using one-way ANOVA with Tukey's multiple comparison test; **p < 0.01; n.s., not significant.

(B) Workflow depicting intracellular flow cytometry of γ-H2AX in single cells from the mammary gland and immunofluorescence staining on the mammary tissue sections following 6 Gy *in vivo* irradiation. Representative flow cytometry histograms of γ-H2AX recruitment in basal, LP, ML, and stromal populations at indicated time points post-irradiation. The dotted line delineates γ-H2AX positivity (γ-H2AX$^+$ cells) from the background (γ-H2AX$^-$ cells) based on the Fc control.

(legend continued on next page)

fluorescence intensity (Figure 2C, bottom) corrected by the nuclear size differences (Figure S2F), indicating overall higher levels of γ-H2AX in luminal progenitors. Altogether, the luminal lineage, and especially the luminal progenitors, exhibited greater susceptibility to DNA damage and mounted a stronger, prolonged DDR than basal cells.

To assess how this DDR difference affects DNA integrity, we quantified DSBs at baseline and upon irradiation at the single-cell level in the three cell populations. Exposure of FACS-purified cells to irradiation (20 Gy) *in vitro* increased tail moments in all three populations at 0.5 h (Figure 2D). Across all time points, luminal progenitors displayed the highest DNA tail moment, followed by basal and mature luminal. By 24 h, most of these DSBs were resolved in basal cells, while luminal cells again lagged. To assess the significance of these differences in functional DDR capacity for mammary epithelial lineages, we enumerated the effects of irradiation-induced DNA damage on their lineage-specific progenitor activity and clonogenic survival through colony-forming cell (CFC) assays with and without irradiation (3 Gy; Figure 2E, left). As anticipated, irradiation itself dampened the progenitor capacity of all primary mammary epithelial cells (Figure 2E, middle), yet the luminal lineage retained relatively more colony formation capacity than the basal lineage, with a reduction of 13% vs. 32%, respectively (Figure 2E, right). Notably, a similar CFC assay performed on lineage-purified human mammary epithelial cells also showed a significant reduction in basal colonies, while marginally affecting the luminal progenitor colonies (Figure 2F). Collectively, luminal progenitors exhibit an enhanced ability to engage DDR, leading to better survival.

Mammary epithelial lineages exhibit differential engagement of HR

We next queried whether this differential DDR capacity of the mammary epithelial linages could be attributed to varying utilization of HR and NHEJ, the two major DSB repair pathways. HR engagement is reflected via nuclear RAD51 foci and NHEJ via phospho-DNA-PKcs (S2056) foci. To precisely quantify protein recruitment to nuclei at the single-cell level in the three mammary populations freshly isolated from mice after irradiation, we employed Amnis imaging flow cytometry (Figures 3A and S3A). At baseline, without insult by irradiation, luminal progenitors already displayed a strikingly higher proportion of cells with nuclear RAD51 positivity (RAD51$^+$) compared with basal cells, in both relative and absolute levels (Figures 3B, 3C, and S3B). Meanwhile, mature luminal cells exhibited the highest baseline phospho-DNA-PKcs$^+$ fractions followed by luminal progenitors and basal cells (Figures 3D, S3C, and S3D). Upon irradiation, the proportion of RAD51$^+$ cells increased in all populations, with the luminal lineage once again showing higher ratios of RAD51$^+$ cells than the basal lineage (Figures 3C and S3B). Conversely, the proportion of phospho-DNA-PKcs$^+$ cells increased to relatively similar amounts in all three populations (Figures 3D and S3C). These data collectively suggest that NHEJ is consistently utilized across all mammary epithelial cells, while HR appeared to be predominantly engaged by the luminal lineage, especially by luminal progenitors.

Lineage-specific HR differences exist irrespective of the cell cycle

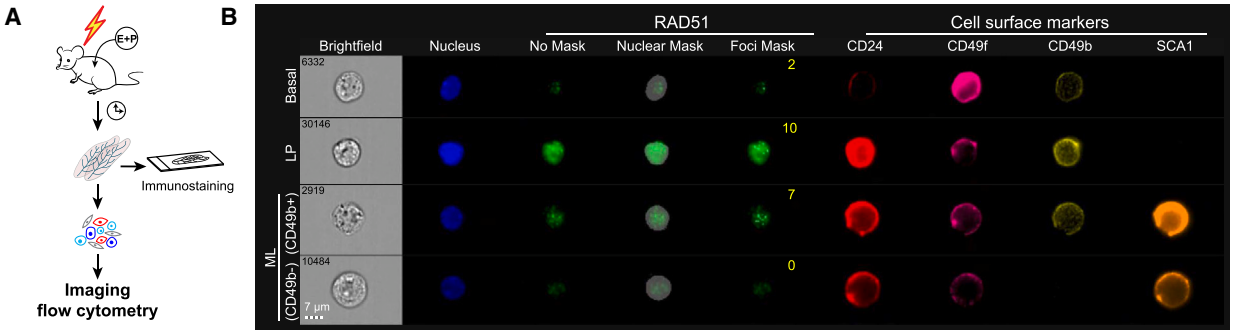
The above data suggested that luminal cells were more susceptible to DSBs than basal cells but able to evade the cytotoxic consequences (Figure 2), which might be related to their effective engagement of repair pathways. However, DSB repair is tightly coordinated with cell cycle: HR is restricted to cells in late S/G2 phases, while NHEJ is predominantly used in the G1 phase.[37] The adult mammary gland undergoes cyclical remodeling in response to circulating ovarian hormones during the reproductive cycle. Thus, throughout our study, ovariectomized mice were treated with estrogen + progesterone pellets to ensure that both basal and luminal progenitor populations were activated and at the maximal proliferative states. We tested whether differential engagement of these two DSB repair pathways in luminal vs. basal cells is influenced by differences in their proliferative behavior. We performed *in situ* immunofluorescence analysis of RAD51 and phospho-DNA-PKcs foci kinetics, along with a proliferation marker, Ki67, in mammary tissue after *in vivo* irradiation. The RAD51 foci were primarily observed in proliferating (Ki67$^+$) cells as early as 0.5 h and progressed into larger, punctate foci by 3–6 h (Figure 3E). We then categorized proliferating/Ki67$^+$ cells into one of the four groups, RAD51-positive or RAD51-negative basal (K14$^+$) and RAD51-positive or RAD51-negative luminal (K14$^-$), to investigate whether the frequency of RAD51$^+$ cells in the two lineages differed. Notably, RAD51 foci were predominantly present in the proliferating luminal but largely absent in the proliferating basal cells ($p < 0.05$, Fisher's exact test; Figure 3F); RAD51$^+$ events in the proliferating luminal cells were approximately twice more frequent than proliferating basal cells (odds ratio 0.48; Figure 3F). Further co-staining with the progesterone receptor (PR), a marker to separate mature

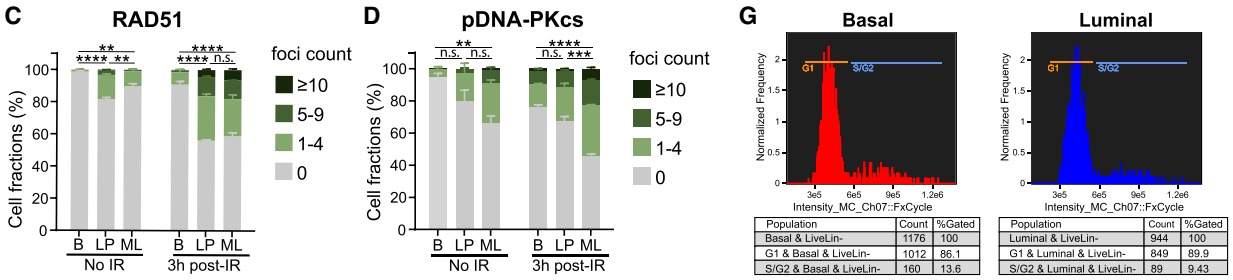
(C) Boxplots showing relative number of γ-H2AX$^+$ cells in each population (n = 4–10 mice per time point) or median fluorescence intensity (MFI) in γ-H2AX$^+$ cells normalized by cell size (FSC-A; see Figure S2F). One-way ANOVA with Tukey's multiple comparison test was performed on each time point; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

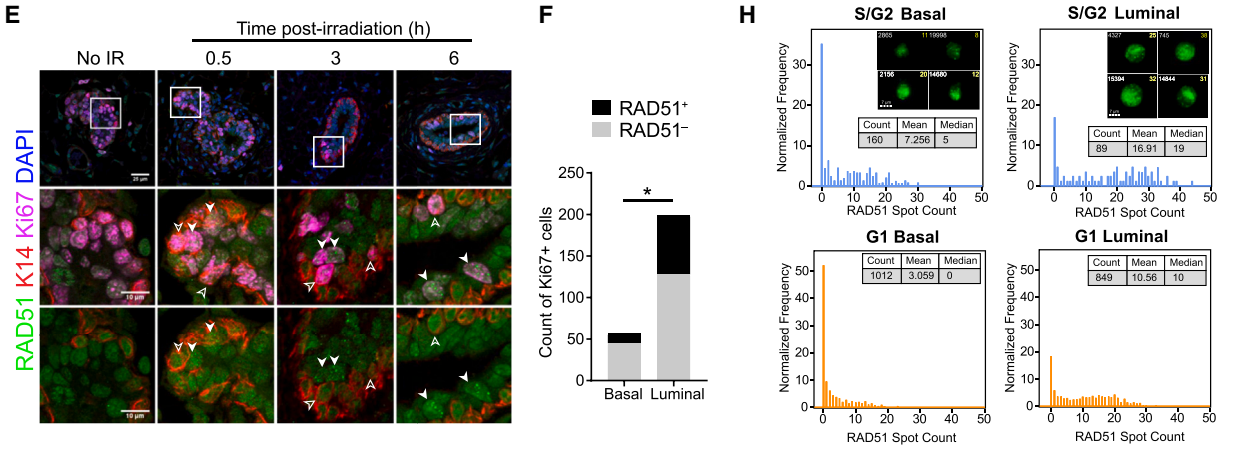
(D) Neutral comet assay following *in vitro* irradiation of 20 Gy. A total of ~40–150 comets were scored. The p values were calculated using one-way ANOVA with Tukey's multiple comparison test; n.s., not significant.

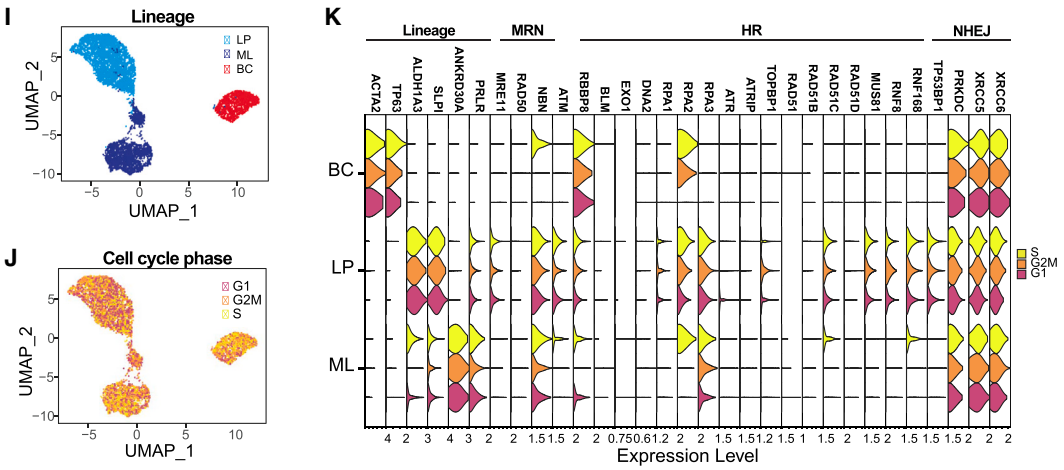
(E) Workflow of colony-forming cell (CFC) assay on FACS-purified basal (red) and total luminal (blue) cells from mouse mammary glands. The seeded cells were irradiated *in vitro* (3 Gy). Representative images of basal and luminal colonies with or without irradiation (IR) after 7 days of culture are shown. Scatterplot showing the absolute counts of basal (red) and luminal (blue) colonies with or without IR (left; n = 13 mice; paired t test) is shown. Colony count of irradiated samples relative to the paired unirradiated sample is also shown on the right (n = 13; unpaired *t* test). All data represent the mean ± SEM; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

(F) Workflow of CFC assay on FACS-purified basal (red) and total luminal (blue) cells from human breast tissue specimens (n = 3). The seeded cells were irradiated *in vitro* (3 Gy). Representative images of basal and luminal colonies with/without IR after 10 days of culture are shown. Scatterplot showing the absolute counts of basal (red) and luminal (blue) colonies with/without IR is shown (Kruskal-Wallis test followed by Dunn's multiple comparison test). All data represent the mean ± SEM; *$p < 0.05$.









(legend on next page)

luminal cells ($K14^{-}PR^{+}$) from luminal progenitors ($K14^{-}PR^{-}$), showed that the enrichment of RAD51 foci was predominant in PR^{-} luminal cells (luminal progenitors), possibly due to their tendency to be more proliferative than PR^{+} (mature) luminal cells (Figures S4A and S4B). Meanwhile, phospho-DNA-PKcs foci were observed in both lineages at all time points, and mostly in $Ki67^{-}$ cells, indicating equivalent NHEJ activity (Figure S4C). To further investigate the observed lineage differences in RAD51 foci recruitment specifically in the S/G2 phase, we conducted Amnis imaging flow cytometry in combination with a cell-cycle dye, FxCycle (Figures 3G and S4D). Cell-cycle distributions were similar between the basal and the luminal populations, with 9.43% of luminal and 13.6% of basal cells in S/G2 phase (Figure 3G). Overall, RAD51 foci recruitment was higher in S/G2 populations compared with cells in G1, as expected. Notably, the median number of RAD51 foci per cell was found to be higher in the luminal population compared with the basal during S/G2 (19 vs. 5 foci; Figure 3H, blue) and, surprisingly, even in G1 (10 vs. 0 foci), where HR repair is expected to be suppressed (Figure 3H, orange). Thus, irradiated luminal cells maintain a comparatively higher HR repair capacity throughout the cell cycle.

