Transcriptional Regulation by Small RNAs at Sequences Downstream from 3' Gene Termini

Xuan Yue, Jacob C. Schwartz, Yongjun Chu, Scott T. Younger, Keith T. Gagnon, Sayda Elbashir, Bethany A. Janowski, and David R. Corey

Table of Contents 1

Supplementary Methods

Treatment of cultured cells with physiologic stimuli. T47D, MCF7, and MDA-MB-2231 cells were cultured at 5 % CO_2 in RPMI media (Sigma) supplemented with 10 % FBS (Atlanta Biologicals), 10 mM HEPES, 0.5 % NEAA, 10 μg/mL insulin, and 1 mM sodium pyruvate (Sigma). For experiments evaluating the effects of physiologic stimuli, T47D or MCF7 cells were cultured in media containing 5 % dextran/charcoal-stripped FBS (serum-stripped media) (Atlanta Biologicals) the day after transfection (day-1). The next day (day-2) the cells were either switched to full media or maintained in stripped media supplemented with 17β-estradiol (Steraloids), EGF (Sigma), or IL1β (Sigma). Cells were harvested for RNA on day-3.

Cellular delivery of duplex RNAs and expression assays. We used RNAi-max (Invitrogen) to deliver duplex RNAs (**Supplementary Table S1**) into MCF7 or T47D cells. Cells were seeded at 150,000 cells per well in a six well plates and were transfected with duplex RNAs two days later. Cells were harvested for RNA three days after transfection for RNA measurements and five days after transfection when measuring protein levels.

Total RNA was isolated using TRI reagent (Sigma). For purification of poly(A) RNA total RNA was isolated from T47D or MCF7 cell pellets with GenElute™ Mammalian Total RNA Miniprep Kit (Sigma), and poly(A) RNA was purified from total RNA with Poly(A) Purist™ mRNA Purification Kit (Applied Biostystems). cDNA was synthesized using Applied Biosystem's High Capacity Reverse Transcription kit. qPCR was performed on a CFX-96 Real Time System (BioRad) using iTaq SYBR Supermix (BioRad) or the TaqMan Universal PCR Master Mix (Applied Biosystems). Primers were designed using primer3 software with the exception of

primers for GAPDH, PR, and BRCA1 mRNA which were supplied by Applied Biosystems. Only those primer sets that show linear amplification over several orders of magnitude were used. RNA was treated with DNaseI (Worthington) prior to cDNA synthesis. cDNA was synthesized using Applied Biosystem's High Capacity Reverse Transcription kit.

For analysis of protein levels, cell pellets were lysed and protein concentrations were quantified using the BCA assay (Pierce). Westerns were performed on protein lysates (30 µg per well for PR and 60 µg per well for BRCA1). Primary antibodies (Ab) included PR-Ab (6A1, Cell Signaling) and BRCA1-Ab (MS110, Calbiochem). β-actin-Ab (Sigma) was used as an internal control and for quantitation. Protein was visualized using HRP-tagged secondary anti-mouse antibody (Jackson Immunolabs) and Super Signal developing solution (Pierce).

Northern blot analysis. All reagents were purchased from Applied Biosystems unless otherwise noted. Poly(A) RNA (5 µg per lane) was denatured with formaldehyde and separated on 1% agarose gels containing formaldehyde. The gel was soaked in 0.05 M NaOH/1.5 M NaCl for 30 min with gentle shaking, rinsed with several changes of DEPC-treated water and soaked in 10XSSC buffer (1.5 M NaCl and 0.15 M sodium citrate) for 40 min. The separated RNA species in the gel were transferred to BrightStar®-Plus positively charged nylon membrane by vacuum transfer apparatus in 10 X SSC buffer for 1.5 h and fixed to the membrane by cross-linking with UV light. After that, the blot was prehybridized in ULTRAhyb™ ultrasensitive hybridization buffer for at least 30 min and then hybridized overnight in ULTRAhyb™ buffer containing 0.1 nM psoralen-biotin labeled RNA probe at 68°C. Stringent washings were performed at 68°C with NorthernMax low/high stringency wash buffers. RNA

species were visualized by BrightStar® BioDetect™ nonisotopic detection kit following the manufacturer's instruction manual and subsequent exposure to film. Sizes of the mRNA species were determined by comparison with two RNA ladders (0.5-9 kb RNA Millenium Markers from Applied Biosystems & 0.5-10 kb RNA ladder from Invitrogen) that were run on adjacent gel lanes and stained with ethidium bromide.

For RNA probes (300-500 bases long), DNA template was generated by PCR with an antisense primer tailed by a T7-RNA-polymerase recognition sequence. After purification, the product was *in vitro* transcribed by MAXIscript® T7 kit, and then labeled by BrightStar® Psoralen-Biotin nonisotopic labeling kit according to the manufacturer's instructions.

Chromatin immunoprecipitation (ChIP). Transfections were performed as described above, and cells were crosslinked with formaldehyde (1%) three days after transfection. Before crosslinking, a small sample was collected for qPCR to confirm either silencing or activation. Cells were recovered by scraping and nuclei isolated. Nuclei were lysed in 1 mL lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1, 1x Roche protease inhibitor cocktail) and sonicated (2 pulses, 40% power, 30 sec). 100 µL of lysate was incubated overnight with 4 µg of monoclonal anti-RNAP2 antibody (Millipore 05-623) or mouse IgG negative control antibody (Millipore 12-371) diluted in 900 µL of immunoprecipitation buffer (0.01% SDS, 1.1% Triton-X, 1.2 mM EDTA, 16.7 mM Tris pH 8.1, 167 mM NaCl, and 1x Roche protease inhibitor cocktail). Antibody was recovered with 60 µL of Protein A/G beads (Calbiochem EMD IP05). Beads were washed with low salt (0.1% SDS, 1% Triton-X, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), high salt (see low salt but with 500 mM NaCl),

LiCl solution (0.25 M LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, and 10 mM Tris-HCl pH 8.1), and Tris-EDTA pH 8 (Ambion AM9863) washes. Protein was eluted with 500 µL of 1 % SDS, 0.1 M $NaHCO_3$ for 30 min at room temperature. Crosslinking was reversed by adding NaCl to 200 mM and heating at 65 ºC for at least 2 hours. Protein was digested by Proteinase K followed by extraction using an equal volume of phenol:chloroform:isoamyl alchohol and centrifugation. DNA in the aqueous layer was precipitated using 1/10 volume sodium acetate/ 2 volumes ethanol. The pellet was resuspended in 50 µL of nuclease-free water. Real time PCR was performed using BioRad iTaq SYBR® Supermix.

RNA immunoprecipitation (RIP). The general anti-AGO antibody was provided by Z. Mourelatos, University of Pennsylvania. MCF7 or T47D cells were grown in 150 cm^2 dishes and transfected with duplex RNAs using RNAiMax (Invitrogen). Cells were crosslinked using 1 % (v/v) formaldehyde solution and harvested. Cells were lysed and nuclei obtained. Anti-AGO2 antibody (Millipore 07-590) or the general anti-AGO antibody was incubated with nuclear lysates overnight. The antibody-treated material was then mixed with Protein A/G Agarose Plus (Santa Cruz) and washed five times. Complex was eluted and crosslinking reversed by adding 200 mM NaCl and heating to 65ºC for two hours. Samples were treated with DNase I, reverse transcribed, and amplified.

Rapid amplification of cDNA ends (RACE). 5'-RACE and 3'-RACE were performed according to the manufacturers protocol using the GeneRacer kit (Invitrogen). This kit includes enzymatic treatments that select for full length RNA with intact 5' caps rather than truncated

Nature Chemical Biology: doi: 10.1038/nchembio.400

products. For 5'-RACE, RNA was treated with phosphatase prior to removal of the cap to prevent cloning of truncated transcripts. For 3'-RACE, cDNA was made using oligo dT primers to allow cloning of the polyadenylated 3' ends. Multiple primer sets (**Supplementary Table S4**) were used to maximize detection of transcripts and reduce the likelihood of bias from any one primer set. We used the Platinum Taq High Fidelity kit (Invitrogen) to produce product for cloning. PCR products were cloned into a PCR-4 Topo vector and sequenced (McDermott sequencing core, UT Southwestern). We sequenced multiple clones from at least two independent experiments to confirm results.

Chromatin conformation capture (3C). 20 million cells were grown and crosslinked in 1% formaldehyde. Cells were recovered by scraping (5 µg genomic DNA). Nuclei were purified using hypotonic lysis and distriubted into 1 million nuclei aliquots. Aliquots were stored at -80°C. An additional 10 million nuclei were recovered without the use of formaldehyde for a no crosslink control. Aliquots were removed from -80ºC storage and resuspended in 500 µL of 1x restriction buffer (RB) and 3% SDS and incubated for 1.5 hr at 37ºC with shaking at 1000 rpm to loosen chromatin. Then Triton-X was added to 1.8% and samples were incubated 1 hr more at 37ºC with shaking to sequester SDS. 300 units of restriction enzyme DpnII were added and incubated overnight at 37ºC with shaking. Next day, SDS was added to 1.6% and samples were incubated at 65ºC for 30 min to inactivate restriction enzyme. 150 µL of sample was saved to check restriction enzyme efficiency.

Samples were diluted to 2 mL volume with 1.2x final concentration of ligase buffer and 1% final concentration of Triton-X. Samples were incubated at 37ºC for 1 hr. Samples were

placed on ice and either 40 units of T4 ligase were added and incubated overnight at 16ºC. The next day samples were incubated for 30 min at room temp. Crosslinks were reversed by adding NaCl to 200 mM and incubating at 65ºC for 2 hrs with proteinase K. Samples were then incubated with RNase A for 45 minutes at 41ºC. Finally DNA was purified by phenol:chloroform extraction and ethanol:sodium acetate precipitation. DNA was resuspended in 50 µL of nuclease free water and 1 µL was used in each PCR reaction.

Primers were designed to span several DpnII cut sites at the 5' end, 3' end, and internal sites. Various combinations of forward and reverse primers were tested. Positive control templates for primers were synthesized as single strand DNA oligonucleotides (Sigma). Product was cloned and sequenced to ensure product was specific. Sequenced products aligned with their respective sites in the genome.

Interferon response assay. T47D and MCF7 cells were transfected with duplex RNAs and mRNA was recovered using TRIzol® (Invitrogen) at 72 h after transfection. As a positive control in parallel, cells were transfected with 0.1 and 0.5 µg/ml of poly(I:C) dsRNA sequence (Sigma). Total RNA from each sample was isolated, treated with deoxyribonuclease I (Worthington Biochemical Corporation) and used to synthesize cDNA with High-capacity cDNA Reverse Transcription Kit (Applied Biosystems). Relative gene expression levels were quantified with Interferon Response Detection Kit (System Biosciences) on BioRad CFX-96™ Real-time System and were scaled to the untreated sample (relative expression=1). Relative expression levels are shown on a log scale.

Nature Chemical Biology: doi: 10.1038/nchembio.400

mRNA stability assay. MCF7 seeded in 6-well plates at 2.5 x 10^5 cells/well and allowed to grow for 24 hours. Cells were transfected with 25 nM of mismatch containing RNA duplexes using oligofectamine RNAi-Max (Invitrogen) at a 2:1 ratio. Each treatment group contained triplicate samples. Medium containing either 1.25 μg/ml actinomycin D (Sigma) or vehicle was added to the wells 52 h after transfection. Cells were harvested at different time points following actinomycin D treatment (4, 9, 24, 30 h). Total RNA was harvested from each sample using TRIzol (Invitrogen) and reverse transcribed using Applied Biosystem's High Capacity Reverse Transcription kit. Quantitative PCR was performed as previously described using primer sets to PR and 18S ribosomal RNA.

Cloning of the PR 3' noncoding RNA for plasmid-based expression. Total RNA from T47D cells was reverse transcribed using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Inc.). The cDNA sequence of the PR 3' noncoding RNA (3' NCR) was amplified by touchdown PCR with Hi Fidelity Platinum Taq (Invitrogen) using the following forward and reverse primers based upon the sequence of the mature PR 3' NCR determined from RACE experiments: 3'NCR-F, GGTTTCTTTCTCCAGGTCCTCACTGGCCATAC; 3'NCR-R, ggatccggGAGAGCTATACTCTTAGTTTGTTATAGTGACTCCA. PCR product was gel-purified and 3'-end adenine overhangs blunted by incubation with PfuTurbo DNA Polymerase (Stratagene) and 0.1 mM dNTPs in 1X PfuTurbo reaction buffer. Blunted PCR product was inserted into the Strataclone Mammalian Expression Vector pCMV-SC (Stratagene). Correct orientation of the 3' NCR cDNA sequence for expression from the CMV promoter in the generated pCMV-SC-3'NCR vector was verified by sequencing.

Nature Chemical Biology: doi: 10.1038/nchembio.400

Plasmid-based overexpression experiments. T47D, MCF7, or MDA-231 cells were seeded in six well plates at 200,000 cells per well. Cells were transfected with plasmid DNA using Lipofectamine-2000 according to the manufacturer's instructions when the cells were 85 % confluent. Cells were harvested for RNA 24 h after transfection.

Branched DNA assay (bDNA). The Quantigene 2.0 bDNA assay was performed essentially as described by the manufacturer (Panomics/Affymetrix). Samples in 96-well plates were lysed in a 200 µL solution containing two parts nuclease-free water and one part lysis mixture. Proteinase K (50 µg/mL) was added to a final volume of 10 µL Proteinase K per milliliter prepared solution. The plates were then incubated at 55°C for 60 m. Specific probes to detect human GAPDH (used as control for normalization) or human PR (mRNA, or other transcripts) were prepared in lysis mixture and blocking buffer was separately added to bDNA capture plates. The probes were designed with the help of and manufactured by Panomics.

Lysed cells were added to the prepared bDNA plates containing probes in a ratio of (80 µL lysed cells to 20 µL probe preps). The plates were incubated overnight at 55°C. Plates were removed to room temperature and washed three times with 300 µL wash solution per well. The following steps consisted of three one-hour incubations in preamplifier, amplifier, and labeling probe respectively. The first two incubations were performed at 55°C, while the labeling probe incubation was at 50°C. Between each step was a set of washes as described above. After a final series of washes, substrate was added. Luminescence was measured in the wells using a spectrophotometer and an integration time of 200 milliseconds. Analysis was done after normalizing the plates with their corresponding GAPDH probes.

Nature Chemical Biology: doi: 10.1038/nchembio.400

SUPPLEMENTARY RESULTS (Tables and Figures)

Supplementary Table 1. Primers employed for synthesis of RNA probes in northern blot analysis.

		Sequence
Probe 1	Forward	5'-TGG TGT TTG GTC TAG GAT GGA-3'
	Reverse	5'-TAA TAC GAC TCA CTA TAG GCA CCA TCC CTG CCA ATA TCT-3'
Probe 2	Forward	5'-CAC CTT GCT CCT CAT TTC TGA-3'
	Reverse	5'-TAA TAC GAC TCA CTA TAG GTT CAA ACC ACC AGC CAA TTT-3'
Probe 3	Forward	5'-TCA GGC AAT CAA GTT GAA ACC-3'
	Reverse	5'-TAA TAC GAC TCA CTA TAG GTT GCC TGC ATC AGT TCC TTA TAG-3'
Probe for Cyclophilin A mRNA	Forward	5'-CAC CGT GTT CTT CGA CAT TG-3'
	Reverse	5'-TAA TAC GAC TCA CTA TAG GTC GAG TTG TCC ACA GTC AGC-3'

Supplementary Table 2. Primers used for qPCR to detect transcript level across PR locus.

Target region		Sequence
33-104	Forward	5'-CTT GTT GTA TTT GCG CGT GT-3'
	Reverse	5'-GCC TCG GGT TGT AGA TTT C-3'
102-200	Forward	5'-TAC AAC CCG AGG CGG CTA-3'
	Reverse	5'-GAA GGG TCG GAC TTC TGC T-3'
171-313	Forward	5'-GTA CGG AGC CAG CAG AAG TC-3'
	Reverse	5'-TCT CTG GCA TCA AAC TCG TG-3'
2693-2813	Forward	5'-ACT GGA TGC TGT TGC TCT CC-3'
	Reverse	5'-CAG GTT GAT CAG TGG TGG AA-3'
3439-3518	Forward	5'-CGG GCA CTG AGT GTT GAA T-3'
	Reverse	5'-CAC CAT CCC TGC CAA TAT CT-3'
3501-3619	Forward	5'-ATA TTG GCA GGG ATG GTG AA-3'
	Reverse	5'-CAA GAC CTC ATA ATC CTG ACC A-3'
3584-3663	Forward	5'-ATG TCT TTT TGT TTT GGT CAG GA-3'
	Reverse	5'-TGA TGT TAT AAA TGT AAG GCT TTC AGA-3'
3599-3703	Forward	5'-TGG TCA GGA TTA TGA GGT CTT G-3'
	Reverse	5'-AAT TAG AAC CTC ACA ATT TTT CTT TT-3'
3739-3846	Forward	5'-TGT TTT GTT TAC CCA TAT TTT CTT GA-3'
	Reverse	5'-TCC TCA CCT ACA TGG TAT GAA A-3'
3940-4028	Forward	5'-CCC TTT GTG TCA ATT ATA TTT CCA A-3'
	Reverse	5'-TGA ATT TCC TCC TCA CAC AAA-3'
4273-4422	Forward	5'-TGA GGT ATT GCG AGT GGA CA-3'
	Reverse	5'-GAG CAA TTG GCA GGA AAG AT-3'
4549-4649	Forward	5'-TGT TTC AGC CAT GCA AAT CT-3'
	Reverse	5'-TTT GGC TGA GTC TCG AAG GT-3'
5904-6052	Forward	5'-CCA CAG GTT TGG CTT TTG TT-3'
	Reverse	5'-CCA TTT GGT GAA GCC ATA TTC-3'
7371-7485	Forward	5'-CCA GAG CCA TGT GCA TAA GA-3'
	Reverse	5'-ATC AGT GGG GAC CAC AGT TG-3'
8139-8266	Forward	5'-TAA GAA TTT GGG GGT GTT GG-3'
	Reverse	5'-AGG AGA ACA AAC CCC TTG GT-3'
8959-9086	Forward	5'-GCT CAG CAG CTT TCA TTG AT-3'
	Reverse	5'-TTC TCC TCC CCC AGA AAA GT-3'
12207-12356	Forward	5'-TGT GCC TGA CAG TTC TCC TG-3'
	Reverse	5'-CCA TGT GCA AAA CAG TCA AAG-3'
12604-12714	Forward	5'-CAC CTT GCT CCT CAT TTC TGA-3'
	Reverse	5'-CCC AGG CAT ACA CAG ATG AA-3'
13271-13382	Forward	5'-TCA GGC AAT CAA GTT GAA ACC-3'
	Reverse	5'-TCA ATG TGA ATA GCC AAA ACA G-3'
14354-14445	Forward	5'-CAT GTG CGT TGA CAT TCA CA-3'
	Reverse	5'-TGA TTC TGA TGG TTG GTG GA-3'
Exon 4/5	Forward	5'-TGC AGG ACA TGA CAA CA CAA-3'
	Reverse	5'-TAC AGC ATC TGC CCA CTG AC-3'
Exon 6/7	Forward	5'-CCT TAC CAT GTG GCA GAT CC-3'
	Reverse	5'-TGT GAG CTC GAC ACA ACT CC-3'

14770-14858	Forward	5’-CAC TGC ACC TGG CCT AAA CT-3’
	Reverse	5’-CCT TGA CCT CCT TTG CTG AA-3’
15496-15627	Forward	5’-TTC AGC AAA GGA GGT CAA GG-3’
	Reverse	5’-CAA TCT GTC CCA TGC AGA AA-3’

Supplementary Table 3. Primers used in RACE.

