**SUPPLEMENTAL MATERIALS AND METHODS**

**Cell lines and Culture conditions**

MCF-7 and T47D were maintained in minimum essential medium Eagle alpha modification (alpha MEM) supplemented with 10% FBS, 10mM HEPES, nonessential amino acids, sodium pyruvate, 2mM L-glutamine, insulin, EGF, and hydrocortisone. 293T cells were cultured in Dulbecco modified Eagle medium (DMEM) supplemented with 10% FBS.

***High-throughput Dose Response Survival Assay***

Cells were seeded at a density of 1,000 cells/well in a 96-well plate and treated with a single drug or drug combinations at the indicated concentrations. Treatment with palbociclib, ribociclib (Selleckchem), and abemaciclib (MedChem Express) was for 6 consecutive days. Treatment with Napabucasin (Selleckchem), TTI-101 (previously C188-9; Tvardi Therapeutics, Inc.), Niraparib (MedChem), Olaparib (MedChem), AZD1775 (ChemieTek), BKM120 (APExBIO), Everolimus (RAD001; APExBIO), Mifepristone (RU486; Selleckchem), Dinaciclib (Merck & Co., Inc.), Roscovitine (ManRos Therapeutics), and SNS-032 (Selleckchem) was for 72 hours. At completion of drug treatment, cultures were continued in drug-free medium (also replaced every other day) until day 12, after which they were stained with 0.5% crystal violet solution.For treatment with antiestrogens, tamoxifen (Sigma) and fulvestrant (Sigma), cells were seeded at a density of 1,000 cells/well in a 96-well plate and cultured in estrogen-free media which is composed of phenol red-free improved minimum essential medium (IMEM) supplemented with 10% charcoal/dextran treated FBS for 24 hours. Cells were then treated with 10nM beta-estradiol (E2) and either tamoxifen or fulvestrant for 48 hours, after which they were stained with 0.5% crystal violet solution.The plates were then solubilized with a solution of 0.1% sodium citrate in 50% ethanol, and absorbance was measured at 570 nm using the Epoch Microplate Spectrophotometer (BioTek Instruments, Inc, Winooski, VT). Values were normalized to those of their no treatment controls and analyzed in GraphPad Prism by non-linear regression to obtain the half-maximal inhibitory concentrations (IC50 values).

***Cell Proliferation Assays***

To determine drug effect on doubling time of cells, cells were seeded at a density of 5,000 and plated in each well of 6-well plates and treated with the indicated agents for either 72 hours or 6 days dependent on drug used and cells were allowed to recover in the absence of drug until 9 days. Cells were then harvested and counted using the BioRad TC20 Automated Cell Counter on days 0, 3, 6, and 9.

To determine drug effect on colony formation (clonogenic assay), 5,000 cells were plated in each well of 6-well plates, treated with the indicated agents for either 72 hours or 6 days dependent on drugs used and allowed to recover in the absence of drug until 12 days. Cells were then washed with 1X PBS and stained with a 0.5% crystal violet solution in 25% methanol for 10 minutes. Plates were then scanned to obtain pictures.

***Flow Cytometry Analysis***

Cell Cycle: Cells were plated (1×105 cells) on 10-cm plates and treated with the indicated agents for 6 days. Following treatment cells were subjected to trypsinization, washed with 1X PBS, and fixed with 3.5 ml ice-cold 1X PBS and 1.5 ml of 95% ethanol at 4 °C overnight. Cells were then stained with propidium iodide (1 mg ml-1) and RNaseA (1 mg ml-1) in 0.1% (v/v) Triton X-100 (Sigma) in 1X PBS at 4 °C overnight and then analyzed with the Beckman Coulter Gallios Flow Cytometer.

CD44/CD24:Cells were plated (1x105 cells) on 10-cm dishes then harvested by trypsinization.Cells were washed 3 times with 1X PBS containing 1% horse serum andresuspended in 10μl PE anti-mouse CD24 (BD Pharmingen), 10 μl APC anti-mouse CD44 (BD Pharmingen), and **3**0μl of the 1% horse serum PBS buffer. The samples were incubated for 20 minutes on ice, washedwith 1% horse serum PBS buffer and then analyzed with the Beckman Coulter Gallios Flow Cytometer.The data was analyzed by the Kaluza software to identify the CD44 high and CD24 lowpopulation of cells, which are termed breast cancer stem cell-like (B-CSC-L). Cells with only a single antibody staining were used to set up the gates foranalysis.

Aldefluor (ALDH):Measurement of B-CSC-L population based on ALDH positivity was performed using the Aldefluor kit (StemCell technologies, Cambridge, MA) according to the manufacturer’s protocol.