Next, to examine the relationship between lineage, DNA repair, and cell cycle in the human breast, we leveraged our published single-cell RNA sequencing (scRNA-seq) dataset[26] and computationally mapped the distinct cell populations (Figure 3I), where single cells were stratified into one of the three cell-cycle phases, G1, S, and G2/M (Figure 3J), based on their gene expression profile. Potential differences in the DSB repair gene expression landscape during the cell cycle were interrogated through critical genes involved in HR and NHEJ and those involved upstream of both pathways (*ATM*, *MRE11*, *RAD50*, *NBN*).[38,39] While NHEJ genes *PRKDC*, *XRCC5*, and *XRCC6* were ubiquitously expressed across mammary lineages and cell cycle, the expression of several critical HR genes, such as *RAD51C*, *RPA1*, *RPA3*, and *TOPBP1*, was restricted to the luminal lineage (Figure 3K). Additional genes essential for effective HR, such as *MUS81*, *RNF8*, and *RNF168*,[40–42] were also largely restricted to the luminal progenitor lineage (Figure 3K). Collectively, the differences in engagement and expression of DSB repair genes between luminal progenitors and basal cells were not attributable to the cell cycle.

Mammary cell lineage is a key factor in PARPi sensitivity

In tumors, HR deficiency is a major therapeutic vulnerability that is successfully being exploited through the use of PARPis. Based on the aforementioned HR disparities observed in normal mammary lineages, we asked whether the different progenitor pools, from which different breast cancer subtypes are thought to arise, also carry an inherently different vulnerability to PARPis. By utilizing the synthetic lethality concept to test intrinsic HR deficiency in the basal lineage, we assessed the effect of the PARPis olaparib and talazoparib on the colony-forming capacity of mammary progenitor populations freshly isolated from unirradiated, wild-type female mice (Figure 4A). We observed strikingly divergent vulnerabilities of luminal vs. basal progenitors: basal colony counts decreased precipitously upon PARPi treatment in a dose-dependent manner, while luminal progenitors were comparatively more resistant (Figures 4B and 4C). We also assessed the dependence of lineage-specific clonogenicity on the NHEJ pathway. Treatment with KU-57788, a DNA-PK inhibitor that targets NHEJ, did not show any differential effects on basal or luminal CFCs (Figure 4D).

To determine whether lineage-specific differences in PARPi vulnerability are also seen in humans, we isolated human mammary epithelial progenitor populations from non-neoplastic, prophylactic mastectomy specimens that were BRCA wild type (Figure 4E). CFC assay performed on FACS-purified mammary epithelial cells showed basal-specific sensitivity to the highly potent PARPi talazoparib but not olaparib (Figures 4F and 4G). We then tested whether the lineage-specific response to PARPi is exacerbated by *BRCA1/2* heterozygosity using 20

Figure 3. HR repair is intrinsically different between luminal progenitors and basal cells following irradiation

(A) Workflow of imaging flow cytometry and tissue immunostaining from the mouse mammary gland 3 h after *in vivo* irradiation (6 Gy).

(B) A panel of representative Amnis ImageStreamX Mark II images displaying RAD51 foci in the basal, luminal progenitor (LP), and mature luminal (ML) populations as marked by a set of established cell-surface markers. Foci counts, determined by IDEAS software, are indicated in yellow in the RAD51 "Foci Mask" column.

(C and D) Bar graphs summarizing proportion of cells displaying 0, 1–4, 5–9, or ≥10 foci of RAD51 (C) or phospho-DNA-PKcs (S2056; p-DNA-PKcs; D) in each population before or 3 h post-irradiation (n = 3–4 mice per treatment group). Data represent the mean ± SEM and p values were determined by one-way ANOVA with Tukey's multiple comparison test on $RAD51^{+}$ or p-$DNA\text{-}PKcs^{+}$ cells (i.e., cells displaying ≥1 focus) across all three mammary populations; $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$; n.s., not significant.

(E) *Z*-projected confocal immunofluorescence images of RAD51 co-stained with basal cytokeratin (K14) and proliferation (Ki67) markers on mammary tissues harvested from virgin female mice with or without *in vivo* irradiation. Filled white arrowheads indicate proliferating/$Ki67^{+}$ $RAD51^{+}$ cells, and hollow arrowheads indicate proliferating/$Ki67^{+}$ $RAD51^{-}$ cells.

(F) Absolute counts and proportion of proliferating/$Ki67^{+}$ cells that exhibit $RAD51^{+}$ (≥2 foci) in basal ($K14^{+}$) or luminal ($K14^{-}$) cells at 3 h post-irradiation by immunofluorescence staining (n = 3 mice). The number of RAD51 foci per nucleus was determined by ImageJ. The p value was calculated by Fisher's exact test; $^{*}p < 0.05$.

(G) Histograms of the DNA content in basal (red) or luminal (blue) cells as determined by Amnis imaging flow cytometry. Exact cell counts in G1 or S/G2 phase are shown in the tables.

(H) Distribution of RAD51 focus count exhibited in irradiated basal or luminal cells from G1 (orange) or S/G2 (blue) phase of the cell cycle. Mean and median values of RAD51 focus count were determined for G1 basal, G1 luminal, S/G2 basal, and S/G2 luminal cell populations.

(I and J) Uniform manifold approximation and projection (UMAP) embedding of four aggregated human breast samples from previously published scRNA-seq.[26] The UMAP embeddings of the three human breast cell populations (basal/BC, n = 1,272; LP, n = 3,364; ML, n = 2,172 cells) are colored by the lineage or cell-cycle phase.

(K) Violin plots depicting the expression levels of lineage markers and DSB repair genes (homologous recombination, HR; non-homologous end joining, NHEJ) in the scRNA-seq dataset,[26] across three mammary epithelial lineages (BC, LP, and ML) stratified by cell-cycle phase.

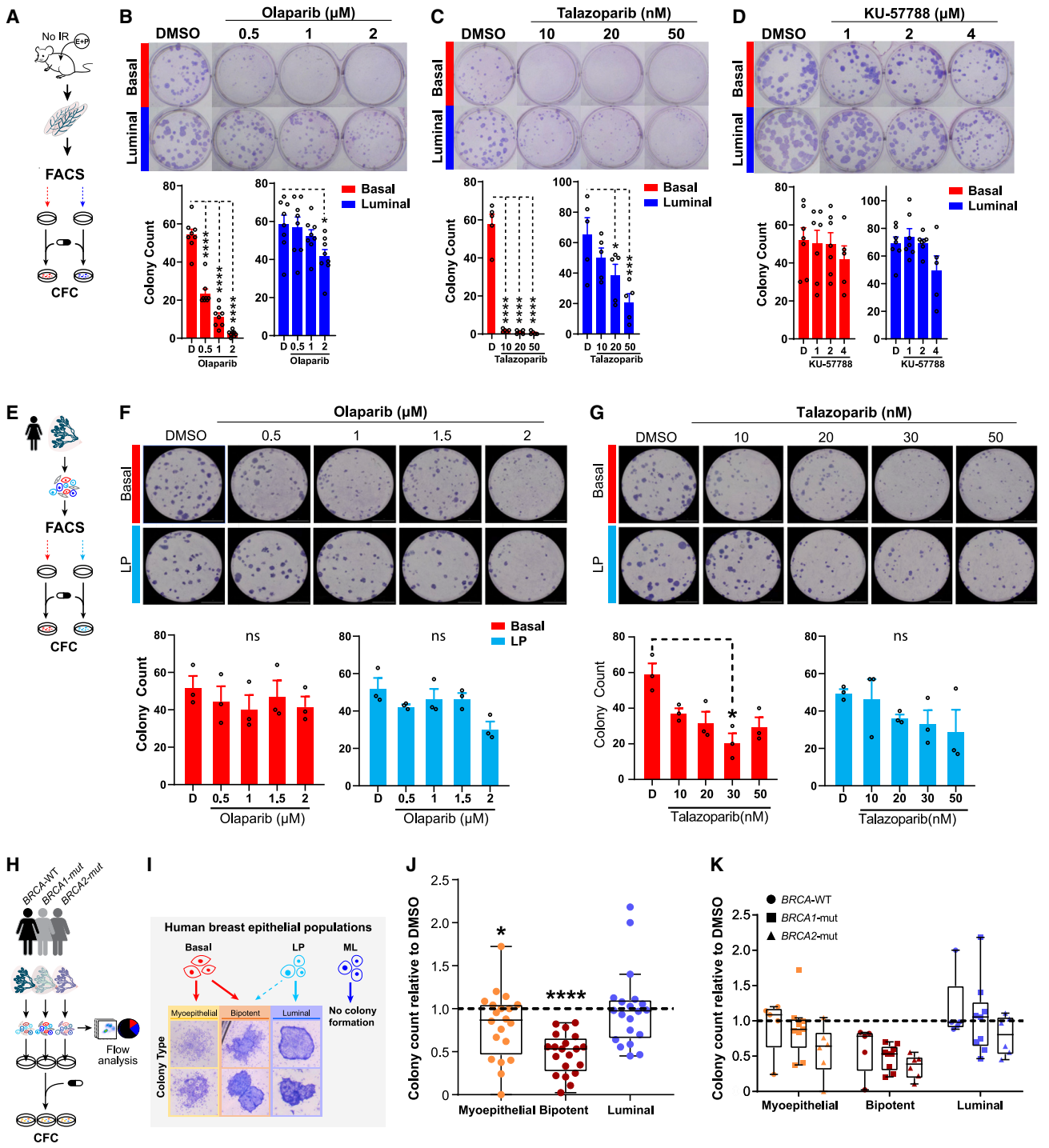


Figure 4. Luminal and basal progenitors exhibit divergent PARPi vulnerability

(A) Workflow of colony-forming cell (CFC) assay from FACS-purified basal (Lin$^-$CD24$^+$CD49f^{hi}; red) and total luminal (Lin$^-$CD24$^+$CD49f^{lo}; blue) cell fractions to evaluate basal and luminal progenitor capacities in response to olaparib, talazoparib, or KU-57788.

(B–D) Representative images of basal and luminal colonies treated with olaparib, talazoparib, KU-57788, or DMSO alone (vehicle control). Bar graphs show the colony count for basal (red) and luminal (blue) populations at the indicated drug concentrations after 7 days of culture (n = 5–8 mice). One-way ANOVA with Dunnett's multiple comparisons test was performed; *p < 0.05, ***p < 0.001, ****p < 0.0001.

(E) Workflow of CFC assay from FACS-purified human basal (Lin$^-$CD24$^+$CD49f^{hi}; red) and luminal progenitor (Lin$^-$CD24$^+$CD49f^{lo}; light blue) fractions to evaluate basal and luminal progenitor capacities in response to olaparib or talazoparib.

(legend continued on next page)

specimens from patients of comparable age (n = 6 BRCA wild type; n = 9 *BRCA1*-mutation carriers; n = 6 *BRCA2*-mutation carriers; Figures 4H and S5A). Dissociated total breast epithelial cells give rise to morphologically distinct colonies that reflect the activity of the respective progenitor populations (Figure 4I). Luminal progenitors form luminal colonies, mature luminal cells lack clonogenicity in conventional CFC culture conditions, while the basal population harbors both myoepithelial progenitors and bipotent progenitors.[43–45] Comparison of patient-paired CFC assays allowed us to quantitatively assess the relative olaparib sensitivity of the different progenitor populations contained within each patient sample (Figures S5B and S5C). In every sample pair, across the wild-type allele and *BRCA1/2* mutation carriers, we observed a reduction in bipotent colony count upon olaparib treatment, along with a mild reduction in myoepithelial colony counts in approximately half of the cases (Figure 4J). On average, this translated to a 13% myoepithelial and 47% bipotent colony abrogation ($p < 0.05$ and $p < 0.0001$, respectively, ratio paired t test), while luminal colonies remained largely resistant (Figure 4J). Colonies from *BRCA1/2* heterozygous mutation carriers displayed overall higher olaparib sensitivity ($p < 0.05$, two-way ANOVA; Figure 4K), yet notably, this effect was largely driven by basal lineage-derived colonies (myoepithelial, bipotent), while the resistance of the luminal lineage toward PARPi was independent of BRCA mutation status ($p < 0.0001$, two-way ANOVA; Figure 4K). Collectively, these data show that cell lineage is a notable determinant of PARPi vulnerability in murine as well as human mammary progenitor populations.

Mammary lineage resemblance correlates with PARPi response of breast cancer cells

We investigated whether the lineage-specific DDR vulnerabilities observed in normal breast progenitors would translate to vulnerabilities in corresponding breast cancers. To this end, we were able to identify four murine breast cancer cell lines that exhibited the basal or luminal lineage characteristics, as assessed by expression of the lineage markers EPCAM and CD49f (Figure S6A), and examined their PARPi sensitivity when grown as xenografts (Figure 5A). This included (1) two cell lines (sg4205 and sg4687) derived from *Pik3ca*^{H1047R} mice with CRISPR-Cas9-mediated *p53* ablation (*Pik3ca*^{H1047R}*p53*$^{-/-}$),[46,47] both exhibiting a basal FACS profile[48,49]; (2) a polyoma middle T (PyMT) model-derived cell line with a luminal FACS profile; and, in addition, (3) NDL3903,[50] a cell line derived from a HER2^{+} breast cancer model with a mixed basal-luminal FACS profile (Figure S6A). After implantation of tumor cell lines into NSG mice, we followed tumor growth under treatment with olaparib, talazoparib, and vehicle control until the endpoint (Figure 5A). While olaparib treatment showed no obvious tumor growth inhibition, talazoparib significantly impeded the growth of tumors formed from cells with the basal phenotype (sg4205 and sg4687), but not the other tumor models (Figure 5B).

We proceeded to interrogate the PARPi sensitivity of human breast cancer cell lines by leveraging the Genomics of Drug Sensitivity in Cancer (GDSC) database, which catalogs drug responses for over 100 anti-cancer agents.[51] First, as a proxy for the likeness of these human breast cancer models to basal, luminal progenitor, or mature luminal cells, we constructed "lineage signatures" based on our comprehensive global mammary proteomes, comprising a total of 6,034 human and 4,695 mouse proteins. These lineage signatures were obtained by identifying proteins enriched in one cell type compared with the other two, with a fold change >5 in human (Figure S6B; Table S2) or fold change >3 in mouse mammary proteomes (Figure S6B; Table S3). The resulting lists consisted of unique and highly expressed proteins that met these specific cutoffs, ensuring an unbiased approach for generating representation of lineage signatures. We then assessed their molecular enrichment across the five major breast cancer subtypes from the METABRIC database. In accordance with the cell-of-origin model in breast cancer,[52,53] our human proteomic basal signature showed the highest enrichment in the mesenchymal "Claudin-low" subtype (Figures 5C and S6B). The luminal progenitor signature was most closely linked to the "basal-like" subtype (Figure 5C), which is thought to arise from luminal progenitors, not basal cells,[54] despite its histopathological classification label.[55] This subtype has also been associated with *BRCA1*-mutated breast cancers.[56,57] The mature luminal signature was highly correlated with the hormone-receptor-expressing "luminal A" subtype (Figure 5C). Mouse mammary proteome-derived signatures exhibited similar patterns (Figures S6B and S6C), consistent with the anticipated associations of breast cancer subtypes. We next compared enrichment of the basal and luminal progenitor signatures in 50 human breast cancer cell lines curated in the GDSC database, selected the top 10 cell lines most enriched in each lineage signature (Figure 5D; Table S4), and correlated

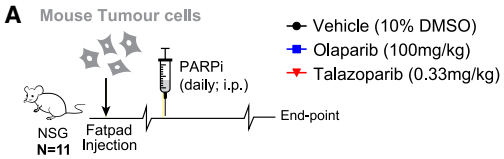
(F and G) Representative images of basal and luminal colonies treated with olaparib, talazoparib, or DMSO alone (vehicle control). Bar graphs show the colony count for basal (red) and luminal (light blue) populations at the indicated drug concentrations after 10 days of culture (n = 3). Kruskal-Wallis test with Dunnett's multiple comparisons test was performed; *$p < 0.05$; ns, not significant.

