## Supplementary Methods

### Immunofluorescence and immunohistochemistry staining

*Human breast, ovary, and fallopian tube tissue*: unstained sections were cut from paraffin blocks of embedded tissue or tissue microarrays. Slides were deparaffinized by immersion in xylene and rehydrated in decreasing concentrations of ethanol. Antigen retrieval was conducted by immersing the slides in citrate buffer (10mM citric acid, 0.05% Tween 20, pH 6.0), heating in a microwave, and cooling for 30 minutes at room temperature. Slides were then washed, blocked for 1 hour with CAS blocking reagent, and incubated with primary γH2AX antibody overnight at 4°C. Slides were then washed and incubated with Alexa Fluor 546 secondary antibody and Hoechst stain for 2 hours. Following this incubation period, all slides were washed and mounted with ProLong Gold until imaging by confocal microscopy.

*Mouse mammary gland*: Similar protocol as described above was conducted with sections from paraffin embedded mouse mammary gland tissue, however following antigen retrieval, staining procedures were conducted in accordance with the manufacturer's protocol for the Mouse on Mouse (M.O.M). immunodetection kit with some modifications. In brief, slides were incubated in M.O.M. Mouse IgG blocking reagent for 1 hour, washed, and incubated with M.O.M. diluent for 5 min followed by incubation with primary γH2AX antibody in M.O.M diluent overnight at 4°C. Slides were washed, incubated with secondary antibody and Hoechst stain in M.O.M diluent for 2 hours, washed and mounted.

*Plated breast epithelial cells*: Primary breast epithelial cells, *BRCA1* and *BRCA2* heterozygous MCF-10A cells, or WT MCF-10A cells were fixed in ice cold methanol for 15 mins at -20°C, washed and blocked with CAS block reagent for 1 hour followed by overnight

incubation with primary γH2AX antibody at 4°C. The next day cells were washed, incubated with secondary antibody and Hoechst for 2 hours and imaged by confocal microscopy.

*ERα IHC and IF colocalization of ERα* and *γH2AX*: IHC was conducted on breast tissue sections from *BRCA1* and *BRCA2* mutation carriers to detect presence of ERα positive epithelial cells. Slides were deparaffinized in xylene and re-hydrated followed by 1 hour antigen retrieval (citrate buffer 0.01 M, pH 6.0) at 95°C and 30 minutes cooling at room temperature. Slides were rinsed and then incubated for 5 minutes with 3% $H_2O_2$ to inhibit endogenous peroxidase activity. Vectastain ABC-HRP (Peroxidase, Anti-Mouse IgG) immunodetection kit (Vector Labs #PK-4002) was then used following the manufacturer's protocol. In brief, after 30-minute incubation in blocking serum, primary ERα antibody (Leica Biosystems #NCL-ER-6F11/2, RRID:AB_876939, 1:50 dilution) was added for 1 hour at room temperature followed by biotinylated secondary antibody incubation for 30 minutes. Slides were washed, developed with DAB substrate kit (Vector Labs #SK-4100) and counterstained with hematoxylin for visualization of ERα positive epithelial cells. In γH2AX co-localization studies by IF, a similar protocol through antigen retrieval was followed as described above. Slides were then blocked for 1 hour with CAS block and incubated with primary ERα and γH2AX (NB100-384, RRID:AB_10002815, 1:300) antibodies overnight at 4°C. Slides were washed and incubated with anti-Mouse Alexa Fluor 546 (Life technologies #A11030, RRID:AB_2534089) and anti-Rabbit Alexa Fluor 488 (Life technologies #A11008, RRID:AB_143165) secondary antibodies with Hoechst.

**RNA-seq studies & computational analysis:**

Total RNA was extracted from samples in all studies by Trizol lysis and RNeasy Mini Kit (Qiagen) following the manufacturer's protocols. Samples were submitted for RNA-seq at the Genomics Resources Core Facility (GRCF, Weill Cornell Medicine). Additional details for each study are described below:

