Supplemental Appendix

Supplemental Methods

Supplemental References

Supplementary Figures 1, 2, 3, 5, 6, 7 and 8

Supplementary Tables 1, 2, 3, 4 and 5

Supplemental Methods

SNP genotyping and copy number analysis

SNP genotyping was performed on a 250K NspI Affymetrix SNP chip array according to the manufacturer's protocol (Affymetrix, Santa Clara, CA). Briefly, 250 ng of genomic DNA was digested with *NspI*, ligated to a universal adaptor, and amplified by polymerase chain reaction (PCR) using a single universal primer. The PCR products were then purified and the subsequent amplicons were quantified, fragmented, labeled, and then hybridized to the 250K *NspI* SNP arrays. The dChip program (1) allows copy number (CN) and loss-of-heterozygosity (LOH) analyses to be performed using user-defined baseline and matched samples. Data originating from the 250K NspI array was normalized to the baseline array using its median signal intensity with the invariant set method. The model-based and mismatch probe (PM/MM) method was used to obtain the signal values after normalization in dChip. To infer CN, the parental MCF-7 sample was used as the baseline readout and reference genome, and the median smoothing method was applied with a window size of 5 SNP's. For the calculation of LOH values, the Hidden Markov Model was applied from within dChip with the LOH call threshold set to 0.95. Copy number values for the TamR clones were normalized to the MCF-7 reference sample. Results of this analysis were then plotted according to relative SNP position and copy-number, and regions of shared CN gains and losses within the TamR clones were filtered from the data set to be further characterized.

Tumor samples

We obtained five samples from the Johns Hopkins Rapid Autopsy study from female patients with documented tamoxifen resistance. Clinical pathologic characteristics are shown in Table S4. Metastatic breast cancer tumor and adjacent non-malignant liver samples were excised from the patients and snap frozen for further analysis. For each patient sample, genomic DNA was isolated and purified from both normal and metastatic tissue, which was then analyzed via quantitative real time PCR (qRT-PCR) using primers specific to the *MACROD2* locus and normalized to an invariant locus found on chromosome 20p. Matched primary and metastatic samples from this series of patients were previously used for construction of tissue microarrays (TMAs) using formalin fixed paraffin embedded tissues (2). These TMAs were used for immunohistochemistry staining as described below. All human samples were collected and processed in accordance with approved Institutional Review Board tissue banking and use protocols.

Tissue microarrays of human primary and metastatic breast carcinomas.

Briefly, each TMA consisted of 99 cores measuring 1.4 mm in diameter, with five to ten cores of tumor sampled from each PBC and MBC, including a core of benign breast lobules as an internal control. The autopsies were performed on patients with widely metastatic breast carcinoma within a 4 hour post-mortem interval, and the multiple site MBCs harvested at autopsy were formalin-fixed or flash frozen and processed identically to surgical breast specimens at our institution. Metastases that were subjected to decalcification were excluded due to potential interference with the immunohistochemistry.

Immunohistochemistry and expression scoring.

The TMAs were labeled by immunohistochemistry for MACROD2 with a custom antibody (1:8000 dilution, clone 6711, Genscript). Briefly, unstained 5- µm sections were cut from paraffin TMA blocks. Slides were deparaffinized by routine techniques and subjected to antigen retrieval with target retrieval solution (Dako). Primary antibody incubation was performed overnight at 4°C. MACROD2 labeling was recorded as a percentage and intensity of nuclear labeling. Intensity of staining was recorded as weak, weak-to-focally moderate, moderate, moderate-to-focally strong and strong. Diffuse (>90%) cytoplasmic labeling of any intensity for MACROD2 was seen in 89% (17/19) of evaluable PBCs and in at least one MBC from 100% (20/20) patients. Furthermore, diffuse, weak cytoplasmic labeling for MACROD2 was seen in normal benign breast lobules, but not in the myoepithelial cells. Weak to strong cytoplasmic labeling was also noted in other normal organ sites included on the autopsy arrays, including smooth muscle, liver, brain, lymphocytes, basal epidermis, and small bowel mucosa. For these reasons, cytoplasmic labeling for MACROD2 was interpreted to be background labeling or background expression, and only nuclear MACROD2 labeling was considered significant for further analysis. Any nuclear staining with greater than weak intensity (i.e., weak-to-focally

moderate, moderate, moderate-to-focally strong, or strong) was considered to be positive. Any case with weak intensity labeling is scored as 0.

Cell culture and reagents

MCF-7, T47D breast cancer cell lines, and MCF-10A cells, were obtained from the American Type Culture Collection (Rockville, MD), and verified by DNA STR fingerprint analysis. MCF-7 and T47D cells were grown and maintained on DMEM with 5% fetal bovine serum (FBS) and 1% penicillin and streptomycin. The establishment of tamoxifen resistant cell lines has been previously described (3). Briefly, the TamR clones were generated by single cell diluting MCF-7 cells which had been chronically grown on 1μM 4-OH-tamoxifen in phenol red free DMEM with 5% charcoal-stripped dextran treated (CD) serum (Hyclone). Growth and drug assays performed using these cell lines were done in phenol red free DMEM with 0.5% CD serum with the addition of either ethanol (vehicle control), 1.25 nM to 12.5 nM of 17-β-estradiol (Sigma), 1μM 4-OH-tamoxifen (Sigma) or a combination of these drugs as indicated. All cells were cultured at 37°C at 5% CO_2.

The non-transformed human breast epithelial cell line MCF-10A (4) and its derivatives were grown in DMEM/F12 (1:1) supplemented with 5% horse serum (Hyclone), EGF at 20 ng/ml, insulin at 10 μg/ml, hydrocortisone at 0.5 μg/ml, and cholera toxin at 0.1 μg/ml (hereafter denoted as ‘‘growth medium’’). All supplements were purchased from Sigma-Aldrich unless otherwise noted.

Reverse Phase Protein Array (RPPA)

Parental MCF-7 cells and the three TamR clones were seeded at $2x10^6$ cells per well in a 6-well plate using assay media (phenol red free DMEM media supplemented with 0.5% CD serum) supplemented with ER ligands or vehicle only controls. The cells were grown for 2 days and then harvested for RPPA lysates using lysis buffer and 4X SDS sample buffer supplied by the 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 buffer (lysis buffer containing 1% SDS). Serial diluted lysates were arrayed on nitrocellulose-coated slides (Grace Biolab) by Aushon 2470 Arrayer (Aushon BioSystems). A total of 5808 array spots were arranged on each slide including the spots corresponding to 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 in reverse phase protein array study. 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 Microvigene (VigeneTech Inc.) to generate spot intensities. Each dilution curve was fitted with a logistic model ("Supercurve Fitting" developed by the Department of Bioinformatics and Computational Biology at MD Anderson Cancer Center, "http://bioinformatics.mdanderson.org/OOMPA"). The protein concentrations of each set of slides were then normalized by median polish.

Overexpression constructs

We sequenced verified a *MACROD2* overexpression vector, pCMV6-*MACROD2*, purchased from Origene (TrueORF RC222689) and a second construct was generated by cloning the full length cDNA of *MACROD2* (NM_080676) into pCDNA3.1(-) (Invitrogen). Constructs were transfected into MCF-7, T47D and MCF-10A cell lines using Fugene 6 (Promega) and Opti-MEM (Invitrogen) as per the manufacturers' protocols. Two days post-transfection, transfected cells were diluted into 96-well plates in 1.5mg/ml G418 (Invitrogen) for MCF-7 and T47D cells and 120 μg/ml G418 for MCF-10A cells. Single cell clones were isolated, expanded, and tested for over-expression of the *MACROD2* via RT-PCR and western blot analysis.

Cell proliferation and growth assays

Cell proliferation was assessed and quantified through cell counting with a Vi-Cell XR cell counter (Beckman Coulter) and the use of the sulforhodamine B (SRB) assays according to established protocols (5). For SRB assays, cells were seeded at 1000 cells per well into 96-well plates using assay medium. This was considered the Day 0 time point. The following day, cells were treated with varying doses of drug (17-β-estradiol, 4-OH-tamoxifen or both) as indicated by the experiment. Assay medium was changed every other day and cell counts were collected at that time until the end of the experiment. Each experiment was performed with five biological replicates, and the experiments were repeated three times.

For cell growth assays with MCF-10A and derivative cell lines, exponentially growing cells were washed with HBSS twice and seeded in growth media as well as "assay media" consisting of DMEM/F12 medium, 1% charcoal dextran-treated fetal bovine serum, insulin at 10 μg/ml, hydrocortisone at 0.5 μg/ml, and cholera toxin at 0.1 μg/ml, without EGF. Cells were seeded at a density of $2x10^4$ cells/well of a 6-well tissue culture dish on day 0. Medium was changed every third day. Cells were harvested on days 0, 2, 4, and 6, and cell numbers were counted using a Beckman Coulter counter. All assays were performed in triplicate and repeated at least three times.