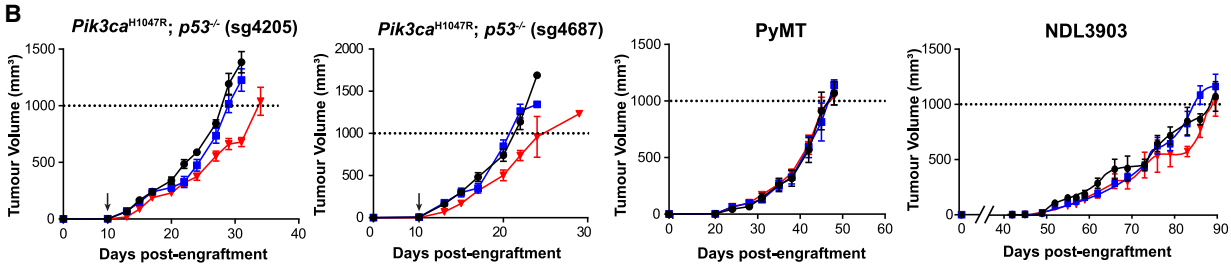
(H) Workflow for flow cytometry and olaparib testing using the CFC assay on 20 primary human breast specimens from prophylactic mastectomy procedures. The specimens are categorized into *BRCA* wild-type (n = 5), *BRCA1*-mutated (n = 9), and *BRCA2*-mutated (n = 6) carriers.

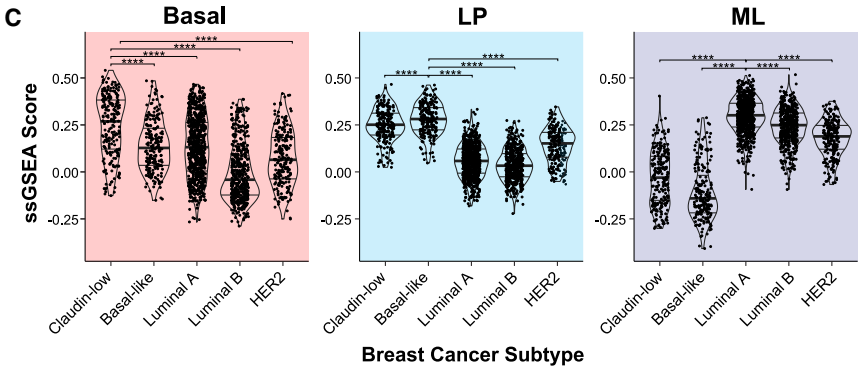
(I) Schematic diagram illustrating human mammary epithelial populations and representative images of three types of colonies that arise from corresponding progenitor activities.

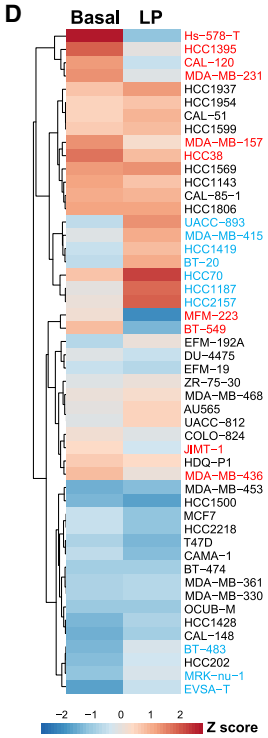
(J) Boxplot summarizing counts of myoepithelial, bipotent, or luminal colonies compared with DMSO vehicle control (horizontal dotted line) in all primary breast specimens (n = 20). The p values for colony growth in olaparib vs. DMSO were calculated by ratio paired t test; *$p < 0.05$, ****$p < 0.0001$.

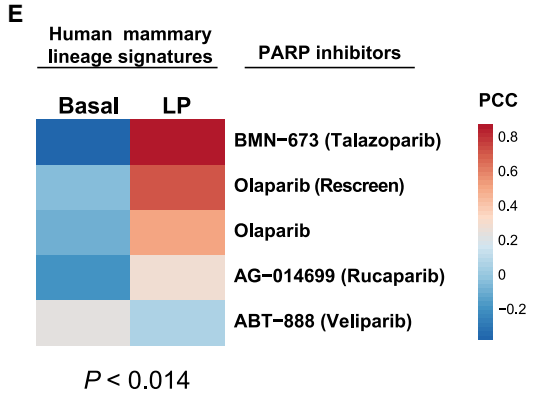
(K) Boxplot summarizing the effect of olaparib on the growth of each colony type, according to BRCA mutation status (*BRCA* wild type, n = 5; *BRCA1* mutated, n = 9; *BRCA2* mutated, n = 6), compared with DMSO control (horizontal dotted line). Whether the two independent factors "colony type" and "BRCA mutation status" affected colony growth was tested by performing two-way ANOVA. The colony type significantly affected colony growth upon olaparib treatment ($p < 0.001$). The BRCA status significantly affected colony growth upon olaparib treatment ($p < 0.05$). No interaction was found between these two factors ($p < 0.9754$). *Post hoc* analysis by Tukey's multiple comparisons test showed no differences in colony growth of each colony type based on BRCA mutation status (p values ranging from 0.24 to 0.98).











(legend on next page)

their drug sensitivity to 23 DNA-damaging agents, including five PARPis (Figures 5E and S6D; Table S4). In accordance with our findings from the normal human mammary epithelium, breast cancer cell lines enriched in basal vs. luminal progenitor signatures exhibited differential responses, where those with a strong basal signature displayed greater PARPi sensitivity (i.e., negatively correlated, $p < 0.014$; Table S4). As variations in PARPi sensitivity can be influenced by PARP1 levels[58] and drug efflux pumps,[59] we examined their expression in the top 10 GDSC cell lines enriched in either basal or luminal progenitor lineage signatures and found no differences in the expression of *PARP1* or the drug efflux pump (*ABCB1*; Figure S6E).

Next, we selected human breast cancer cell lines for xenograft assays from the top 10 cell lines in the GDSC database that consistently showed basal or luminal characteristics at both molecular and phenotypic levels. These criteria included (1) clear enrichment of the respective lineage signatures, (2) confirmed expression profiles of established lineage markers via immunofluorescence (K14, K18, ER, and PR) and flow cytometry (EPCAM and CD49f; Figure S7A), and (3) demonstrated reliable ability to form tumors in NSG mice (Figure S7B). Consequently, MDA-MB-231 and HCC1187 were the only cell lines that met these criteria and were subsequently engrafted in mice for daily treatment with either olaparib or vehicle control until the endpoint (Figure 6A). HCC1187 tumors, which have luminal lineage characteristics, showed no sensitivity to treatment, while olaparib significantly impeded the growth of basal characteristic MDA-MB-231 tumors (multiple t test; FDR <0.01; Figure 6B). Furthermore, flow cytometric analysis confirmed that the established MDA-MB-231 tumors maintained the expression of the mesenchymal stem cell marker CD44 and low expression of human epithelial cell marker (h-EpCAM), while HCC1187 lacked CD44 and expressed h-EpCAM highly (Figure 6C), in concordance with characteristics of tumors derived from a basal vs. a luminal origin, respectively. Altogether, these data indicate that the intrinsic PARPi sensitivity identified here in normal basal cells persists in a variety of human and murine breast cancer models.

Finally, we applied our lineage signatures to transcriptomes of 47 human breast cancer patient-derived xenograft (PDX) models[60] and correlated these with treatment response to talazoparib (0.33 mg/kg). Examination of the association between individual PDXs and their enrichment scores of the three lineage signatures revealed distinct tumor segregation (Figure 6D; Table S5). We observed a group of PDXs that were mainly enriched in basal or luminal progenitor signatures, while some exhibited enrichment in both. In addition, PDXs enriched in mature luminal signatures were notably distinct (Figure 6D). To investigate talazoparib sensitivity, we examined the relationship between individual PDXs lineage scores and their response to talazoparib. We observed a significant correlation between talazoparib sensitivity and the basal-lineage-enriched PDXs ($p < 0.007$), but not with luminal progenitor (Figure 6E; Table S5). Intriguingly, an inverse trend was noted with the mature luminal signature (Figure 6E). Collectively, these data demonstrate that the mammary lineage signatures can be applied to breast cancer xenografts and that these signatures are associated with PARPi sensitivity *in vivo*.

DISCUSSION

Understanding the key biological differences between the mammary epithelial lineages can help us better understand and target vulnerabilities of breast cancers that arise from the different cells of origin within the epithelium. This study reveals that mammary epithelial cell lineages have striking differences in their ability to engage DNA damage repair processes and sensitivity to PARPis, and the differential PARPi sensitivity was reflected in breast cancer models that resembled the molecular identity of these lineages.

In an unsupervised analysis across human and mouse, differences in DNA damage repair dominated the proteomic landscape of luminal vs. basal cells in the mammary epithelium. This unexpected finding challenged the notion that DSB repair capacity is uniform across normal somatic cells. Certain DDR and DSB repair variances have been reported in the mammary epithelium due to telomere-associated,[30] oncogene-mediated,[61] and endonuclease-induced factors.[62] Our series of studies centered on *in situ* and single-cell analyses, as well as clonogenic assays, and demonstrated that cell lineage is a previously unknown intrinsic determinant of DSB repair pathway choice in the normal mammary gland with consequences for PARPi sensitivity. Luminal progenitors were naturally more susceptible to generating DSBs at homeostasis and upon ionizing radiation, as evidenced by the high numbers of γ-H2AX-positive cells. Following irradiation, the luminal lineage utilized HR repair, whereas the basal lineage had limited capacity. Remarkably, even when considering the cell-cycle phases, lineage-rooted HR disparities persisted. This fundamental observation in the normal mammary gland translated into a selective PARPi vulnerability of the basal lineage over the luminal lineage. In clonogenic assays, this manifested as differential radiation sensitivity of the two lineages in both mouse and human, as well as distinct

Figure 5. Normal mammary lineage signatures correlate with breast cancer PARPi sensitivity

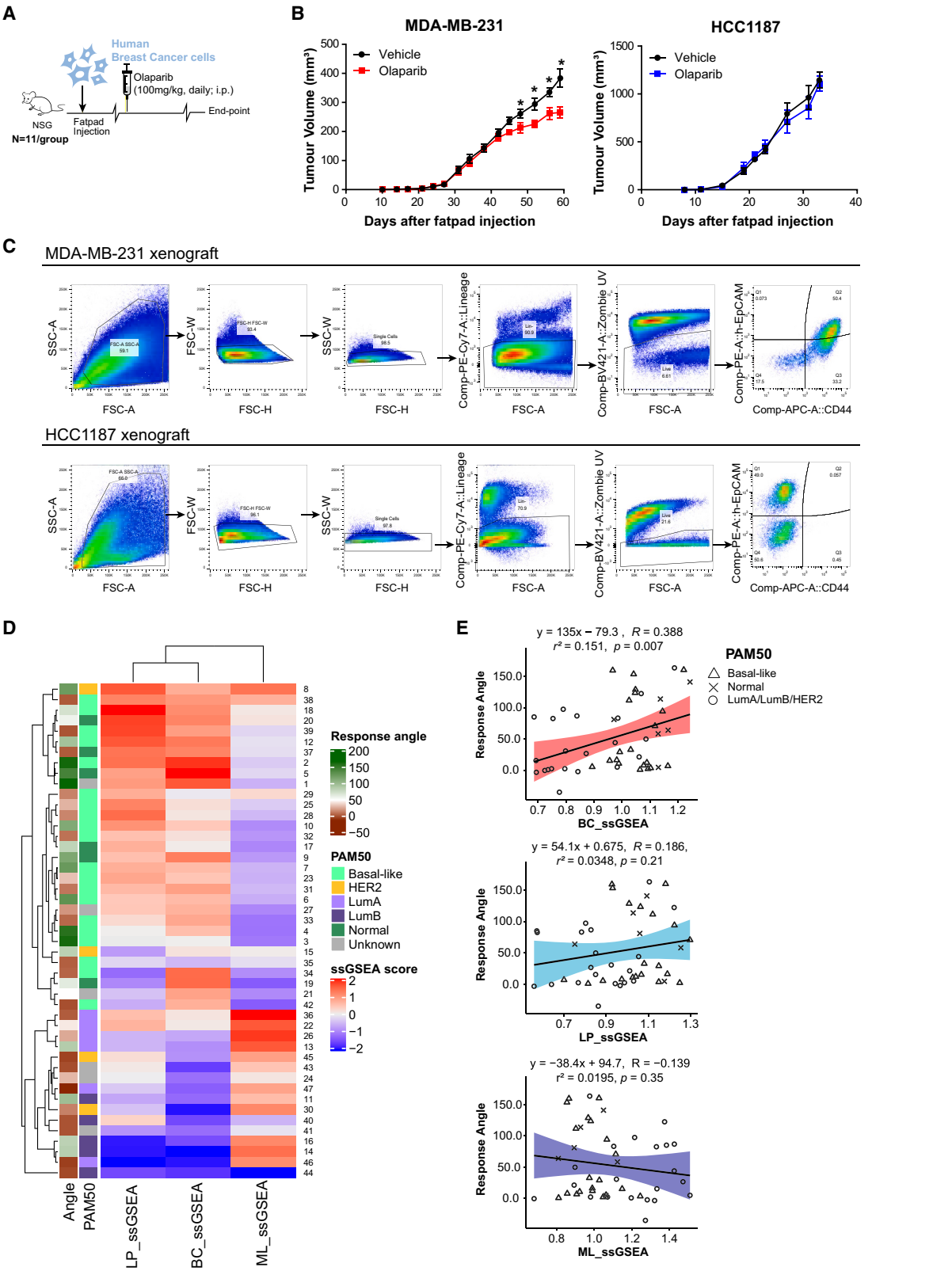
(A) Schematic diagram illustrating the mouse mammary tumor xenograft assay.