*Human breast tissue*: To assess differences in breast tissue gene expression in *BRCA* mutation carriers with BMI ≥ 25 (*n*=64) vs relative to carriers with BMI < 25 (*n*=67), snap frozen non-tumorous breast tissue was obtained from women who had previously undergone prophylactic or therapeutic mastectomy surgery at Memorial Sloan Kettering Cancer Center (MSKCC). In the therapeutic mastectomy cases, only normal breast tissue from quadrant free of tumor or contralateral breast, as determined by a pathologist, was collected and snap frozen. Sequencing libraries were constructed at the GRCF following the Illumina TrueSeq Stranded Total RNA library preparation protocol with ribosomal RNA depletion. The libraries were sequenced with paired-end 51 bp on the Illumina HiSeq4000 sequencing platform. Raw sequenced reads were pseudoaligned to the human reference genome (UCSC hg19) using the RNA-seq quantification program Kallisto as previously described (*67*) and transcript abundance was quantified to obtain and raw counts. Normalization of the RNA-seq raw read counts and pairwise statistical analysis were performed using the DESeq2 package (version 1.30.1)(*68*). Differentially expressed genes (DEG) were determined by comparing cases with BMI ≥ 25 relative to cases with BMI < 25 or DNA damage Q4 relative to Q1. The tissue RNA-seq heatmap plot (**Fig. 2A**) was generated with R pheatmap package, using the DESeq2 normalized CPM values.

DEGs with Log2FC of >0.3 and <-0.3 and *P* value<0.05 were uploaded to Ingenuity Pathway Analysis (IPA, Qiagen) for data visualization and analysis. The estrogen metabolism heatmap (**Fig. 2C**) was generated by standardizing values relative to data across cases for each gene returning a normalized value from a distribution characterized by the mean and standard deviation of each gene.

*Breast epithelial organoids*: To assess differences in breast epithelial cell gene expression in *BRCA* mutation carriers with BMI ≥ 25 (*n*=9) relative to carriers with BMI < 25 (*n*=10) and in women wildtype (WT) for *BRCA* with BMI ≥ 25 (*n*=8) relative to women with BMI < 25, breast epithelial organoids were isolated from patients. Sequencing libraries were constructed at the Genomics Resources Core Facility (GRCF, Weill Cornell Medicine) following the Illumina TruSeq Stranded mRNA library preparation protocol (Poly-A selection and Stranded RNA-Seq). The libraries were sequenced with paired-end 50 bp on the Illumina HiSeq4000 sequencing platform. All reads were independently aligned with STAR_2.4.0f1 (*69*) for sequence alignment against the human genome sequence build hg19, downloaded using the University of California Santa Cruz (UCSC) genome browser and SAMTOOLS v0.1.19 (*70*) for sorting and indexing reads. Normalization of the RNA-seq raw read counts, and pairwise statistical analysis were performed as described above. DEGs were determined by comparing the cases with BMI ≥ 25 relative to cases with BMI <25. All DEGs with *P* value <0.05 were uploaded to IPA for data visualization and analysis.

*MCF-10A cells*: To assess gene expression changes associated with treatment with obese breast adipose CM, MCF-10A cells carrying a heterozygous *BRCA1* mutation or WT for *BRCA1* were treated with breast adipose CM derived from women with BMI ≥ 25 or BMI < 25 (*n*=3/group). RNA was extracted after 24-hour treatment for RNA-seq. Sequencing libraries

were constructed at the GRCF following the Illumina TruSeq Stranded mRNA library preparation protocol (Poly-A selection and Stranded RNA-Seq). The libraries were sequenced with paired-end 50 bp on the Illumina HiSeq4000 sequencing platform. All reads were independently aligned with STAR_2.4.0f1 (*69*) for sequence alignment against the human genome sequence build hg19, downloaded using the UCSC genome browser and SAMTOOLS v0.1.19 (*70*) for sorting and indexing reads. Cufflinks (2.0.2) (*71*) was used to estimate the expression values (FPKMS), and GENCODE v19 (*72*) GTF file for annotation. The gene counts from htseq-count and DESeq2 Bioconductor package (*68*) were used to identify DEGs. All DEGs with *P* value <0.05 were uploaded to IPA for data visualization and analysis.

*Mouse mammary fat pads*: To assess differences in mammary fat pad gene expression in association with obesity, mammary fat pads were harvested from female *Brca1*+/- mice fed high-fat diet-fed or low-fat diet-fed (*n*=6/group) followed by RNA extraction for RNA-seq. Sequencing libraries were constructed at the GRCF following the Illumina TruSeq Stranded mRNA library preparation protocol (Poly-A selection and Stranded RNA-Seq). The libraries were sequenced with paired-end 50 bp on the Illumina NovaSeq 6000 sequencing platform. The raw sequencing reads in were processed through bcl2fastq 2.19 (Illumina) for FASTQ conversion and demultiplexing. After trimming the adaptors with cutadapt (version1.18), RNA reads were aligned and mapped to the GRCm38 mouse reference genome by STAR (Version2.5.2) (*69*), and transcriptome reconstruction was performed by Cufflinks (Version 2.1.1). The abundance of transcripts was measured with Cufflinks in Fragments Per Kilobase of exon model per Million mapped reads (FPKM). Normalization of the RNA-seq raw read counts and pairwise statistical analysis were performed using the DESeq2 package (version

1.30.1) (*68*). DEGs were determined by comparing the high-fat diet group relative to the low-fat diet group.