Antibodies and immunoblotting

Whole cell lysates for use in western blotting were generated by seeding cells in assay media containing ethanol, 17-β-estradiol, 4-OH-tamoxifen or a combination of drug as indicated. Cells were trypsinized after 48 hours of drug treatment and then were heat-denatured in Laemmli sample buffer. Western blotting was performed as previously described (6, 7). Briefly, prepared cell lysates were loaded equally into 4-12% Bis-Tris polyacrylamide gels (Invitrogen), resolved by SDS-PAGE, and then transferred to PVDF membranes (Invitrogen). Following transfer, membranes were blocked in 10% non-fat milk, incubated overnight with primary antibody at 4ºC, washed with TBST (1% Tween-20), and incubated with secondary antibodies. The primary antibodies used in this study include anti-*MACROD2* rabbit antibody (custom antibody 6711, Genscript), anti-ER alpha mouse antibody (2512; Cell Signaling), and anti-GAPDH mouse antibody (6C5) (ab8245; Abcam). All experiments were repeated at least 3 times with representative blots shown.

Mouse xenografts

Xenografts were performed using 10-week-old athymic nude female mice (Taconic), which were randomly distributed into equal groups (5 mice per group) for each experiment. Mice were injected subcutaneously in the right flank with $2x10^6$ cells of TamR1, TamR1-sh3, TamR1-sh5, TamR2, TamR2-sh3, TamR2-sh5, TamR3, TamR3-sh3 or TamR3-sh5. Cells were prepared in a mixture of 80% growth factor reduced Matrigel and 20% 1X PBS. Mice were not supplemented with estrogen pellets. Tumor volume (mean ± SD; mm^3) measurements were taken of the palpable tumors every week until week 5 when the mice were sacrificed. Tumors were then excised and fixed in 10% formalin and embedded in paraffin for further analysis by immunohistochemistry. The National Institutes of Health Guide for the Care and Use of Laboratory Animals was followed in all experiments.

Quantitative real time RT-PCR analysis

Samples used for total RNA extraction were seeded in assay media with indicated treatment. Total RNA was isolated through a previously described Trizol/Chloroform method (8) and purified using the RNeasy kit (Qiagen, Valencia, CA). cDNA was synthesized using a First Strand cDNA synthesis kit (Amersham Biosciences, UK) following the manufacturers protocol. The cDNA templates were tested in triplicate and included a no reverse transcriptase control. Quantitative real-time PCR was carried out using SYBR Green (Molecular probes) in conjunction with an iCycler (Bio-Rad, Hercules, California). Primers used for quantitative real time PCR are listed in Table S5.

RNA interference

Short-hairpin RNA constructs were designed to target the macro domain of *MACROD2* and were cloned into lentiviral vectors, and used to create stable shRNA expressing cell lines as previously described (9). Target sequences used for shRNA knock-down are listed in Table S5.

Chromatin immunoprecipitation

Chromatin immunopreciptation (ChIP) was conducted using the Millipore EZ ChIP kit according to the manufacturer's protocol (Millipore, Billerica, MA). Briefly, parental MCF-7 and TamR clones were grown for 48 hours to 80% confluency in assay media (phenol-red free DMEM containing 0.5% CD serum). Then ethanol, 125nM 17-β-estradiol, or 1μM 4-OH-tamoxifen was added to the cells for 45 minutes before beginning the ChIP protocol. The crosslinking step was accomplished by fixing cells for 10 minutes at room temperature in 18.5% formaldehyde, after which the media was removed and the cells were washed with ice-cold PBS containing a protease-inhibitor cocktail. The cells were then lysed, pelleted, and kept on ice in preparation for sonication, which sheared the chromatin into 200-1000 base pair fragments as determined by gel electrophoresis. The sheared and crosslinked protein-DNA complexes were then incubated overnight with 1 ug

of the following immunoprecipitating antibodies: anti-RNA polymerase (positive control), anti-Mouse IgG (negative control), anti-*MACROD2* rabbit antibody, anti-ER alpha mouse antibody (2512; Cell Signaling), anti-p300 rabbit antibody (N-15, Santa Cruz, sc-584) and anti-Rabbit IgG (negative control). Protein G Agarose beads were then utilized to isolate the protein/DNA complexes, after which the crosslinking was reversed through incubation with 5M NaCl at 65°C overnight. RNase A and Proteinase K were added to each sample to digest unwanted RNA and protein, respectively. The genomic DNA was then recovered using spin columns and further purified using ethanol precipitation. Quantitative real time PCR was performed using the purified genomic DNA fragments with control primers and promoter specific primers designed for the reported genes.

In silico data analysis

For replication of the association between *MACROD2* overexpression and overall survival, we obtained gene expression and copy number data from the METABRIC study. Briefly, we extracted normalized expression measures from two probes (manufacturer ids ILMN_1677485 ILMN_1725633) that interrogate *MACROD2* on the Illumina HT 12 array platform and copy number estimates (log ratio scale) for the *MACROD2* gene from the Affymetrix 6.0 genotyping array. For the expression measures, we computed a z-score by subtracting the median expression across all samples (log scale) and dividing by the across-subject standard deviation. An indicator for overexpression was defined independently for each probe using a z-score cutoff at which the empirical quantiles were more extreme than the theoretical standard normal quantiles under the null of no over-expression. Empirical quantiles of probe ILMN_1725633 were approximately normal and not further evaluated. Copy number amplification was assessed using a cutoff of 2 median absolute deviations from the modal copy number (approximately 13 percent of participants have estimates exceeding this threshold). Differences in the Kaplan-Meier survival curves among participants with and without the genetic marker for overexpression/amplification were assessed by a log rank test.

Statistical analysis

All statistical analyses were performed using GraphPad InStat software unless otherwise indicated (La Jolla, CA). A *P* value of less than 0.05 was considered significant.

Supplemental References

1. Zhao X*, et al.* (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 64(9):3060-3071.
2. Cimino-Mathews A*, et al.* (2012) Androgen receptor expression is usually maintained in initial surgically resected breast cancer metastases but is often lost in end-stage metastases found at autopsy. *Human pathology* 43(7):1003-1011.
3. Abukhdeir AM*, et al.* (2008) Tamoxifen-stimulated growth of breast cancer due to p21 loss. *Proc Natl Acad Sci U S A* 105(1):288-293.
4. Soule HD*, et al.* (1990) Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res* 50(18):6075-6086.
5. Vichai V & Kirtikara K (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc* 1(3):1112-1116.
6. Higgins MJ*, et al.* (PIK3CA mutations and EGFR overexpression predict for lithium sensitivity in human breast epithelial cells. *Cancer Biol Ther* 11(3):358-367.
7. Konishi H*, et al.* (2007) Knock-in of mutant K-ras in nontumorigenic human epithelial cells as a new model for studying K-ras mediated transformation. *Cancer Res* 67(18):8460-8467.
8. Rio DC, Ares M, Jr., Hannon GJ, & Nilsen TW (Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb Protoc* 2010(6):pdb prot5439.
9. Lauring J*, et al.* (2008) The multiple myeloma associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood* 111(2):856-864.