(B) Xenograft growth curves depicting the response of "basal" (*Pik3ca*^{H1047R}; *p53*$^{-/-}$) or "luminal" (PyMT and NDL3903) mouse mammary tumor cell lines to *in vivo* treatment with olaparib or talazoparib (n = 11 mice per treatment group).

(C) Violin plots illustrating the single-sample GSEA (ssGSEA) enrichment score for each of the three human-proteome-defined mammary lineage signatures across five major breast cancer molecular subtypes (Claudin-low, basal-like, luminal A/B, HER2$^+$) from the METABRIC gene expression dataset. Unpaired t test was used, and the p values were adjusted for multiple testing; ****$p < 0.0001$.

(D) Heatmap showing unsupervised hierarchical clustering of 50 human breast cancer cell-line gene expression profiles based on human basal and luminal progenitor (LP) proteome-defined lineage signatures. The top 10 cell lines that display the largest differential enrichment based on ssGSEA scores between basal and LP signatures are highlighted in red or blue, respectively.

(E) Heatmap of Pearson correlation coefficient (PCC) values representing the correlations between human basal and LP signature scores and drug sensitivity (based on normalized half-maximal inhibitory values, or IC_{50}) to the five PARP inhibitors from the GDSC database.



(legend on next page)

synthetic lethality of basal progenitors to PARPis compared with the luminal progenitors. Importantly, this lineage-dependent PARPi vulnerability was also observed in normal human breast tissues and in prophylactic mastectomy samples from *BRCA1* and *BRCA2*-mutation carriers, where olaparib preferentially targeted bipotent clonogenicity known to arise from the basal lineage. This uncovers mammary epithelial cell lineage as a novel, non-mutational determinant of PARPi response in the normal human breast.

The paradox of luminal progenitors exhibiting higher levels of DSBs both at baseline and following irradiation despite having the capacity to engage HR repair requires further study. Other distinguishing features inherent to mammary lineages, such as the epigenomes and cell metabolism, may play a role in this phenomenon.[26,28] For instance, a mechanism for cell survival upon DNA damage involves a metabolic switch to oxidative phosphorylation through PARP1-dependent depletion of NAD^+.[63] Since luminal progenitors are enriched in oxidative phosphorylation, while basal cells mainly rely on glycolysis,[26] the increased utilization of oxidative phosphorylation within luminal progenitors may render the luminal lineage to be more permissive in acquiring higher levels of DSBs, possibly via higher reactive oxygen species (ROS) accumulation, while continuing to survive.[27] Moreover, while ovarian hormones induce cellular expansion of both lineages, considerable proliferative heterogeneity has been reported in the mammary epithelium.[64] Lineage-dependent metabolic programs and cyclical proliferation within the mammary gland may underlie differential replication stress in the two lineages,[65] which can also be a contributing factor to DDR and downstream signaling operating differently in the two mammary lineages.

The impact of mammary lineage identity on PARPi response not only remained consistent across mouse and human primary mammary epithelium, but also extended to breast cancer patient-derived or cell-line-derived xenografts. In these human and mouse breast cancer models, tumors representing the basal lineage features displayed PARPi sensitivity *in vivo*, while luminal-like cancer models remained resistant. This finding opens exciting possibilities for patient selection strategies for this highly successful class of targeted drugs. For instance, identification of patient subgroups whose tumors exhibit a basal lineage identity may serve as an approach to pinpointing additional PARPi monotherapy responders, beyond germline BRCA testing. Indeed, both *in vitro* PARPi response data that tested 13 PARPis on 12 breast cancer cell lines[16,17] and PARPi clinical trials[18–22,66,67] show that BRCA1/2 mutation alone does not entirely predict PARPi sensitivity nor clinical benefit, while a subset of non-BRCA mutation carriers also exhibited a response.

To further potentiate PARPi efficacy, numerous combinatorial therapies are currently being tested, both experimentally and in clinical settings, including with chemotherapy, immunotherapy, endocrine therapy, and other DDR-targeted therapy.[15,68–71] Our findings further suggest that luminal progenitors, which are considered the putative cell of origin for *BRCA1*-mutated breast cancers,[56,57] are the least PARPi sensitive and therefore may persist following PARPi monotherapy. Thus, strategies that simultaneously deplete more than one type of breast cancer precursor cell/cell of origin may prove useful, and combining PARPis with known epigenetic and/or metabolic inhibitors that preferentially target luminal progenitors warrants further investigation.

Basal and luminal epithelial cells of the breast are highly specialized, with prominent lineage identities that pertain to fundamental cellular features, such as metabolic[26,27] and epigenetic[28,29] programs as well as telomere biology.[30] To date, several global profiling datasets on purified mammary subpopulations are available, including transcriptomes, epigenomes, and methylomes, and have revealed a wealth of lineage-rooted distinctions,[28,29,72] emphasizing the divergent nature of these cells and their potential as therapeutic targets specific to the respective cell of origin of breast cancers. The importance of cell-lineage-dependent DNA damage repair likely extends to other organ systems and cancer types. Overall, this study demonstrates selective DNA repair preferences in different breast cancer precursor cell populations within a single tissue and suggests a potential avenue for identifying PARPi-responsive patient subpopulations.

Limitations of the study

There are several limitations inherent in our study. First, given that our study centers on primary mammary cells, where mammary cells must retain their lineage fidelity in tissue culture, the feasibility of measuring functional HR competency through HR reporter assays has been a challenge. Therefore, we have been unable to directly measure differences in DNA damage repair activity. In addition, although cell cycle was taken into consideration, our study cannot exclude the possibility that differences in cell-cycle dynamics between mammary lineages did not have an impact on the DDR pathways. We also have yet to experimentally verify a causal relationship between the mammary lineage-associated differential HR competency and PARPi vulnerability. Finally, translating the concept of

Figure 6. Tumor sensitivity to PARPis is determined by the mammary lineage signatures

(A) Schematic diagram illustrating the human breast cancer tumor xenograft assay.

(B) Growth curve of olaparib-treated MDA-MB-231 or HCC1187 cell-line-derived xenografts compared with the DMSO control group (black, n = 11 per treatment group). All data are represented as the mean ± SEM. The p values were determined by multiple t test; *q < 0.001.

(C) Flow cytometric analysis on single-cell suspensions of freshly dissociated full-size MDA-MD-231 or HCC1187 xenograft tumors. Human lineage-negative ($CD45^+CD31^+$ depleted), live (Zombie UV^-) cells were profiled by the human-specific epithelial marker h-EpCAM and the mesenchymal marker CD44.

(D) Heatmap illustrating ssGSEA enrichment scores of 47 PDX breast cancer models for the three proteome-defined lineage signatures (basal, BC; luminal progenitor, LP; mature luminal, ML). PAM50 breast cancer subtyping is annotated.

(E) Regression plots illustrating the talazoparib sensitivity (“response angle”) against each mammary lineage signature enrichment score (ssGSEA). Each dot represents one PDX model, and response angle for each model was determined as the inverse tangent of the angle between the regression line slope of the mean control and treatment curves.[60]

heightened PARPi response of the basal over the luminal lineage to tumor biology is not straightforward. Genomic alterations and cancer cell evolution can impinge on normal lineage programs, confounding relationships between specific primary cell subpopulations and tumor subtypes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
  - Primary human breast specimen
  - Primary mammary single-cell suspensions
  - Primary mouse mammary tumour cell-lines
- METHOD DETAILS
  - Irradiation
  - γ-H2AX intracellular flow cytometry
  - Imaging flow cytometry
  - Fluorescence-activated cell sorting (FACS)
  - Human mammary epithelial flow cytometry
  - Neutral comet assay
  - *In situ* immunofluorescence staining
  - Confocal microscopy image acquisition
  - *In situ* foci counting
  - Colony-forming cell (CFC) assay
  - Small molecule inhibitors/*In vitro* drug testing
  - Gene set enrichment analysis (GSEA)
  - Enrichment map visualization
  - Heatmap of mouse DDR protein abundance
  - *In vivo* engraftment of human breast cancer and mouse mammary tumor-derived cell lines
  - *In vivo* drug testing
  - Human DDR proteomic analysis
  - Volcano plot and pathway analysis on the human DDR proteins
  - Generation of human and mouse mammary lineage signatures
  - Enrichment of signatures in breast cancer subtypes
  - Enrichment in human breast cancer cell-lines
  - Correlation to breast cancer cell-line drug sensitivity screening
  - Enrichment of signatures in human breast cancer PDX
  - DDR gene expression in human single-cell RNAseq data
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.113256.

ACKNOWLEDGMENTS

This work was supported by grants to R.K and T.K. from the Canadian Cancer Society (CCS), Canadian Institutes of Health Research (CIHR), and Terry Fox Research Institute (TFRI), and to D.W.C. from the Canadian Institutes of Health Reserach (CIHR).

AUTHOR CONTRIBUTIONS

H.K. designed and performed the majority of the experiments and data analyses. K.A. performed bioinformatics on METABRIC, GDSC drug sensitivity, and human PDX datasets and assisted in *in vivo* drug injections. T.K. and M.W. generated mouse and human global mammary proteomes. L.P. and M.A.P. guided bioinformatics analyses. C.W.M. interrogated human breast scRNA-seq. A.K. and B.Z. performed human CFC assays on sorted cells. D.S., K.N.A., and E.R.L. provided mouse mammary tumor models. A.S.M., B.H.-K., M.J.E., and D.W.C. tested talazoparib responses in breast cancer PDXs and computed "response angles." S.N. and P.D.W. assisted in *in vivo* engraftment. M.P. operated the Amnis ImageStreamX Mark II instrument and guided post-acquisition data analyses. M.J.M. and S.N. assisted in processing primary human breast specimens and human CFC assays. S.H. and H.K.B. acquired the human breast specimens. H.K., P.T., B.T.G., P.D.W., and R.K. wrote and edited the manuscript. R.K. supervised the study.

DECLARATION OF INTERESTS

D.W.C. reports consultancy and advisory relationships with AstraZeneca, Daiichi Sankyo, Exact Sciences, Eisai, Gilead, GlaxoSmithKline, Inflex, Inivata/NeoGenomics, Lilly, Merck, Novartis, Pfizer, Roche and Saga; research funding to their institution from AstraZeneca, Guardant Health, Gilead, GlaxoSmithKline, Inivata/NeoGenomics, Knight, Merck, Pfizer, ProteinQure and Roche; and holds a holds a patent (US62/675,228) for methods of treating cancers characterized by a high expression level of spindle and kinetochore associated complex subunit 3 (ska3) gene.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: February 3, 2023
Revised: September 2, 2023
Accepted: September 28, 2023
Published: October 16, 2023

REFERENCES

1. Tomasetti, C., and Vogelstein, B. (2015). Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science *347*, 78–81. https://doi.org/10.1126/science.1260825.
2. Tomasetti, C., Li, L., and Vogelstein, B. (2017). Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. Science *355*, 1330–1334. https://doi.org/10.1126/science.aaf9011.
3. Joshi, P. a, di Grappa, M. a, and Khokha, R. (2012). Active allies: hormones, stem cells and the niche in adult mammopoiesis. Trends Endocrinol. Metabol. *23*, 299–309. https://doi.org/10.1016/j.tem.2012.04.002.
4. Joshi, P. a, Jackson, H.W., Beristain, A.G., Di Grappa, M. a, Mote, P. a, Clarke, C.L., Stingl, J., Waterhouse, P.D., and Khokha, R. (2010). Progesterone induces adult mammary stem cell expansion. Nature *465*, 803–807. https://doi.org/10.1038/nature09091.
5. Nielsen, F.C., van Overeem Hansen, T., and Sørensen, C.S. (2016). Hereditary breast and ovarian cancer: New genes in confined pathways. Nat. Rev. Cancer *16*, 599–612. https://doi.org/10.1038/nrc.2016.72.
6. Knijnenburg, T.A., Wang, L., Zimmermann, M.T., Chambwe, N., Gao, G.F., Cherniack, A.D., Fan, H., Shen, H., Way, G.P., Greene, C.S., et al. (2018). Genomic and Molecular Landscape of DNA Damage Repair Deficiency

across The Cancer Genome Atlas. Cell Rep. *23*, 239–254.e6. https://doi.org/10.1016/j.celrep.2018.03.076.