Heatmaps showing comparisons between RNA-seq data from mouse and human mammary tissue were generated using IPA. Briefly, mouse raw read counts were standardized relative to data across all mouse samples for each gene returning a normalized value from a distribution characterized by the mean and standard deviation of each gene. Top canonical pathways affected based on DEGs from human cases with BMI ≥ 25 relative to cases with BMI < 25 were then compared to standardized values from mouse samples.

**Generation of *Brca1*+/- mice**

We generated global *Brca1* heterozygous mice (*Brca1*+/-) by crossing *Brca1*$^{flox\ 5\text{-}13/flox\ 5\text{-}13}$ (*73, 74*) with CMV-Cre mice (JAX strain #006054, (*75*)) to produce global *Brca1*$^{D5\text{-}13/+}$ mice. Genotyping was performed using PCR with primers as previously described (*73*). In brief, primers F: 5'- TATCACCACTGAATCTCTACCG -3' and R: 5'-TCCATAGCATCTCCTTCT AAAC -3' were used, whereby deletion of exons 5-13 was detected following amplification of a 594 nucleotide (nt) product. Primers F: 5'-TATCACCACTGAATCTCTACCG -3' and R: 5'-GAC CTCAAACTCTGAGATCCAC -3' were used to detect wild type *Brca1* alleles with product at 390nt. No offspring were found to be homozygous knockout for *Brca1* suggesting that complete loss of *Brca1* is embryonically lethal. These mice were backcrossed to a C57BL/6 strain over 20 times. Genetic testing of mice following backcrossing was performed at Charles River Genetic Testing Services and demonstrated 98.6% fidelity with C57BL/6 inbred strains.

**Generation of *BRCA2* heterozygous MCF-10A cell line**

We used CRISPR-Cas9 gene editing to generate an isogenic MCF-10A cell line carrying a clinically-relevant heterozygous *BRCA2* mutation (6174insT). The forward (A) and reverse (B) sgRNA cloning primers were as follows:

A: CACCGGCCAAACGAAAATTATGGC

B: AAACGCCATAATTTTCGTTTGGCC

These primers were annealed and cloned following standard procedures using the BsmBI/EcoRI site of the *pLenti-U6-sgRNA-tdTomato-P2A-Blas* (LRT2B) vector. WT MCF-10A cells stably expressing an optimized *pLenti-Cas9-P2A-Puro* lentiviral vector were transduced with sgRNAs and underwent Blasticidin selection followed by isolation of single clones using the limiting dilution assays, as described previously (*76, 77*). Candidate colonies expressing TdTomato were submitted for Sanger sequencing (Genewiz) to identify clones exhibiting a heterozygous *BRCA*2 mutation (**fig. S5**). Analysis of sequencing data showed insertion of a thymine (T) nucleotide 5670bp from the start of the open reading frame (ATG) that led to an early stop codon at amino acid 1889. This mutation was maintained upon re-sequencing after multiple passages. The clonal population was expanded and utilized for downstream in vitro studies.