Primers used to identify the 3’-termini of PR mRNA	
5’-GGC ACT GGC TGG TAA CAG ATG CAA AAC TG-3’	
5’-TTG GCA AGA GAT GCA GGG AAT CTT TCT CAT-3’	
Primers used to identify the 3’ noncoding sense transcript	
A	5’-CACTGTGTAGTTGGTTTCAACTTGATTGCCTGA-3’
B	5’-TGATTGCCTGAGAATCACTCTTTGCTTTGCTA-3’
C	5’-GCAAAGCAAAGAGTGATTCTCAGGCAATCAAG-3’
D	5’-GCCTAAATTCTATAAGGAACTGATGCAGGCAAACC-3’
E	5’-AGGGAGCACTGGTGAGCAGTAGGTTGAAGA-3’
F	5’-CCATTTCTTGCTGGCTTAGCACATTCCTCA-3’
G	5’-CAGTGACATATAGTGACACAAGGGAAAAGTCTCA-3’
H	5’-TCCAAATTGCTCACAAATAACTGGTCATGGA-3’

Supplementary Table 4. RNAs targeting PR or BRCA1.

Name	Target site location	Sequence
RNAs complementary to the PR promoter		
PR-9	-9/+10	UGUCUGGCCAGUCCACAGCTT
PR-11	-11/+8	GCUGUCUGGCCAGUCCACATT
RNA complementary to the PR 3'-UTR		
PR2526	+2526/+2544	AGUUGUGCUGCCCUUCCAUTT
PR3593	+3593/+3611	CAUAAUCCUGACCAAAACATT
RNAs complementary beyond the PR 3'-UTR		
PR13063	+13063/+13082	UGUAUUGAGGUUUUAGAUGCTT
PR13485	+13485/+13503	CCUAGUAAUGAAACCAAUGTT
PR13515	+13515/+13534	GUAUAUUUCUAGAGCUAUACTT
PR13580	+13580/+13598	GUUUGCCUGCAUCAGUUCCTT
RNA complementary to the BRCA1 protein-encoding RNA		
siBRCA	+1416/+1436	UACAUCAGCUACUUUGGCAUUTT
RNAs complementary beyond the BRCA1 3'-UTR		
BRCA7209	+7209/+7227	AAAGGCUCUGAGAAAGUCGTT
BRCA7324	+7324/+7342	UUUGUUUUGGCAGCAACAGTT
BRCA7851	+7851/+7869	CAUCUGUGGAUUAAGCAUGTT
BRCA7988	+7988/+8007	CUCCAUUUUCUCUAUCUUCCTT
BRCA8870	+8870/+8888	GUGGAAAAGUGGGAGGACATT
BRCA9077	+9077/+9095	UGUUGAUGAUUCUGGUUGCTT
Mismatch-Containing RNAs		
MM		UCUCUCGCCAGUGCACACCTT
PR13485_MM3		CCUAGUAAU<u>C</u>AAA<u>G</u>CA<u>U</u>UGTT
PR13515_MM3		GUAUAUUUC<u>A</u>AGA<u>C</u>CUA<u>A</u>ACTT
PR13515_MM3B		GU<u>U</u>UAU<u>A</u>UCU<u>U</u>GAGCUAUACTT
PR13515_MM4		GUAU<u>U</u>UUUC<u>A</u>AGA<u>C</u>CUA<u>A</u>ACTT
PR13580_MM3		G<u>A</u>UUG<u>G</u>CUG<u>G</u>AUCAGUUCCTT
PR13580_MM4		GUUU<u>A</u>CCU<u>C</u>CAU<u>G</u>AGU<u>A</u>CCTT

RNA sequences are listed 5' to 3'. Only one strand of the duplex RNA is shown, corresponding to + strand of genomic DN
of target site locations is relative to the +1 transcription start site for PR (NM_000926.4) and BRCA1a (NM_007294.2), resp

Supplementary Table 5. Primers used to detect PR transcripts.

Target region		Sequence
5'-noncoding antisense transcript	Forward	5'-GGA GGA GGC GTT GTT AGA AA-3'
	Reverse	5'-GAA GGG TCG GAC TTC TGC T-3'
PR Intron	Forward	5'-CAA AAA GGG TCC GGT GTA GA-3'
	Reverse	5'-AGG CAC TGC TCC ACT GTC TT-3'
3'-noncoding sense transcript	Forward	5'-TTC AGA CTA CAT TGG TTT CAT TAC TAG G-3'
	Reverse	5'-TTG CCT GCA TCA GTT CCT TAT AG-3'

Primer pair amplifying PR protein-encoding mRNA was designed by Applied Biosystems.

Supplementary Table 6. Primers used in CHIP.

Primers used in RNAP2 CHIP at PR promoter	
Forward	5'-GGA GGA GGC GTT GTT AGA AA-3'
Reverse	5'-GAA GGG TCG GAC TTC TGC T-3'
Primers used in RNAP2 CHIP at BRCA1 promoter (Figure 8D)	
Forward	5'-GCG CGG GAA TTA CAG ATA AA-3'
Reverse	5'-TAC CCA GAG CAG AGG GTG AA-3'

Supplementary Table 7. Primers used in Figure 4.

Target transcript		Sequence
PR mRNA	Forward	5'-TGC AGG ACA TGA CAA CA CAA-3'
	Reverse	5'-TAC AGC ATC TGC CCA CTG AC-3'
3' noncoding transcript (3' NCR)	Forward	5'-TAA GGA ACT GAT GCA GGC AAA-3'
	Reverse	5'-AAG CCA AAA ATC CTC CCA AG-3'
5' noncoding transcript (5' NCR)	Forward	5'-CCT AGA GGA GGA GGC GTT GT-3'
	Reverse	5'-ATT GAG AAT GCC ACC CAC A-3'

Supplementary Table 8. Primers used in RIP (Figure 5).

Target Transcript	Direction	Sequence
5' noncoding RNA	Forward	5'-GGA GGA GGC GTT GTT AGA AA-3'
	Reverse	5'-GAA GGG TCG GAC TTC TGC T-3'
3' noncoding RNA	Forward	5'-CAGACTACATTGGTTTCATTACTAGG-3'
	Reverse	5'-GCATTTATTTTATTTACTAAAGGAGCA-3'
3' noncoding RNA	Forward	5'-TAA GGA ACT GAT GCA GGC AAA-3'
	Reverse	5'-AAG CCA AAA ATC CTC CCA AG-3'
3' noncoding RNA	Forward	5'-TTC AGA CTA CAT TGG TTT CAT TAC TAG G-3'
	Reverse	5'-TTG CCT GCA TCA GTT CCT TAT AG-3'
3' noncoding RNA	Forward	5'-CCT AGA GGA GGA GGC GTT GT-3'
	Reverse	5'-ATT GAG AAT GCC ACC CAC A-3'

Supplementary Table 9. Primers used in 3C analysis.

Primers used in 3C analysis of PR

T2	5'-AGTTTAGGCCAGGTGCAGTG-3'
T1	5'-TAGCTGATTTGGGCCAGTTT-3'
E4	5'-TGTTAATGAGCATTGAACCAGA-3'
E3	5'-ATCATCTGCCCCTGTTGAAA-3'
E2	5'-TCCTTACCTGTGGGAGCTGT-3'
E1	5'-GGAGAACTCCCCGAGTTAGG-3'
F2	5'-GAAGGGTCGGACTTCTGCT-3'
F1	5'-TACAACCCGAGGCGGCTA-3'
P	5'-TCTAACAACGCCTCCTCCTC-3'

Primers used in 3C analysis of BRCA1

D1	5'-GGTGCATATAAAATCCTCAGGC-3'
D2	5'-GCTCCTCAGCGCCCGGTC-3'
D3	5'-GTATTCTTTGACGGGGGGTAGG-3'
D4	5'-GCTCCATCACTTGAAATGGC-3'
D5	5'-CACTGCCCTGTGCTATGTCAA-3'
D7	5'-CCAGCCTGGGTGACAGAGC-3'
D9	5'-GGCTTGGCCTCAAGAGAATAGCTG-3'
D10	5'-CCGGTCATGGTGGTGGACA-3'
D11	5'-CATGAGGCCAAAATAAAGGTGTTCG-3'

Supplementary Table 10. Primers used to define BRCA1 mRNA.

Target region		Sequence
5'-promoter	Forward	5'-GCG CGG GAA TTA CAG ATA AA-3'
	Reverse	5'-TAC CCA GAG CAG AGG GTG AA-3'
BRCA1 pre-mRNA	Forward	5'-ACC TGC GAA ATC CAG AAC AA-3'
	Reverse	5'-TGT GCT GAG CAA GGA TCA TAA-3'
3'-flanking region	Forward	5'-TTC AGA CTA CAT TGG TTT CAT TAC TAG G-3'
	Reverse	5'-TTG CCT GCA TCA GTT CCT TAT AG-3'

Primer pair amplifying BRCA1 mRNA was designed by Applied Biosystems.

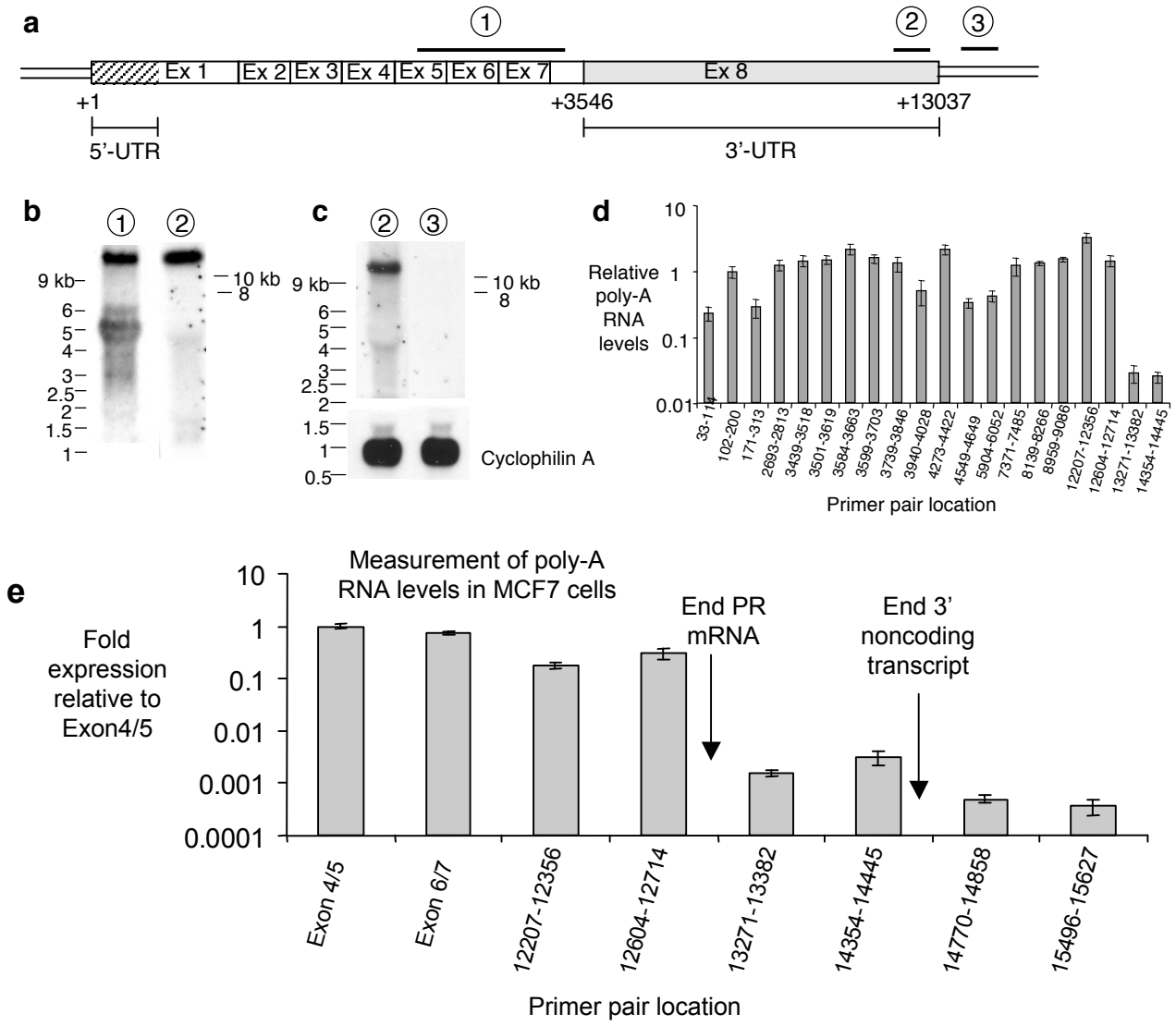
Supplementary Table 11. Primers used in Supplementary Figure S23a.

Target region		Sequence
6889-6957	Forward	5'-ATG ACA GAT CCC ACC AGG AA-3'
	Reverse	5'-CAA CAG GGA GCA AAG GAA AA-3'
7197-7275	Forward	5'-ACT TAG GCC AGC GAC TTT C-3'
	Reverse	5'-TAC TAT CAT TAC CCC CAT TT-3'
7319-7405	Forward	5'-ATT TCC TGT TGC TGC CAA AA-3'
	Reverse	5'-GGC CTC ATG AAA CCT GAA AC-3'
7402-7464	Forward	5'-GGC CAA AAT AAA GGT GTT CG-3'
	Reverse	5'-AGA ACT GCA AGG ACC CAG AG-3'

Supplementary Table 12. Primers used in Supplementary Figure S23b.

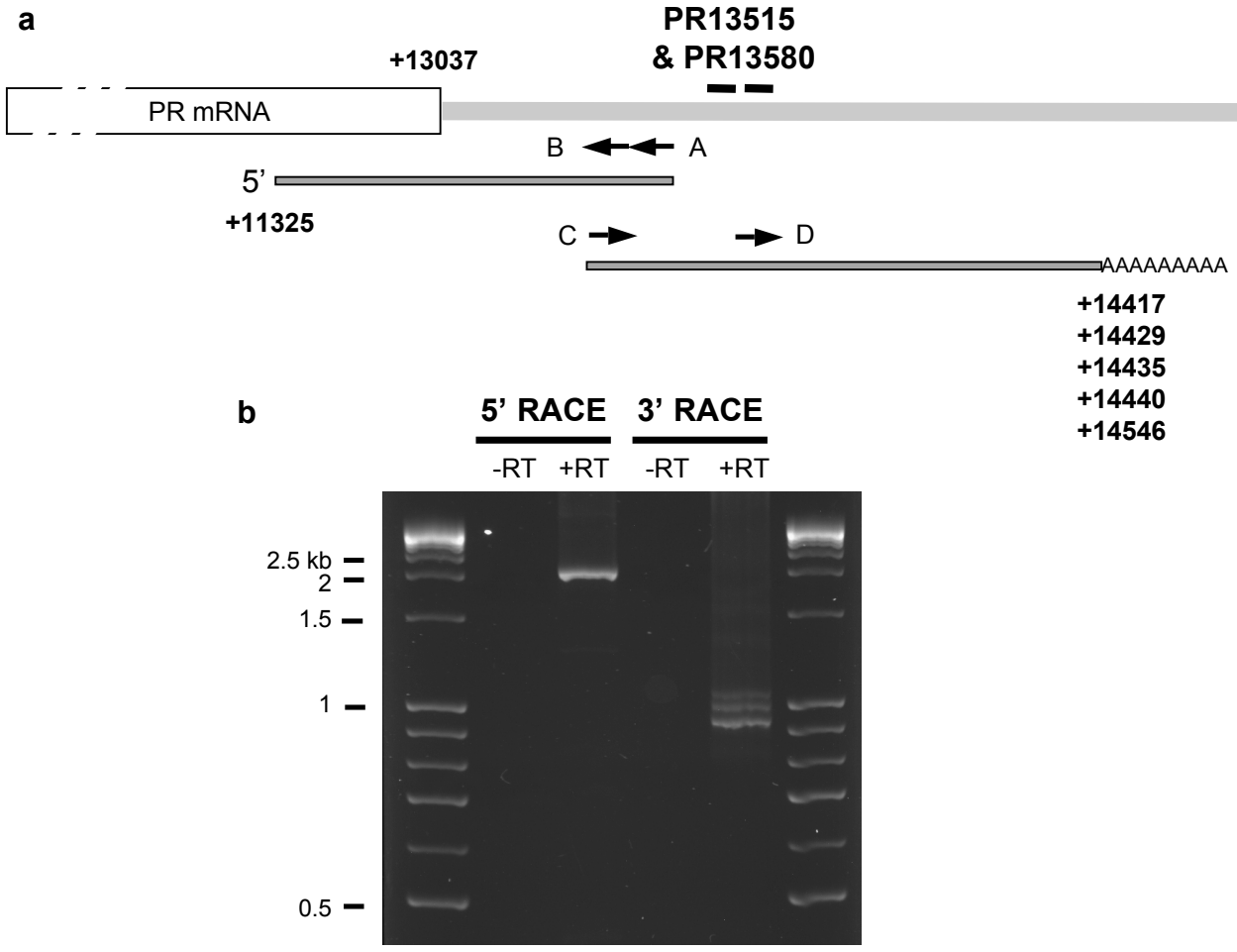
Primer	Sequence
A	5'-ATG ACA GAT CCC ACC AGG AA-3'
B	5'-GTG CAA GGG CAG TGA AGA CT-3'
C	5'-ACT TAG GCC AGC GAC TTT C-3'
D	5'-ATT TCC TGT TGC TGC CAA AA-3'
E	5'-GGC CAA AAT AAA GGT GTT CG-3'

Supplementary Figure S1.



Supplementary Figure S1. Characterization of PR mRNA. (**a**) Schematic of PR mRNA predicted by GenBank and locations of probes for Northern analysis. (**b**) Northern analysis of PR mRNA comparing results using probes that detect PR mRNA (probe 1) or targeting the 3' termini of PR mRNA (probe 2), (**c**) Northern analysis of PR mRNA using probe 2 or a probe immediately downstream of the potential mRNA terminus (probe 3). (**d**) and (**e**) qPCR showing levels of poly-A RNA in T47D and MCF7 cells respectively detected from the PR transcription start site (+1) past the most downstream annotated terminus of PR mRNA. Notation indicates target region for PCR primers. Data is the resultant of triplicate independent experiments. Primer set Exon4/5 targets the boundary of exon 4 and 5 in PR mRNA. Primer set Exon6/7 targets the boundary of exon 6 and 7 in PR mRNA.

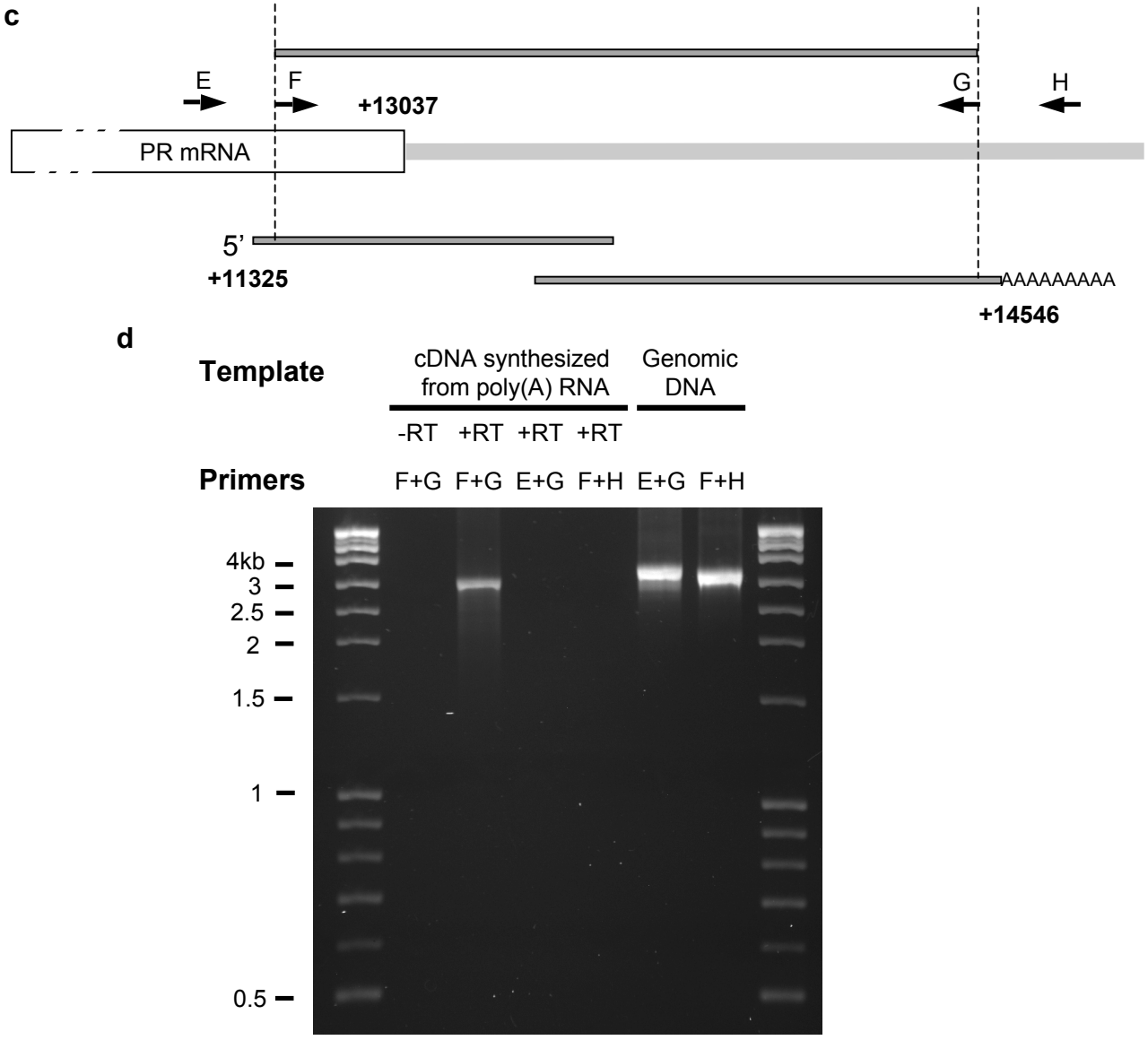
Supplementary Figure S2.