7. Davies, H., Glodzik, D., Morganella, S., Yates, L.R., Staaf, J., Zou, X., Ramakrishna, M., Martin, S., Boyault, S., Sieuwerts, A.M., et al. (2017). HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat. Med. *23*, 517–525. https://doi.org/10.1038/nm.4292.
8. Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N.J., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917–921. https://doi.org/10.1038/nature03445.
9. Bryant, H.E., Schultz, N., Thomas, H.D., Parker, K.M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N.J., and Helleday, T. (2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature *434*, 913–917. https://doi.org/10.1038/nature03443.
10. Ashworth, A., and Lord, C.J. (2018). Synthetic lethal therapies for cancer: what's next after PARP inhibitors? Nat. Rev. Clin. Oncol. *15*, 564–576. https://doi.org/10.1038/s41571-018-0055-6.
11. Robson, M., Im, S.-A., Senkus, E., Xu, B., Domchek, S.M., Masuda, N., Delaloge, S., Li, W., Tung, N., Armstrong, A., et al. (2017). Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. N. Engl. J. Med. *377*, 523–533. https://doi.org/10.1056/NEJMoa1706450.
12. Robson, M.E., Tung, N., Conte, P., Im, S.A., Senkus, E., Xu, B., Masuda, N., Delaloge, S., Li, W., Armstrong, A., et al. (2019). OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer. Ann. Oncol. *30*, 558–566. https://doi.org/10.1093/annonc/mdz012.
13. Litton, J.K., Rugo, H.S., Ettl, J., Hurvitz, S.A., Gonçalves, A., Lee, K.-H., Fehrenbacher, L., Yerushalmi, R., Mina, L.A., Martin, M., et al. (2018). Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. N. Engl. J. Med. *379*, 753–763. https://doi.org/10.1056/NEJMoa1802905.
14. Litton, J.K., Hurvitz, S.A., Mina, L.A., Rugo, H.S., Lee, K.-H., Gonçalves, A., Diab, S., Woodward, N., Goodwin, A., Yerushalmi, R., et al. (2020). Talazoparib versus chemotherapy in patients with germline BRCA1/2-mutated HER2-negative advanced breast cancer: final overall survival results from the EMBRACA trial. Ann. Oncol. *31*, 1526–1535. https://doi.org/10.1016/j.annonc.2020.08.2098.
15. Groelly, F.J., Fawkes, M., Dagg, R.A., Blackford, A.N., and Tarsounas, M. (2023). Targeting DNA damage response pathways in cancer. Nat. Rev. Cancer *23*, 78–94. https://doi.org/10.1038/s41568-022-00535-5.
16. Keung, M.Y., Wu, Y., Badar, F., and Vadgama, J.V. (2020). Response of Breast Cancer Cells to PARP Inhibitors Is Independent of BRCA Status. J. Clin. Med. *9*, 940. https://doi.org/10.3390/jcm9040940.
17. Sanders, M.E., Bauer, J.A., Chen, X., Lehmann, B.D., Chakravarthy, A.B., Shyr, Y., Pietenpol, J.A., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J. Clin. Invest. *121*, 2750–2767. https://doi.org/10.1172/JCI45014.2750.
18. Coleman, R.L., Oza, A.M., Lorusso, D., Aghajanian, C., Oaknin, A., Dean, A., Colombo, N., Weberpals, J.I., Clamp, A., Scambia, G., et al. (2017). Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet *390*, 1949–1961. https://doi.org/10.1016/S0140-6736(17)32440-6.
19. Mirza, M.R., Monk, B.J., Herrstedt, J., Oza, A.M., Mahner, S., Redondo, A., Fabbro, M., Ledermann, J.A., Lorusso, D., Vergote, I., et al. (2016). Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. N. Engl. J. Med. *375*, 2154–2164. https://doi.org/10.1056/NEJMoa1611310.
20. Gelmon, K.A., Tischkowitz, M., Mackay, H., Swenerton, K., Robidoux, A., Tonkin, K., Hirte, H., Huntsman, D., Clemons, M., Gilks, B., et al. (2011). Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: A phase 2, multicentre, open-label, non-randomised study. Lancet Oncol. *12*, 852–861. https://doi.org/10.1016/S1470-2045(11)70214-5.
21. Tutt, A., Robson, M., Garber, J.E., Domchek, S.M., Audeh, M.W., Weitzel, J.N., Friedlander, M., Arun, B., Loman, N., Schmutzler, R.K., et al. (2010). Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet *376*, 235–244. https://doi.org/10.1016/S0140-6736(10)60892-6.
22. Sandhu, S.K., Schelman, W.R., Wilding, G., Moreno, V., Baird, R.D., Miranda, S., Hylands, L., Riisnaes, R., Forster, M., Omlin, A., et al. (2013). The poly(ADP-ribose) polymerase inhibitor niraparib (MK4827) in BRCA mutation carriers and patients with sporadic cancer: a phase 1 dose-escalation trial. Lancet Oncol. *14*, 882–892. https://doi.org/10.1016/S1470-2045(13)70240-7.
23. Lord, C.J., and Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. Science *355*, 1152–1158. https://doi.org/10.1126/science.aam7344.
24. Nguyen, Q.H., Pervolarakis, N., Blake, K., Ma, D., Davis, R.T., James, N., Phung, A.T., Willey, E., Kumar, R., Jabart, E., et al. (2018). Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. Nat. Commun. *9*, 2028. https://doi.org/10.1038/s41467-018-04334-1.
25. Chen, W., Morabito, S.J., Kessenbrock, K., Enver, T., Meyer, K.B., and Teschendorff, A.E. (2019). Single-cell landscape in mammary epithelium reveals bipotent-like cells associated with breast cancer risk and outcome. Commun. Biol. *2*, 306. https://doi.org/10.1038/s42003-019-0554-8.
26. Mahendralingam, M.J., Kim, H., McCloskey, C.W., Aliar, K., Casey, A.E., Tharmapalan, P., Pellacani, D., Ignatchenko, V., Garcia-Valero, M., Palomero, L., et al. (2021). Mammary epithelial cells have lineage-rooted metabolic identities. Nat. Metab. *3*, 665–681. https://doi.org/10.1038/s42255-021-00388-6.
27. Kannan, N., Nguyen, L.V., Makarem, M., Dong, Y., Shih, K., Eirew, P., Raouf, A., Emerman, J.T., and Eaves, C.J. (2014). Glutathione-dependent and -independent oxidative stress-control mechanisms distinguish normal human mammary epithelial cell subsets. Proc. Natl. Acad. Sci. USA *111*, 7789–7794. https://doi.org/10.1073/pnas.1403813111.
28. Casey, A.E., Sinha, A., Singhania, R., Livingstone, J., Waterhouse, P., Tharmapalan, P., Cruickshank, J., Shehata, M., Drysdale, E., Fang, H., et al. (2018). Mammary molecular portraits reveal lineage-specific features and progenitor cell vulnerabilities. J. Cell Biol. *217*, 2951–2974. https://doi.org/10.1083/jcb.201804042.
29. Pellacani, D., Bilenky, M., Kannan, N., Heravi-Moussavi, A., Knapp, D.J.H.F., Gakkhar, S., Moksa, M., Carles, A., Moore, R., Mungall, A.J., et al. (2016). Analysis of Normal Human Mammary Epigenomes Reveals Cell-Specific Active Enhancer States and Associated Transcription Factor Networks. Cell Rep. *17*, 2060–2074. https://doi.org/10.1016/j.celrep.2016.10.058.
30. Kannan, N., Huda, N., Tu, L., Droumeva, R., Aubert, G., Chavez, E., Brinkman, R.R., Lansdorp, P., Emerman, J., Abe, S., et al. (2013). The luminal progenitor compartment of the normal human mammary gland constitutes a unique site of telomere dysfunction. Stem Cell Rep. *1*, 28–37. https://doi.org/10.1016/j.stemcr.2013.04.003.
31. Alfonso-Pérez, T., Baonza, G., and Martin-Belmonte, F. (2021). Breast cancer has a new metabolic Achilles' heel. Nat. Metab. *3*, 590–592. https://doi.org/10.1038/s42255-021-00394-8.
32. Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993–997. https://doi.org/10.1038/nature04496.
33. Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.-L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84–88. https://doi.org/10.1038/nature04372.

34. Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I.A., Avril, S., Prater, M., Eirew, P., Caldas, C., Watson, C.J., and Stingl, J. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. *14*, R134. https://doi.org/10.1186/bcr3334.

35. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

36. de Jager, M., Wyman, C., van Gent, D.C., and Kanaar, R. (2002). DNA end-binding specificity of human Rad50/Mre11 is influenced by ATP. Nucleic Acids Res. *30*, 4425–4431. https://doi.org/10.1093/nar/gkf574.

37. Hustedt, N., and Durocher, D. (2016). The control of DNA repair by the cell cycle. Nat. Cell Biol. *19*, 1–9. https://doi.org/10.1038/ncb3452.

38. Sishc, B.J., and Davis, A.J. (2017). The Role of the Core Non-Homologous End Joining Factors in Carcinogenesis and Cancer. Cancers *9*, 81. https://doi.org/10.3390/cancers9070081.

39. Sun, Y., McCorvie, T.J., Yates, L.A., and Zhang, X. (2020). Structural basis of homologous recombination. Cell. Mol. Life Sci. *77*, 3–18. https://doi.org/10.1007/s00018-019-03365-1.

40. Patel, P.S., Abraham, K.J., Guturi, K.K.N., Halaby, M.-J., Khan, Z., Palomero, L., Ho, B., Duan, S., St-Germain, J., Algouneh, A., et al. (2021). RNF168 regulates R-loop resolution and genomic stability in BRCA1/2-deficient tumors. J. Clin. Invest. *131*, e140105. https://doi.org/10.1172/JCI140105.

41. Ghamrasni, S. el, Cardoso, R., Li, L., Guturi, K.K.N., Bjerregaard, V.A., Liu, Y., Venkatesan, S., Hande, M.P., Henderson, J.T., Sanchez, O., et al. (2016). Rad54 and Mus81 cooperation promotes DNA damage repair and restrains chromosome missegregation. Oncogene *35*, 4836–4845. https://doi.org/10.1038/onc.2016.16.

42. Li, L., Guturi, K.K.N., Gautreau, B., Patel, P.S., Saad, A., Morii, M., Mateo, F., Palomero, L., Barbour, H., Gomez, A., et al. (2018). Ubiquitin ligase RNF8 suppresses Notch signaling to regulate mammary development and tumorigenesis. J. Clin. Invest. *128*, 4525–4542. https://doi.org/10.1172/JCI120401.

43. Eirew, P., Stingl, J., and Eaves, C.J. (2010). Quantitation of human mammary epithelial stem cells with in vivo regenerative properties using a subrenal capsule xenotransplantation assay. Nat. Protoc. *5*, 1945–1956. https://doi.org/10.1038/nprot.2010.148.

44. Stingl, J., Eaves, C.J., Zandieh, I., and Emerman, J.T. (2001). Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast Cancer Res. Treat. *67*, 93–109. https://doi.org/10.1023/A:1010615124301.

45. Fridriksdottir, A.J., Kim, J., Villadsen, R., Klitgaard, M.C., Hopkinson, B.M., Petersen, O.W., and Rønnov-Jessen, L. (2015). Propagation of oestrogen receptor-positive and oestrogen-responsive normal human breast cells in culture. Nat. Commun. *6*, 8786. https://doi.org/10.1038/ncomms9786.

46. Loganathan, S.K., Schleicher, K., Malik, A., Quevedo, R., Langille, E., Teng, K., Oh, R.H., Rathod, B., Tsai, R., Samavarchi-Tehrani, P., et al. (2020). Rare driver mutations in head and neck squamous cell carcinomas converge on NOTCH signaling. Science *367*, 1264–1269. https://doi.org/10.1126/science.aax0902.

47. Langille, E., Al-Zahrani, K.N., Ma, Z., Liang, M., Uuskula-Reimand, L., Espin, R., Teng, K., Malik, A., Bergholtz, H., Ghamrasni, S.E., et al. (2022). Loss of Epigenetic Regulation Disrupts Lineage Integrity, Induces Aberrant Alveogenesis, and Promotes Breast Cancer. Cancer Discov. *12*, 2930–2953. https://doi.org/10.1158/2159-8290.CD-21-0865.

48. Meyer, D.S., Brinkhaus, H., Müller, U., Müller, M., Cardiff, R.D., and Bentires-Alj, M. (2011). Luminal expression of PIK3CA mutant H1047R in the mammary gland induces heterogeneous tumors. Cancer Res. *71*, 4344–4351. https://doi.org/10.1158/0008-5472.CAN-10-3827.

49. Adams, J.R., Xu, K., Liu, J.C., Agamez, N.M.R., Loch, A.J., Wong, R.G., Wang, W., Wright, K.L., Lane, T.F., Zacksenhaus, E., and Egan, S.E. (2011). Cooperation between Pik3ca and p53 mutations in mouse mammary tumor formation. Cancer Res. *71*, 2706–2717. https://doi.org/10.1158/0008-5472.CAN-10-0738.

50. Al-Zahrani, K.N., Abou-Hamad, J., Cook, D.P., Pryce, B.R., Hodgins, J.J., Labrèche, C., Robineau-Charette, P., de Souza, C.T., Bell, J.C., Auer, R.C., et al. (2020). Loss of the Ste20-like kinase induces a basal/stem-like phenotype in HER2-positive breast cancers. Oncogene *39*, 4592–4602. https://doi.org/10.1038/s41388-020-1315-3.

51. Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell *166*, 740–754. https://doi.org/10.1016/j.cell.2016.06.017.

52. Visvader, J.E. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev. *23*, 2563–2577. https://doi.org/10.1101/gad.1849509.

53. Fu, N.Y., Nolan, E., Lindeman, G.J., and Visvader, J.E. (2020). Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. Physiol. Rev. *100*, 489–523. https://doi.org/10.1152/physrev.00040.2018.

54. Visvader, J.E., and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy : current status and perspectives. Genes Dev. *28*, 1143–1158. https://doi.org/10.1101/gad.242511.114.targeted.

55. Badve, S., Dabbs, D.J., Schnitt, S.J., Baehner, F.L., Decker, T., Eusebi, V., Fox, S.B., Ichihara, S., Jacquemier, J., Lakhani, S.R., et al. (2011). Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. Mod. Pathol. *24*, 157–167. https://doi.org/10.1038/modpathol.2010.200.

56. Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913. https://doi.org/10.1038/nm.2000.

57. Molyneux, G., Geyer, F.C., Magnay, F.-A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A., et al. (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell *7*, 403–417. https://doi.org/10.1016/j.stem.2010.07.010.

58. Murai, J., Huang, S.Y.N., Das, B.B., Renaud, A., Zhang, Y., Doroshow, J.H., Ji, J., Takeda, S., and Pommier, Y. (2012). Trapping of PARP1 and PARP2 by clinical PARP inhibitors. Cancer Res. *72*, 5588–5599. https://doi.org/10.1158/0008-5472.CAN-12-2753.

59. Henneman, L., van Miltenburg, M.H., Michalak, E.M., Braumuller, T.M., Jaspers, J.E., Drenth, A.P., de Korte-Grimmerink, R., Gogola, E., Szuhai, K., Schlicker, A., et al. (2015). Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. Proc. Natl. Acad. Sci. USA *112*, 8409–8414. https://doi.org/10.1073/pnas.1500223112.

60. Mer, A.S., Ba-Alawi, W., Smirnov, P., Wang, Y.X., Brew, B., Ortmann, J., Tsao, M.-S., Cescon, D.W., Goldenberg, A., and Haibe-Kains, B. (2019). Integrative Pharmacogenomics Analysis of Patient-Derived Xenografts. Cancer Res. *79*, 4539–4550. https://doi.org/10.1158/0008-5472.CAN-19-0349.

61. Morel, A.P., Ginestier, C., Pommier, R.M., Cabaud, O., Ruiz, E., Wicinski, J., Devouassoux-Shisheboran, M., Combaret, V., Finetti, P., Chassot, C., et al. (2017). A stemness-related ZEB1-MSRB3 axis governs cellular pliancy and breast cancer genome stability. Nat. Med. *23*, 568–578. https://doi.org/10.1038/nm.4323.

62. Kass, E.M., Lim, P.X., Helgadottir, H.R., Moynahan, M.E., and Jasin, M. (2016). Robust homology-directed repair within mouse mammary tissue is not specifically affected by Brca2 mutation. Nat. Commun. *7*, 13241. https://doi.org/10.1038/ncomms13241.

63. Murata, M.M., Kong, X., Moncada, E., Chen, Y., Imamura, H., Wang, P., Berns, M.W., Yokomori, K., and Digman, M.A. (2019). NAD+ consumption

by PARP1 in response to DNA damage triggers metabolic shift critical for damaged cell survival. Mol. Biol. Cell *30*, 2584–2597. https://doi.org/10.1091/mbc.E18-10-0650.