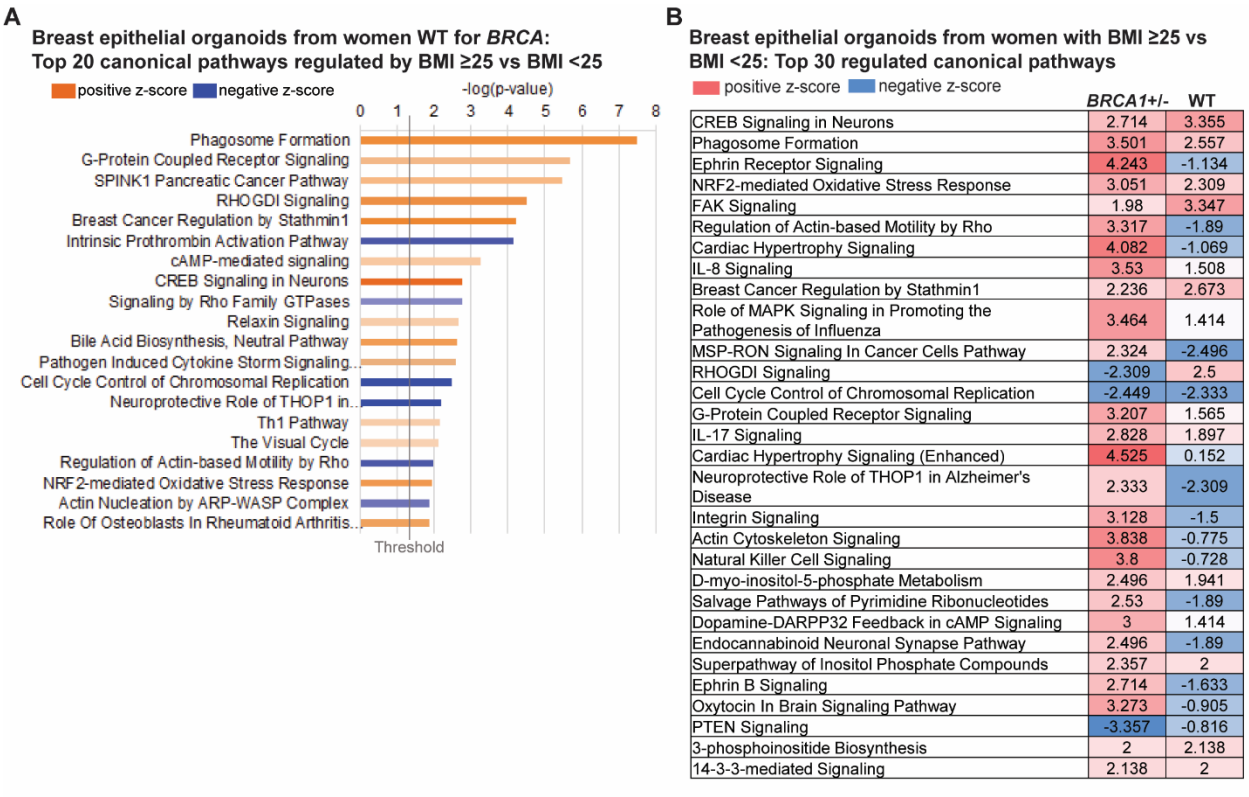
**Quantitative steroid analysis in breast explants**

Steroid profiling was done based on methods previously described (*78*). In brief, snap frozen breast tissue explants were lyophilized and homogenized and 10 mg of individual samples were used in each assay. After spiking with 20 μL of an internal standard mixture (2,2,4,6,6,21,21,21-

$d_8$-17α-hydroxyprogesterone, 1 μg/mL; 2,2,3,4,4,6-$d_6$-dehydroepiandrosterone and 2,2,4,6,6,17α,21,21,21-$d_9$-progesterone, 0.5 μg/mL; 16,16,17-$d_3$-testosterone and 9,11,12,12-$d_4$-cortisol, 0.25 μg/mL; 2,4,16,16-$d_4$-17β-estradiol, 0.2 μg/mL) and mixed with 1 mL of 0.2 M phosphate buffer (pH 7.2), the sample was pulverized using a TissueLyser (Qiagen) at 25 Hz for 10 min with three zirconia beads (3.0 mm I.D., Toray Industries) and centrifuged at 12,000 rpm for 10 min twice. The combined supernatant was loaded onto a preconditioned solid-phase extraction (SPE) cartridge (Oasis HLB; 3 mL, 60 mg; Wasters), the SPE cartridge was washed with 1 mL of water twice and eluted with 2 mL of methanol and 2 mL of 90% methanol. The combined eluate was evaporated under a nitrogen stream at 40°C. The sample was dissolved in 1 mL of 0.2 M acetate buffer (pH 5.2) and 100 μL of 0.2% ascorbic acid solution. It was then extracted with 2.5 mL of methyl *tert*-butyl ether twice. The organic solvent was evaporated under a nitrogen stream and further dried in a vacuum desiccator with $P_2O_5$/KOH for at least 30 min. Finally, the dried residue was derivatized with 50 μL of *N*-methyl-*N*-trifluorotrimethylsilylacetamide/ammonium iodide/dithioerythritol (500:4:2, v/w/w) at 60°C for 20 min. Then 2 μL of the final mixture was injected into the GC-MS system.