Supplemental Figures and Legends

Figure S1

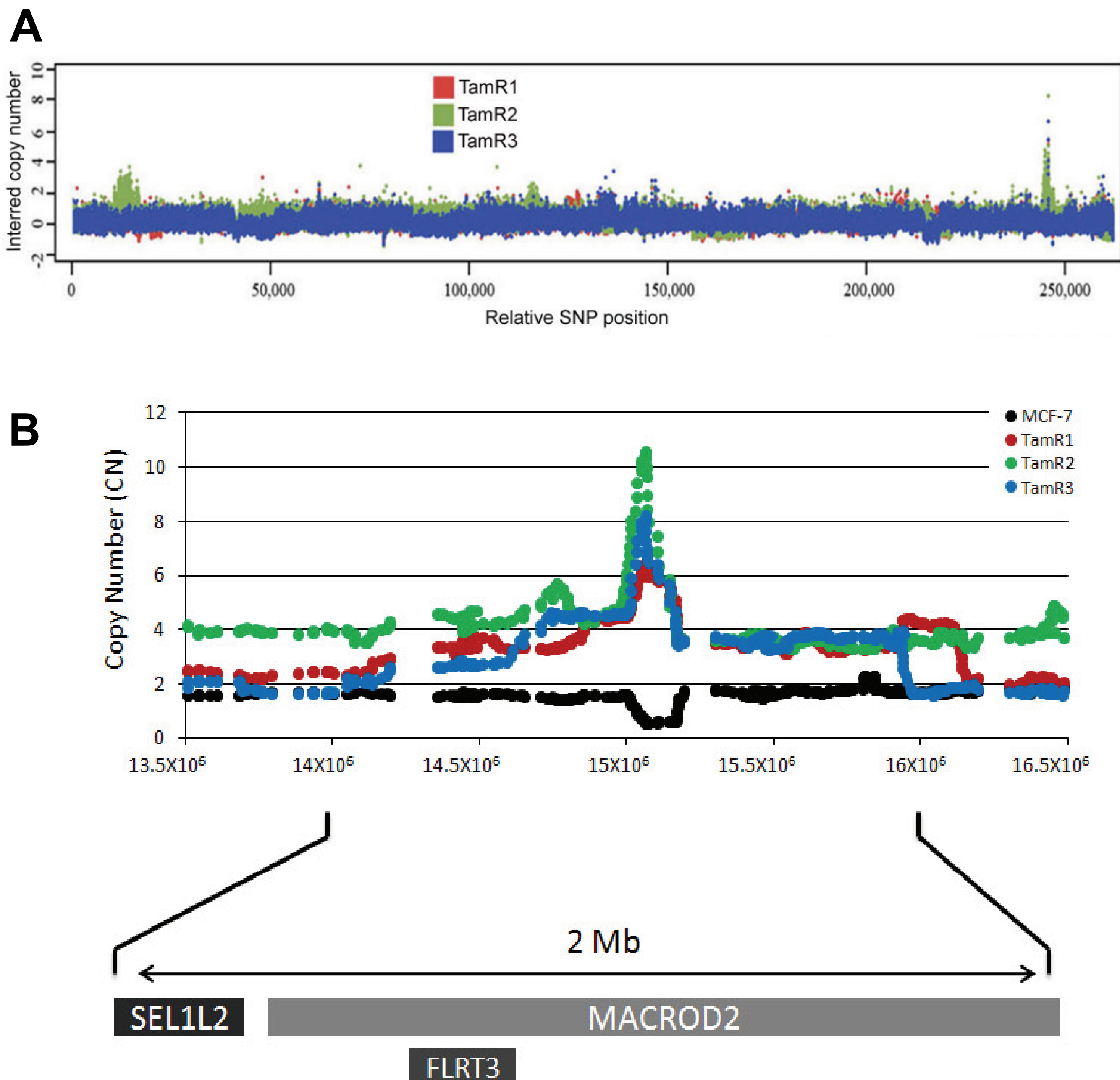


Figure S1. Increased copy number of *MACROD2* in tamoxifen-resistant MCF-7 cell lines.
A) SNP genotyping was performed using the Affymetrix 250K NspI array. Copy number analysis was performed with the dChip program. The three tamoxifen-resistant clones (TamR) were normalized to the parental MCF-7 cell line, and relative SNP position and normalized inferred copy number values are plotted. **B)** A focal amplification was found on Chr20p12.1 with average inferred copy number ranging from 4-12 copies in the TamR clones. This 2 Mb region of shared copy number increase contains three genes: *MACROD2, SEL1L2*, and *FLRT3*.

Figure S2

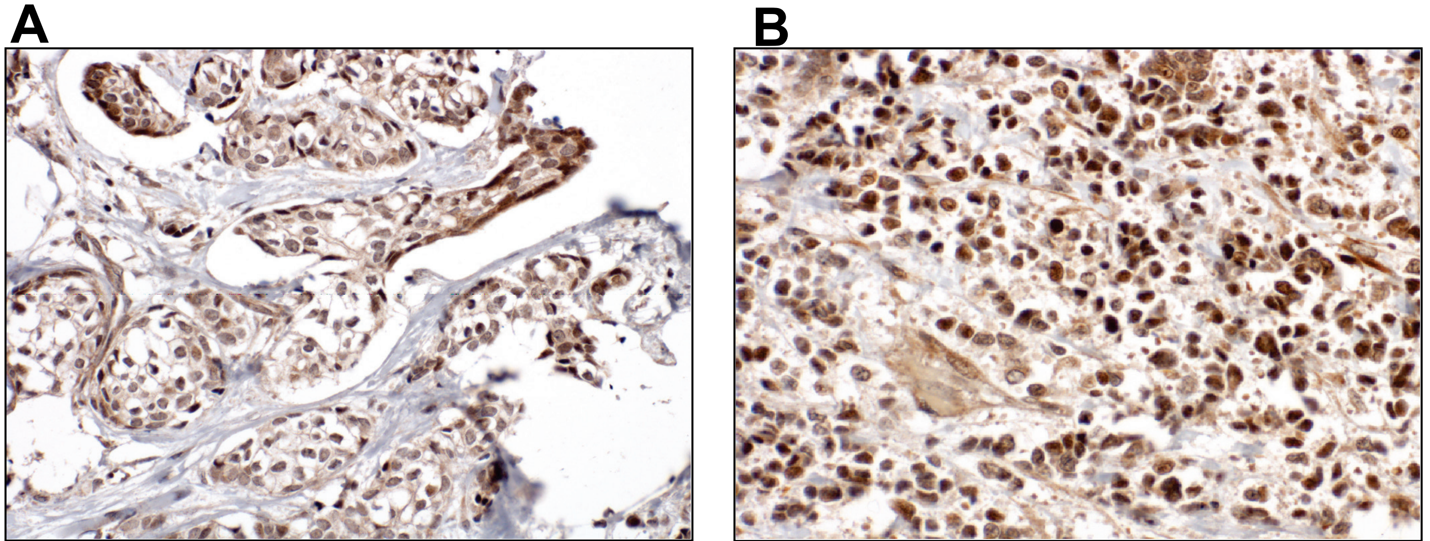


Figure S2. MACROD2 is expressed in primary and metastatic breast cancers. FFPE samples were used for MACROD2 immunohistochemical labeling as described in the text. Shown is patient 5's **A)** primary breast cancer scored as positive for nuclear MACROD2 labeling and **B)** a metastatic lesion in a fallopian tube from patient 5, scored as positive for nuclear MACROD2 labeling. Magnification 400x.