64. Shehata, M., Waterhouse, P.D., Casey, A.E., Fang, H., Hazelwood, L., and Khokha, R. (2018). Proliferative heterogeneity of murine epithelial cells in the adult mammary gland. Commun. Biol. *1*, 111. https://doi.org/10.1038/s42003-018-0114-7.

65. Zhang, Y., Wu, L., Wang, Z., Wang, J., Roychoudhury, S., Tomasik, B., Wu, G., Wang, G., Rao, X., and Zhou, R. (2022). Replication Stress: A Review of Novel Targets to Enhance Radiosensitivity-From Bench to Clinic. Front. Oncol. *12*, 838637. https://doi.org/10.3389/fonc.2022.838637.

66. González-Martín, A., Pothuri, B., Vergote, I., DePont Christensen, R., Graybill, W., Mirza, M.R., McCormick, C., Lorusso, D., Hoskins, P., Freyer, G., et al. (2019). Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N. Engl. J. Med. *381*, 2391–2402. https://doi.org/10.1056/NEJMoa1910962.

67. Swisher, E.M., Lin, K.K., Oza, A.M., Scott, C.L., Giordano, H., Sun, J., Konecny, G.E., Coleman, R.L., Tinker, A.v., O'Malley, D.M., et al. (2017). Rucaparib in relapsed, platinum-sensitive high-grade ovarian carcinoma (ARIEL2 Part 1): an international, multicentre, open-label, phase 2 trial. Lancet Oncol. *18*, 75–87. https://doi.org/10.1016/S1470-2045(16)30559-9.

68. Veneris, J.T., Matulonis, U.A., Liu, J.F., and Konstantinopoulos, P.A. (2020). Choosing wisely: Selecting PARP inhibitor combinations to promote anti-tumor immune responses beyond BRCA mutations. Gynecol. Oncol. *156*, 488–497. https://doi.org/10.1016/j.ygyno.2019.09.021.

69. Zimmermann, M., Bernier, C., Kaiser, B., Fournier, S., Li, L., Desjardins, J., Skeldon, A., Rimkunas, V., Veloso, A., Young, J.T.F., et al. (2022). Guiding ATR and PARP inhibitor combinations with chemogenomic screens. Cell Rep. *40*, 111081. https://doi.org/10.1016/j.celrep.2022.111081.

70. Asim, M., Tarish, F., Zecchini, H.I., Sanjiv, K., Gelali, E., Massie, C.E., Baridi, A., Warren, A.Y., Zhao, W., Ogris, C., et al. (2017). Synthetic lethality between androgen receptor signalling and the PARP pathway in prostate cancer. Nat. Commun. *8*, 374. https://doi.org/10.1038/s41467-017-00393-y.

71. Pacey, S., Linch, M.D., Kynaston, H., Warren, A., Freeman, A., Tysoe, R., Shaw, G., Narahari, K., Kumar, S., Martin, H., et al. (2019). A study into the pharmacodynamic biomarker effects of olaparib (PARP Inhibitor) ± degarelix (GnRH antagonist) given prior to radical prostatectomy (RP) CANCAP03. J. Clin. Oncol. *37*, 35. https://doi.org/10.1200/JCO.2019.37.7_suppl.35.

72. Tharmapalan, P., Mahendralingam, M., Berman, H.K., and Khokha, R. (2019). Mammary stem cells and progenitors: targeting the roots of breast cancer for prevention. EMBO J. *38*, e100852. https://doi.org/10.15252/embj.2018100852.

73. Stingl, J., Emerman, J.T., and Eaves, C.J. (2005). Enzymatic dissociation and culture of normal human mammary tissue to detect progenitor activity. Methods Mol. Biol. *290*, 249–263. https://doi.org/10.1385/1-59259-838-2:249.

74. Filby, A., Perucha, E., Summers, H., Rees, P., Chana, P., Heck, S., Lord, G.M., and Davies, D. (2011). An imaging flow cytometric method for measuring cell division history and molecular symmetry during mitosis. Cytometry A. *79*, 496–506. https://doi.org/10.1002/cyto.a.21091.

75. Ortyn, W.E., Hall, B.E., George, T.C., Frost, K., Basiji, D.A., Perry, D.J., Zimmerman, C.A., Coder, D., and Morrissey, P.J. (2006). Sensitivity measurement and compensation in spectral imaging. Cytometry A. *69*, 852–862. https://doi.org/10.1002/cyto.a.20306.

76. Light Microscopy Core Facility/Duke University Count Nuclear Foci - ImageJ. https://microscopy.duke.edu/guides/count-nuclear-foci-ImageJ.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse CD45-PE-Cy7	eBioscience	Cat#25-0451-82; RRID:AB_2734986
Anti-mouse CD31-PE-Cy7	eBioscience	Cat#25-0311-81; RRID:AB_2734962
Anti-mouse Ter119-PE-Cy7	eBioscience	Cat#25-5921-82; RRID:AB_469661
Anti-mouse CD45-biotin	eBioscience	Cat#13-0451-82; RRID:AB_466446
Anti-mouse CD31-biotin	eBioscience	Cat#13-0311-82; RRID:AB_466420
Anti-mouse Ter119-biotin	eBioscience	Cat#13-5921-81; RRID:AB_466796
Anti-mouse CD24-PerCP-eFluor® 710	eBioscience	Cat#46-0242-80; RRID:AB_1834426
Anti-mouse CD24-APC-eFluor® 780	eBioscience	Cat#47-0242-82; RRID:AB_10853172
Anti-mouse CD24-FITC	BD Biosciences	Cat#553261; RRID:AB_394740
Anti-mouse CD326 (EpCAM)-APC-eFluor® 780	eBioscience	Cat#47-5791-82; RRID:AB_2573986
Anti-human/mouse/bovine CD49f-APC	R&D Systems	Cat#FAB13501A; RRID:AB_357016
Anti-human/mouse CD49f-FITC	BioLegend	Cat#313606; RRID:AB_345300
Anti-human/mouse CD49f-PE-Cy7	BioLegend	Cat#313622; RRID:AB_2561705
Anti-mouse CD49b-PE	BioLegend	Cat#103506; RRID:AB_313029
Anti-mouse Ly-6A/E (Sca-1)-APC-Cy7	BioLegend	Cat#108126; RRID:AB_10645327
Anti-mouse Ly-6A/E(Sca-1)-PE-CF594	BD Biosciences	Cat#562730; RRID:AB_2737751
Anti-human CD45-PE-Cy7	BioLegend	Cat#304016; RRID:AB_314404
Anti-human CD31-PE-Cy7	BioLegend	Cat#303118; RRID:AB_2247932
Anti-human CD326 (EpCAM)-PE	BioLegend	Cat#324206; RRID:AB_756080
Anti-phospho-histone H2A.X (Ser139) (20E3)- Alexa Fluor® 488	CST	Cat#9719S; RRID:AB_561076
Anti-phospho-Histone H2A.X (Ser139) (20E3)	CST	Cat#:9718; RRID:AB_2118009
Anti-RAD51 (H-92)	SantaCruz	Cat#sc-8349; RRID:AB_2253533
Anti-RAD51	Proteintech	Cat#14961-1-AP; RRID:AB_2177083
Anti-DNA-PKcs (phospho S2056)	Abcam	Cat#ab18192; RRID:AB_869495
Anti-Keratin 14	BioLegend	Cat#906001; RRID:AB_2565055
Anti-Ki67	BioLegend	Cat#652402; RRID:AB_11204254
Anti-progesterone receptor	LifeSpan Biosciences Inc.	Cat#LS-B5236; RRID:AB_10851510
Alexa Fluor® 488-conjugated rabbit IgG Isotype (Fc)	CST	Cat#4340S; RRID:AB_561545
Streptavidin-eFluor 450™	eBioscience	Cat#48-4317-82; RRID:AB_10359737
Streptavidin-Alexa Fluor® 594	BioLegend	Cat#405240; RRID:AB_2337250
Goat anti-rat-AlexaFluor® 647	Thermo Scientific	Cat#A21247; RRID:AB_141778
Goat anti-rabbit AlexaFluor® 488	Thermo Scientific	Cat#A11008; RRID:AB_143165
Goat anti-chicken-Cy™3 AffiniPure	Jackson ImmunoResearch	Cat#103-165-155; RRID:AB_2337386
Biological samples		
Human breast tissues (prophylactic mastectomy)	Princess Margaret Cancer Centre/UHN	N/A
Chemicals, peptides, and recombinant proteins		
Collagenase (for human breast specimen)	Sigma	Cat#C9891
Hyaluronidase (for human breast specimen)	Sigma	Cat#H3884
Collagenase/hyaluronidase (for mouse mammary glands)	STEMCELL Technologies	Cat#07912
Ammonium chloride	STEMCELL Technologies	Cat#07850

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dispase	STEMCELL Technologies	Cat#07913
Mouse EpiCult media	STEMCELL Technologies	Cat#05610
Human EpiCult-B	STEMCELL Technologies	Cat#05601
0.25% trypsin-EDTA	STEMCELL Technologies	Cat#07901
Cold water fish skin gelatin	Sigma	Cat#G7765
DNase I	Sigma	Cat#D4513
hFGF	STEMCELL Technologies	Cat#02634
bovine EGF	STEMCELL Technologies	Cat#78006
Heparin	STEMCELL Technologies	Cat#07980
ROCKi	Millipore	Cat#SCM075
Antibiotic-Antimycotic Solution	Wisent	Cat#450-115-EL
Cholera toxin	Sigma	Cat#C8052
Insulin	Thermo Fisher	Cat#12585014
hEGF	STEMCELL Technologies	Cat#78006.1
Hydrocortisone	STEMCELL Technologies	Cat#78003.1
Collagen	STEMCELL Technologies	Cat#04902
Matrigel	Corning	Cat#356231
Olaparib	Selleck Chemicals, MedChemExpress	Cat#S1060, Cat#HY-10162
Talazoparib	MedChemExpress	Cat#HY-16106
KU-57788	MedChemExpress	Cat#HY-11006
2-hydroxypropyl-β-cyclodextrin	Sigma	Cat#H107-100G
Zombie UV Dye	BioLegend	Cat#423107
Live/Dead Fixable Red	Invitrogen	Cat#L23102
FxCycle™ Violet	Invitrogen	Cat#F10347
DRAQ5™	Thermo Scientific	Cat#62251
Reveal Decloaker, RTU	BioCare Medical	Cat# RV1000
Anti-fade ProLong Gold with DAPI	Thermo Scientific	Cat# P36935
SYBR® Gold gel staining solution	Thermo Fisher	Cat#S-11494
Critical commercial assays		
Neutral CometAssay® Reagent kit	Trevigen	Cat#4250-050-K
CometAssay® Electrophoresis System II	Trevigen	Cat#4250-050-ES
Decloaking Chamber™ Pro	BioCare Medical	Cat#DC2002
Deposited data		
Human mammary proteomics data	Mahendralingam et al.[26]	MassIVE: MSV000087042
Mouse mammary proteomics data	Casey et al.[28]	MassIVE: MSV000079330
Human mammary single-cell RNA-seq data	Mahendralingam et al.[26]	GEO: GSE168660
Experimental models: Cell lines		
MDA-MB-231	ATCC	Cat#HTB-26; RRID:CVCL_0062
Hs 578T	ATCC	Cat#HTB-126; RRID:CVCL_0332
HCC1395	ATCC	Cat#CRL-2324; RRID:CVCL_1249
BT-549	ATCC	Cat#HTB-122; RRID:CVCL_1092
HCC1187	ATCC	Cat#CRL-2322; RRID:CVCL_1247
EFM-192A	DSMZ	Cat#ACC 258; RRID:CVCL_1812
EVSA-T	DSMZ	Cat#ACC 433; RRID:CVCL_1969
Pik3ca^{H1047R}; *p53*$^{-/-}$ sg4205	Loganathan et al.[46]; Langille et al.[47]	N/A
Pik3ca^{H1047R}; *p53*$^{-/-}$ sg4687	Loganathan et al.[46]; Langille et al.[47]	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
polyoma middle T (PymT)	A gift from D. Schramek	RRID:CVCL_AQ09
NDL3903	Al-Zahrani et al.[50]	N/A
Experimental models: Organisms/strains		
Mice: FVB/NJ	Jackson Laboratory, Charles River Laboratories	Cat#001800, Cat#CRL:207; RRID:IMSR_JAX:001800, RRID:IMSR_CRL:207
Mice: NSG	UHN/In-house breeding facility (Cancer Stem Cell Colony)	Cat#JAX:5557
Software and algorithms		
PRISM (v 9.2.0)	GraphPad	https://www.graphpad.com/
IDEAS (v 6.2)	Amnis Corporation/EMD Millipore	https://www.emdmillipore.com/CA/en/20150212_144049
FACSDiva (v 6.1.3)	BD Biosciences	RRID:SCR_001456; https://www.bdbiosciences.com/en-ca/products/software/instrument-software/bd-facsdiva-software
FlowJo (v 7.6.5, v 10.3)	BD Biosciences	RRID:SCR_008520; https://www.flowjo.com/solutions/flowjo/downloads
FIJI (is just ImageJ; v 1.52i)	Open source	RRID:SCR_002285; https://fiji.sc/
GSEA (v 4.0)	Broad Institute	RRID:SCR_003199; https://www.gsea-msigdb.org/gsea/index.jsp
Enrichment Map App (v 3.5.1)	Cytoscape	RRID:SCR_016052; https://apps.cytoscape.org/apps/enrichmentmap
ssGSEA; GSVA R package (v 1.30.0)	Bioconductor	https://bioconductor.org/packages/devel/bioc/vignettes/GSVA/inst/doc/GSVA.html
biomaRt R package (v 2.38.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
ComplexHeatmap R package (v 2.8.0)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
ComBat; sva R package (v 3.30.1)	Bioconductor	https://bioconductor.org/packages/release/bioc/html/sva.html
circlize R package (v 0.4.12)	CRAN	https://cran.r-project.org/web/packages/circlize/index.html
CellCyclceScoring; Seurat R package (v 4.3.0)	CRAN	https://cran.r-project.org/web/packages/Seurat/index.html
VlnPlot; Seurat R package (v 4.3.0)	CRAN	https://cran.r-project.org/web/packages/Seurat/index.html
pheatmap R package (v 1.0.12)	CRAN	https://cran.r-project.org/web/packages/pheatmap/index.html
Komet Software (v 7.0)	Andor Technology	N/A
BioTek Gen5 (v 3.12)	Agilent Technologies Inc.	https://www.agilent.com/en/product/cell-analysis/cell-imaging-microscopy/cell-imaging-microscopy-software/biotek-gen5-software-for-imaging-microscopy-1623226
All original code for bioinformatics analysis	This paper	https://doi.org/10.5281/zenodo.836855

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rama Khokha (rama.khokha@uhn.ca).