In samples with no quantifiable estrogens, the estrogen assay with improved analytical detectability (*79*) was performed from an additional 10 mg of homogenized breast tissues. For the calibration sets, steroid-free breast tissue was freshly prepared 1 day before the experiment. Human breast samples (50 mg) were pulverized in 1 mL methanol/chloroform (1:1, v/v) with 4 zirconia beads at 25 Hz for 5 min followed by centrifugation twice at 12,000 rpm for 3 min, and then supernatants were discarded. The remaining tissue sample was washed with l mL of chloroform/0.6 M methanolic HCl (1:1, v/v), sonicated for 5 min, and centrifuged at 12,000 rpm for 3 min three times. All supernatants were also discarded. To eliminate residual methanolic

HCl, 1 mL of 20% ethanol was added for washing five times. Tissue samples were then frozen at -80 °C until used. No steroid was detectable in GC-MS chromatogram.




| | *BRCA1+/-* | WT |
|---|---|---|
| CREB Signaling in Neurons | 2.714 | 3.355 |
| Phagosome Formation | 3.501 | 2.557 |
| Ephrin Receptor Signaling | 4.243 | -1.134 |
| NRF2-mediated Oxidative Stress Response | 3.051 | 2.309 |
| FAK Signaling | 1.98 | 3.347 |
| Regulation of Actin-based Motility by Rho | 3.317 | -1.89 |
| Cardiac Hypertrophy Signaling | 4.082 | -1.069 |
| IL-8 Signaling | 3.53 | 1.508 |
| Breast Cancer Regulation by Stathmin1 | 2.236 | 2.673 |
| Role of MAPK Signaling in Promoting the Pathogenesis of Influenza | 3.464 | 1.414 |
| MSP-RON Signaling In Cancer Cells Pathway | 2.324 | -2.496 |
| RHOGDI Signaling | -2.309 | 2.5 |
| Cell Cycle Control of Chromosomal Replication | -2.449 | -2.333 |
| G-Protein Coupled Receptor Signaling | 3.207 | 1.565 |
| IL-17 Signaling | 2.828 | 1.897 |
| Cardiac Hypertrophy Signaling (Enhanced) | 4.525 | 0.152 |
| Neuroprotective Role of THOP1 in Alzheimer's Disease | 2.333 | -2.309 |
| Integrin Signaling | 3.128 | -1.5 |
| Actin Cytoskeleton Signaling | 3.838 | -0.775 |
| Natural Killer Cell Signaling | 3.8 | -0.728 |
| D-myo-inositol-5-phosphate Metabolism | 2.496 | 1.941 |
| Salvage Pathways of Pyrimidine Ribonucleotides | 2.53 | -1.89 |
| Dopamine-DARPP32 Feedback in cAMP Signaling | 3 | 1.414 |
| Endocannabinoid Neuronal Synapse Pathway | 2.496 | -1.89 |
| Superpathway of Inositol Phosphate Compounds | 2.357 | 2 |
| Ephrin B Signaling | 2.714 | -1.633 |
| Oxytocin In Brain Signaling Pathway | 3.273 | -0.905 |
| PTEN Signaling | -3.357 | -0.816 |
| 3-phosphoinositide Biosynthesis | 2 | 2.138 |
| 14-3-3-mediated Signaling | 2.138 | 2 |

**Fig. S1. Elevated BMI is associated with distinct changes in gene expression in breast epithelial cells isolated from women who are WT for *BRCA* compared with breast epithelial cells from *BRCA* mutation carriers**

**(A)** IPA analysis of RNA-seq gene expression data showing activation of the top 20 canonical pathways regulated in primary breast epithelial organoids from women WT for *BRCA* with BMI ≥ 25 (*n*=8) relative to women with BMI < 25 (*n*=11) with an absolute value z-score of > 1. The length of the bars are determined by the Fisher's Exact Test *P* value with entities that have a -log(p-value) > 1.3 shown. (**B**) Top 30 canonical pathways regulated in breast epithelial organoids from women with BMI ≥ 25 relative to BMI < 25, comparing activation z-scores in *BRCA*+/- mutation carriers with non-carriers. Activation z-scores are color coded with gradations of red representing a positive z-score and gradations blue representing a negative z-score. Significantly regulated pathways as defined by -log(p-value) >1.3 are shown.