Supplementary Figure S2. **Characterization of a noncoding transcript that overlaps the 3'-termini of PR mRNA.** The RACE assay was used to detect the transcription start site and polyadenylation sites of 3' noncoding transcripts. We used primers recognizing sequences close to the complementary sequences of 3' inhibitory and activating agRNAs. RACE products were analyzed on a 2% agarose gel and sequenced. (**a**) Location of primers (A,B,C, and D) relative to PR mRNA. Data is the representative of triplicate independent determinations. (**b**) Agarose gel analysis of RACE products. For 5' RACE use of primer A (5'-CACTGTGTAGTTGGTTTCAACTTGATTGCCTGA-3') and nested primer B (5'-TGATTGCCTGAGAATCACTCTTTGCTTTGCTA-3') detects a transcription start site 11325 nucleotides downstream from the PR mRNA +1 start site. For 3' RACE use of primer C (5'-GCAAAGCAAAGAGTGATTCTCAGGCAATCAAG-3') and nested primer D (5'-GCCTAAATTCTATAAGGAACTGATGCAGGCAAACC-3') detects multiple polyadenylation sites 1400-1500 nucleotides downstream from the 3' end of PR mRNA. RACE experiments did not detect antisense transcripts. Total RNA used in RACE was treated with DNase prior to reverse transcription. +RT: Reverse transcriptase added. -RT: No reverse transcriptase added. The results from sequencing RACE products are shown in part (**e**).

Supplementary Figure S2 continued on next page

Supplementary Figure 2 (cont.)



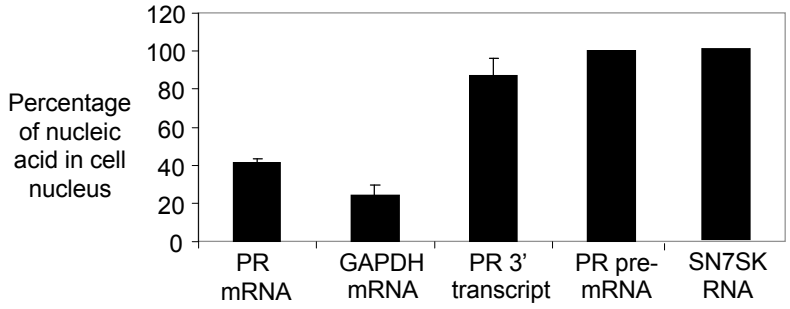
Supplementary Figure S2 (cont). (**c**) Locations of primers for RT-PCR. (**d**) RT-PCR with primer F (5'-CCATTTCTTGCTGGCTTAGCACATTCCTCA-3') and primer G (5'-CAGTGACATATAGTGACACAAGGGAAAAGTCTCA-3') on poly(A) RNA from T47D cells revealed an unspliced transcript from +11325 transcription start site to +14546 polyadenylation site detected by RACE in (**a**). No transcript was detected from primers E to G or from primers F to H, even though these primer sets could amplify genomic DNA efficiently. Primer E (5'-AGGGAGCACTGGTGAGCAGTAGGTTGAAGA-3') recognizes sequences 200nt upstream of +11325 transcription start site. Primer H (5'-TCCAAATTGCTCACAAATAACTGGTCATGGA-3') recognizes sequences 150nt downstream from +14546 polyadenylation site. Poly(A) RNA was DNase-treated prior to reverse transcription. +RT: Reverse transcriptase added. -RT: No reverse transcriptase added. (**e**) (next page) Sequence of the noncoding transcript that overlaps the 3' end of PR mRNA was deduced on the basis of sequencing results from (**a**) and (**b**).

e

GTTTCTTTCTCCAGGTCCTCACTGGCCATACACCAGTCCCTTGTTAGTTATGCCTGGTCATAGACC
CCCGTTGCTATCATCTCATATTTAAGTCTTTGGCTTGTGAATTTATCTATTCTTTCAGCTTCAGCAC
TGCAGAGTGCTGGGACTTTGCTAACTTCCATTTCTTGCTGGCTTAGCACATTCCTCATAGGCCCAG
CTCTTTTCTCATCTGGCCCTGCTGTGGAGTCACCTTGCCCCTTCAGGAGAGCCATGGCTTACCAC
TGCCTGCTAAGCCTCCACTCAGCTGCCACCACACTAAATCCAAGCTTCTCTAAGATGTTGCAGACT
TTACAGGCAAGCATAAAAGGCTTGATCTTCCTGGACTTCCCTTTACTTGTCTGAATCTCACCTCCT
TCAACTTTCAGTCTCAGAATGTAGGCATTTGTCCTCTTTGCCCTACATCTTCCTTCTTCTGAATCAT
GAAAGCCTCTCACTTCCTCTTGCTATGTGCTGGAGGCTTCTGTCAGGTTTTAGAATGAGTTCTCAT
CTAGTCCTAGTAGCTTTTGATGCTTAAGTCCACCTTTTAAGGATACCTTTGAGATTTAGACCATGTT
TTTCGCTTGAGAAAGCCCTAATCTCCAGACTTGCCTTTCTGTGGATTTCAAAGACCAACTGAGGAA
GTCAAAAGCTGAATGTTGACTTTCTTTGAACATTTCCGCTATAACAATTCCAATTCTCCTCAGAGCA
ATATGCCTGCCTCCAACTGACCAGGAGAAAGGTCCAGTGCCAAAGAGAAAAACACAAAGATTAAT
TATTTCAGTTGAGCACATACTTTCAAAGTGGTTTGGGTATTCATATGAGGTTTTCTGTCAAGAGGG
TGAGACTCTTCATCTATCCATGTGTGCCTGACAGTTCTCCTGGCACTGGCTGGTAACAGATGCAA
AACTGTAAAAATTAAGTGATCATGTATTTTAACGATATCATCACATACTTATTTTCTATGTAATGTTT
TAAATTTCCCCTAACATACTTTGACTGTTTTGCACATGGTAGATATTCACATTTTTTTGTGTTGAAGT
TGATGCAATCTTCAAAGTTATCTACCCCGTTGCTTATTAGTAAAACTAGTGTTAATACTTGGCAAGA
GATGCAGGGAATCTTTCTCATGACTCACGCCCTATTTAGTTATTAATGCTACTACCCTATTTTGAGT
AAGTAGTAGGTCCCTAAGTACATTGTCCAGAGTTATACTTTTAAAGATATTTAGCCCCATATACTTC
TTGAATCTAAAGTCATACACCTTGCTCCTCATTTCTGAGTGGGAAAGACATTTGAGAGTATGTTGA
CAATTGTTCTGAAGGTTTTTGCCAAGAAGGTGAAACTGTCCTTTCATCTGTGTATGCCTGGGGCTG
GGTCCCTGGCAGTGATGGGGTGACAATGCAAAGCTGTAAAAACTAGGTGCTAGTGGGCACCTAA
TATCATCATCATATACTTATTTTCAAGCTAATATGCAAAATCCCATCTCTGTTTTTAAACTAAGTGTA
GATTTCAGAGAAAATATTTTGTGGTTCACATAAGAAAACAGTCTACTCAGCTTGACAAGTGTTTTAT
GTTAAATTGGCTGGTGGTTTGAAATGAATCATCTTCACATAATGTTTTCTTTAAAAATATTGTGAATT
TAACTCTAATTCTTGTTATTCTGTGTGATAATAAAGAATAAACTAATTTCTATATCTCTCTTTATTAAT
GAATTATAGCATCTAAAACCTCAATACAATTACATACAAGACACACACTAATCATCAGTAGTACATT
CTTTACCAGTCATTTAAAATTTATAACCAACATTTCAATTTGTACAATACTGTATGTGGCATAAGGT
GAGATATTTATATGGAAGATTTGGCATTATAGAGAAAATATCCTTGACTGGGTATGCATTTTAGCAA
AGCAAAGAGTGATTCTCAGGCAATCAAGTTGAAACCAACTACACAGTGTTTCAATCAGAAAGACAA
AATACAATCAACTGACATCTAGTGAGATTCAATAATATACTGTTTTGGCTATTCACATTGATTAAAAA
GTTTGGTATTATACAAGAATTTTCATATGAAATTTAATGCATTTCACATTAAGGTGAATGATATTCTA
CTTGTAATAACAAAGTATTTCAGACTACATTGGTTTCATTACTAGGAATATAATTTAGTATAGCTCTA
GAAATATACAAATATGCTCCTTTAGTAAATAAAATAAATGCCTAAATTCTATAAGGAACTGATGCAG
GCAAACCCTAAAATGGGGGCTCAGCTTGGGAGGATTTTTGGCTTAATTCAGAAAAGAATTCAAGA
GGGAACCCACAGTGAAAGAAGCCAAGTTTATTGGAGCAACAGCATCCAGCAAAATGGCTACTCCA
CAGGCAGAGTAGCCCTCGTGGGTTGCTGGCTAGCTATATGTATACCAACTCTTAATTATATGCTAA
ATATGAGGTCTGTTATTCACAGATTTTCTGGAAAAGCTGCAGGGAGTTCTTGGAACTATATAACTT
AATTTCTGGGTGTTCCCATGGCATTTGTAAAGTGTCATGGTGCTGGTGCAGTGTCTCATAGCCTG
CAGATGCATTATAATTTCTAGTCCTAGCTGATTTGGGCCAGTTTCTTAGCTACATCCTGTTTTTGAT
CAGCAGGGTCATGAAAACAAGTCCTGGTGATCTTTTACCTCAGAACCATGTTAGGTCTTGGAGAC
ACAAAGATAAATGAGTGGAACAGGATTCTTGCTGTAAGTTTGAGTTGATACAATGCCACATATTATT
GTTTGAAATGTCAAAATACTTGTTTCTTATTGAATCAACTAGATTTGGAATAGACTGGAAAATCTGG
AAAGCTTTATTGGATCATTTTTCCTATCACTTAAACACTATTTTTTCCTGGTTAATACCATGTTTTCG
GTTTTTAAAACATGCCACCATATTCAGATTTACAAATGTAAATACGGTTGAGAAACTTCGTATTTAC
CTATGAAACAACATGCTGTTTTTTATAAATACTTGATTCTGATGGTTGGTGGAGTTTTTCCAATCAA
AGTTAACTAATGAATAAAGAGAAAATGTGGCAAGTGGCTGTGAATGTCAACGCACATGAGACTTTT
CCCTTGTGTCACTATATGTCACTGTATGGAGTCACTATAACAAACTAAGAGTATAGCTCTCTCTACA
C T A A T C A T G A A T A A A T A C T T A A T A T C A A A A A A A A A A A A A A A A A A A A A A A A

The alternative polyA sites are underlined.

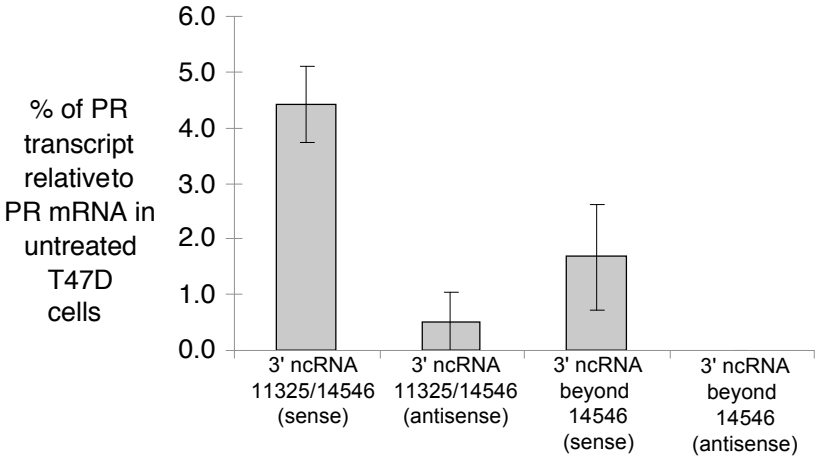
Supplementary Figure S3



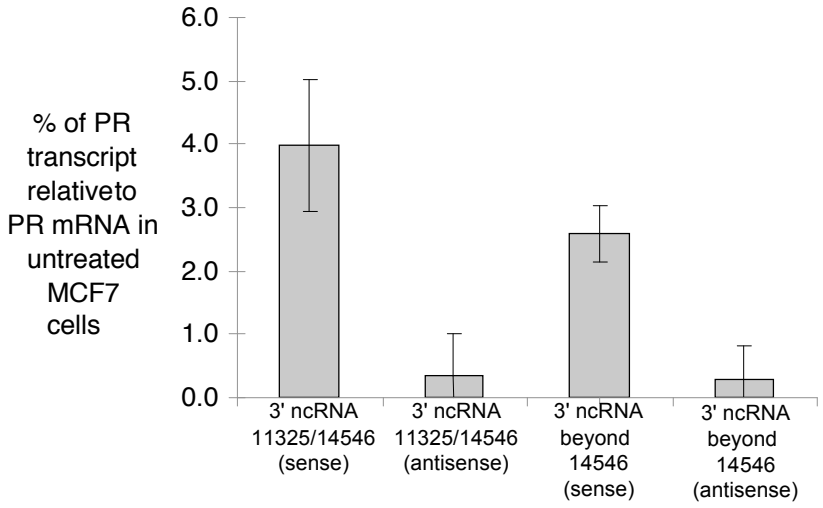
Supplementary Figure S3. **qPCR measurement of nuclear localization of 3' noncoding transcript.** SN7SK RNA is a nuclear RNA and is used as a standard.

Supplementary Figure S4 B-DNA assay

a



b



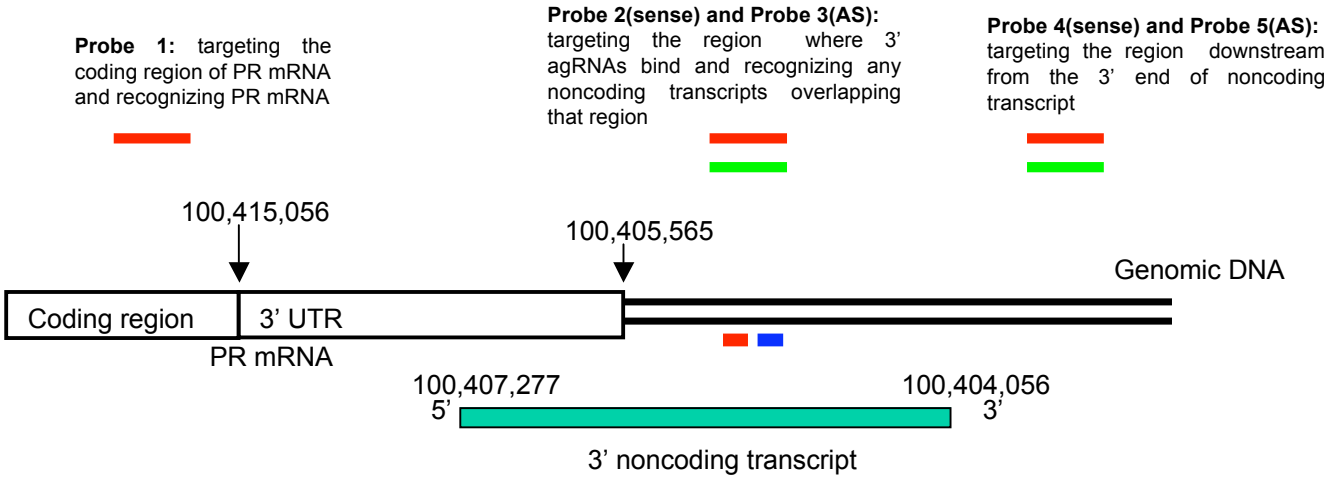
Supplementary Figure S4. **bDNA assay of transcript levels at the PR locus**. Transcript levels in (**a**) T47D cells and (**b**) MCF7 cells measured by using bDNA prove sets complementary to the 3' +11325/+14546 noncoding transcript (sense orientation and antisense orientation) or complementary to the region beyond the +11325/+145463' 3' noncoding transcript (sense orientation and antisense orientation. (**c)** (next page) Schematic of target locations for bDNA probe sets. (**D**-**H**) (next pages) Probe sets for bDNA assay. CE: Capture extender. LE: Label extender. BL: Blockers. Data is an average of four experiments.