Materials availability

This study did not generate new unique materials.

Data and code availability

- The human and mouse mammary proteomes used in this study are openly accessible on the MassiVE repository, with the accession codes MSV000087042 and MSV000079330 respectively. The human mammary single-cell RNAseq dataset was acquired from NCBI GEO under the accession number GSE168660. These three datasets are also accessible through an interactive web application at https://github.com/kazeera/MEC-explorer.
- All original code used for bioinformatics analysis has been deposited on Zenodo and is also accessible via GitHub at https://github.com/kazeera/Kim-et-al-2023. The DOI is provided in the key resources table.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

8-10-week-old virgin female FVB/NJ mice (wild-type) were subcutaneously implanted with a 14-day slow-release pellet containing 0.14 mg 17β-estradiol plus 14 mg progesterone (E+P; Innovative Research of America). This implantation took place one week after bilateral ovariectomy. Upon the completion of a 14-day E+P treatment period, these mice were sacrificed and three pairs of mammary glands were collected (including second and third thoracic, as well as fourth inguinal). If required, mice underwent whole-body irradiation prior to sacrifice. For the xenograft experiments, 4-week-old virgin female NSG mice were employed. All mice were housed in a standard, controlled environment with a 12h light/dark cycle, maintained at a room temperature at 21–23 °C, humidity levels between 30-60%, and provided with standard chow diet *ad libitum.* All procedures were conducted in compliance with the Canadian Council for Animal Care guidelines under protocols approved by Animal Care Committee of the Princess Margaret Cancer Centre, Toronto, Ontario, Canada.

Primary human breast specimen

Human breast tissues were collected from women who had undergone prophylactic mastectomies within 24-48 hours of surgery at the Princess Margaret Cancer Centre under full informed consent and in accordance with Institutional Research Ethics Board approval. The breast specimen underwent the following processing[28,73]: the excised breast tissues were minced with blades and digested in DMEM/F12 media with 15 mM HEPES plus 2% BSA, 1% penicillin–streptomycin, 5 μg/mL insulin, 300 U/mL collagenase and 100 U/mL hyaluronidase shaking gently at 37 °C for 16-18 h. Next day, dissociated tissues underwent a series of centrifugation steps: 80g for 30 sec, 200g for 4 min, and 450g for 5 min to collect epithelial-, endothelium-, and fibroblast-rich fractions, respectively. The resulting three fractions were cryopreserved separately in liquid nitrogen. Only the epithelial-rich fractions were used to prepare single-cell suspensions for FACS and/or human CFC assays.

Primary mammary single-cell suspensions

Mouse

Single-cell suspensions of mouse mammary glands were prepared by mincing freshly harvested mammary glands and then incubating them in DMEM/F12 media with 300 U/mL collagenase plus 100 U/mL hyaluronidase at 37°C for 1.5 hours.[32] The resulting mammary organoids were serially treated with ammonium chloride, 0.25% trypsin-EDTA and 5 U/mL dispase plus 50 μg/ml DNase I in Hank's Balanced Salt Solution (HBSS; Gibco, 14025092) supplemented with 2% FBS (Gibco, 12483020). The resulting single-cells were filtered through a 40 μM cell strainer (Fisher Scientific, 22-363-547).

Human

Single-cell suspensions of human breast tissues were prepared from cryopreserved epithelial-rich fractions. These vials were thawed in 37°C water bath and serially digested with 0.25% trypsin-EDTA and then 5 U/mL dispase + 50 μg/ml DNase I in HBSS supplemented with 2% FBS.[28,73] The resulting cells were filtered through a 40 μm cell strainer.

Primary mouse mammary tumour cell-lines

Four primary mouse mammary tumour cell-lines: 1) *Pik3ca*^{H1047R}; *p53*$^{-/-}$ sg4205, 2) *Pik3ca*^{H1047R}; *p53*$^{-/-}$ sg4687,[46,47] 3) a polyoma middle T, and 4) NDL3903 were cultured in DMEM/F12 supplemented with MEGS (Gibco, S0155) and 10% FBS. All cell-lines were incubated at 37°C, 5% CO_2.

Human breast cancer cell-lines

MDA-MB-231, Hs 578T, EFM-192A, and EVSA-T were cultured in DMEM media + 10% FBS. HCC1395, BT-549, and HCC1187 were cultured in RPMI media + 10% FBS. All cell-lines were incubated at 37°C, 5% CO_2. All 7 human breast cancer cell-lines used in this study were purchased from either ATCC or DSMZ.

Human breast cancer PDX modeling

Drug response and gene expression data (RNAseq) for 47 breast cancer PDX models was generated,[60] and their response to BMN-673/talazoparib was measured. Each PDX was generated by serial subcutaneous transplant of fresh donor tumor fragments in NOD scid mice. Once tumour volumes reached ~125-250 mm^3, talazoparib (0.33 mg/kg) was administered via oral gavage daily until the mice reached the end point. The sensitivity to talazoparib was determined by calculation of the "response angle", calculating the inverse tangent between the regression line slope of mean control (untreated) and treatment tumour growth curves after 30 days.[60] All experiments were conducted under an approved REB protocol at the UHN, and included data are those available up to December 2020.

METHOD DETAILS

Irradiation

In vivo

Mice were exposed to a single dose of 6 Gray (Gy) using a Cs-137 irradiator (Gammacell® 40 Exactor). The mice were subsequently sacrificed at a designated time-point after irradiation.

In vitro

Primary mouse mammary epithelial cells were seeded in a 6-well cell culture plate (Greiner, 657160). The cells were then irradiated at a single dose of 3 Gy (Gammacell® 40 Exactor) the next day. For neutral comet assay, sorted primary mouse mammary epithelial cells that had been sorted were seeded in a 96-well plate (Greiner, 655090), and irradiated at a single dose of 20 Gy the following day.

γ-H2AX intracellular flow cytometry

Freshly dissociated mouse mammary single cells were stained with a viability dye (Zombie UV, 1:100) and a cocktail of cell surface markers to exclude mouse blood cell lineage-positive (Lin^+) cells, including: CD45 (hematopoietic, 1:800), CD31 (endothelial, 1:200), Ter119 (erythrocyte, 1:100) cells. Then, Anti-CD24-PerCP-eFluor® 710 (1:400), anti-CD49f-APC (1:100), anti-CD49b-PE (1:250), anti-Ly-6A/E (Sca-1)-APC/Cy7 (1:500) were used to segregate basal, LP, ML, and stromal populations. After the cell surface staining, the cells were fixed in 4% PFA for 10 min at room temperature, washed in PBS, and stored overnight at 4°C in HBSS+2% FBS. The next day, the cells were permeabilized with 0.1% Triton-X/PBS for 5 min at room temperature and stained with Alexa Fluor® 488-conjugated anti-phospho-histone H2A.X (Ser139; γ-H2AX, 1:200) or concentration matched Alexa Fluor® 488-conjugated rabbit IgG Isotype (Fc) control. Flow cytometry acquisition was performed using BD LSRFortessa™ and analyzed with FlowJo software.

Imaging flow cytometry

Staining

Freshly dissociated mouse single cells were stained with a fixable viability Zombie UV Dye (1:100) and a cocktail of cell surface markers including: biotin-conjugated CD45 (1:800), CD31 (1:200), Ter119 (1:100) which were subsequently labelled with secondary conjugate streptavidin-eFluor 450™ (1:500). Concurrently, the cells were stained with anti-CD24-APC-eFluor® 780 (1:400), anti-CD49f-PE/Cy7 (1:100), anti-CD49b-PE (1:250), anti-Ly-6A/E(Sca-1)-PE-CF594 (1:500). After PBS washing, the surface-stained cells were fixed in 4% PFA for 10 min at room temperature, washed in PBS, and stored overnight at 4°C. The next day, the cells were permeabilized with 0.1% Triton-X/PBS for 5 min at room temperature, washed in PBS and stained with anti-RAD51 (H-92; 1:50) or anti-DNA-PKcs (phospho S2056; 1:1,500) which were subsequently labelled with goat anti-rabbit Alexa Fluor® 488-conjugated secondary antibody (1:200). Nuclear DNA was stained in 2.5 μM DRAQ5 in PBS.

For the cell cycle analysis, Alexa Fluor® 594 Streptavidin (1:400) was used to label the Lin^+ primary antibodies as indicated above. Concurrently, the cells were stained with Live/Dead Fixable Red (1:200), anti-CD326 (EpCAM)-APC-eFluor® 780 (1:200), anti-CD49f-PE/Cy7 (1:100). After the fix/permeabilization step as mentioned above, the cells were stained with anti-RAD51 (Proteintech, 1:400) which were subsequently labelled with goat anti-rabbit Alexa Fluor® 488-conjugated secondary antibody (1:200). The cells were re-suspended in FxCycle™ Violet according to their manufacturer's protocol. Imaging flow cytometry was performed on an ImageStream®^{X} Mark II (Excitation lasers: 405, 488, 561, 592, 642nm; MilliporeSigma) with INSPIRE® software.

Image acquisition

Stained cells were resuspended in a volume of 50 μl PBS + 1% FBS (or FxCycle/PBS for the cell cycle analysis) in a 1.5 mL low retention microfuge tube (Sigma, T4816). Samples were then acquired on a 5 laser 12 channel ImageStream®^{X} Mark II imaging flow cytometer at 60X magnification following ASSIST calibration (Amnis Corporation). A bright-field (BF) area lower limit of 50 μm^2 was used to eliminate debris and calibration beads during sample acquisition, while samples were collected in a series of 50×10^3 event raw image files (.rif). For single stained compensation controls, BF illumination was turned off and approximately 3000 events within the positive signal fraction were acquired. An initial compensation matrix was generated by loading the single stained raw image files into the IDEAS compensation wizard with further refinements to the compensation matrix made as necessary through manual adjustment.[74,75] Once generated the compensation matrix was then applied to the sample raw image files to create compensated image files (.cif) which were then analyzed.

High-throughput DSB repair foci counting

Analysis was carried out using masks and features as defined in the IDEAS reference manual (version 6.0). The analysis workflow including IDEAS formatted axis feature/mask descriptors in parenthesis was as follows: (a) Gate on focused cells using the gradient RMS feature and default M01 mask in the BF channel (Gradient RMS_M01_Ch01). (b) Gate on single cells by plotting features aspect ratio vs area within the M01 mask in the BF channels (Area_M01 vs. Aspect Ratio_M01). (c) Gate on circular cells by plotting features aspect ratio vs circularity within the M01 mask in the BF channels (Circularity_M01 vs. Aspect Ratio_M01). (d) Gate on the viable lineage negative (Lin^-) population by plotting the intensity feature of the viability/lineage-negative stain vs. area (both viability dye and lineage negative panel are detected in the same channel (Intensity_MC_Ch07 Lin neg/viability vs. Area_M01). Lin^- gate was determined by using fluorescence minus one (FMO) control. (e) The stromal, luminal and basal populations were distinguished by plotting the intensity features of CD49f vs. CD24 (Intensity_MC_Ch06_CD49f vs. Intensity_MC_Ch12_CD24). (f) Further subdivision of the luminal population into luminal progenitor (LP) and mature luminal (ML) populations was achieved by plotting the intensity of CD49b vs. Sca1 (Intensity_MC_Ch03_CD49b vs. Intensity_MC_Ch04_Sca1). (g) The number of RAD51 or phospho-DNA-PKcs foci in basal, LP, and ML populations were then quantified using the "Spot Count" feature based on three different user-defined masks that define nucleus by DRAQ5 signal and repair foci by signal/intensity and size. The following feature/mask parameters were used to count RAD51 or phospho-DNA-PKcs foci in single cells. RAD51: Spot Count_Morphology(M11, Ch11) And Intensity(M02, Ch02, 1200-4095) And Spot(M02, Ch02, Bright, 9.5, 1, 0)_4. phospho-DNA-PKcs: Spot Count_Morphology(M11, Ch11) And Intensity(M02, Ch02, 1000-4095) And Spot(M02, Ch02, Bright, 9.5, 1, 0)_4.

For cell cycle analysis, the DNA content distribution (linear scale) was visualized by plotting the intensity of FxCycle stain (Intensity_MC_Ch07). Within basal ($Lin^-EpCAM^+CD49f^{lo}$) or luminal ($Lin^-EpCAM^+CD49f^{hi}$) population, S/G2 cells were defined from the end of first peak until the end of the second peak of the DNA content histogram. G1 cells span the first peak of the histogram. Then, the following feature/mask parameters were used to count RAD51 foci at each cell cycle phase: Spot Count_Morphology(M07, Ch07) And Intensity(M02, Ch02, 550-4095) And Spot(M02, Ch02, Bright, 2, 2, 0) And Peak(M02, Ch02, Bright, 1)_4.

Fluorescence-activated cell sorting (FACS)

Mouse

Freshly dissociated mouse mammary epithelial cells were stained with the following antibodies: PE/Cy7-conjugated lineage antibodies as described above, and various combinations of antibodies were employed: anti-CD24-APC-eFluor® 780 (1:400), anti-CD24-PerCP-eFluor® 710 (1:400), or anti-CD24-FITC (1:400), and anti-CD49f-APC (1:100) or anti-CD49f-FITC (1:100) to distinguish between the total basal and total luminal cell populations. Further refinement of LP and ML populations was achieved by incorporating CD49b-PE (1:250), anti-Ly-6A/E (Sca-1)-APC/Cy7 (1:500) antibodies. Dead cells were excluded with DAPI (5 mg/mL; 1:10,000). Cell sorting was performed on a BD FACSAria II (v8.0.1) and FACSDiva.

Human

Freshly dissociated human mammary epithelial cells were stained with antibodies: anti-human CD45-PE/Cy7 (1:200), anti-human CD31-PE/Cy7 (1:50), anti-human CD326 (EpCAM)-PE (1:50) and anti-CD49f-FITC (1:100). Dead cells were excluded with DAPI (5 mg/mL; 1:10,000). Cell sorting was performed on a BD FACSAria II (v8.0.1) and FACSDiva.

Human mammary epithelial flow cytometry

The same anti-human antibody panel was utilized, as detailed in the FACS section above. Dead cells were gated out using Zombie UV Dye (1:100). Flow cytometry analysis was conducted using BD LSR II or BD LSRFortessa and analyzed with FlowJo software.