Annexin V:Apoptotic cells were measured by using the Alexa Fluor 488 Annexin V Dead Cell Apoptosis kit (Invitrogen, Waltham, MA) according to the manufacturer’s protocol.

***Western Blot Analysis***

Protein expression was analyzed by western blotting using whole cell lysates lysed in RIPA buffer and the following antibodies: ER(6F11; VP-E613), PR (#8757), AR (Clone441; M356201-2), FOXA1 (ab23738), Cyclin E (HE12; sc247), Rb (4H1; #9309), p-Rb (Ser807/811; #9308), CDK4 (C-22; sc260), CDK6 (C-21; sc177), CDK2 (D-12; sc6248), p-CDK2 (Thr160; #2561), mTOR (7C10; #2983), p-mTOR (Ser2448; #5536), S6 (54D2; #2317), p-S6 (Ser235/236; #2211), E-cadherin (24E10; #3195), N-cadherin (#4061), Vimentin (R28; #3932), STAT3 (D3Z2G; #12640), pY-STAT3 (Tyr705; #9145), Vinculin (V9131), and actin (C4; MAB1501) following standard procedures.

***Quantitative RT-PCR***

Total RNA was isolated from cell cultures with TRIzol reagent according to the manufacturer’s protocol (Invitrogen). Two micrograms of the RNA samples were reverse-transcribed using a cDNA Synthesis Kit (Applied Biosystems). RT-PCR was done with aliquots of cDNA samples mixed with SYBR Green Master Mix (Sigma). Reactions were carried out in triplicate. The fold difference in transcripts was calculated using the ΔΔCt method with GAPDH as a control. The following primers pairs were used:

**ER:** 5’TGAAAGTGGGATACGAAAAGAC-3’; 5’CAGGATCTCTAGCCAGGCACAT-3’

**PgR:** 5’ACCCGCCCTATCTCAACTACC-3’; 5’AGGACACCATAATGACAGCCT-3’

**pS2:** 5’GTACACGGAGGCCCAGACAGA-3’; 5’-AGGGCGTGACACCAGGAAA-3’

**GREB1**: 5’ATCAGCTGCTCGGACTTGCTG-3’; 5’TGAGCTCCGGTCCTGACAGATG-3’

**GATA3:** 5’TGTCTGCAGCCAGGAGAGC-3’; 5’ATGCATCAAACAACTGTGGCCA-3’

 **NRIP1**: 5’GCTGGGCATAATGAAGAGGA-3’; 5’CAAAGAGGCCAGTAATGTGCTATC-3’

**AR:** 5’CCTGGCTTCCGCAACTTACAC-3’; 5’GGACTTGTGCATGCGGTACTCA-3’

**FOXA1**: 5’CATCACCATGGCCATCCA-3’; 5’GGTCCATGATCCACTGGTAGATC-3’

**GR**5′CAGCATGCCGCTATCGAAA-3′; 5′ACTCCTGTAGTGGCCTGCTGAA-3′

**AhR:** 5′CTCAGCCTCCCGAGTTGCT-3′; 5′-TTCCCAGGTTCAGGCTATTCTC-3′

**E-cadherin**: 5’GTCAGTTCAGACTCCAGCCC-3’; 5’AAATTCACTCTGCCCAGGACG-3’

**Vimentin:** 5’GAGAACTTTGCCGTTGAAGC-3’; 5’GCTTCCTGTAGGCAATC-3’

**N-cadherin:** 5’ACAGTGGCCACCTACAAGG-3’; 5’CCGAGATGGGGTTGATAATG-3’

**Snail**: 5’CACACGCTGCCTTGTGTCT-3’; 5’GGTCAGCAAAAGCACGGTT-3’

**Twist:** 5’GGAGTCCGCAGTCTTACGAG-3’; 5’TCTGGAGGACCTGGTAGAGG-3’

**Slug:** 5’GGGGAGAAGCCTTTTTCTTG-3’; 5’TCCTCATGTTTGTGCAGGAG-3’

**CD133**: 5’TGCTGCTTGTGGAATAGACAGAATG-3’; 5’AGGAAGGACTCGTTGCTGGTGAA-3’

**CD44:** 5’CGGACACCATGGACAAGTTT-3’; 5’GAAAGCCTTGCAGAGGTCAG-3’

**FoxC2**: 5’GCCTAAGGACCTGGTGAAGC-3’; 5’TTGACGAAGCACTCGTTGAG-3’

**ALDH1:** 5’CTGTGTTCCAGGAGCCGAAT-3’; 5’AGCATCCATAGTACGCCACG-3’

**OCT4**: 5’GAGCAAAACCCGGAGGAGT-3’; 5’TTCTCTTTCGGGCCTGCAC-3’

**SOX2:** 5’ACACCAATCCCATCCACACT-3’; 5’GCAAACTTCCTGCAAAGCTC-3’

**NANOG:** 5’GCTTGCCTGCTTTGAAGCA-3’; 5’TTCTTGACCGGGACCTTGTC-3’

**IL-6:** 5’GAGAAAGGACATGTAACAAGAGT-3’; 5’GCGCAGAATGAGATGAGTGT-3’

**Stat3:** 5’AAAGCAGCAAAGAAGGAGGC-3’; 5’CTGGCCGACAATACTTTCCG-3’

**SOCS3:** 5’GGAGTTCCTGGACCAGTACG-3’; 5’TTCTTGTGCTTGTGCCATGT-3’

**Stat5a:** 5'GTAAGGCTGTGTACACTGACAC-3'; 5'CATAGGGTTCACAGAGAGTCTG-3'

**Stat5b:** 5'CTCCAAATACTACACACCAGTTC-3'; 5'-TCGCACTCCCTTCGCTGGTG-3'

**RAD51**: 5’CAACCCATTTCACGGTTAGAGC-3’; 5’TTCTTTGGCGATAGGCAACA-3’

**BRCA1**: 5’ACCTTGGAACTGTGAGAACTCT-3’; 5’TCTTGATCTCCCACACTGCAATA-3'

**BRCA2**: 5’ACAGAACCAATAAGCCATGTGG-3’; 5’TGCAAACATTAACGCAGCTTC-3

***Immunofluorescence (H2AX and Rad51)***

To measure DNA damage repair using γH2AX and Rad51 staining, cells were plated at 1x104 cells/well in an eight-well chamber slide. After 72 hours, cells were fixed in 4% paraformaldehyde for 20 minutes before staining. For immunofluorescence staining, fixed cells were rinsed in PBS and permeabilized for 5 minutes in 0.1% Triton X-100. The cells were then blocked in 1% BSA for 1 hour at room temperature. Primary antibodies prepared in 1% BSA were added to the cells followed by a 2-hour incubation at room temperature in a moist chamber. Primary antibodies used were those against γH2AX and RAD51. A secondary antibody was added after the cells were washed thoroughly in PBS. Cells were incubated with secondary antibodies tagged with Alexa Fluor dyes (goat-anti-mouse-Alexa Fluor-488 and goat-anti-rabbit-Alexa Fluor-594; Invitrogen) for 1 hour at room temperature. After being rinsed and washed thoroughly with PBS, slides were mounted using VECTASHIELD mounting medium (Vector Laboratories) containing DAPI and sealed. Cells were visualized using a fluorescence microscope.

***Migration Assay***

Cells were plated at a concentration of 0.5 million per well in a 6-well dish and grown until they reached 90% confluency. A scratch was made in the wells with a pipette tip, and cells were allowed to grow under normal culture conditions for 48 hours. Pictures were taken at 0hr, 12hr, 24hr, and 48hr to measure the migration capacity of the cells. The distance moved by the cells divided by time was used to obtain the percentage of wound closure.

***Mammosphere formation Assay***

To generate primary mammospheres, cells were seeded in 6-well ultra-low attachment plates (10,000 cells/ml) in serum-free MEM supplemented with 20 ng/ml bFGF, 20 ng/ml EGF and B27 (Invitrogen) and incubated for 5-7 days. For secondary mammosphere assay, cells from primary mammospheres were dispersed with 0.05% trypsin, seeded in 6-well ultra-low attachment plates (7,500 cells/ml) in mammosphere media and incubated for 5-7 days. Mammospheres with at least a size of 100μm were counted with an automated colony counter (Oxford Optronix, Oxford, UK) following MTT staining (Sigma Aldrich, St. Louis, MO).

***Mammary Acini Morphogenesis Assay***

Eight-chambered glass slides were coated with 35 μl growth factor reducued (GFR) matrigel per well and left to solidify for 15 minutes. 1X105 cells were suspended in assay media (DMEM/F12, 2% horse serum, insulin, EGF, and hydrocortisone. Cells were mixed with assay media containing 4% GFR matrigel 1:1 and 400 μl was added to each chamber. Cells were cultured for 15 days and assay media was replaced every 4 days. The cells were fixed with 2% paraformaldehyde, washed with PBS:glycine, blocked for 1hour with blocking buffer containing 10% goat serum and then incubated with primary antibody (E-cadherin) overnight at 4 °C. Primary was removed, cells were washed, and probed with secondary antibody coupled with Alexa flour dye for 1hour. Slides were mounted using VECTASHIELD mounting medium (Vector Laboratories) containing DAPI and sealed. Cells were visualized using a confocal microscope to obtain z-stacked images.