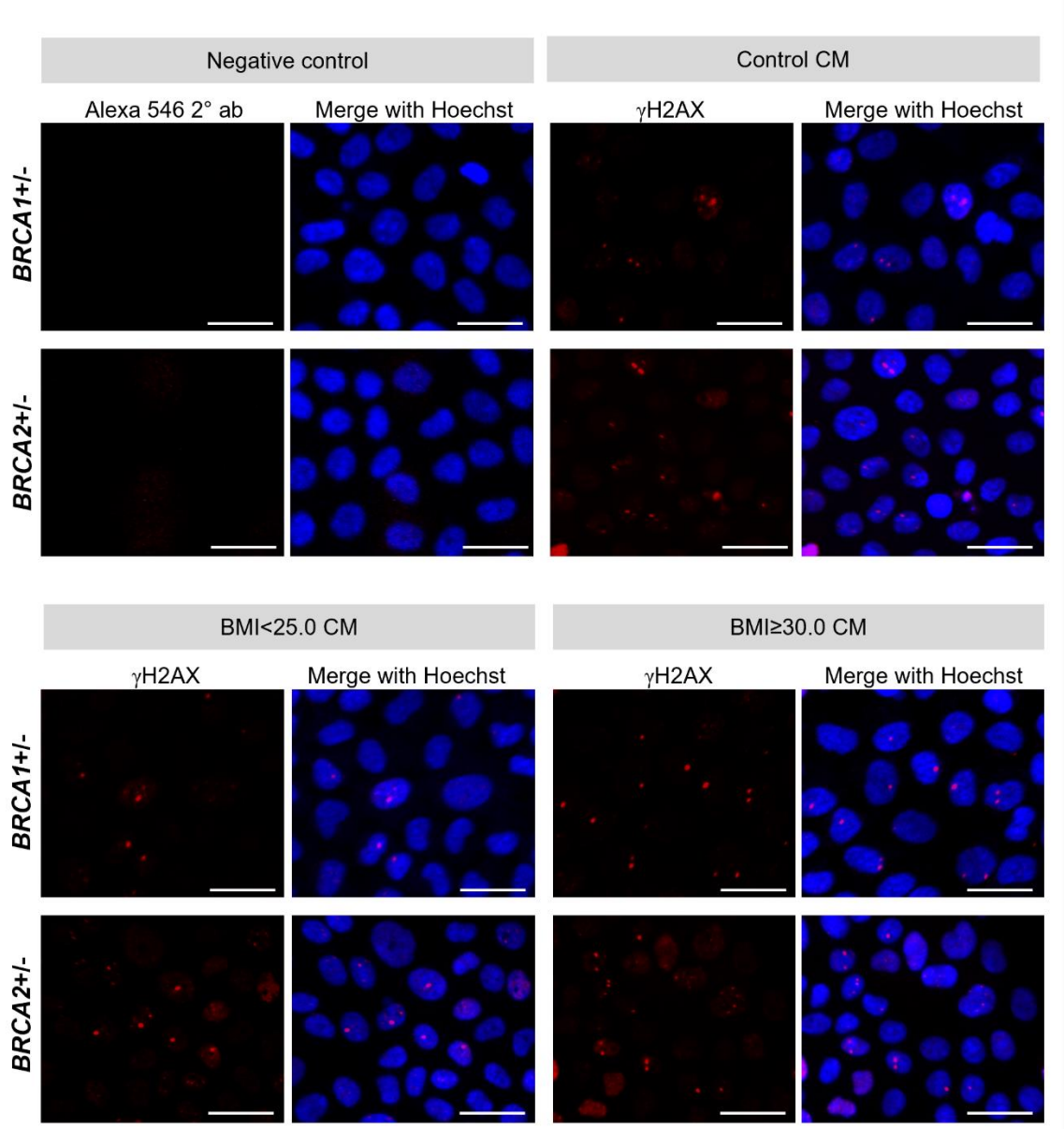



**Fig. S2. Representative confocal images of γH2AX foci immunofluorescence staining in *BRCA1*+/- and *BRCA2*+/- MCF-10A cells.** Top half shows representative γH2AX foci (red) in control conditions, including negative control (cells incubated with secondary antibody only and no primary antibody) and control conditioned media (CM) treatment (cells treated with media used for collection of CM which had not been exposed to conditioning by adipose explants). Foci are shown merged with nuclei stained with Hoechst (blue). Bottom half shows representative

γH2AX staining in cells treated with CM derived from women with BMI < 25 or ≥ 30. *BRCA1*+/- and *BRCA2*+/- MCF-10A cells are shown. Scale bar = 30 μM.