Neutral comet assay

Freshly FACS-sorted basal, LP and ML cells from adult virgin female mice were resuspended in PBS at 20,000 cells/mL and were mixed with 0.7% low-melting point agarose gel at 1:10 ratio. Each population mixture was laid on a glass slide pre-coated with 1% normal melting point agarose and covered with an 18x18 mm coverslip. The coverslips were carefully removed after gel solidification at 4°C, and the slides were then processed according to the manufacturer's recommendation (Trevigen Neutral CometAssay®). Briefly, slides were immersed in Lysis solution at 4°C overnight and then immersed in 1X Neutral Electrophoresis buffer for 30 min. The slides were transferred to the CometAssay electrophoresis unit filled with chilled 1X Neutral Electrophoresis buffer and electrophoresed at 21 V for 7 min at 4°C. The slides were then fixed in DNA Precipitation Solution followed by 70% ethanol for 30 min each at room temperature. The slides were dried at room temperature and stained with SYBR® Gold gel staining solution (1:10,000). The Olive tail moment was measured by using Komet Software.

For *in vitro* irradiation experiment, 500-2000 freshly FACS-sorted cell populations were seeded in a 96-well plate in 100 μl of culture media: DMEM/F12 (1:1) supplemented with 10% FBS, 5 μg/ml insulin, 10 ng/ml hEGF, 10 ng/ml cholera toxin, 1.8×10^4 M adenine, 0.5 μg/ml hydrocortisone and 10 μM ROCK inhibitor. Next day, the cells were irradiated at 20 Gy and incubated at 37°C, 5% CO_2. When a time-point was reached, cells were trypsinized, washed in PBS, mixed with 0.7% low-melting point agarose gel at 1:10 ratio, and the rest of the neutral comet assay was performed as above. Agilent BioTek Cytation 5 was used for automated comet assay imaging. Acquired comet images were analyzed using BioTek Gen5 software following the manufacturer's protocol (https://www.agilent.com/cs/library/applications/automated-comet-assay-dual-mask-5994-2595EN-agilent.pdf).

In situ immunofluorescence staining

Freshly harvested mouse mammary glands were fixed in 4% PFA at 4°C overnight and subsequently stored in 70% EtOH before being sent off to Toronto Center for Phenogenomics for paraffin-embedding and tissue sectioning services. Tissue-section slides underwent deparaffinization and rehydration prior to antigen retrieval in Decloaking Chamber™ Pro for 30 minutes at 121°C in Reveal Decloaker (commercial pH 6.0 citrate buffer). The sections were blocked with a blocking buffer (5% goat serum, 1% glycerol, 0.1% BSA, 0.1% cold water fish skin gelatin) in PBS for 1 h at room temperature. The sections were stained with a cocktail of primary antibodies that were diluted in the same blocking buffer overnight at 4°C. Next day, the tissue sections were washed in 0.1% Tween-20/PBS and then incubated with secondary antibodies diluted in blocking buffer for 1 h at room temperature. The sections were mounted in Anti-fade ProLong™ Gold with DAPI. Primary antibodies used are as follows: anti-Keratin 14 (K14; 1:400), anti-Ki67 (1:400), anti-phospho-Histone H2A.X (Ser139; 1:400), anti-RAD51 (H-92, 1:50 or 14961-1-AP, 1:400), anti-DNA-PKcs (phospho S2056; 1:400), anti-progesterone receptor (PR; 1:400). Secondary antibodies (1:200-1:500) used are as follows: anti-rat conjugated to AlexaFluor® 647, anti-chicken-Cy™3 AffiniPure, and anti-rabbit-AlexaFluor® 488.

Confocal microscopy image acquisition

All tissue immunofluorescence images were acquired on a Zeiss LSM 700 confocal microscope using a 40X oil-immersion objective lens. At least three or four Z-planes were acquired per field of view.

In situ foci counting

Z-projected confocal immunofluorescence images were analyzed using FIJI/ImageJ.[76] RAD51 foci were counted by examining nuclear (DAPI), RAD51 (AF488), and Ki67 (AF647) masks corresponding to their detection channels. The same enumerated nuclear mask was then applied to the RAD51 foci mask or the Ki67 mask to count RAD51 foci and assess Ki67 intensity (the sum of all pixel values per nucleus divided by the nuclear area) for each nucleus. Nuclei exhibiting a Ki67 intensity greater than 50th percentile were categorized as ''proliferating''. Meanwhile, based on the K14 (Cy3) channel, each identified nucleus was manually annotated as ''basal'' (K14^{+}) or ''luminal'' (K14^{-} cells surrounded by K14^{+} cells). Among the proliferating cells, basal or luminal cells displaying the total of at least 2 RAD51 foci were counted. All analyzed images represent maximum intensity Z-projections derived from 3-4 acquired Z-planes.

Colony-forming cell (CFC) assay

Mouse

FACS-purified mouse total basal and luminal cells were plated in 6-well plate with 2x10^{5} irradiated (50 Gy) NIH-3T3 feeder cells per well in Mouse EpiCult media supplemented with 5% FBS, 20 ng/mL hFGF, 10 ng/mL bovine EGF, 4 μg/mL heparin, 5 μM ROCKi, Antibiotic-Antimycotic Solution (1:100) and incubated in a low oxygen (5%) incubator at 37°C.[28] On Day 7, grown colonies were fixed with 1:1 acetone:methanol (v/v) and stained with Giemsa (Fisher, 264-983) according to their manufacturer's protocol. The total number of colonies were manually counted under a Leitz dissecting microscope.

Human

A total of 1000 dissociated breast cells along with 2x10^{5} irradiated NIH 3T3 cells in Human EpiCult-B media supplemented with 5% FBS were plated in a 60 mm culture dish (Greiner, 82050-546) that had been pre-coated with collagen and incubated for 1 hour at 37°C. The seeded cells were subsequently cultured in a low oxygen environment (5%) for 10 days. On the following day, the media was changed to serum (FBS)-free conditions and maintained in this state until the endpoint.[28,43,73] For drug treatments, compounds or DMSO were introduced during this media change. On the 11th day, colonies were fixed, stained, and counted following the same procedure as the mouse CFCs. As for the lineage-sorted human CFCs, isolated human mammary single cells were seeded at a density of 500 cells per well in a 6-well plate and underwent the same processes of culture, fixation, staining, and counting, as outlined earlier.

Small molecule inhibitors/*In vitro* drug testing

All inhibitors (olaparib/AZD-2281, talazoparib/BMN-673, and KU-57788/NU7441) prepared in DMSO were added to wells/dishes to ensure that the final DMSO concentration did not exceed 0.1% (v/v) in each well/dish. The drugs were added on the following day after cell seeding in all CFC assays.

Gene set enrichment analysis (GSEA)

Protein expression profiles[28] of basal, LP, and ML populations from adult female mice treated with estrogen plus progesterone were included in the pathway analysis using GSEA,[35] and the cell populations exposed to estrogen alone were excluded from this analysis. Each mammary population's protein expression profiles were compared to those of the other two populations (referred to as "Rest") to identify unique upregulated or downregulated biological pathways in that population. The Gene Matrix Transposed (.gmt) or Mouse_GOBP_AllPathways_no_GO_iea_August_01_2017_UniProt.gmt file was accessed from the Bader Lab gene sets website (http://download.baderlab.org/EM_Genesets/current_release/) and used for GSEA and enrichment map visualization. GSEA parameters were defined as suggested by the Bader Lab GSEA Tutorial website (https://enrichmentmap.readthedocs.io/en/docs-2.2/Tutorial_GSEA.html#step-1-generate-gsea-output-files).

Enrichment map visualization

GSEA results were visualized by using the Enrichment Map App in Cytoscape. Only the upregulated gene-sets or pathways that have met the cut-off of FDR<0.01 were clustered using the MCL cluster algorithm in the clusterMaker App. Then, each cluster was manually annotated with a common biological theme. A node (circle) represents a single pathway defined by publicly available databases curated by the Bader Lab. An edge (line) represents the extent of shared genes between the two nodes.

Heatmap of mouse DDR protein abundance

Protein abundance values of the genes curated in the "DNA double-strand break repair" pathway (R-HSA-5693532; Reactome Database ID Release 63) was queried in our previously published mouse mammary proteomic dataset.[28] The heat map created using the pheatmap R package shows the z-scores of LFQ-adjusted IBAQ values in basal, luminal progenitor (LP), and luminal mature (LM) cell types from estrogen plus progesterone samples.

In vivo engraftment of human breast cancer and mouse mammary tumor-derived cell lines

Each cell-line suspended in PBS was mixed with Matrigel in 1:1 v/v ratio resulting in a concentration that a 10 μl sample of the cell-Matrigel mixture contained 0.5M HCC1187, 0.5M MDA-MB-231, 0.3M *Pik3ca*^{H1047R}; *p53*$^{-/-}$ sg4205/sg4687, 0.15M PyMT, or 1.5M NDL3903 cells. This volume was then injected directly into the right inguinal mammary gland of 6-7-week-old virgin female NSG mice using a Hamilton syringe. Xenograft tumours were monitored twice a week starting 7 days after engraftment. Tumour dimensions were measured with a Vernier caliper two times a week and tumour volume (mm^3) was estimated by 0.5 × (minimum diameter in mm)^2 × (maximum diameter in mm) from day 7 post-injection until the end of the study. Mice were sacrificed when endpoints were reached (tumour volume >1500 mm^3, cumulative clinical score >8 or a max score for any animal condition).

In vivo drug testing

The powder form of olaparib or talazoparib was dissolved in DMSO and stored at -80°C. A fresh aliquot of 100 mg/mL olaparib or 0.495 mg/mL talazoparib stock solution was diluted 1:10 in the vehicle solution (10% w/v 2-hydroxypropyl-β-cyclodextrin in PBS) at the time of treatment. Once tumors reached a volume of 100-200 mm^3, mice were randomized into two groups and were treated with either vehicle control (10% DMSO in the vehicle solution) or 100 mg/kg olaparib or 0.33 mg/kg talazoparib daily via intraperitoneal injections until they reached endpoint.

Human DDR proteomic analysis

The abundance values of proteins in the three mammary epithelial populations: basal (BC), luminal progenitor (LP), and mature luminal (ML) were extracted from our previously published human mammary epithelial proteome.[26] The proteomic dataset contains log2-transformed iBAQ-adjusted LFQ values representing protein abundance which were adjusted for batch effects using the ComBat function in the 'sva' R package. Only samples from premenopausal patients were taken into account (n=6 for each BC, LP, ML cell types). Among the 276 curated human DDR genes,[6] 127 proteins were detected in the mammary proteomic dataset based on matching by gene symbol.

Volcano plot and pathway analysis on the human DDR proteins

Of the 276 curated human DDR genes,[6] 127 proteins were detected in the mammary proteomic dataset based on matching by gene symbol. Of 127 DDR proteins, proteins were defined as highly expressed or enriched in one cell type if they met the fold-change (FC) and statistical significance cut-offs compared to the other two cell types. Enriched proteins had a log2FC > 0 and a p-value < 0.05 (paired t-test; p-value was adjusted for multiple testing via FDR). These enriched proteins were visualized in a volcano plot. To identify the top 5 "upregulated" and "downregulated" proteins, their rank values derived from the summation of p-value and fold-change ranks. The protein with the lowest p-value was given a rank of 1, while the highest earned the rank of 'r', which represents the number of total proteins (in this case, 127). For upregulated proteins, the protein with the most positive log2FC value was given a rank of 1, while the least positive (log2FC > 0) value was given a rank of r. For downregulated, the protein with the most negative log2FC value was given a rank of 1, while the least negative (log2FC < 0) value was given a rank of r. The gene names of all enriched proteins were labelled and visualized in the volcano plots.

Generation of human and mouse mammary lineage signatures

The global human and mouse mammary proteomes consisted of 6034 and 4695 annotated proteins, respectively. Basal, luminal progenitor and luminal mature lineage signatures were established by identifying proteins enriched in one cell lineage compared to the other two. This enrichment was defined as a fold-change>5 in humans or fold-change>3 in mice, with a p-value<0.05 determined through one-way ANOVA in conjunction with Tukey's test.

Enrichment of signatures in breast cancer subtypes

Human and mouse lineage signatures in breast cancer expression profiles from METABRIC via single-sample Gene Set Expression Analysis (ssGSEA) using the 'GSVA' R package. The scores were categorized by PAM50 plus Claudin-low subtypes for each signature and assessed for significance using a one-way ANOVA and Tukey's multiple comparisons test.

Enrichment in human breast cancer cell-lines

Enrichment of human or mouse lineage signatures in 50 human breast cancer cell-lines from the Genomics of Drug Sensitivity in Cancer database was determined using the ssGSEA algorithm. For mouse signatures, gene symbols were converted to human homologs using the 'biomaRt' R package. The top cell-lines were determined based on the most substantial differences in ssGSEA enrichment scores between basal and luminal progenitor signatures.

Correlation to breast cancer cell-line drug sensitivity screening

Pearson correlations were performed to assess the association between enrichment scores for each cell type and IC50 values of 23 DDR-related drugs categorized under 'DNA replication' or 'Genome integrity' in human breast cancer cell-lines from the Genomics of Drug Sensitivity in Cancer portal.[51] We limited our analysis to the top ten cell-lines enriched for each basal and luminal progenitor signature. Pearson's correlation coefficients and p-values were computed in R. The IC50 data was sourced from: https://www.cancerrxgene.org/gdsc1000/GDSC1000_WebResources/Home.html

Enrichment of signatures in human breast cancer PDX

Enrichment of human mammary lineage signatures in the transcriptomes of 47 human breast cancer patient-derived xenograft models was determined using the ssGSEA algorithm. The heatmap depicting enrichment scores with annotations for response to talazorapib (ie. response angle) and PAM50 subtype was created using the R packages ComplexHeatmap and circlize. Linear models were fitted using the lm (linear regression) function and QR factorization method in R. Pearson correlation was used to measure the linear dependence between ''response angle'' and ssGSEA values for each lineage. Point shapes depicted PAM50 subtypes: basal (Basal-like), Normal (Claudin-low) and 'other' (HER2+, Luminal A, Luminal B, unknown).

DDR gene expression in human single-cell RNAseq data

Core DDR/HR genes were referenced from Knijnenburg et al.[6] Cell cycle phases were determined using the CellCyclceScoring function and visualized in the volcano plots using the VlnPlot function within Seurat.

QUANTIFICATION AND STATISTICAL ANALYSIS

Descriptive statistical analyses were conducted using Prism 9. Fisher's Exact test and Pearson's correlation were performed using R. Details regarding sample sizes, the specific statistical tests employed, and corresponding p-values for each figure, are provided in the figure legends, the results section, and/or the Methodology section.