Figure S4 continues on the next four pages

Nature Chemical Biology: doi: 10.1038/nchembio.400

c

PR(NM_000926) contains 8 exons and is on the minus strand of genomic DNA, spanning from Chr11:100,505,754 to Chr11:100,405,565



Activating 3' agRNA, GTATAGCTCTAGAAATATAC, genomic DNA location: 100,405,085-100,405,065

Inhibiting 3' agRNA, GGAACTGATGCAGGCAAAC, genomic DNA location: 100,405,025-100,405,005

d Probe set 1:Complementary to the coding region of PR mRNA

bDNA Probe sets 1

CE is BLUE , LE is RED and BL is green

>gi|160358783|ref|NM_000926| Homo sapiens progesterone receptor (PGR), mRNA. [Homo sapiens] **(ProbeSet1)**

agtccacagctgtcactaatcggggtaagccttgttgtatttgtgcgtgtgggtggcattctcaatgagaactagcttca
cttgtcatttgagtgaaatctacaacccgaggcggctagtgctcccgcactactgggatctgagatcttcggagatgact
gtcgcccgcagtacggagccagcagaagtccgacccttcctgggaatgggctgtaccgagaggtccgactagcccccaggg
ttttagtgaggggcagtggaactcagcgagggactgagagcttcacagcatgcacgagtttgatgccagagaaaaagtc
gggagataaaggagccgcgtgtcactaaattgccgtcgcagccgcagccactcaagtgccggacttgtgagtactctgcg
tctccagtcctcggacagaagttggagaactctcttggagaactccccgagttaggagacgagatctcctaacaattact
actttttcttgcgctccccacttgccgctcgctgggacaaacgacagccacagttcccctgacgacaggatggaggccaa
gggcaggagctgaccagcgccgccctccccgcccccgacccaggaggtggagatccctccggtccagccacattcaaca
cccactttctcctccctctgcccct

e Probe 2 (sense): targeting the region complementary to 3' agRNAs

bDNA Probe sets 2.2

CE is BLUE , LE is RED and BL is green
>gi|83641890|ref|GS00342|(ProbeSet2.2),
tattctgtgtgataataaagaataaactaatttctatatctctctttattaatgaattatagcatctaaaacctcaatac
aattacatacaagacacacactaatcatcagtagtacattctttaccagtcatttaaaatttataaccaacatttcaatt
tgtacaatactgtatgtggcataaggtgagatatttatatggaagatttggcattatagagaaaatatccttgactgggt
atgcattttagcaaagcaaagagtgattctcaggcaatcaagttgaaaccaactacacagtgtttcaatcagaaagacaa
aatacaatcaactgacatctagtgagattcaataatatactgttttggctattcacattgattaaaaagtttggtattat
acaagaattttcatatgaaatttaatgcatttcacattaaggtgaatgatattctacttgtaataacaaagtatttcaga
ctacattggtttcattactaggaatataatttagtatagctctagaaatatacaaatatgctcctttagtaaataaaata
aatgcctaaattctataaggaactgatgcaggcaaaccctaaaatgggggctcagcttgggaggatttttggcttaattc
agaaaagaattcaagagggaacccacagtgaaagaagccaagtttattggagcaacagcat

f Probe 3 (antisense): targeting the region complementary to 3' agRNAs.

bDNA Probe sets 3.2

CE is BLUE , LE is RED and BL is green
>gi|83641890|ref|GS00343| (ProbeSet3.2),
atgctgttgctccaataaacttggcttctttcactgtgggttccctcttgaattcttttctgaattaagccaaaaatcct
cccaagctgagcccccattttagggttgcctgcatcagttccttatagaatttaggcatttatttttatttactaaagga
gcatatttgtatatttctagagctatactaaattatatttcctagtaatgaaaccaatgtagtctgaaatactttgttatt
acaagtagaatatcattcaccttaatgtgaaatgcattaaatttcatatgaaaattcttgtataataccaaacttttaa
tcaatgtgaatagccaaaacagtatattattgaatctcactagatgtcagttgattgtattttgtctttctgattgaaac
actgtgtagttggtttcaacttgattgcctgagaatcactctttgctttgctaaaatgcatacccagtcaaggatatttt
ctctataatgccaaatcttccatataaatatctcaccttatgccacatacagtattgtacaaattgaaatgttggttata
aattttaaatgactggtaaagaatgtactactgatgattagtgtgtgtcttgtatgtaattgtattgaggttttagatgc
tataattcattaataaagagagatatagaaattagtttattctttattatcacacagaata

g Probe 4 (sense): targeting the region downstream from the 3' end of noncoding transcript

bDNA Probe sets 4.2

CE is BLUE , LE is RED and BL is green

>gi|83641890|ref|GS00344|(Probeset4.2),

tatagctctctctacactaatcatgaataaatacttaatatcaaagatcaagtttttttaaagctatattttaatggcagg
aggttgtatttacctaatcagatgaagtttaatcccatcttataaatgattgaatccaaacaagtagaataactataata
tgcagagatccatgaccagttatttgtgagcaatttggaattcataaagcttaaacaacttaggtactctatgaaaaagc
aaagtatttttatgtctgaattatgtcagctacccgtaattatgctgtgatataagaaagcagtatagtaagaatatatg
ctctggagctacaccttctcaatctcatttcaggctgtaccacttgctaactatgtgaccttgcaacatactttcactct
ctgccttagttttcttatctgtaaaatggatataacaatagtacttgtttcatagggttgttataaggatgaaatcagta
taagaaaaataaacttagagcagtgtgtgacatagcaagcactattttggtgtttcatccaaattaagtggctttttcct
taatacataaatgattaaaggtacaaaagatattaaacttttttttttttttttgagatggagtcttgctctgtcaccc
aggctggagtgcagtggcacgatttcggctcactgcaacctccacttcccgggctcaagcgattctcttgcctcagcctc
ccaagtagctgggactacaggtgtacaccaccaggcccagctaatttttgtatttttagtagagacagggtttcaccatg
ttggccaggatggtctcaatctcctgaccttgtgatctgcctgccttggcctcccaaagttctgggattacaggtgtgaa
ccactgcacctggcctaaacttttatattatataaactagtaaaaataatttttaatattttaggtcatgggaaattcag
cagcatagtaattatcaaagtagtaaatttttttcagcaaaggaggtcaaggatagactgttacaagtgaaaaaaatagga
tattaatcaacaaaattctcaatctgaatagccacaatttttcagtttctgtacctgaaatagtttctgcatgggacaga
ttggactaaccaaatttatctcatttttatgactctagcctaactaccctccccattaaataatgtttttaagtatatac
aaagggaaaaaaacaacttttgaattatgagatatactgaataaaacaataatgaaaaagctatttataaaatagttaa
gattagattataaatctaaagttg

h Probe 5 (antisense): targeting the region downstream from the 3' end of noncoding transcript

bDNA Probe sets 5.2

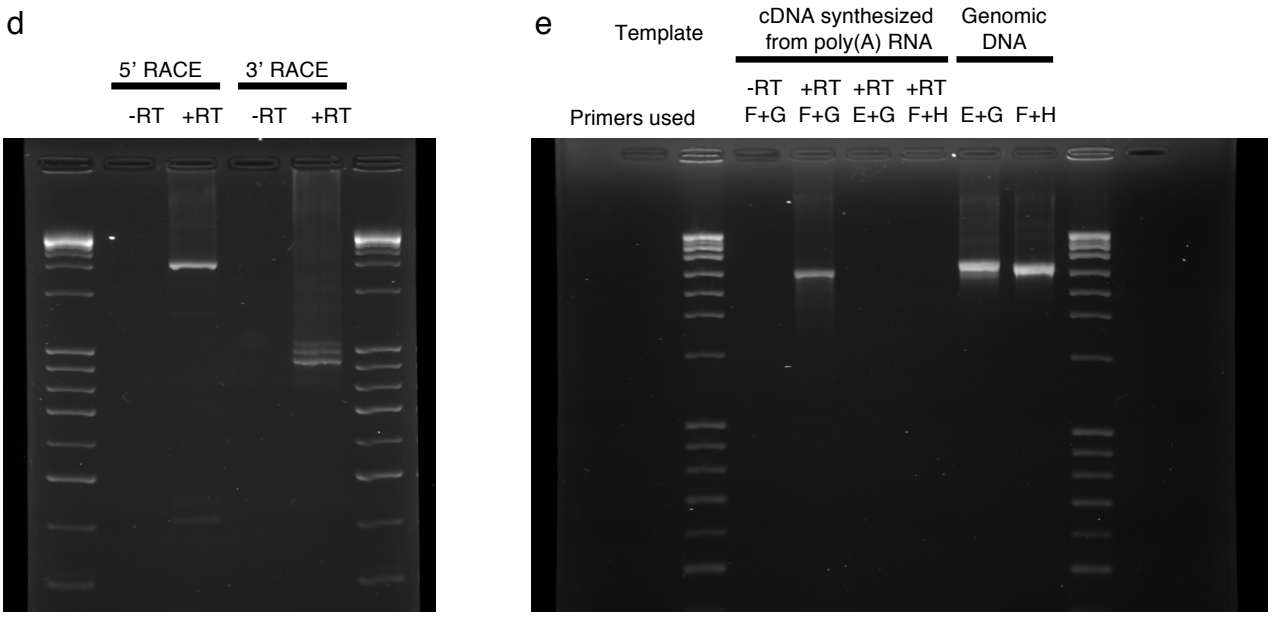
CE is BLUE , LE is RED and BL is green

>gi|83641890|ref|GS00345|(Probeset5.2),

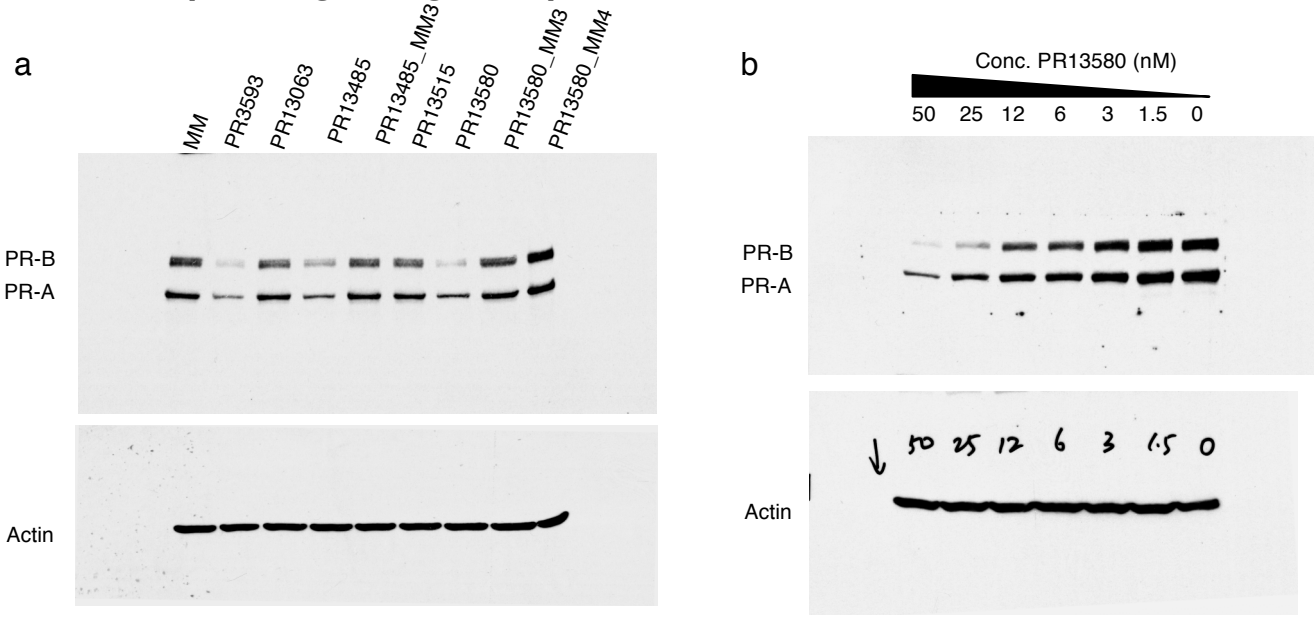
caactttagatttataatctaatcttaactattttataaatagctttttcattattgttttattcagtatatctcataat
tcaaaaagttgttttttttccctttgtatatacttaaaaacattatttaatggggagggtagttaggctagagtcataaaa
atgagataaatttggttagtccaatctgtcccatgcagaaactatttcaggtacagaaactgaaaaattgtggctattca
gattgagaattttgttgattaatatcctatttttttcacttgtaacagtctatccttgacctcctttgctgaaaaaattt
actactttgataattactatgctgctgaatttcccatgacctaaaatattaaaaattatttttactagtttatataatat
aaaagtttaggccaggtgcagtggttcacacctgtaatcccagaactttgggaggccaaggcaggcagatcacaaggtca
ggagattgagaccatcctggccaacatggtgaaaccctgtctctactaaaaatacaaaaattagctgggcctggtggtgt
acacctgtagtcccagctacttgggaggctgaggcaagagaatcgcttgagcccgggaagtggaggttgcagtgagccga
aatcgtgccactgcactccagcctgggtgacagagcaagactccatctcaaaaaaaaaaaaaaaaaagtttaatatctttt
gtaccttaatcatttatgtattaaggaaaaagccacttaatttggatgaaacaccaaaatagtgcttgctatgtcacac
actgctctaagtttatttttcttatactgatttcatccttataacaaccctatgaaacaagtactattgttatatccatt
ttacagataagaaaactaaggcagagagtgaaagtatgttgcaaggtcacatagttagcaagtggtacagcctgaaatga
gattgagaaggtgtagctccagagcatatattcttactatactgctttcttatatcacagcataattacgggtagctgac
ataattcagacataaaaatactttgctttttcatagagtacctaagttgtttaagctttatgaattccaaattgctcaca
aataactggtcatggatctctgcatattatagttatttctacttgtttggattcaatcatttataagatgggattaaactt
catctgattaggtaaatacaacctcctgccattaaaatatagctttaaaaaacttgatctttgatattaagtatttattc
atgattagtgtagagagagctata

Supplementary Figure S5. Full length gels..

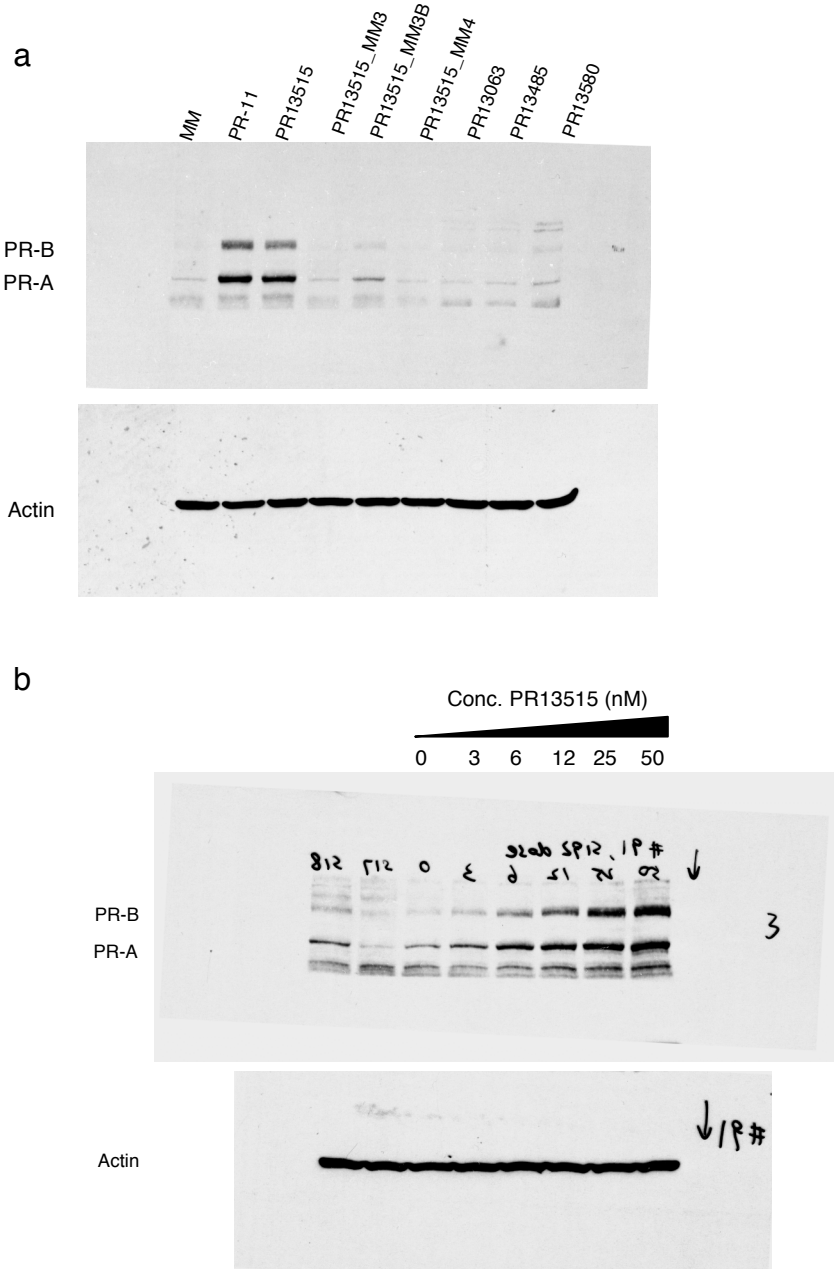
Gels corresponding to Figure 1 parts d and e



Gels corresponding to Figure 2 parts a and b



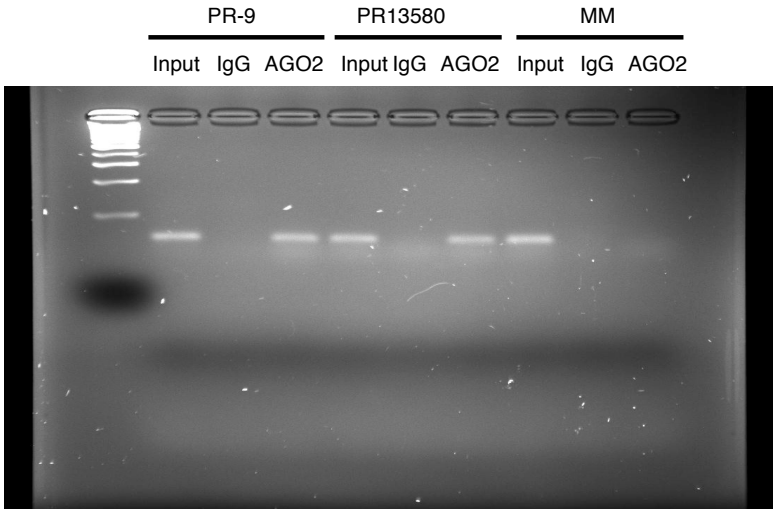
Gels corresponding to Figure 3 parts a and b



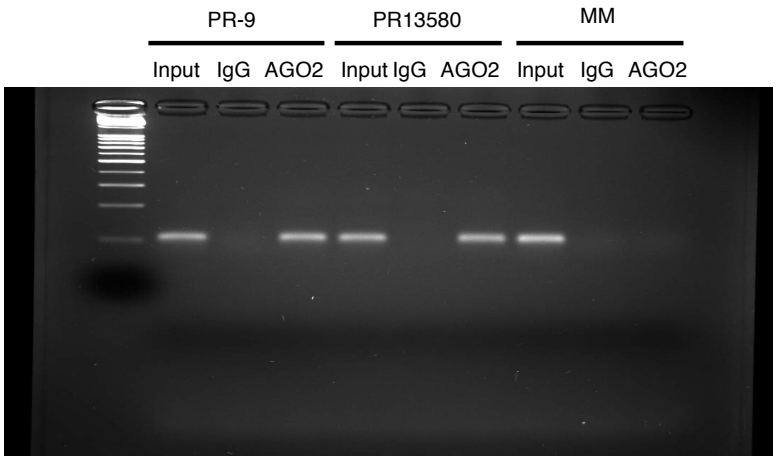
Supplementary Figure 5. Full length gels (cont.).

Gels corresponding to Figure 5 parts a,b,c,e and g.