***shRNA knockdown of STAT3 and RB***

Stable shRNA knockdown cells were generated using the following Dharmacon pGIPZ-shRNA for STAT3: V2LHS\_88502, V3LHS\_641818, V3LHS\_376016 and RB: V2LHS\_130606 purchased from the shRNA and ORFeome core facility at MD Anderson Cancer Center. To generate lentivirus expressing shRNA, HEK 293T cells were transfected with pCMVdeltaR8.2, pMD2.G (produced by the Didier Trono laboratory and made available through the Addgene repository) and pGIPZ vector (scrambled shRNA or shRNA against gene of interest) using polyethylenimine transfection reagent according to the manufacturer’s protocol. After 48 hours of transfection, the virus-containing medium was collected, filtered through a 0.45-μm filter, and added to the cells of interest in the presence of 8 μg/mL of polybrene (Millipore). GFP expression was confirmed and the lentivirus-infected cells were either GFP sorted and/or selected with 1 μg/mL puromycin (InvivoGen, San Diego, CA).

***siRNA knockdown of FOXA1***

Transfections were performed by reverse transfection using Lipofectamine 3000, 50nM FOXA1 siRNA (SMARTpool: ON-TARGETplus, L-010319-00-0005) or non-targeting siRNA (ON-TARGETplus, D-001810-10-05). Cells were cultured for either 48 or 72 hours and used for western blot and CD44/CD24 flow cytometry analyses.

***Kinase assay***

Cell extracts were used for immunoprecipitation with either Cyclin E or CDK2 antibodies. The immunoprecipitates were incubated with kinase assay buffer containing 60 μM cold ATP, 5 μCi of [32P]ATP, and 1 μg of GST-Rb (Santa Cruz Biochemicals) in a final volume of 30 μl at 37°C for 30 min. The products of the reactions were analyzed by SDS-PAGE gel. The gels were then stained, destained, dried, and exposed to X-ray film.

***Reverse Phase Protein Array (RPPA)***

RPPA was performed at the Functional Proteomics RPPA Core Facility at MD Anderson Cancer Center. Cellular proteins were denatured by 1% SDS (with Beta‐mercaptoethanol) and diluted in five 2‐fold serial dilutions in dilution lysis buffer. Serial diluted lysates were arrayed on nitrocellulose‐coated slides (Grace Bio Lab) by Aushon 2470 Arrayer (Aushon BioSystems). Total 5808 array spots were arranged on each slide including the spots corresponding to serial diluted: 1) “Standard Lysates”; 2) positive and negative controls prepared from mixed cell lysates or dilution buffer, respectively. Each slide was probed with a validated primary antibody plus a biotin‐conjugated secondary antibody. Only antibodies with a Pearson correlation coefficient between RPPA and western blotting of greater than 0.7 were used for RPPA. Antibodies with a single or dominant band on western blotting were further assessed by direct comparison to RPPA using cell lines with differential protein expression or modulated with ligands/inhibitors or siRNA for phospho‐ or structural proteins, respectively. The signal obtained was amplified using a Dako Cytomation–Catalyzed system (Dako) and visualized by DAB colorimetric reaction. The slides were scanned, analyzed, and quantified using a customized‐software to generate spot intensity.

Each dilution curve was fitted with a logistic model (“Supercurve Fitting” developed by the Department of Bioinformatics and Computational Biology at the MD Anderson Cancer Center, “http://bioinformatics.mdanderson.org/OOMPA”). This fits a single curve using all the samples (i.e., dilution series) on a slide with the signal intensity as the response variable and the dilution steps as independent variable. The fitted curve is plotted with the signal intensities – both observed and fitted ‐ on the y‐axis and the log2‐concentration of proteins on the x‐axis for diagnostic purposes. The protein concentrations of each set of slides were then normalized for protein loading. Correction factor was calculated by: 1) median‐centering across samples of all antibody experiments; and 2) median‐centering across antibodies for each sample.