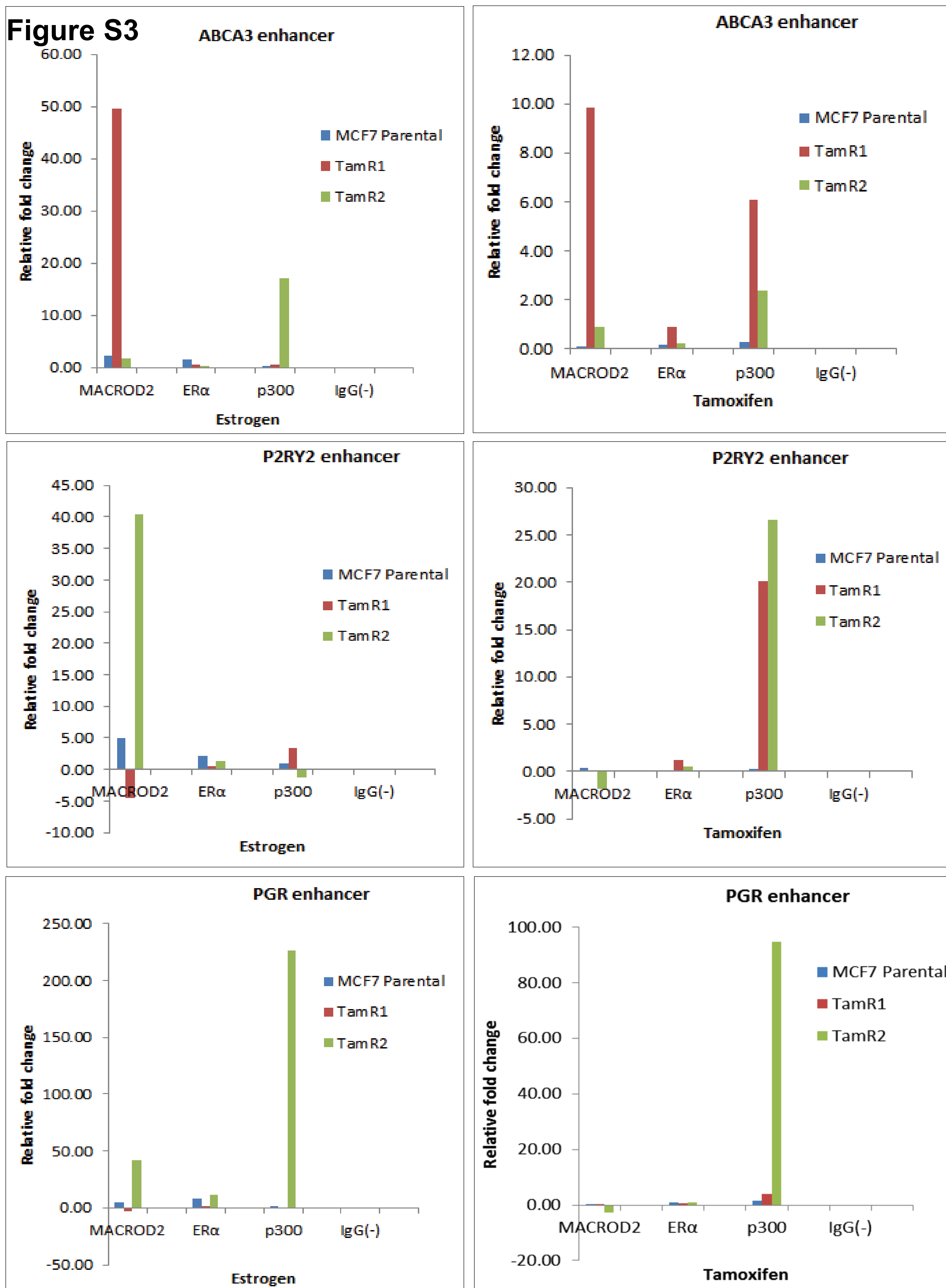
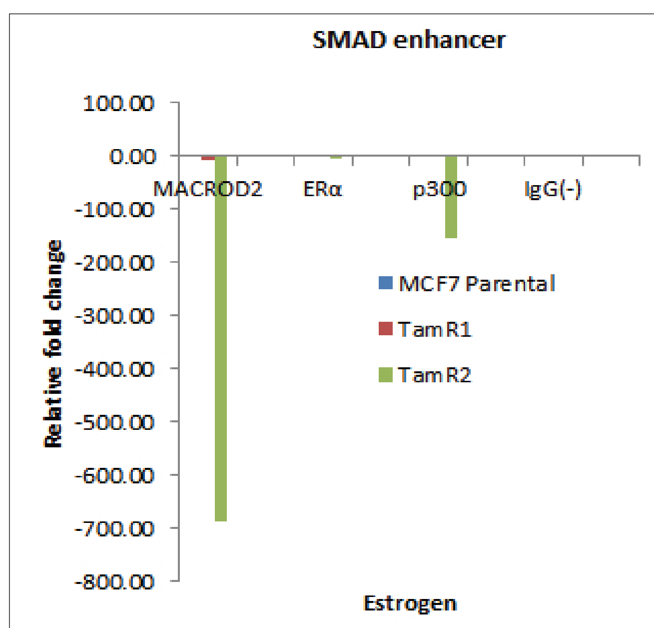
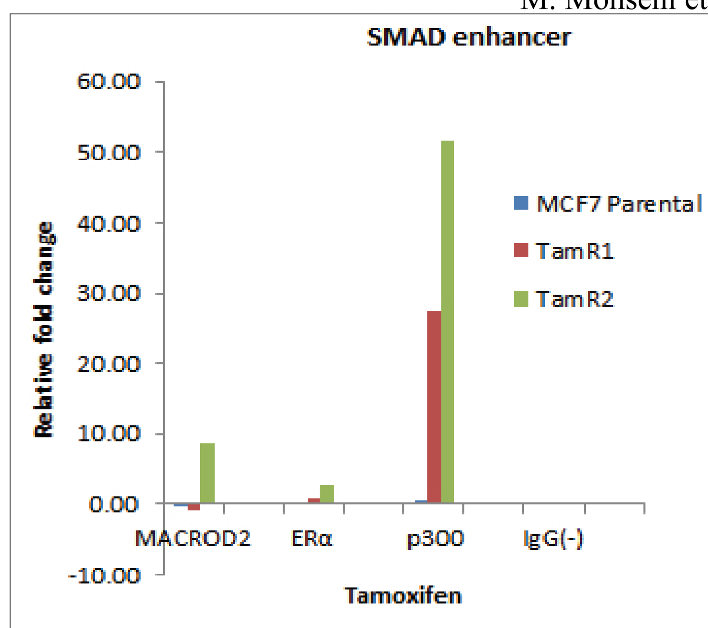
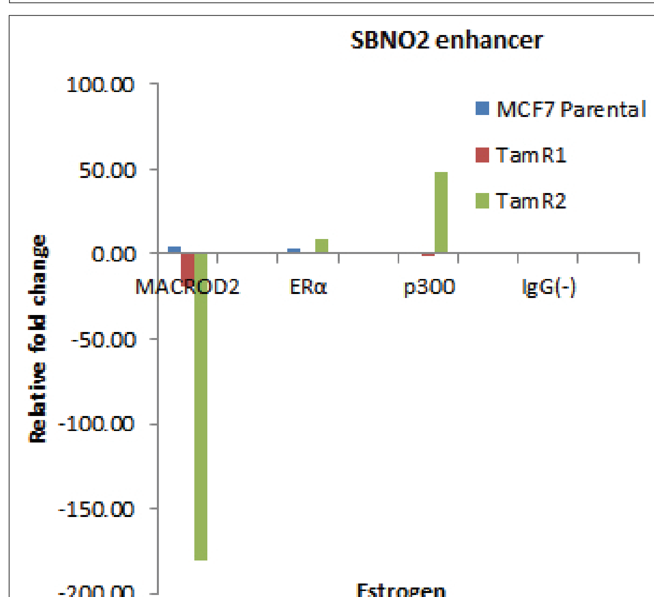
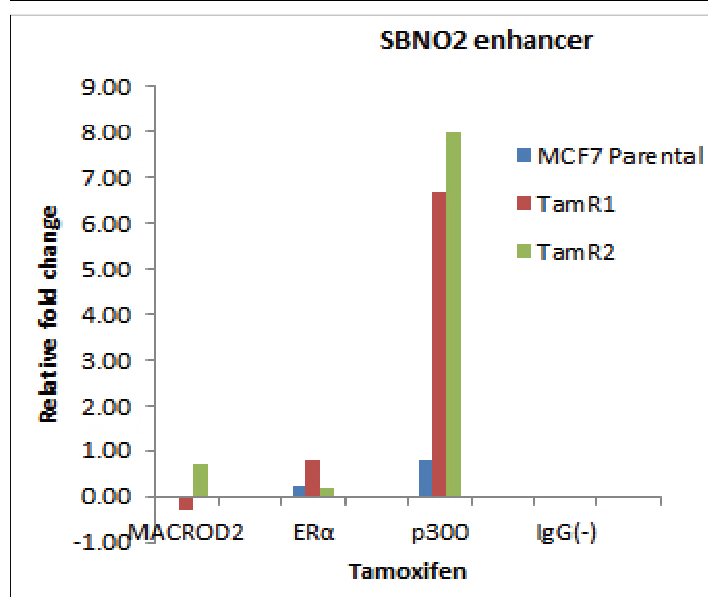


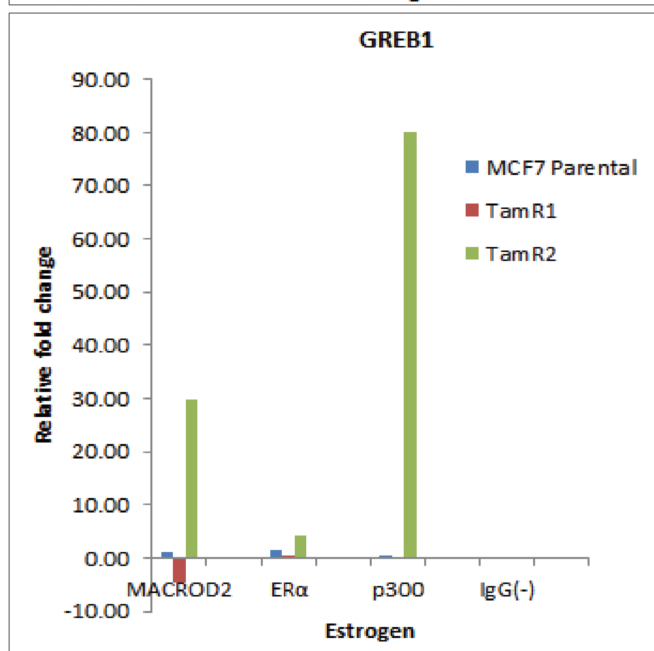
Figure S3
ABCA3 enhancer
Relative fold change
MCF7 Parental
TamR1
TamR2
MACROD2
ERα
p300
IgG(-)
Estrogen
ABCA3 enhancer
Tamoxifen
P2RY2 enhancer
Estrogen
P2RY2 enhancer
Tamoxifen
PGR enhancer
Estrogen
PGR enhancer
Tamoxifen

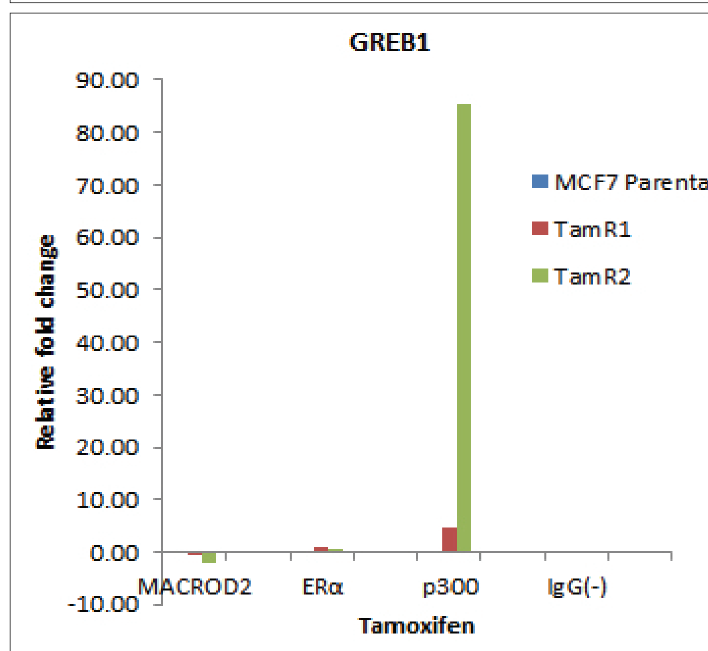

SMAD enhancer
100.00
0.00
-100.00
-200.00
-300.00
-400.00
-500.00
-600.00
-700.00
-800.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Estrogen