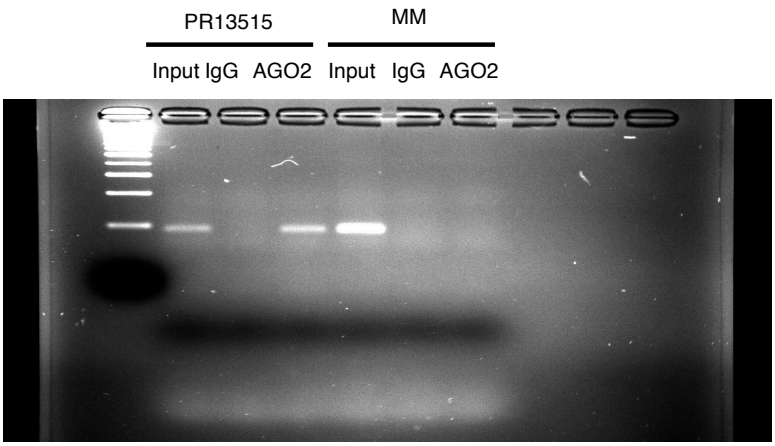
a & g
amplification of 3' noncoding transcript



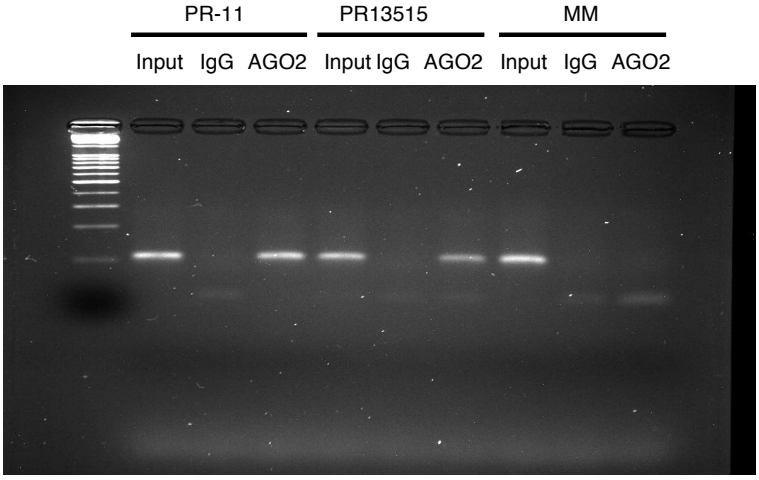
c & e
Amplification of 5' noncoding transcript



b
Amplification of 3' noncoding transcript

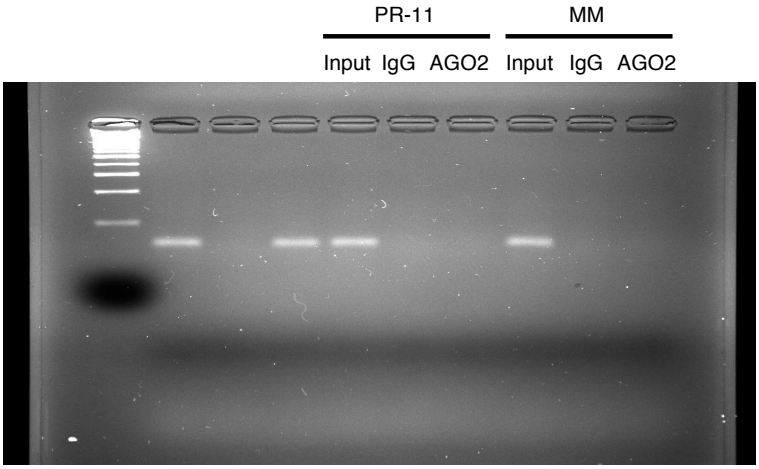


d & f
amplification of 5' noncoding transcript

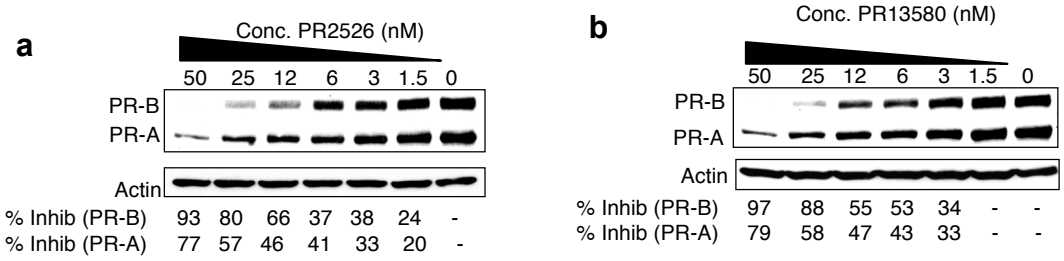


Gels corresponding to Figure 5 parts d,f, and h.

h
amplification of 3' noncoding transcript

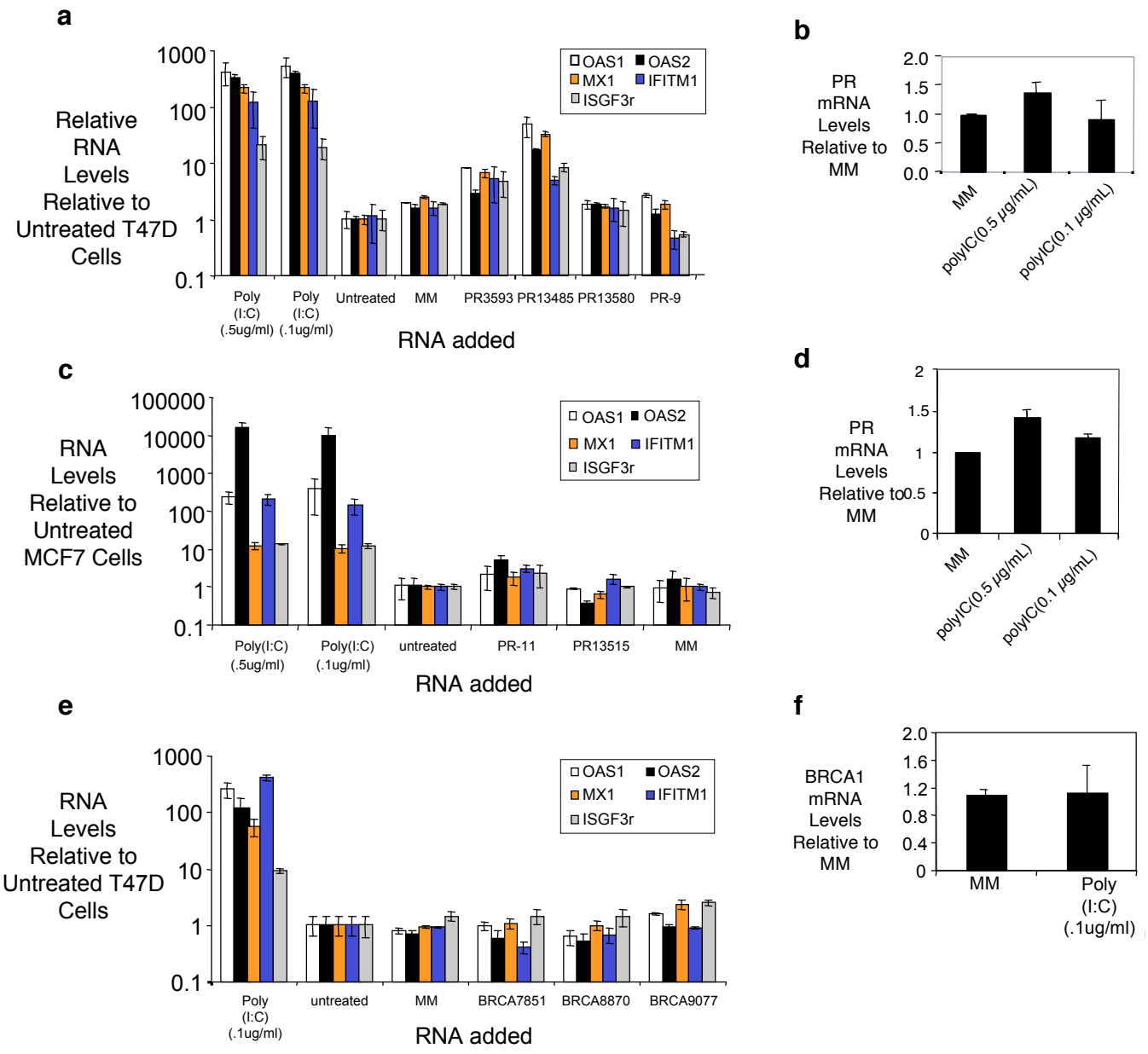


Supplementary Figure S6



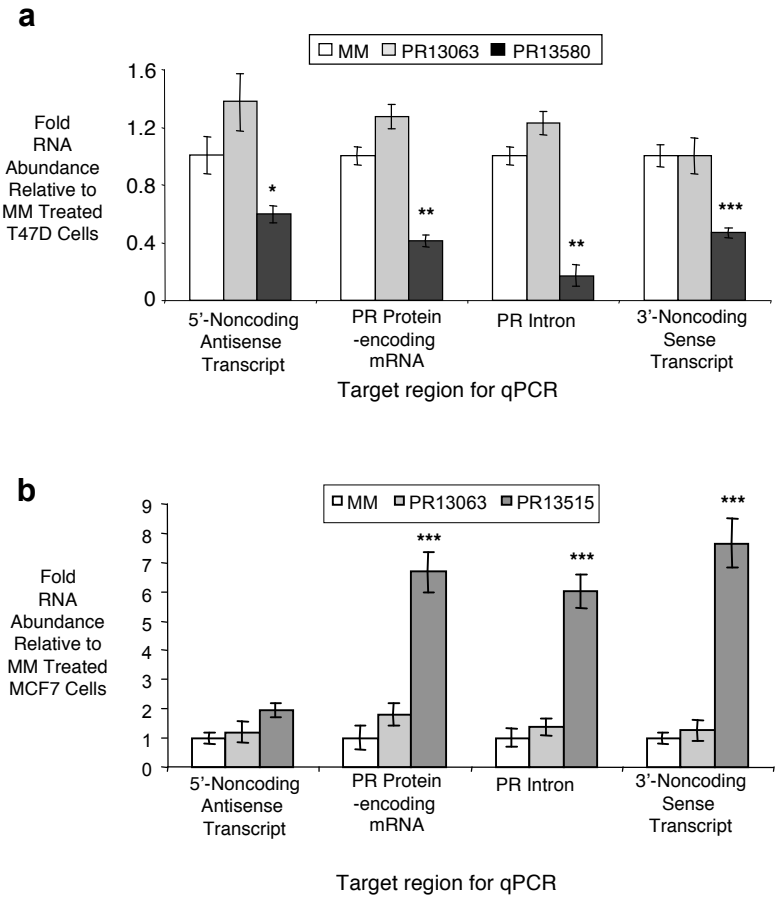
Supplementary Figure S6. **Data for IC_{50} values for siRNA PR2526 and 3' agRNA PR13580.** Western analysis of protein levels. (**a**) Inhibition of PR expression upon addition of increasing concentrations of PR2526, a duplex RNA complementary to PR mRNA (the gel is representative of three replicates). (**b**) Inhibition of PR expression upon addition of increasing concentrations of 3' agRNA PR13580 (the gel is representative of two replicates). Data was quantified and used to calculate IC_{50} values. Data from the dose resonse experiments was fit to the following model equation, $y=100*x^m/(n^m+x^m)$, where y is percent inhibition of PR and x is concentration of RNA duplex. m and n are fitting parameters, where n is taken as the IC_{50} value.

Supplementary Figure S7.



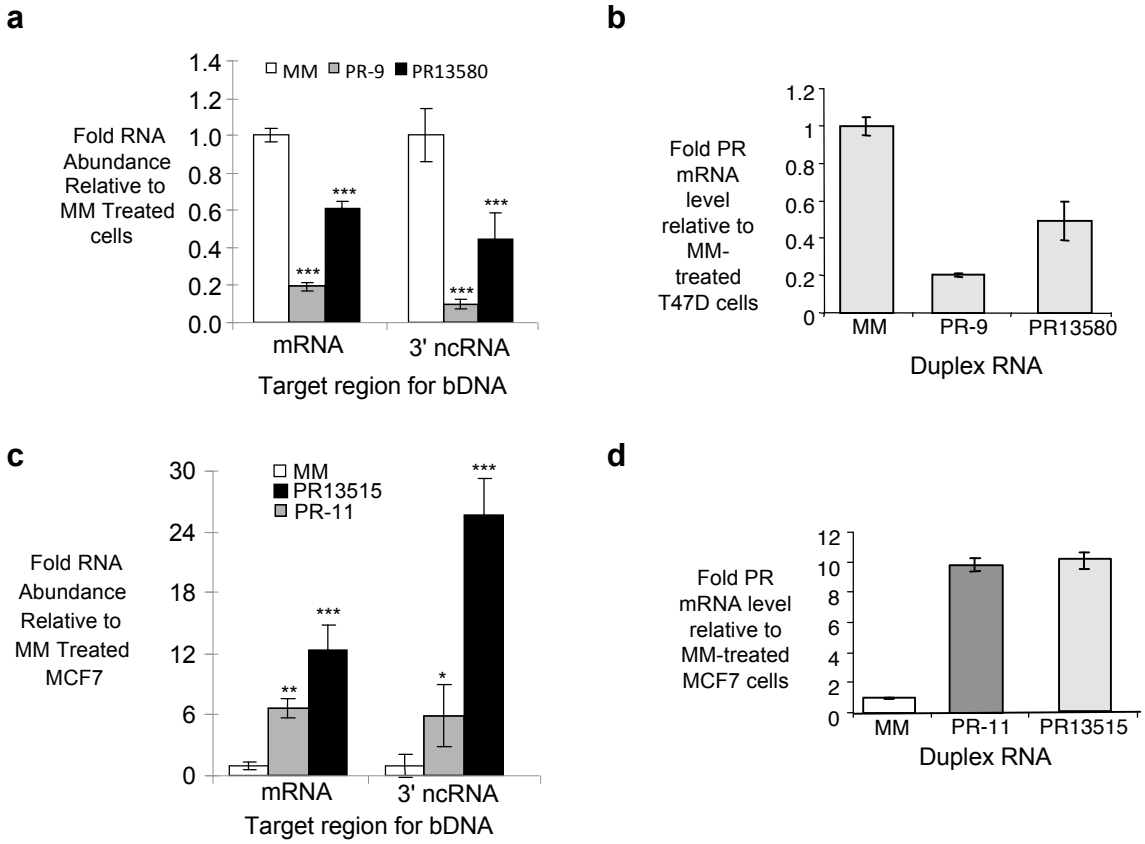
Supplementary Figure S7. PR or BRCA1 gene expression is not altered through the interferon response. All data are from transfections of T47D or MCF7 breast cancer cells. (**a**) qPCR measurement of mRNA levels of various interferon-responsive genes in T47D cells upon addition of poly (I:C) RNA or duplex RNAs after a 72 hour incubation with transfected nucleic acid. (**b**) qPCR measurement of PR mRNA levels 72 hours after addition of poly (I:C) RNA. (**c**) qPCR measurement of mRNA levels of various interferon responsive genes in MCF7 cells upon addition of poly (I:C) RNA or duplex RNAs after a 72 hour incubation with transfected nucleic acid. (**d**) qPCR measurement of PR mRNA levels in MCF7 cells upon addition of poly (I:C) RNA for 72 hours. (**e**) qPCR measurement of mRNA levels of various interferon-responsive genes in T47D cells upon addition of poly (I:C) RNA or duplex RNAs after a 72 hour incubation with transfected nucleic acid. (**f**) qPCR measurement of BRCA1 in T47D cells upon addition of poly (I:C) RNA for 72 hours. Data is from triplicate independent experiments.

Supplementary Figure S8



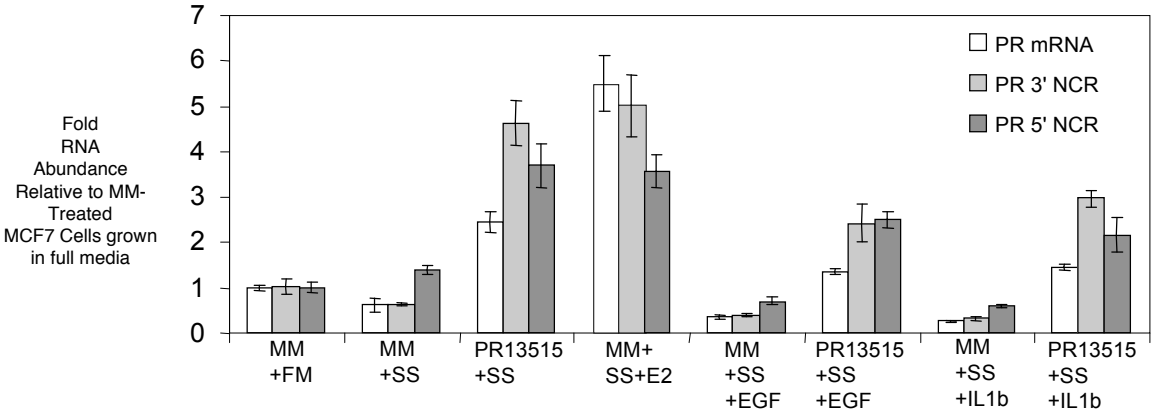
Supplementary Figure S8. Measurement of transcript levels at the PR locus in T47D or MCF7 cells upon treatment with inactive RNA PR13063, inhibitory 3' agRNA PR13580, or activating 3' agRNA PR13515. PR13063 is a duplex RNA that does not modulate PR expression when tested in MCF7 or T47D cells (**Figures 2 and 3**). PR13580 inhibits PR expression in T47D cells (**Figure 2**) and PR13515 activates PR expression in MCF7 cells (**Figure 3**). qPCR showing levels of transcripts are unaffected by inactive 3' agRNA PR13063 in (**a**) T47D cells and (**b**) MCF7 cells. T47D cells were transfected at 50 nM. MCF7 cells were transfected at 25 nM. All error bars are standard deviation. *p*-values were calculated using the two tailed unpaired Student's T-test with equal variances. ***$p<0.005$, **$p<0.01$, *$p<0.05$ as compared to cells treated with duplex RNA MM (mismatch control). Data is from triplicate independent experiments.

Supplementary Figure S9



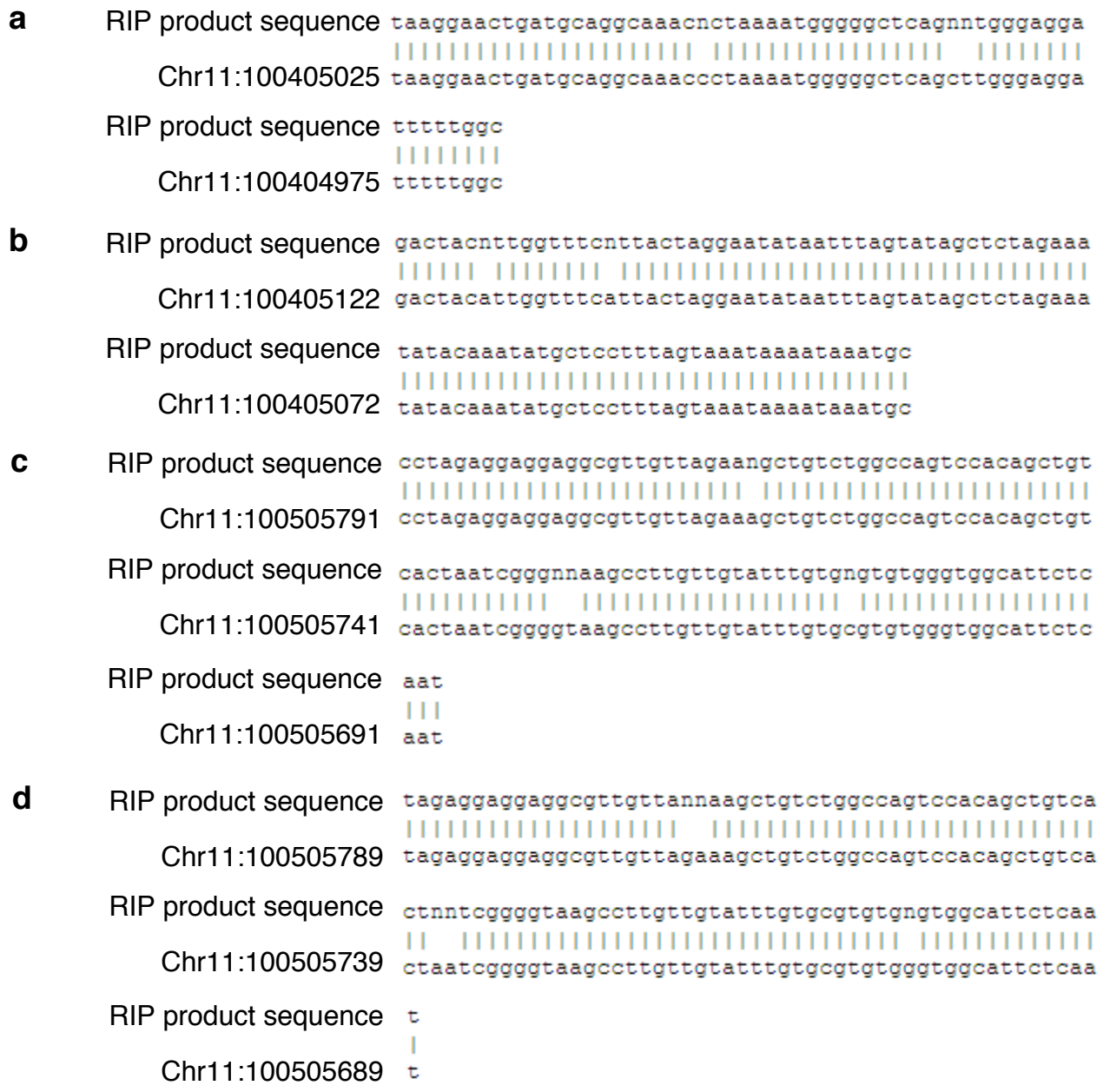
Supplementary Figure S9. Effect of agRNAs on transcript levels measured by the bDNA assay. bDNA assay showing effects of adding inhibitory agRNAs PR-9 or PR13580 on levels of (**a**) PR mRNA or the 3' noncoding transcript in T47D cells. (**b**) qPCR performed in parallel showing effects of adding inhibitory agRNAs PR-9 or PR13580 on levels of PR mRNA in T47D cells. (**c**) bDNA assay showing effects of adding activating agRNAs PR-11 or PR13515 on levels of PR mRNA or the 3' noncoding transcript in MCF7 cells. (**d**) qPCR performed in parallel showing effects of adding activating agRNAs PR-11 or PR13515 on levels of PR mRNA in T47D cell. agRNAs were transfected at 25 nM. All error bars are standard deviation. *p*-values were calculated using two way ANOVA (Graphpad Prism 4 software). ***$p<0.005$, **$p<0.01$, *$p<0.05$ as compared to cells treated with duplex RNA MM. Data is an average of four experiments.