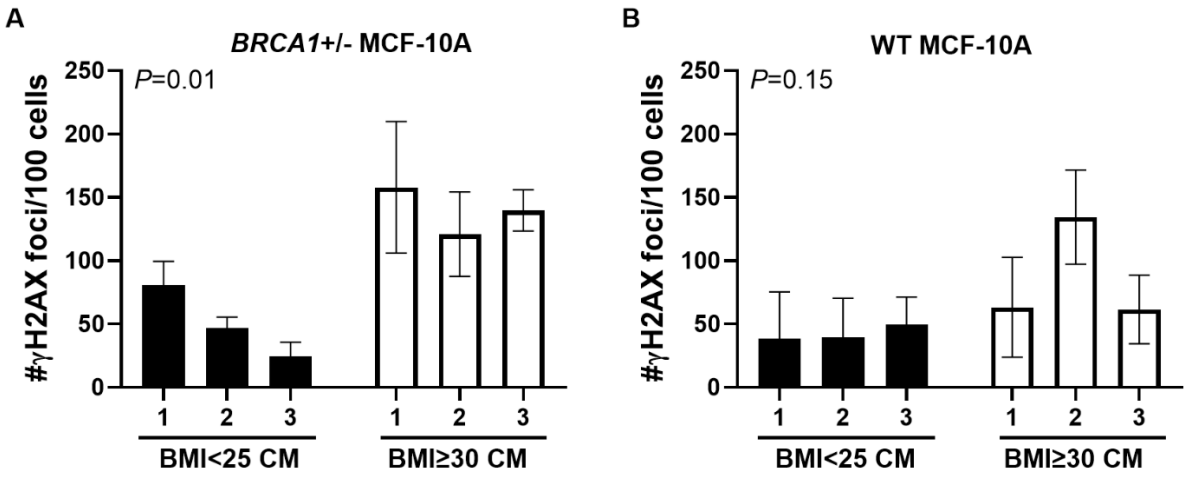



**Fig. S3. Breast adipose conditioned media from women with BMI ≥ 30 stimulates more DNA damage in *BRCA1*+/- MCF-10A cells compared to conditioned media from women with BMI < 25.**

**(A)** *BRCA1*+/- MCF-10A cells and (**B**) wildtype (WT) MCF-10A cells were treated with breast adipose conditioned media (CM) collected from women with BMI< 25 (*n*=3) or BMI ≥ 30 (*n*=3). DNA damage was assessed by IF (number (#) of γH2AX foci/100 cells) after 24 hours treatment. *P* value represents a comparison of average DNA damage in cells treated with CM from breast tissue from women with BMI < 25 vs BMI ≥ 30 by Student's T test. Data is presented as mean +/- SD.

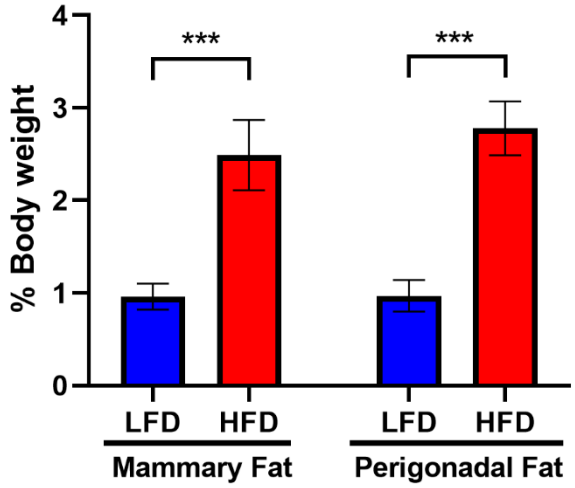



**Fig. S4. *Brca1*+/- mice fed high-fat diet have greater accumulation of body fat compared to *Brca1*+/- mice fed low-fat diet**

Wet weight of inguinal mammary fat pads (subcutaneous fat) and perigonadal fat (visceral fat) at the time of euthanasia, expressed as % of whole-body weight in female *Brca1*+/- mice fed low-fat diet (LFD) or high-fat diet (HFD) for 22 weeks ($n$=12/group). Student's T test was used to determine significant differences between low-fat diet (LFD) and high-fat diet (HFD) mice. Data is presented as mean +/- SD, ***$P$<0.001.