SMAD enhancer
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Tamoxifen


SBNO2 enhancer
100.00
50.00
0.00
-50.00
-100.00
-150.00
-200.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Estrogen


SBNO2 enhancer
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Tamoxifen


GREB1
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Estrogen


GREB1
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
Relative fold change
MACROD2
ERα
p300
IgG(-)
MCF7 Parental
TamR1
TamR2
Tamoxifen

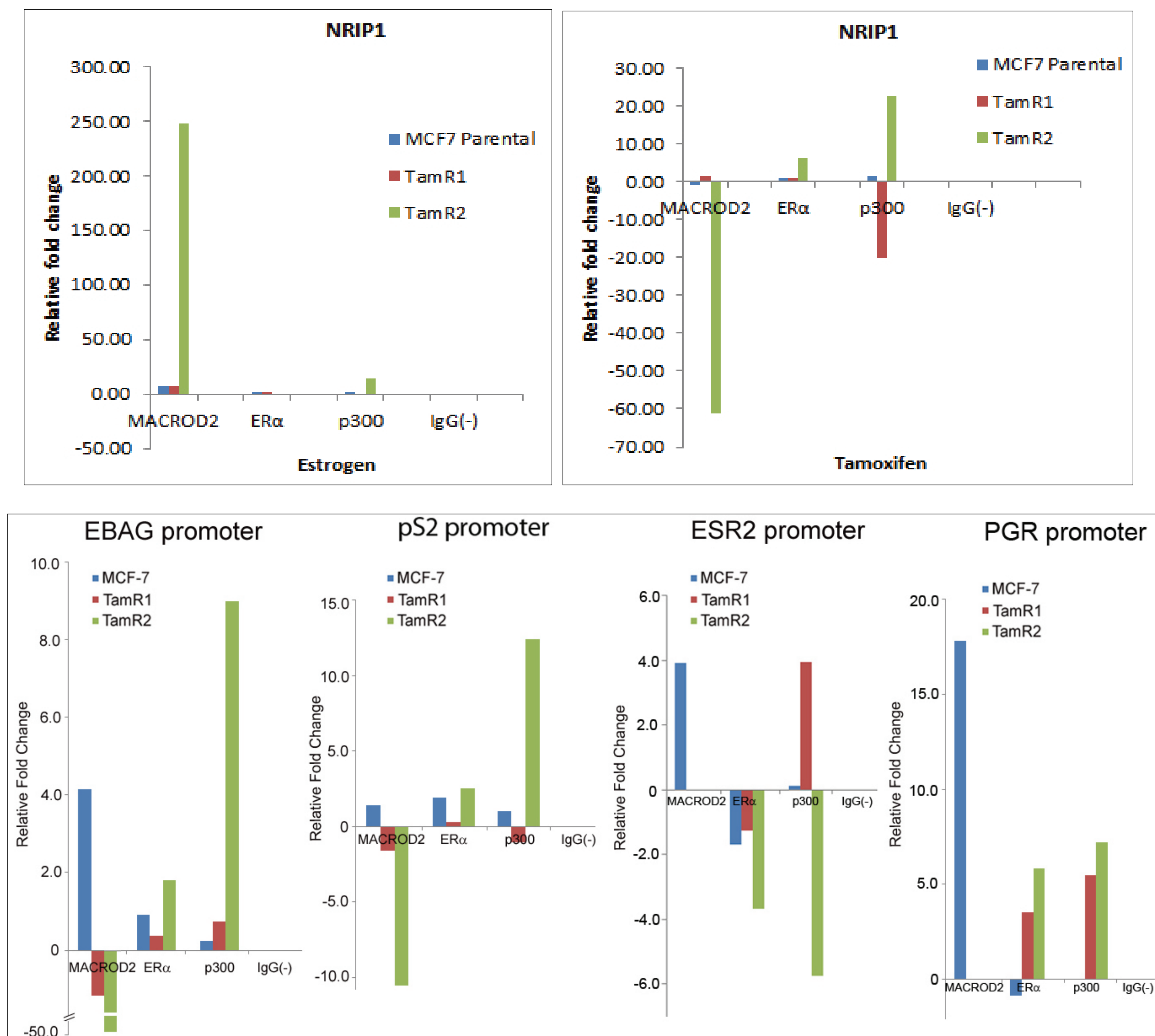
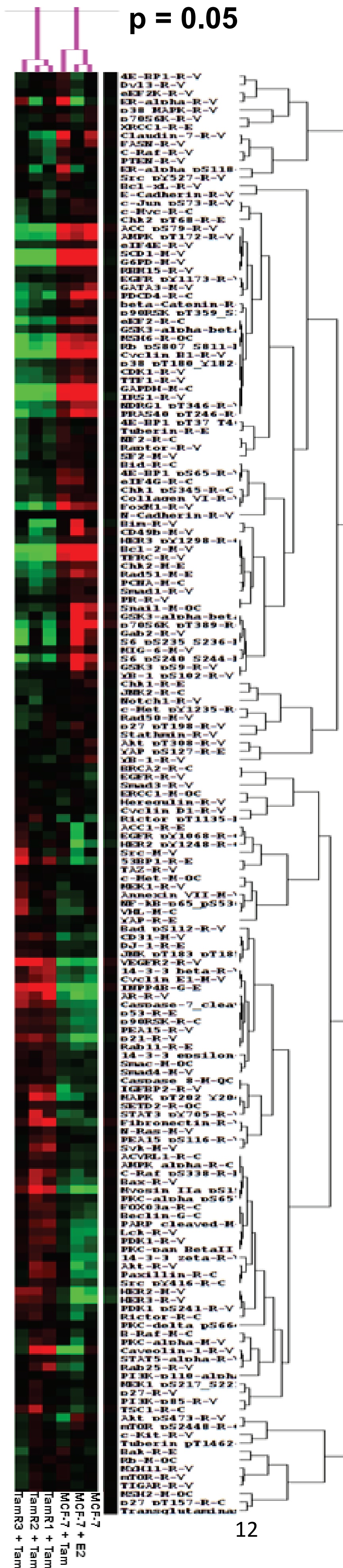


Figure S3. MACROD2 increases p300 binding to cis-regulatory elements in ER responsive genes. ChIP was performed as described in Methods and the text using antibodies against MACROD2, ER, p300 and an irrelevant antibody negative control (IgG(-)). Results are representative of duplicate samples comparing quantitative real time PCR results after controlling for input DNA and then analyzing relative fold differences between drug (estrogen or tamoxifen) and vehicle control for genomic enhancer regions. Bottom panel represents experiments performed for Fig.5B in the text, but comparing estrogen to vehicle control for promoter regions of the genes shown. Primers used for quantitative PCR are shown in Table S5.