Supplementary Figure S10.



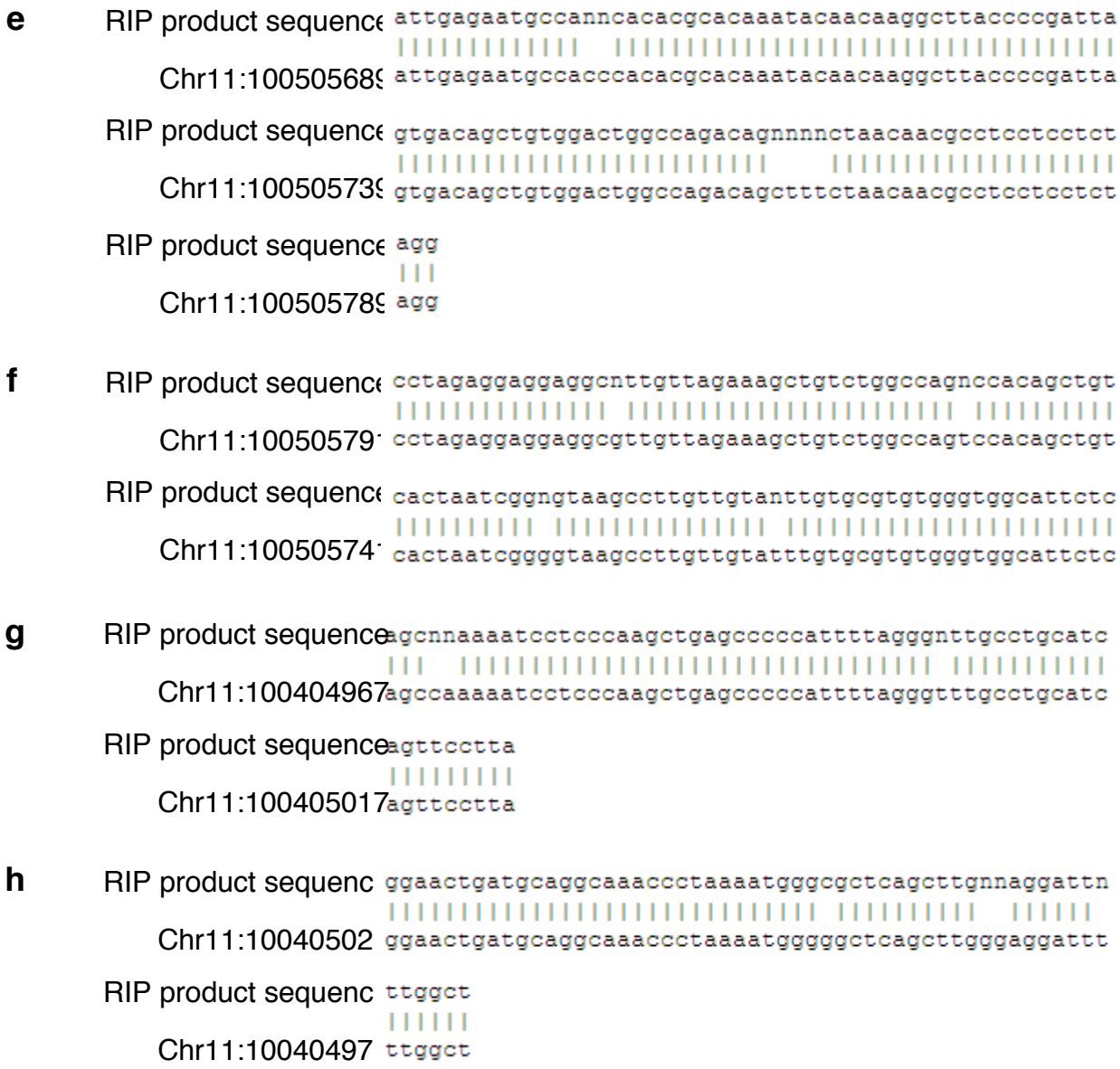
Supplementary Figure S10. Effect of combining physiologic stimuli with addition of agRNAs on expression of PR mRNA, the 3' noncoding RNA, and the 5' noncoding RNA. Data from **Figure 5D** plotted without PR13515+FM and MM+SS+PR13515 data to allow easier comparison of other conditions. SS: serum-stripped media. FM: full media. E2: 17ß estradiol treatment (100 nM). IL1ß: interleukin 1ß treatment (10 ng/mL). EGF: epidermal growth factor treatment (100 ng/mL). Data is derived from triplicate independent experiments.

Supplementary Figure S11.



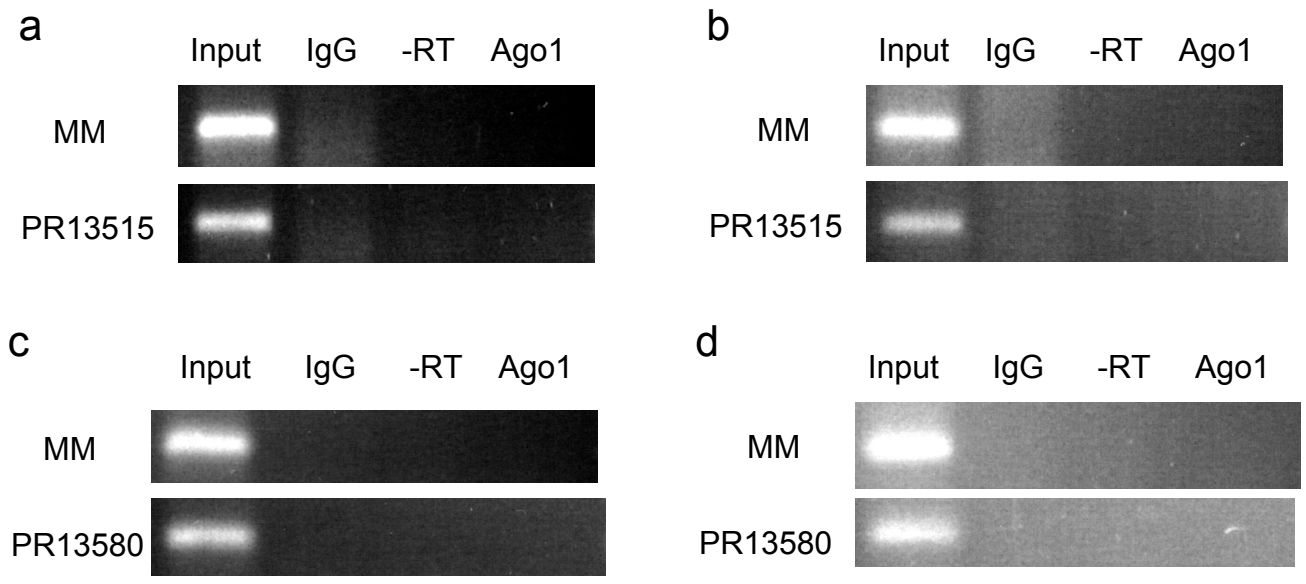
Supplementary Figure S11. Characterization of RIP products. Sequences of the amplified products from Q-PCR after RIP using anti-AGO2 antibody from T47D or MCF7 cells treated with inhibitory or activating agRNAs shown in Figure 6. Parts **a-h** correspond to Parts **a-h** in Figure 5. Continued on next page.

Supplementary Figure 11 (cont.)



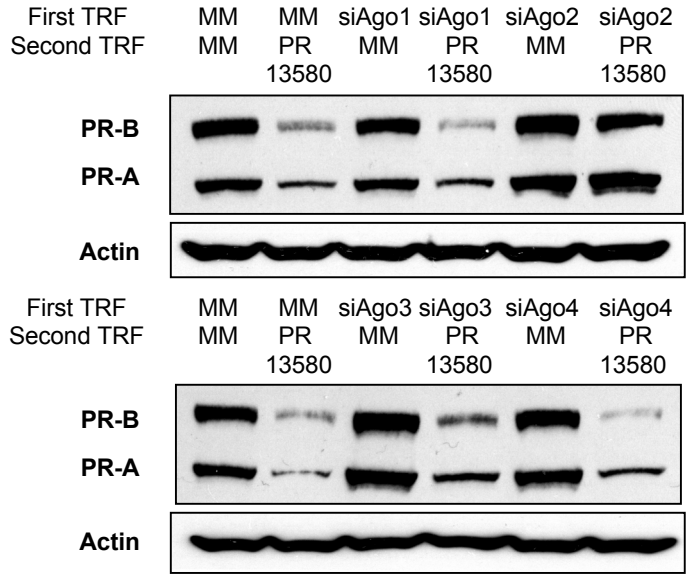
Supplementary Figure S11. Characterization of RIP products. Sequences of the amplified products from Q-PCR after RIP using anti-AGO2 antibody from T47D or MCF7 cells treated with inhibitory or activating agRNAs. Parts (**a-h**) correspond to Parts (**a-h**) in **Figure5**.

Supplementary Figure 12.



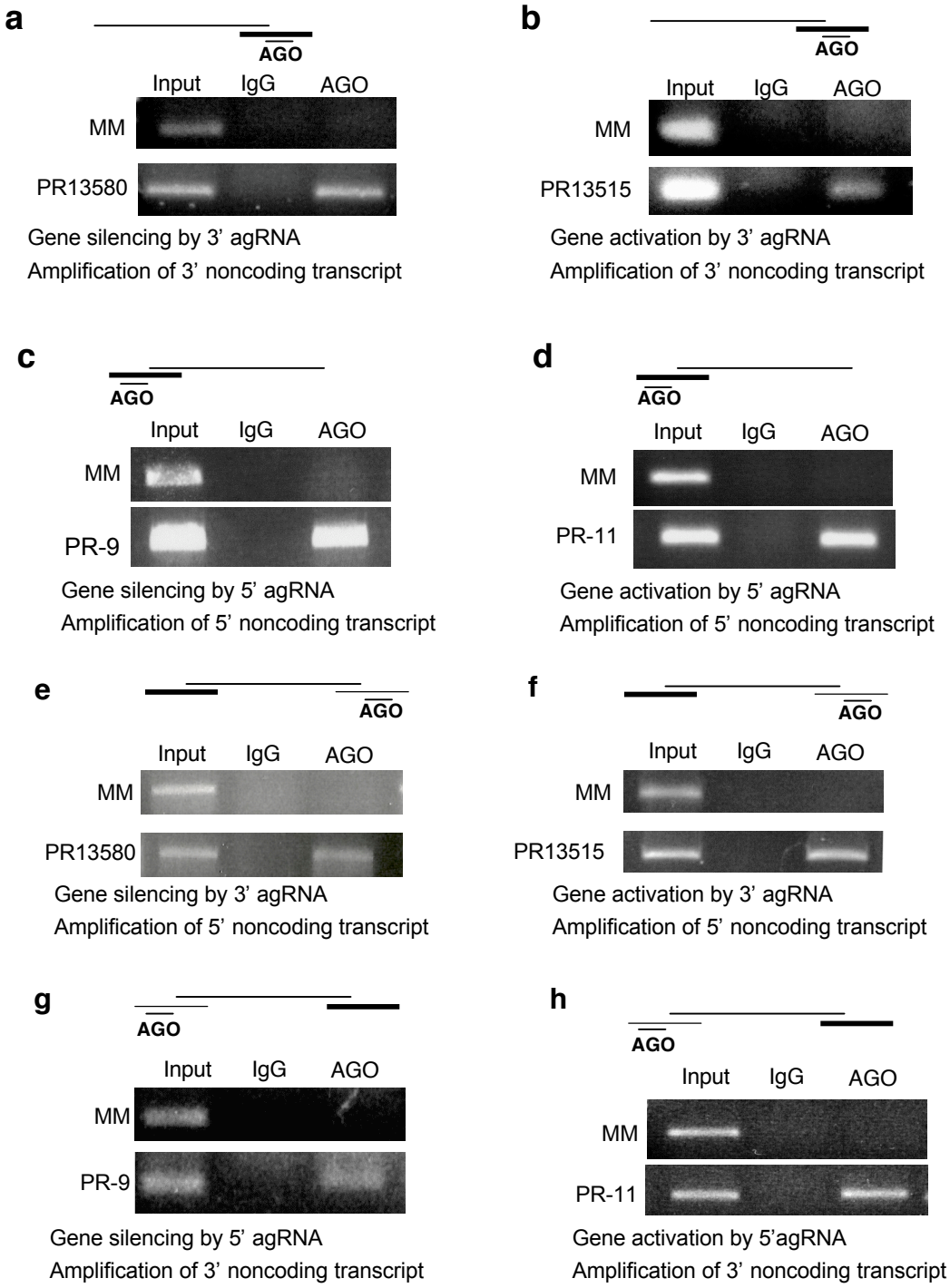
Supplementary Figure S12. RIP examining the association of Ago1 with PR 3' noncoding transcripts in nuclear lysates from T47D or MCF cells. Association of AGO1 with (**a**) PR 3' noncoding transcript or (**b**) 5' noncoding transcript after addition of PR13515 MCF7 cells. Association of AGO1 with (**c**) PR 3' noncoding transcript or (**d**) 5' PR noncoding transcript after addition of PR13580 to T47D cells. MM: negative control RNA duplex. IgG: negative control antibody. -RT: no reverse transcriptase added.

Supplementary Figure 13.



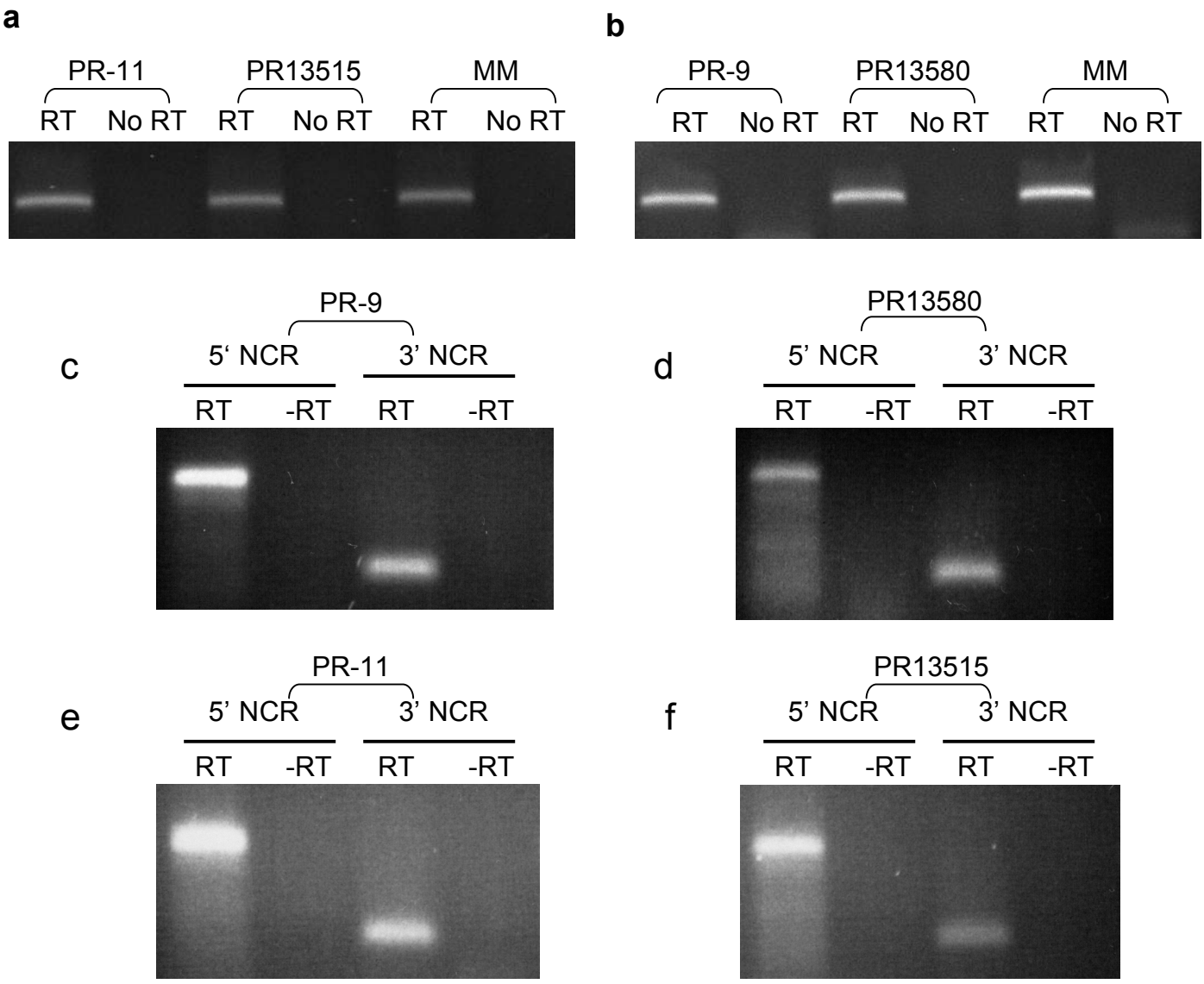
Supplementary Figure S13. Effect of silencing AGOl, AGO2, AGO3, or AGO4 expression on activity of PR13580. Duplex RNAs (complementary to the respective mRNAs) designed to inhibit expression of AGO1, AGO2, AGO3, or AGO4 were transfected into T47D cells at day 0. PR 13580 was transfected into cells at day 3, and cells were harvested for analysis of PR protein at day 7. All RNAs were added at 25 nM. Western analysis of PR protein levels after a double transfection assay using siAGO1, siAGO2, siAGO3, or siAGO4 in the initial transfection MM:mismatch-containing control duplex. TRF: transfection.

Supplementary Figure S14.



Supplementary Figure S14. Effect of 3' or 5' agRNAs on recruitment of AGO protein to the 3' or 5' noncoding transcripts at the PR locus. Unlike the experiments shown in Figure 4, which used an antibody that recognizes AGO2, these experiments used an antibody that recognizes all four AGO variants indiscriminately. Effect of adding (**a**) inhibitory RNA PR13580 to T47D cells or (**b**) activating RNA PR13515 to MCF7 cells on recruitment of AGO protein to the 3' noncoding transcript. Effect of adding (**c**) inhibitory RNA PR-9 to T47D cells or (**d**) activating RNA PR-11 to MCF7 cells on recruitment of AGO protein to the 5' noncoding transcript. Effect of adding (**e**) inhibitory RNA PR13580 to T47D cells or (**f**) activating RNA PR13515 to MCF7 cells on co-immunoprecipitaton of AGO protein with the 5' noncoding transcript. Effect of adding (**g**) inhibitory RNA PR-9 to T47D cells or (**h**) activating RNA PR-11 to MCF7 cells on co-immunoprecipitation of AGO protein with the 3' noncoding transcript. The scheme above each gel depicts PR mRNA, the 3' and/or 5' noncoding transcripts, and AGO bound agRNA. The heaviest line represents the transcript being amplified. Data is representative of four independent experiments.