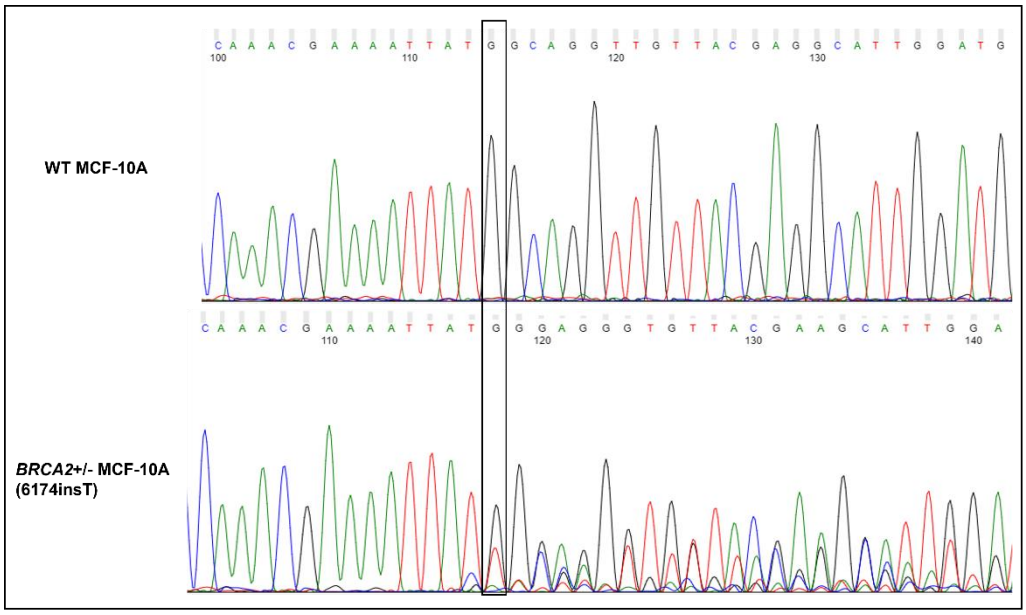



**Fig. S5. Generation of MCF-10A cells carrying a *BRCA*2 heterozygous mutation**

CRISPR-Cas9 gene editing was utilized to generate an isogenic MCF-10A cell line carrying a heterozygous *BRCA*2 mutation (6174insT). Candidate clones were submitted for Sanger sequencing (Genewiz) and a clone exhibiting a heterozygous mutation as determined by the sequencing chromatogram shown in the figure was selected for downstream *in vitro* studies.

**Table S6. Genes involved in "Repair of DNA" regulated in *BRCA1+/-* MCF-10A cells treated with breast adipose CM derived from women with obesity (BMI ≥ 30) compared to women with BMI < 25**

| Gene | Log2FC | *P* value |
|---|---|---|
| TOP2A | -2.971 | 0.0216534 |
| EXO1 | -2.407 | 0.0041363 |
| TNF | -2.366 | 0.0331524 |
| NEIL3 | -1.56 | 0.0454292 |
| BLM | -1.554 | 0.0240665 |
| RAD54L | -1.55 | 0.0265702 |
| BRCA2 | -1.362 | 0.0200827 |
| XRCC2 | -1.228 | 0.0213886 |
| BRIP1 | -1.196 | 0.0160386 |
| BRCA1 | -1.082 | 0.016198 |
| FANCA | -0.927 | 0.0374675 |
| FANCD2 | -0.868 | 0.0456246 |
| TRIP13 | -0.813 | 0.0072051 |
| FANCM | -0.638 | 0.0156575 |
| ANKRD32 | -0.479 | 0.0115143 |
| MDM2 | -0.435 | 0.0095298 |
| FOS | -0.416 | 0.0424863 |
| HMGA2 | -0.384 | 0.0298825 |
| RRM2B | -0.383 | 0.0158232 |
| USP45 | -0.372 | 0.0033442 |
| UCHL5 | -0.369 | 0.0016047 |
| OTUD6B | -0.335 | 0.0271729 |
| RAD18 | -0.334 | 0.0322693 |
| HMGB1 | -0.325 | 0.0245722 |
| TCEA1 | -0.319 | 0.0165788 |
| DEK | -0.304 | 0.0111833 |
| PMS1 | -0.29 | 0.0165185 |
| PHF6 | -0.271 | 0.0112097 |
| HNRNPU-AS1 | -0.269 | 0.0298011 |
| RIF1 | -0.267 | 0.0259465 |
| TDG | -0.259 | 0.0321275 |
| USP15 | -0.255 | 0.0472664 |
| BCLAF1 | -0.254 | 0.0267277 |
| RB1 | -0.253 | 0.0264751 |
| SMCHD1 | -0.245 | 0.0040162 |
| MRE11A | -0.235 | 0.0092542 |
| FAM175A | -0.224 | 0.0329318 |
| RAD21 | -0.22 | 0.0433965 |
| CHEK1 | -0.211 | 0.0422347 |
| THOC1 | -0.206 | 0.0421476 |
| WRN | -0.191 | 0.0306967 |
| EIF3F | 0.157 | 0.0450842 |

| | | |
|---|---|---|
| PRKCZ | 0.259 | 0.032568 |
| PIK3C2B | 0.29 | 0.0040939 |
| CBX8 | 0.338 | 0.0150821 |
| CEBPA | 0.438 | 0.0047412 |
| CLU | 0.613 | 0.0099281 |

**Tables S1-S5 and S7-S8 can be found in Data File S1 (multi-tab Excel file)**