Figure S4

A



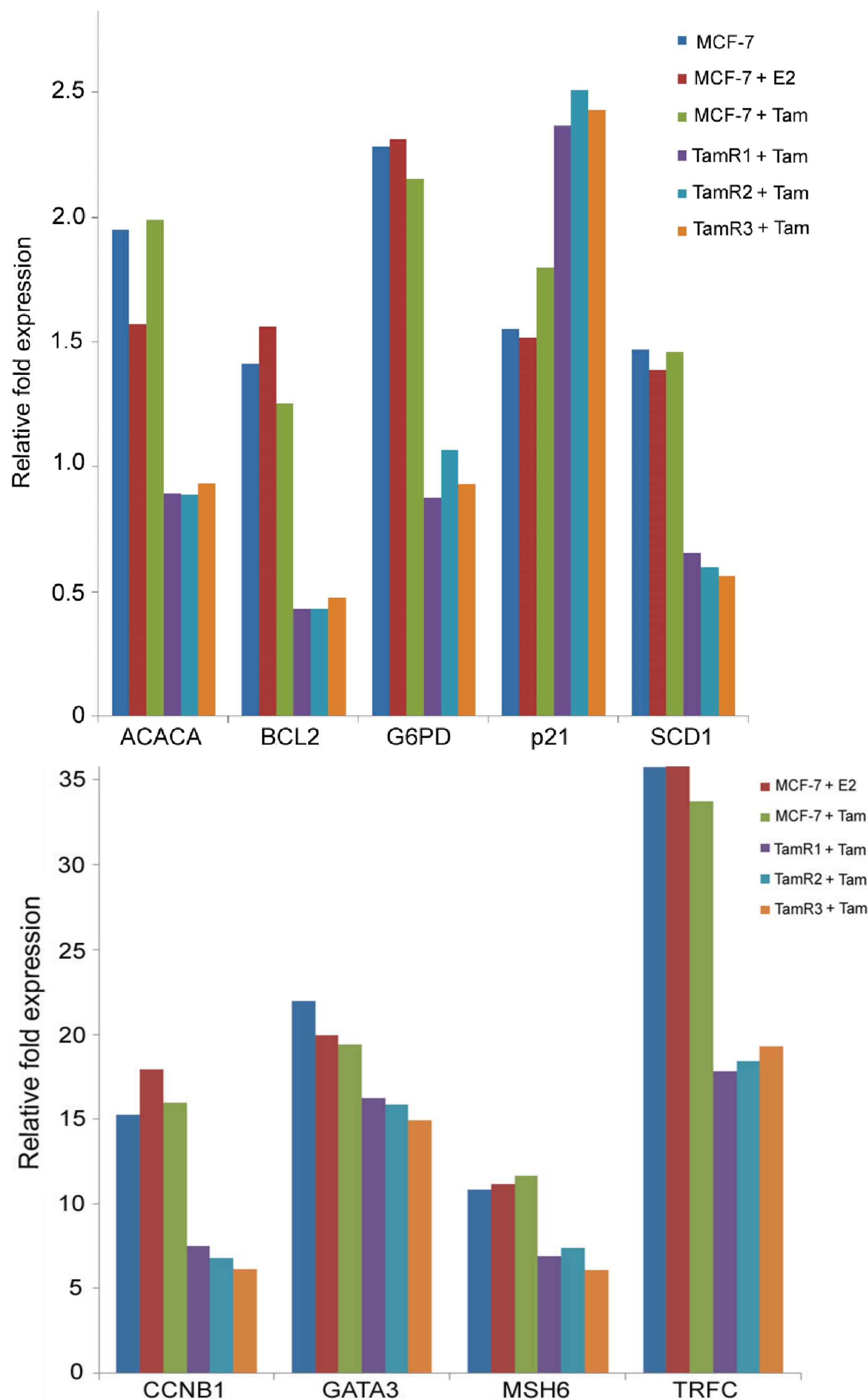


Figure S4. Differential protein expression between MCF-7 and TamR clones. A) Reverse phase protein array (RPPA) shows different global expression pattern for TamR clones compared to parental MCF-7 cells. Parental MCF-7 cells were grown in either assay media (phenol red-free media with 0.5% CD serum) and co-cultured with vehicle (ethanol), estrogen, or tamoxifen as described in the text. The three TamR clones were grown in assay media supplemented with tamoxifen. Whole cell lysates were prepared and analyzed as per the RPPA Core Facility at MD Anderson Cancer Center. **B)** Examples of differentially expressed genes: ACACA: acetyl-cCoA carboxylase alpha, BCL2: B-Cell CLL/Lymphoma 2, G6PD: glucose-6-phosphate dehydrogenase, p21: cyclin-dependent kinase inhibitor 1A, SCD1: stearoyl-CoA desaturase 1, CCNB1: cyclin B1, GATA3: GATA binding protein 3, MSH6: mutS homolog 6, TRFC: transferrin receptor.

Figure S5

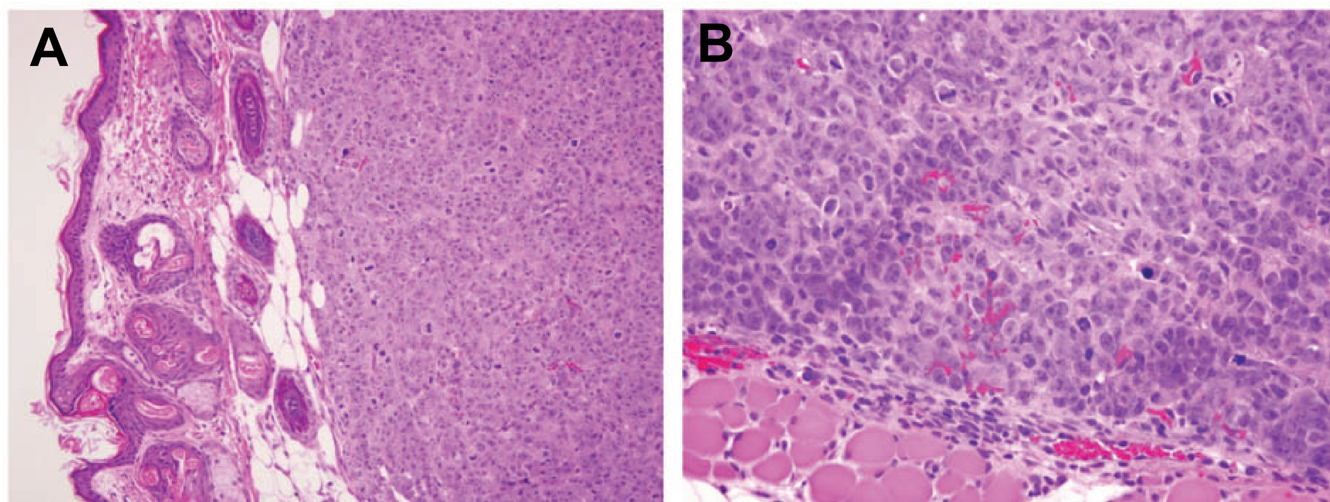


Figure S5. Hematoxylin and eosin (H&E) staining of TamR xenografts. Tumors from TamR cell lines grown as xenografts in athymic female nude mice were dissected after 42 days post inoculation, formalin fixed and paraffin embedded and then stained using H&E. **A)** 20X magnification and **B)** 40X magnification demonstrating loss of glandular structures and increased number of mitotic figures.