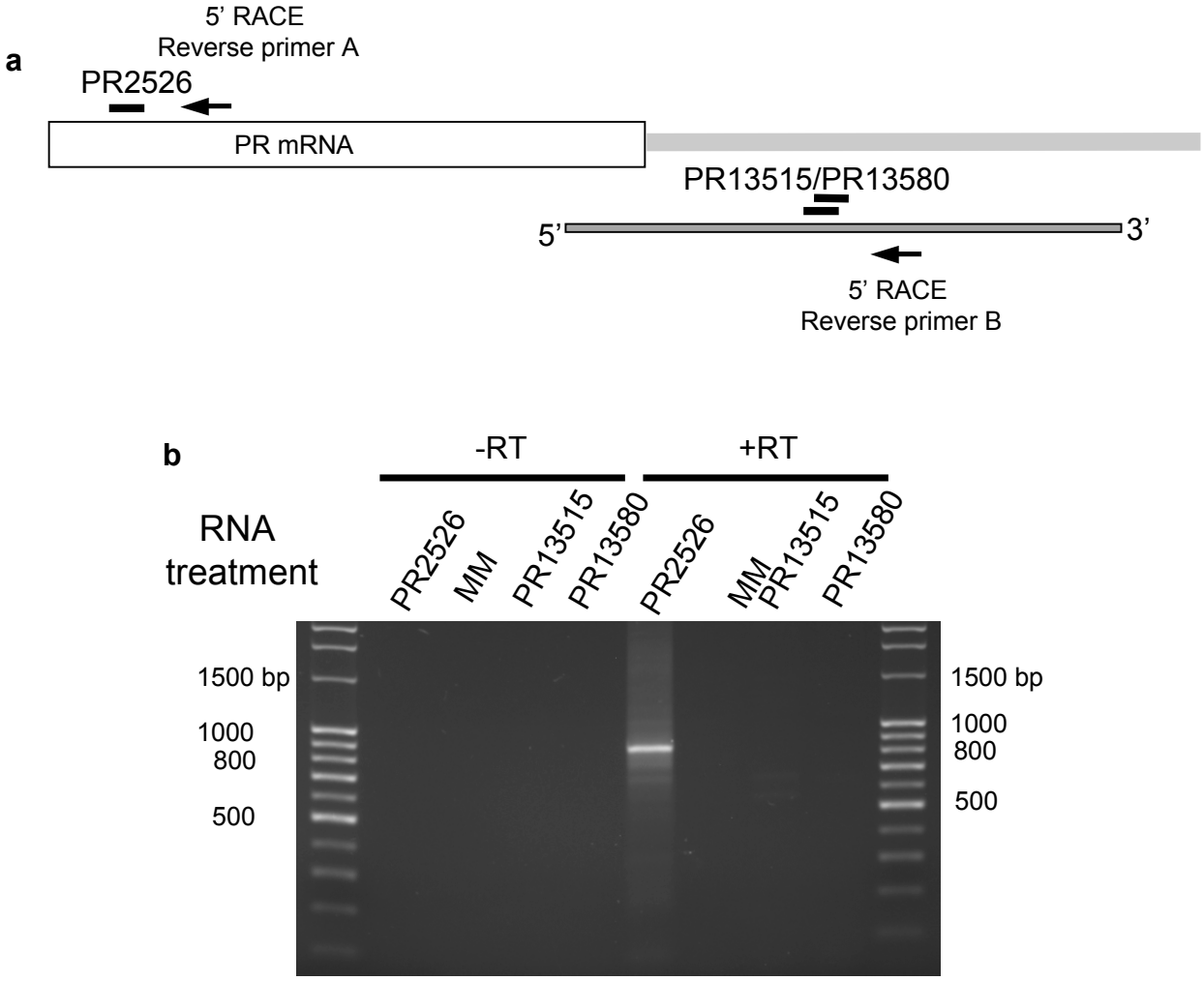
Supplementary Figure S15.



Supplementary Figure S15. The RIP assay detects RNA in the nucleus, not DNA. **When reverse transcriptase (RT) is added, amplified product is observed. When no reverse transcriptase is added (no RT), no amplified product was detected.** Detection of RNA in nuclear extracts after addition of RNAs to (**a**) MCF7 cells; (**b**) T47D cells. “Input” samples are the crude samples after Dnase treatment and prior to immunoprecipitation and are analyzed here. We chose to analyze input samples as a stringent test because the input samples contain much more genomic DNA than the corresponding sample after immunoprecipitation. If DNA can be amplified, the input samples would be much more likely to reveal that fact. The primer set covering PR promoter from -37 to +66 (relative to TSS) was used. The PCR product identities have been confirmed by sequencing. RT: Reverse transcriptase added. No RT: No reverse transcriptase added. The Input sample was 5 % of a 100 µL nuclear fraction (~250,000 cell equivalents).

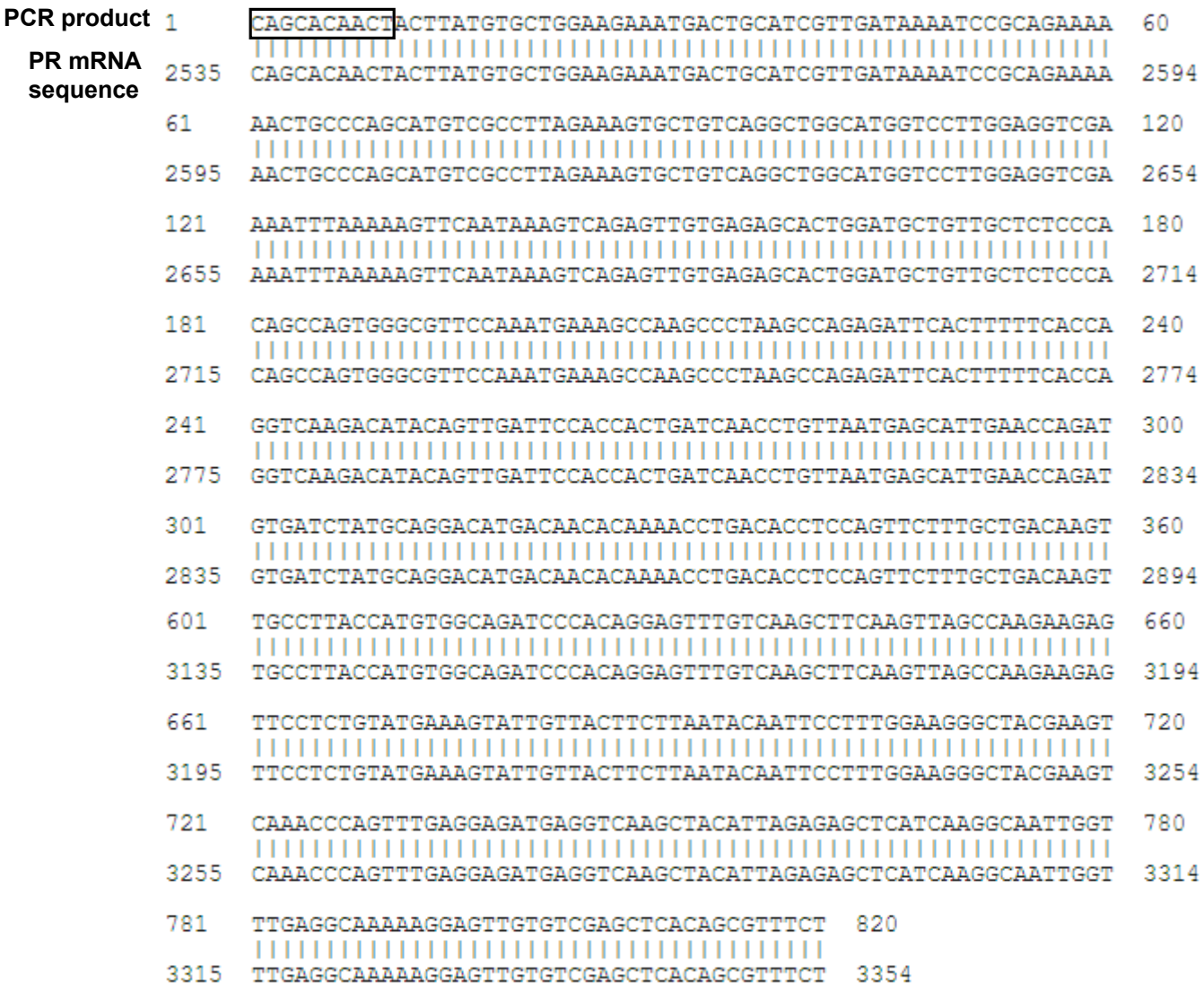
We also examined amplification of samples after immunoprecipitation using anti-AGO2 antibody. Data corresponds with data shown in Figure 4. When no reverse transcriptase (-RT) is added, no amplified product is detected. T47D or MCF7 cells were transfected with (**c**) PR-9, (**d**) PR13580, (**e**) PR-11 or (**f**) PR13515, and immunoprecipitated with anti-Ago2 antibody. The primer set identifying 5’ noncoding antisense transcript (5’ NCR) or 3’ noncoding sense transcript (3’ NCR) was used.

Supplementary Figure S16.

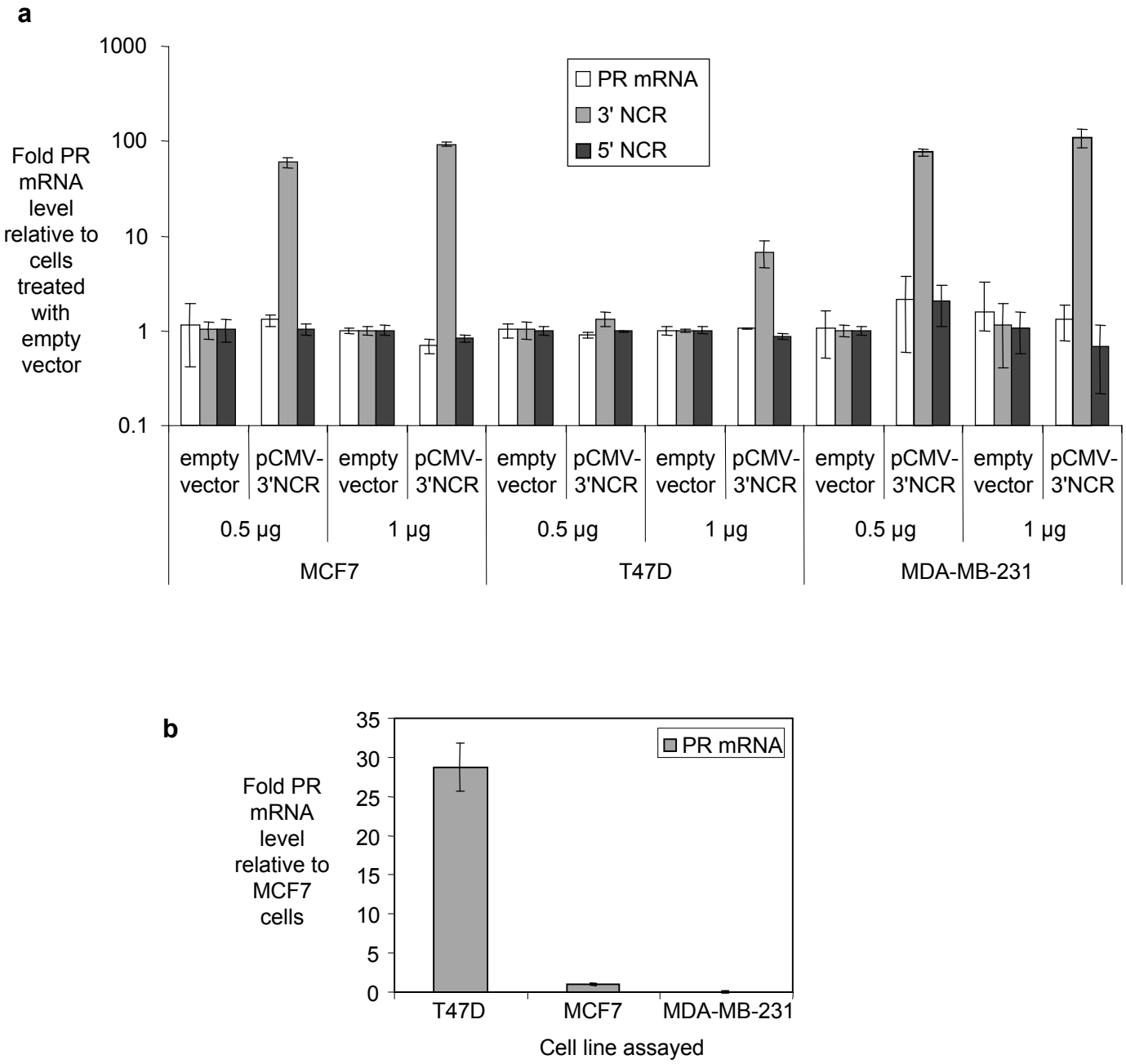


Supplementary Figure S16. **5'-RACE showing no detection of cleavage products within the 3' noncoding transcript after addition of activating agRNA PR13515 or inhibitory agRNA PR13580**. (**a**) Location of primers for 3' and 5' RACE and location of duplex RNAs PR2526, PR13515, and PR13580. (**b**) Results of 5'-RACE with primer B (5'-TCAACTCAAACTTACAGCAAGAATCCTGTTCCACTC-3') downstream from the recognition sites of PR13515 and PR13580 on 3' noncoding sense transcript didn't detect any cleavage site. As a positive control, the cleavage site by PR2526, an siRNA targeting PR mRNA, was detected with the downstream primer A (5'-AGAAACGCTGTGAGCTCGACACAACTCC-3'). (**c**) (next page) The sequence of PCR product from positive control sample treated with siRNA PR2526 was aligned to PR mRNA sequence by BLAST. The part of PR2526 sequence contained in the PCR product is highlighted.

c

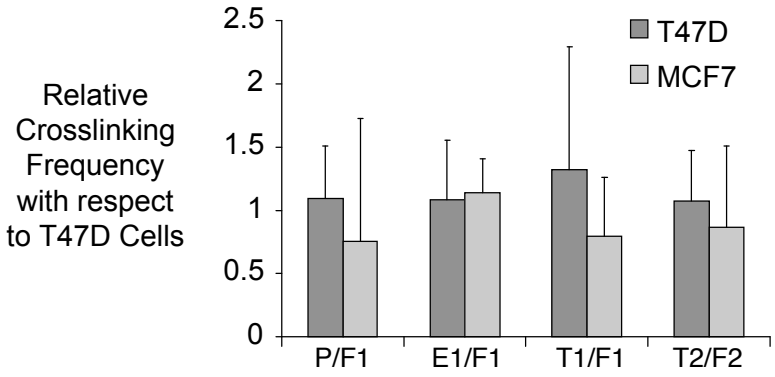

PCR product
PR mRNA sequence
1 CAGCACAACTACTTATGTGCTGGAAGAAATGACTGCATCGTTGATAAAATCCGCAGAAAA 60
2535 CAGCACAACTACTTATGTGCTGGAAGAAATGACTGCATCGTTGATAAAATCCGCAGAAAA 2594
61 AACTGCCCAGCATGTCGCCTTAGAAAGTGCTGTCAGGCTGGCATGGTCCTTGGAGGTCGA 120
2595 AACTGCCCAGCATGTCGCCTTAGAAAGTGCTGTCAGGCTGGCATGGTCCTTGGAGGTCGA 2654
121 AAATTTAAAAAGTTCAATAAAGTCAGAGTTGTGAGAGCACTGGATGCTGTTGCTCTCCCA 180
2655 AAATTTAAAAAGTTCAATAAAGTCAGAGTTGTGAGAGCACTGGATGCTGTTGCTCTCCCA 2714
181 CAGCCAGTGGGCGTTCCAAATGAAAGCCAAGCCCTAAGCCAGAGATTCACTTTTTCACCA 240
2715 CAGCCAGTGGGCGTTCCAAATGAAAGCCAAGCCCTAAGCCAGAGATTCACTTTTTCACCA 2774
241 GGTCAAGACATACAGTTGATTCCACCACTGATCAACCTGTTAATGAGCATTGAACCAGAT 300
2775 GGTCAAGACATACAGTTGATTCCACCACTGATCAACCTGTTAATGAGCATTGAACCAGAT 2834
301 GTGATCTATGCAGGACATGACAACACAAAACCTGACACCTCCAGTTCTTTGCTGACAAGT 360
2835 GTGATCTATGCAGGACATGACAACACAAAACCTGACACCTCCAGTTCTTTGCTGACAAGT 2894
601 TGCCTTACCATGTGGCAGATCCCACAGGAGTTTGTCAAGCTTCAAGTTAGCCAAGAAGAG 660
3135 TGCCTTACCATGTGGCAGATCCCACAGGAGTTTGTCAAGCTTCAAGTTAGCCAAGAAGAG 3194
661 TTCCTCTGTATGAAAGTATTGTTACTTCTTAATACAATTCCTTTGGAAGGGCTACGAAGT 720
3195 TTCCTCTGTATGAAAGTATTGTTACTTCTTAATACAATTCCTTTGGAAGGGCTACGAAGT 3254
721 CAAACCCAGTTTGAGGAGATGAGGTCAAGCTACATTAGAGAGCTCATCAAGGCAATTGGT 780
3255 CAAACCCAGTTTGAGGAGATGAGGTCAAGCTACATTAGAGAGCTCATCAAGGCAATTGGT 3314
781 TTGAGGCAAAAAGGAGTTGTGTCGAGCTCACAGCGTTTCT 820
3315 TTGAGGCAAAAAGGAGTTGTGTCGAGCTCACAGCGTTTCT 3354

Supplementary Figure S17.



Supplementary Figure S16. **Effect of overexpressing the 3' noncoding transcript on expression of transcripts at the PR locus.** (**a**) qPCR analysis showing the effect of the 3' noncoding transcript on PR mRNA and the 5' antisense noncoding transcript. (**b**) Relative expression levels of PR in MCF7, T47D, and MDA-MB-231 cells. 3'NCR: 3' noncoding RNA. 5'NCR: 5'noncoding RNA. All experiments were done in triplicate except for the T47D and MDA-MB-231 empty vector transfections which were done in duplicate.

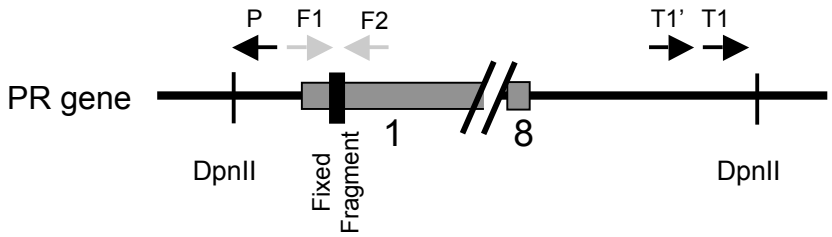
Supplementary Figure S18.



Supplemental Figure S18. 3C analysis of the PR locus: Comparing looping in MCF7 versus T47D cells. Crosslinking frequencies are relative to detection of a fixed fragment within genomic DNA for untreated T47D and MCF7 cells. Primer E1 amplifies PR exon 1. Primers T1 and T2 amplify sequences beyond the terminus of PR mRNA. Primers P, E1, F1, F2, T1, and T2 are described in **Figure 7A**. Data are the result of triplicate independent experiments.