**Whole Exome Sequencing (WES)**

DNA was extracted from cell lysates using QIAamp DNA Mini Kit (Qiagen) following manufacturer’s protocol. The purity of the DNA was analyzed using the Epoch Microplate Spectrophotometer (BioTek Instruments, Inc, Winooski, VT). Isolated DNA was submitted to Admera Health for WES using the Illumina platform. Sequencing read quality was inspected using the FastQC software ([http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/%29). Reference genome hg19 and Burrows-Wheeler Aligner (BWA) was used for sequence alignment. The alignment file was then sorted, indexed and converted to mpileup file using SAMtool. Duplicate reads were marked and removed to eliminate potential false positive variant call resulted from such duplicate reads. The VarScan platform was used as the variant caller. This robust heuristic and statistic model allows for accurate variant calling detecting SNVs, indels and copy number variations. The variant filtering was performed with both VarScan and VCFtools with the following filtering criteria: minimum read depth at the position: 8, minimum supporting reads at the position: 2, minimum variant allele frequency threshold: 0.2, minimum base quality at the position: 15, minimum mapping quality: 1, percentage of read supporting on one strand to be ignored (strand bias): 90%, p-value threshold: 0.01. Variants were further annotated using SnpEff ([http://snpeff.sourceforge.net/)](http://snpeff.sourceforge.net/%29).

**RNA sequencing (RNA-seq) and Gene Set Enrichment Analysis (GSEA)**

Total RNA was isolated from cell cultures using an RNeasy Kit with DNase treatment according to the manufacturer's protocol (Qiagen). RNA samples were submitted to the Sequencing and Microarray Facility at MD Anderson Cancer Center. RNA samples were Poly(A) selected, and libraries for RNA-seq were prepared using mRNA-Seq Sample Prep Kit (Illumina, California, USA) according to the manufacturer’s instructions. Single-end RNA-seq of 36-bp read length was performed using Illumina HiSeq 2000 according to the standard protocol. FASTQ sequence files were obtained, and the RNA-seq tags were aligned to the human reference genome. Functional analysis of the differentially expressed transcripts was performed using gene set enrichment analysis (GSEA). A circos plot was generated with the top 12 upregulated and 12 downregulated pathways where the edges represent the –log10 p-values from the Fisher’s exact test. Fisher’s exact test was used to assess the significance of the number of genes shared by two pathways as represented by the connecting lines in the circos plot.

**Immunohistochemical (IHC) Analysis of Patient Samples**

***CLIA-approved immunohistochemistry of ER, PR and pY-STAT3***

IHC was performed on formalin-fixed, paraffin-embedded (FFPE) tissue sections in our clinical IHC laboratory, which is certified under the provisions of the United States Clinical Laboratory Improvement Act and accredited by the College of American Pathologists. The protocol has been validated for use on 4m FFPE tissue sections. This protocol includes de-paraffinization (30 minutes at 72°C) and rehydration with antigen retrieval performed at 100°C for 30 minutes with citrate buffer, pH 6.0. Endogenous peroxidase is blocked with 3% peroxide for 5 minutes. Primary antibodies (ERα, clone 6F11, Novacastra, 1:35 dilution; PR, clone 1E2, Ventana (pre-diluted); and pY-STAT3, clone Tyr705, Cell Signaling, 1:100 dilution) are applied for 15 minutes. Post primary antibody detection is carried out using a commercial polymer system (Bond Polymer Refine Detection, Leica) for 8 minutes, and stain development is achieved by incubation with DAB and DAB Enhancer (Leica). A positive batch control is added to every IHC run and on-control to each slide. The evaluation of ER was interpreted as positive for any nuclear staining ≥1% of tumor cells. For pY-STAT3, percentage positivity of nuclear staining was evaluated.

For IHC for Cyclin E and RB, 5µm thick sections of FFPE tumor tissues were deparaffinization and rehydrated. Antigen retrieval was performed with citric acid based buffer at pH 6.0 with the use of a hot plate in a metal container for 20 minutes. The slides were then incubated with 3% hydrogen peroxide and methanol to block endogenous peroxidase activity and nonspecific protein-protein interactions, respectively. After 1 hour blocking for nonspecific staining, the sections were incubated overnight at 4°C in a humid chamber with mouse monoclonal antibody for cyclin-E (Clone E-4, Santa Cruz Biotechnology, 1:200 dilution) or mouse monoclonal antibody for RB (Clone 4H1, Cell Signaling Technology, 1:100 dilution). The antibody was detected with a VECTASTAIN Elite ABC kit (PK6102; Vector Laboratories, Burlingame, CA) followed by staining with DAB substrate (Vector Laboratories) and counterstaining with hematoxylin (DAKO), and then were mounted. All washing steps were performed with PBS alone first and then with PBS with 0.1% Tween. The evaluation of cyclin-E was performed according to cytoplasmic staining (LMW-E) of tumor cells as positive or negative.