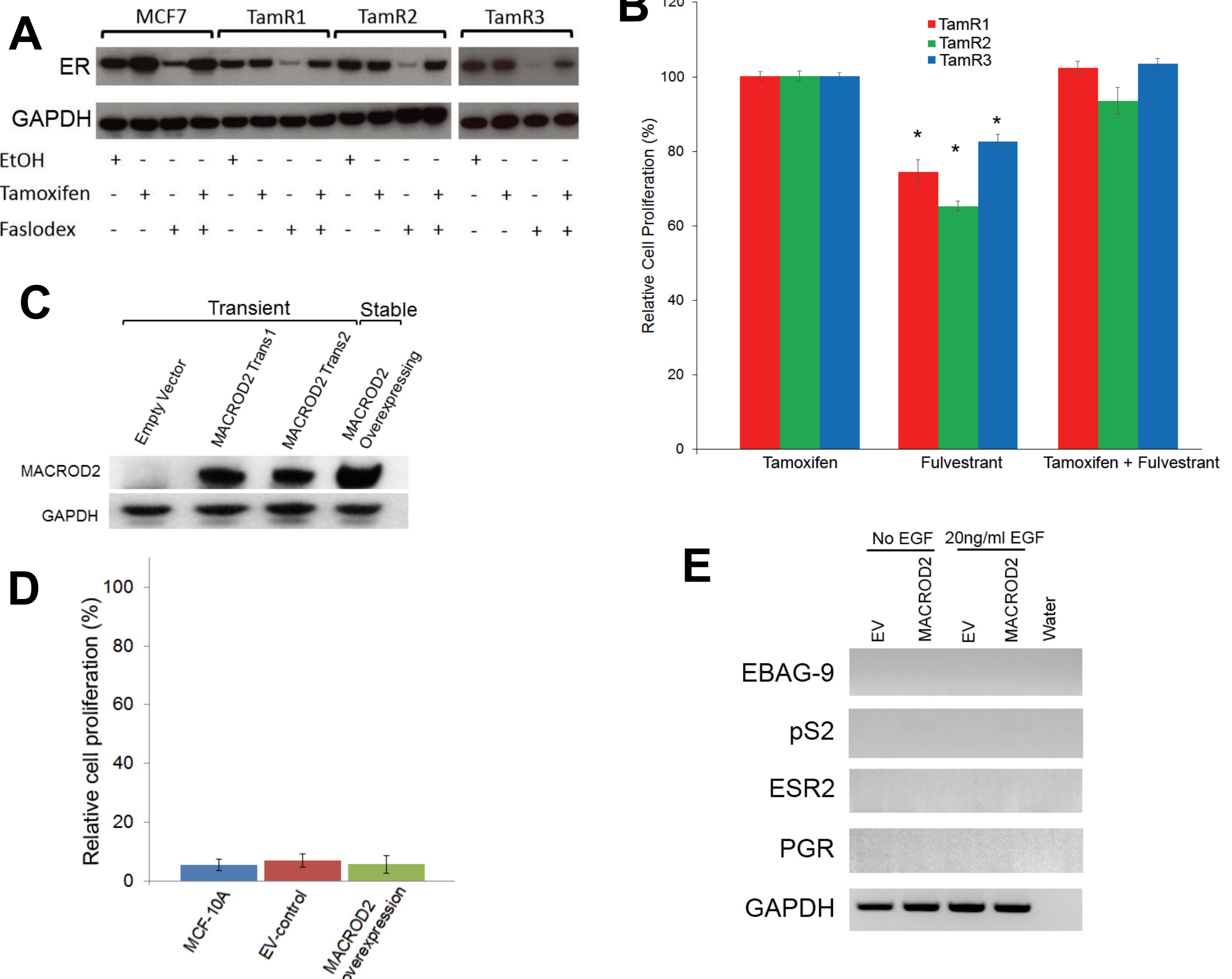


Figure S6. MACROD2 is dependent on ER for mediating cell proliferation. **A)** MCF-7 and TamR clones were exposed to vehicle control (etoh), fulvestrant (100 nM) and/or tamoxifen (1µM) for 48 hours, and lysates harvested and used for western blot with an anti-ER antibody. GAPDH is shown as a loading control. **B)** TamR clones were exposed to tamoxifen, fulvestrant or both and after 6 days in culture cells were counted and results plotted as a percentage relative to growth in tamoxifen. Results are averaged from triplicate samples with standard error shown, and are representative of four independent experiments. *$p<0.05$ relative to tamoxifen controls. **C)** ER/PR/HER2 negative non-cancerous human MCF-10A cells were transfected with an empty vector control as well as a MACROD2 overexpression plasmid vector, transiently and to generate stable clones. Shown is a western blot demonstrating MACROD2 overexpression in MCF-10A cells. GAPDH is shown as a loading control. **D)** Parental MCF-10A, empty vector (EV) control and a stable MACROD2 overexpression clone were seeded at equal density and grown in growth media with 20ng/ml EGF and assay media with no EGF to assess growth factor independent cell proliferation. The assay was carried out for 6 days and cells harvested and counted as described in the text. Results are shown as the average of three replicates with standard error, and are representative of three independent experiments. **E)** MCF-10A cells stably transfected with empty vector (EV) or overexpressing MACROD2 (MACROD2) were seeded in assay media (no EGF) or growth media (20ng/ml EGF) for 48 hours and cells harvested for RNA used for RT-PCR as described in the text. Results are representative of three independent experiments. GAPDH is shown as a loading control.

Figure S7

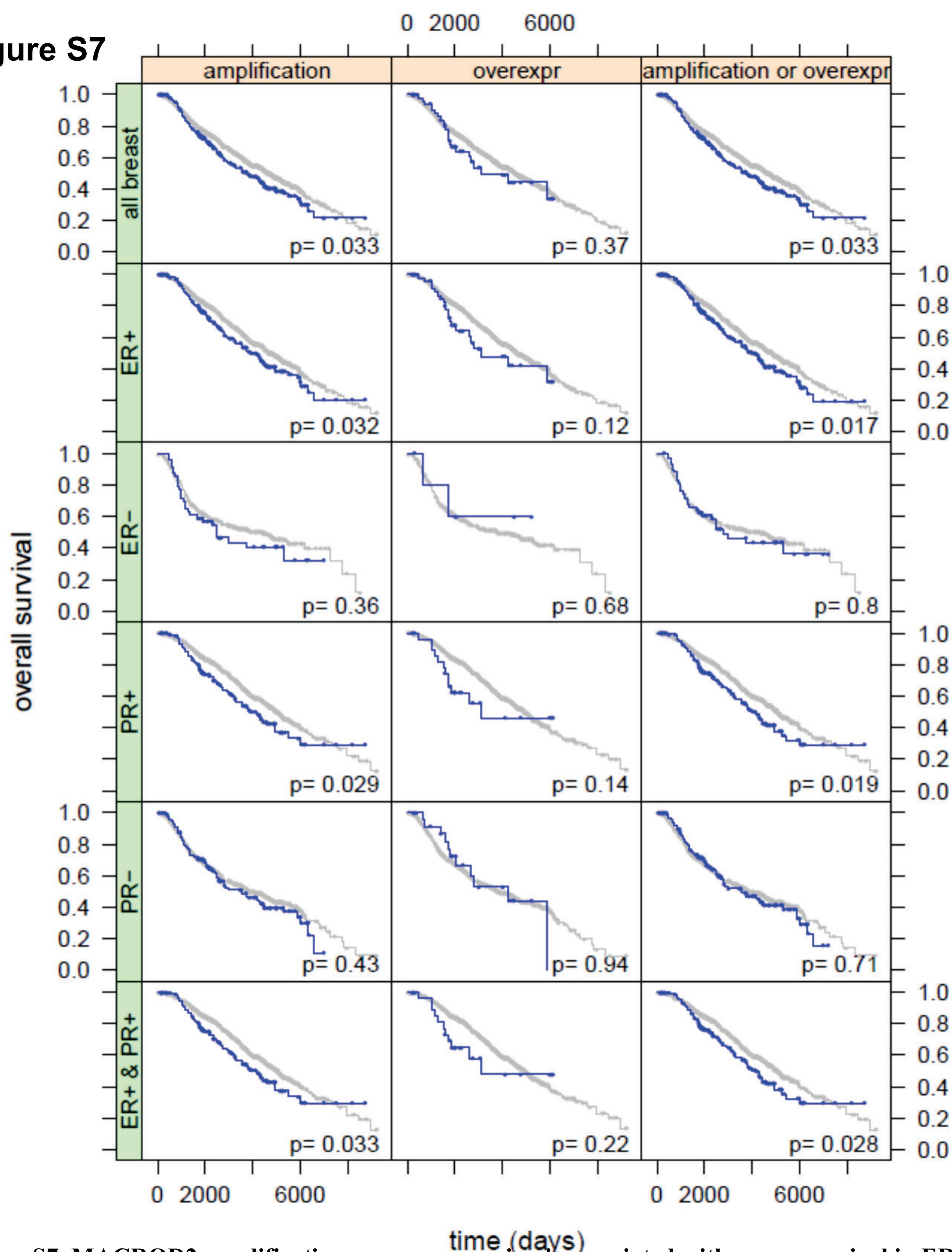


Figure S7. MACROD2 amplification or overexpression is associated with worse survival in ER positive breast cancer. Results are shown for all breast cancer (all breast) and presence or absence of estrogen receptor-α (ER) and/or progesterone receptor (PR). *MACROD2* amplification, overexpression (overexpr) singly or combined were assessed as per Methods. Kaplan Meier survival curves for subjects with (blue) and without (gray) the genomic marker for MACROD2 upregulation (columns) within strata of known clinical factors affecting overall survival (rows). P-values from a log rank test comparing the survival curves were computed using R version 3.1 (http://www.R-project.org/) and are indicated in each panel.