Supplementary Figure S19.

a



b

T47D - cells

P / T1'

```
TCCNNAAGAA CCTGCTATTG AGAGTAGCAT TCAGAATAAC GGGTGGAAAT  50
GCCAACTCCA GAGTTTCAGA TCAAAAACAG GATGTAGCTA AGAAACTGGC  100
CCAAATCAGC TAGGACTAGA AATTATAATG CATCTGCAGG CTATGAGACA  150
CTGCACCAGC AC
```

F2 / T1

```
        GG NNCGGANTTC TGCTGGCTCC GTACTGCGGG CGACAGTCAT  50
CTCCGAAGAT CAAAAACAGG ATGTAGCTAg GACTGGCCCA AATCAGCTA
```

c

MCF7 - cells

P / T1'

```
TGTGNTGGTG NAGTGTCTCA TAGCCTGCAG ATGCATTATA ATTTCTAGTC  50
CTAGCTGATT TGGGCCAGTT TCTTAGCTAC ATCCTGTTTT TGATCTGAAA  100
CTCTGGAGTT GGCATTTCCA CCCGTTATTC TGAATGCTAC TCTCAATAGC  150
AGGTTCTTTG GGATGGAA
```

F2 / T1'

```
TTGAGGNNCG GANTTCTGCT GGCTCCGTAC TGCGGGCGAC AGTCATCTCC  50
GAAGATCTCA GATCAAAAAC AGGATGTAGC TAAGAAACTG GCCCAAATCA  100
GCTAGGACTA GAAATTATAA TGCATCTGCA GGCTATGAGA CACTGCACCA  150
GCAC
```

F2 / T1

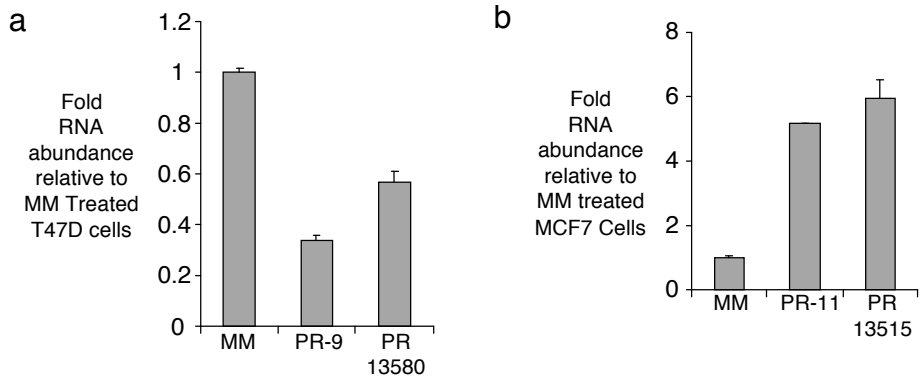
```
CGGANTTCTG CTGGCTCCGT ACTGCGGGCG ACAGTCATCT CCGAAGATCA  50
CCAGGACTTG TTTTCATGAC CCTGCTGATC AAAAACAGGA TGTAGCTAAG  100
AAACTGGCCC AAATCAGCTA
```

F1 / T1

```
TTTACACCCG AGGCGGCTAG TNNTCCCGCA CTACTGGGAT CAAAAACAGG  50
ATGTAGCTAA GAAACTGGCC CAAATCAGCT A
```

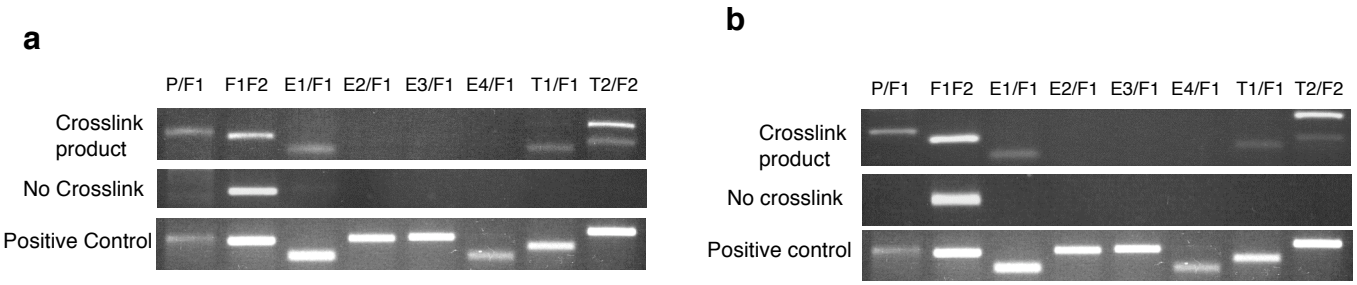
Supplementary Figure S19. PCR product was sequenced for some combinations of primers detecting gene loops between the 5' and 3' ends of PR. (**a**) Product was obtained using four possible combinations of primers between two cut sites at the PR 5' end and one cut site and the PR 3' end. (**b**) Two PCR products were obtained in T47D cells and correctly aligned with their genomic targets at the 5' and 3' ends of PR. Blue text aligned with the 5' end and red text aligned with the 3' end. The DpnII cut site used is bold and underlined. (**C**) Four PCR products were obtained in MCF7 cells and correctly aligned with their genomic targets at the 5' and 3' ends of PR. Coloring is as in part B. Note for F2 / T1 the use of a nearby DpnII cut site for ligation. The alternative site is underlined but not bold. Alignments were made using BLAT from UCSC genome brower at genome.ucsc.edu.

Supplementary Figure S20.



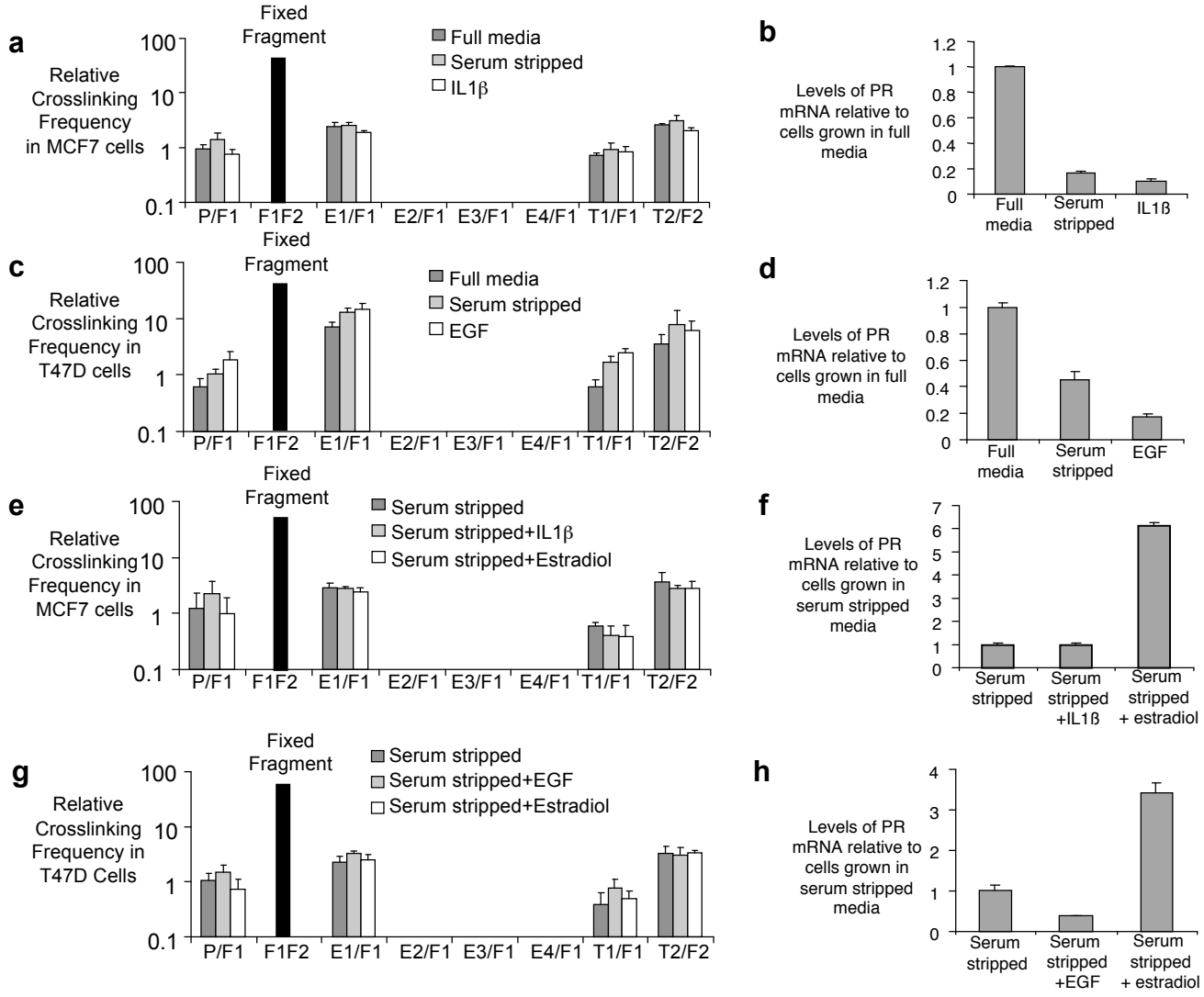
Supplementary Figure S20. Modulation of PR expression in cells used for 3C analysis. (**a**) T47D cells. (**b**) MCF7 cells. Data corresponds to data shown in Figure 6 b and c.

Supplementary Figure S21.



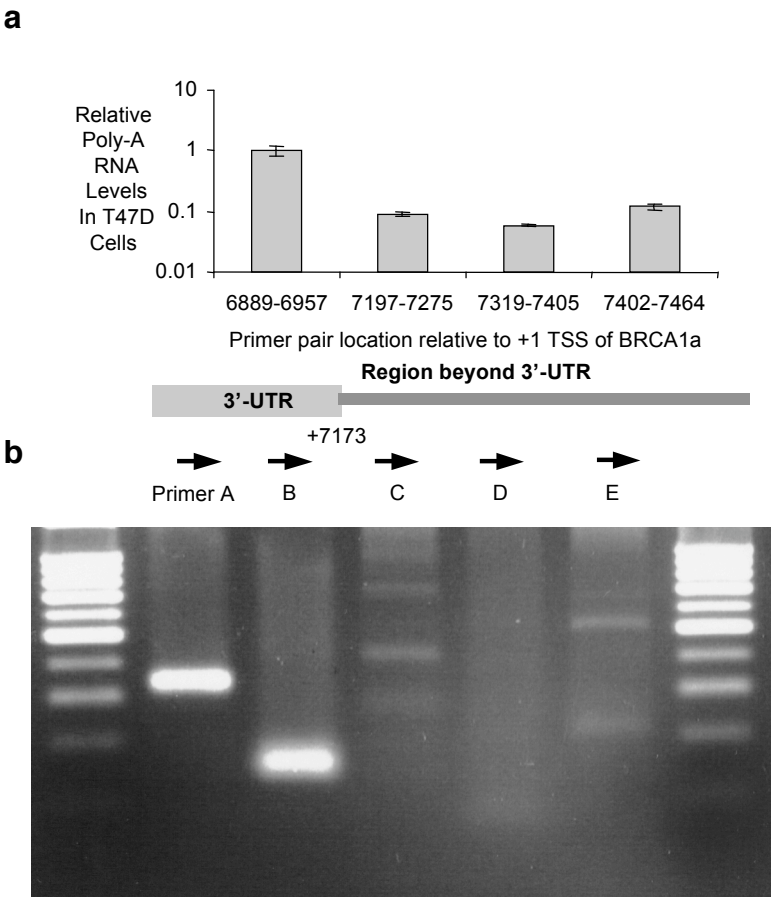
Supplementary Figure S21. Analysis of 3C products corresponding to data in Figure 7b and 7c. (**b**) Gel electrophoresis showing 3C products from T47D cells. (**b**) Gel electrophoresis showing 3C cata from MCF7 cells. Data corresponds to data shown in Figure 7 b and c. Primers E1, E2, E3, and E4 amplify sequences within PR exons 1-4. Primers T1 and T2 amplify sequences beyond the terminus of PR mRNA. F1/F2= Fixed fragment. The fixed fragment is a normalization control derived from genomic DNA by primers complementary to sequences within exon 1.

Supplementary Figure S22.



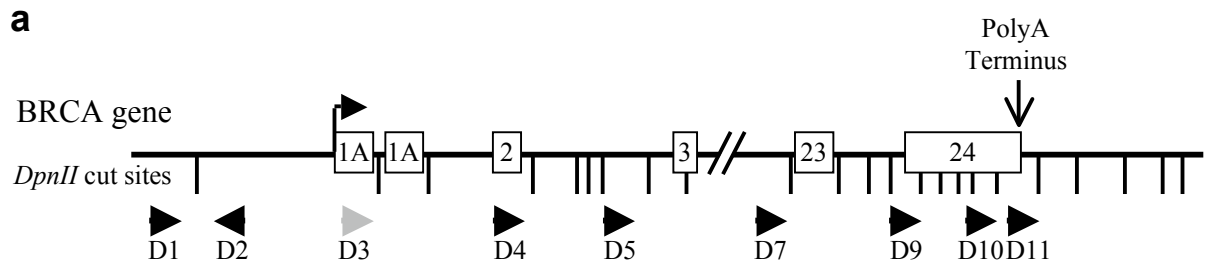
Supplementary Figure S22. **Effect of physiologic stimuli on gene looping.** 3C analysis for MCF7 (**a**) or T47D (**c**) cells upon treatment with full media, serum-stripped media, IL1ß (MCF7, 10 ng/mL), or epithelial growth factor (EGF, T47D 100 ng/mL). 3C analysis for MCF7 (**e**) or T47D (**g**) cells upon treatment with serum-stripped media or serum-stripped media supplemented with IL1ß (MCF7, 10 ng/mL), or epithelial growth factor (EGF, T47D, 100 ng/mL), or estradiol (both T47D or MCF7, 100 nM). Parts (**b**), (**d**), (**f**), and (**h**) show the effects of treatments on PR mRNA levels as monitored by qPCR. FF= Fixed Fragment. As described by Baylin and coworkers,the "Fixed Fragment" is a normalization control derived from genomic DNA by primers complementary to sequences within exon 1. The bar represents performance of the normalization control (a demonstration that genomic DNA can be amplified by primers F1 and F2), not an absolute value. Primer combinations are those shown in **Figure 6a**. Data are from 3-4 independent experiments.

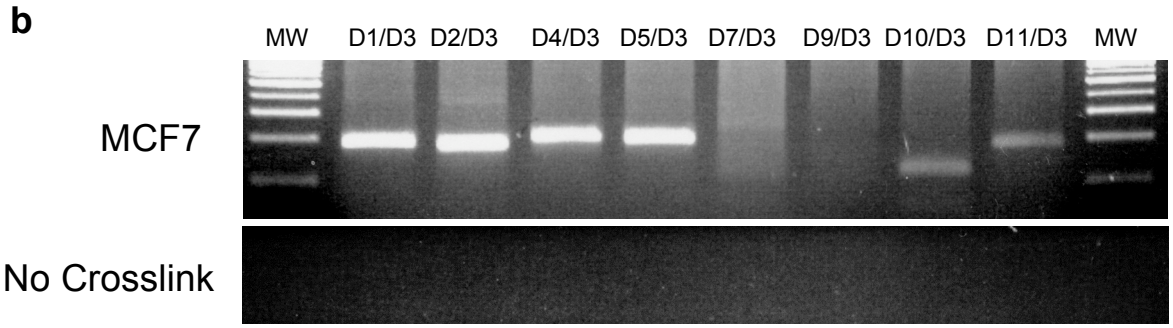
Supplementary Figure S23.

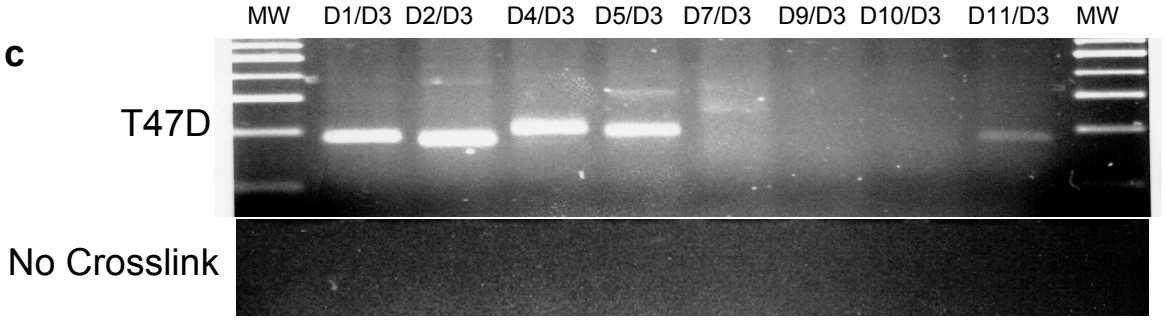


Supplementary Figure S23. Measurement of poly-A RNA levels in T47D cells using different primer sets surrounding the 3' terminus of BRCA1 mRNA. (**a**) qPCR showing levels of poly-A RNA detected from the BRCA1transcription start site past the most downstream annotated terminus of PR BRCA1. (**b**) 3' RACE using primers upstream and downstream of the termination site for BRCA1 revealed no evidence for a longer BRCA1 messenger RNA transcript. Primers A and B detected the known BRCA1 termination and primers C, D, and E downstream of that site did not detect amplified product. The additional faint bands derived from use of primers C could not be amplified in sufficient quantities to enable cloning, sequencing, and identification. They are either nonspecific products or are at too low an abundance to be detected by RACE. Data in (**a**) are from triplicate independent experiments.

Supplementary Figure S24.







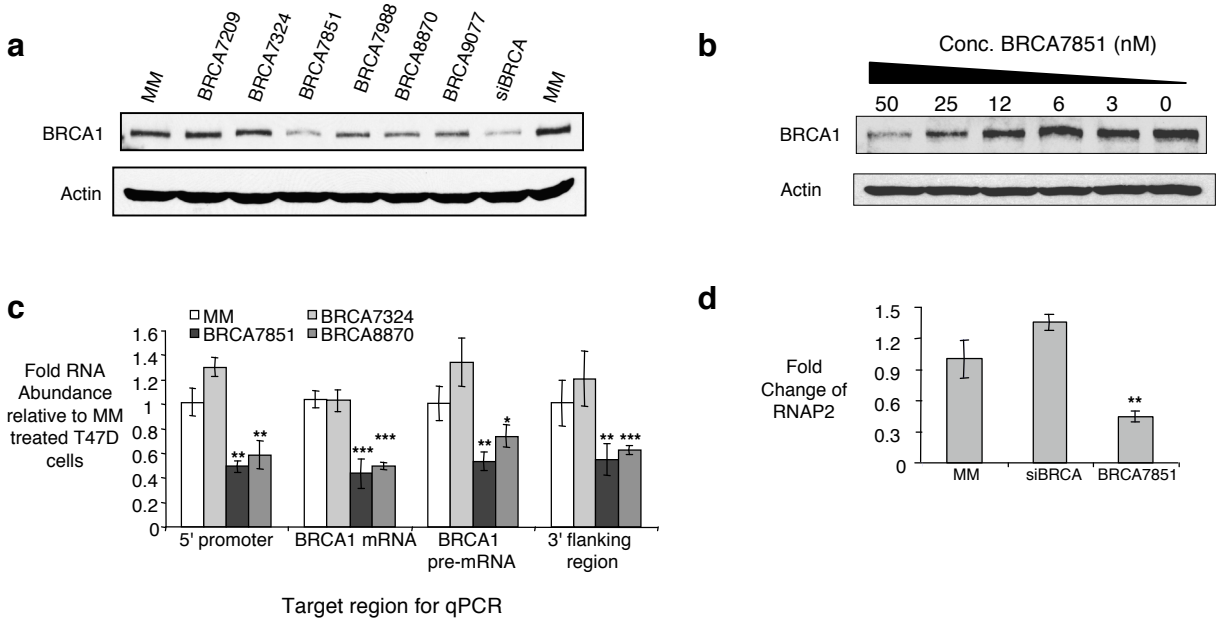
d

D3 / D11

```
tctttgacgggggnaggggcggaacctgagaggcgtaaggcgttgtgaa  50
ccctggggaggggggcagtttgtaggtcgcgagggaagcgctgaggatct  100
tagtcctagaactgcaaggacccagagcctctagaagggaacacgccctg  150
cgaacacctttattttggcctcatgaa
```

Supplementary Figure S24. 3C analysis of the BRCA1 gene. (**a**) Primer sets used as described by Brown and coworkers (Tan-Wong et al., 2008). Primer D3 was a fixed reverse primer. (**b**) Agarose gel analysis of 3C products in MCF7 cells. In contrast to previous results (Tan-Wong et al., 2008) we oberved product when using primer D4 but not with primer D9. (**c**) 3C analysis for T47 D cells. (**d**) 3C product D11/D3 was cloned and sequenced. The product aligns with sequences at the 5' and 3' termini of the BRCA1 gene. Control experiments using non-crosslinked samples produced no PCR product. No Crosslink: no crosslinker added to sample prior to ligation and amplification. Data are representative of duplicate independent experiments.

Supplementary Figure S25



Supplementary Figure S25. Inhibition of BRCA1 expression by agRNAs targeting sequences beyond the 3'-UTR. Western analysis showing levels of BRCA1 protein after addition of (**a**) six agRNAs complementary to sequences beyond the 3' polyadenylation site for BRCA1 mRNA and siBRCA1 complementary to BRCA1 mRNA. (**b**) increasing concentrations of BRCA7851. (**c**) qPCR analysis of RNA levels using primer sets designed to detect noncoding RNA at the BRCA1 promoter, BRCA1 mRNA, BRCA1 pre-mRNA, and noncoding RNA beyond the 3' terminus of BRCA1 mRNA. (**d**) Presence of RNAP2 at the BRCA1 transcription start site evaluated by ChIP. All error bars are standard deviation. ***p<0.005, **p<0.01, and *p<0.05 as compared to cells treated with RNA MM. p-values were calculated using the two tailed unpaired Student's T-test with equal variances. All error bars represent standard deviation. Duplex RNAs were added to cells at 25 nM unless otherwise noted. Data are from duplicate or triplicate independent experiments.