Figure S8

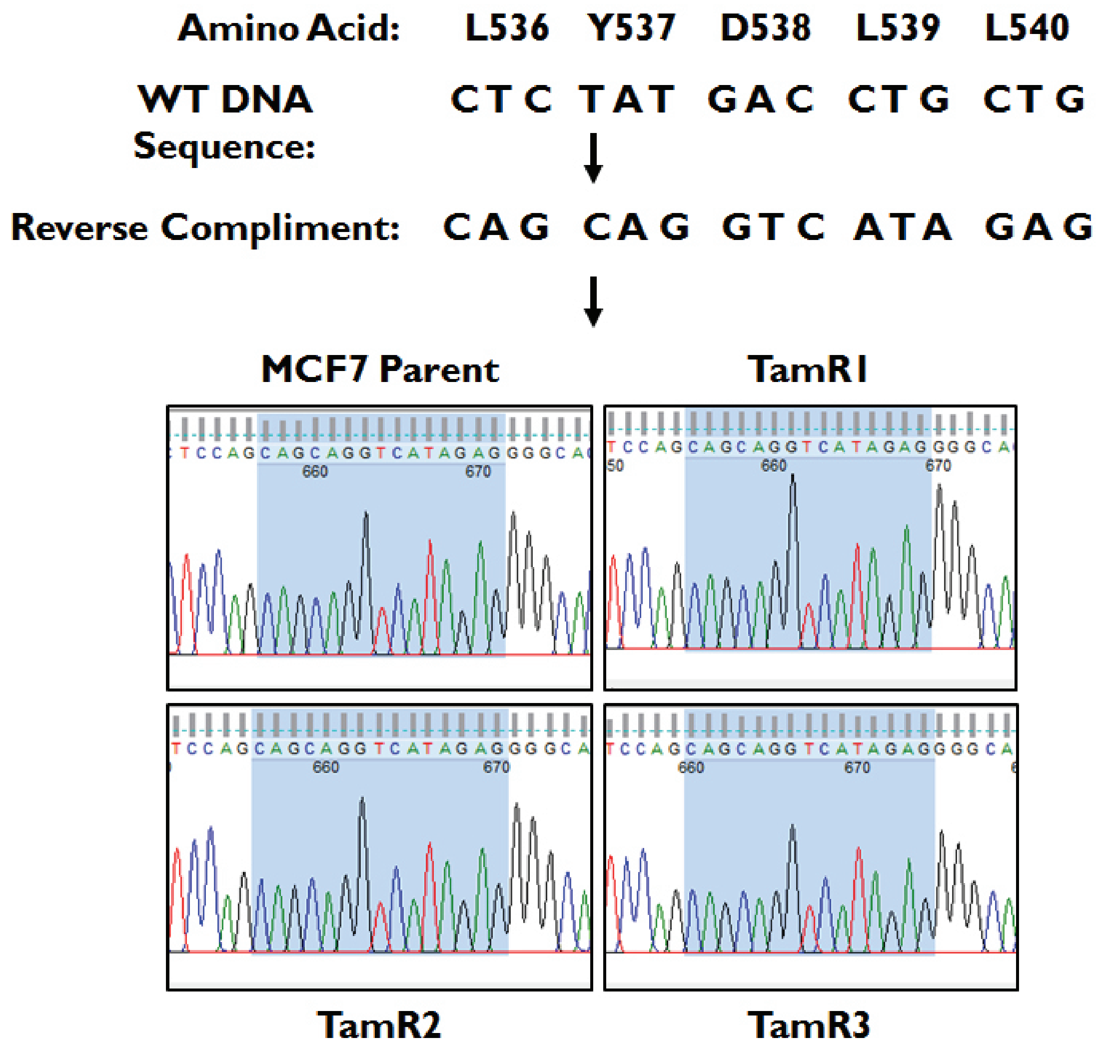


Figure S8: TamR clones do not have ESR1 LBD mutations. The complete ligand binding domain of TamR clones was PCR amplified and Sanger sequenced to verify the absence of ESR1 mutations. Shown are sequencing traces and the corresponding sequences of the Y537 and D538 hotspot codons.

Supplementary Tables

Table S1. Shared copy number alterations in tamoxifen resistant cell lines

Chromosome	Size of CNV	Locus	Loss/ Gain	Avg CN (Diploid =2)	Gene (s)
2	19 Kb	2q36.3	Loss	1.3	*COL4A4*
	757 Kb	2q37.1	Loss	1.3	*GIGYF2, NGEF, INPP5D, USP40*
	2.3 Mb	2q37.1	Loss	1.4	*UGT1A8, AGAP1*
	282 Kb	2q37.3	Loss	1.2	*HDAC4*
	43 Kb	2q37.3	Loss	1.2	*NDUFA10*
	9 Kb	2q37.3	Loss	1.3	*ANKMY1*
	98 Kb	2q37.3	Loss	1.2	*FARP2*
4	27 Kb	4q34.3	Loss	0.5	*NCRNA00290*
	25 Kb	4q35.1	Loss	1.3	*ODZ3*
9	49 Kb	9p24.1 - 9p23	Loss	1.3	*PTPRD*
10	3 Kb	10q23.2	Loss	1.3	*WAPAL*
	73 Kb	10q25.2 - 10q25.3	Loss	1.2	*TCF7L2*
11	12 Kb	11p14.2	Gain	3.3	*ANO3*
16	993 bp	16p13.3	Loss	1.5	*LMF1*
	10 Kb	16q23.1 - 16q23.2	Gain	3.2	*WWOX*
20	2 Mb	20p12.1	Gain	5.9	*MACROD2, SEL1L2, FLRT3*

Table S2. Summary of MACROD2 labeling in patient samples

Case	Type	Grade	PBC Nuclear MACROD2 (%, intensity)	MBC Nuclear MACROD2 (%, intensity)	
Patient 3	Luminal	IDC, Grade 2	0	LIVER	75 M
				LIVER	0
				PERITONEUM x2	75 M
				OMENTUM x2	75 M
				LUNG	100 MS
				SOFT TISSUE	30 S
				OVARY	25 WM
				BRAIN	25 M
				LYMPH NODE	25 M
Patient 5	Luminal	IDC, Grade 2	75 WM	LIVER	50 WM
				LIVER x4	75 M
				PERICARDIUM	100M/ 25S
				LUNG SEROSA	100W/25M
				LUNG	60 M
				DURA	75 M/25S
				PANCREAS	50 M
				FALLOPIAN	100WM/30S
				ADRENAL	75 M
				BRAIN	30 M/100W
Patient 6	Luminal	IDC, Grade 2	0-25M	LIVER x3	0
				PLEURA	75 M
				LUNG	75 S
				COLON	30 M
				EPICARDIUM	75M/10S
				LYMPH NODE	25 WM
Patient 8	Luminal	IDC, Grade 3	0	LYMPH NODE	100 M
				LIVER x4	25 M
				LUNG	50 WM
Patient 10	Luminal	IDC, Grade 3	0-25M	LUNG x4	10 M
				LUNG x2	0
				LUNG	30 M
				LYMPH NODE x3	0

Abbreviations: PBC, primary breast carcinoma; MBC, metastatic breast carcinoma; IDC, invasive ductal carcinoma; W, weak; M, moderate; S, strong.

60% (3/5) of PBC had positive nuclear MACROD2 labeling, and 100% (5/5) matched MBC had positive nuclear MACROD2 labeling. Two cases (Patient 3 and 8) had negative PBC and positive MBC. Two cases (Patients 6 and 10) showed heterogeneous labeling in the PBC and MBC, with some sites showing negative MACROD2 labeling and some sites showing positive labeling.

Table S3. MCF-7 cells do not form tumors without estrogen in nude mice.
Mice were inoculated as described in methods and evaluated weekly x 5 weeks.

Experiment	Number of animals inoculated	Tumor Formation
1	5	0
2	5	0

Table S4. Clinical pathological characteristics of patient samples

Patient ID	Primary			Tam Tx
	ER	PR	HER2	
3	+	+	+	Mets developed at end of 5 yr course
5	+	-	N/A	Mets while on Tam adjuvant Rx
6	+	+	N/A	Recurred on Tam
8	+	+	+	Developed Tam resistant disease
10	+	+	-	Tam for 1 year, developed Mets in bone

All patients had invasive ductal carcinoma. Two cases were Elston grade III of III, and the remaining 3 were grade II. The cohort was composed of three luminal cases (ER+ in both the PBC and MBC) and two luminal-loss cases (ER+) in the PBC with loss of ER or PR expression in the MBC). The metastatic sites at autopsy were multiple, with the number of sites per patient ranging from 6-15. All patients were refractory to multiple rounds of chemotherapy and hormonal therapy at the time of autopsy.

Table S5. Primers used in this study

qPCR Primers

Genomic Region	*Forward (5' - 3')*	*Reverse (5' - 3')*
Invariant Chr20	TTTCCATCCTTTCCCAAGC	AGAAGTGTCTGGGCCTCTCA
Chr20p12.1 Set 1	CATTCAGTTCTTCCTTGGCAAC	TGATGTGAGATAGAGGGTCCAA
Chr20p12.1 Set 2	AGAAAGGCATGATATCCCCATA	TCTACATGTAATACCTGGAACCTCA

qRT-PCR Primers

Gene	*Forward (5' - 3')*	*Reverse (5' - 3')*
MACROD2	TTGGCTCTGCTCTTCCATTT	GCCAAGAATCACCATGAGGT
SEL1L2	TGAGACTAACGGAAAGACCTGA	CATCTCCTCTTTCTGCCAAA
FLRT3	TGCTGGGATTCCTTCAGATT	TTTGGGAGGTTGGTAGGAAA
ACTB	GTCAGAAGGATTCCTATGTG	GCCTGGATAGCAACGTACATG

shRNA Sequences

shRNA	*Target Sequence (5' - 3')*
Control shRNA	ATGCTATAGTCAATGCCGCAAAT
shRNA-3	GAGGTAGATGCTATAGTCAATGC
shRNA-5	TGCTATAGTCAATGCCGCAAATG

ChIP Primers

Gene	*Forward (5' - 3')*	*Reverse (5' - 3')*
EBAG9-promoter	CACCTTCCTTTCGCCCTT	ATTGTGGAAAAATCGGCCTG
pS2-promoter	GGCCATCTCTCACTATGAATCACT	GGATTTGCTGATAGACAGAGACGA
ESR2-promoter	TCACTGAGCTGGTGTGAGGA	CTGGAAATGGAAACCGTCAT
PGR-promoter	GGCGACACAGCAGTGGGGAT	TCTCCTCCCTCTGCCCCTATATTC
PGR-enhancer	TTCCCAGAGGTTTTCACAGA	TTACACAGGCAGGACGACTT
SMAD-enhancer	TGGGTCCAAGGACAGATGTA	ACTCTCTGCATTGGTGAAGC
SBNO2-enhancer	GGAGGATAAACAGGGAGAA	TCCCAGTCCATCTATCCTCA
P2RY2-enhancer	CCATCAAAGCTGTTGCTTCT	CCAGGATAGTGCCAGTGAAC
ABCA3-enhancer	CACCTTCCATCTGTCCAAAG	CAACCCTGAGGTTTGGGAAC
GREB1	TAGGCTTCAAGAGGACCACA	AGCAGCAAAACTGCATAGGA
NRIP1	TCCCCTTCACCCCACAACAC	TCCCCTTCACCCCACAACAC