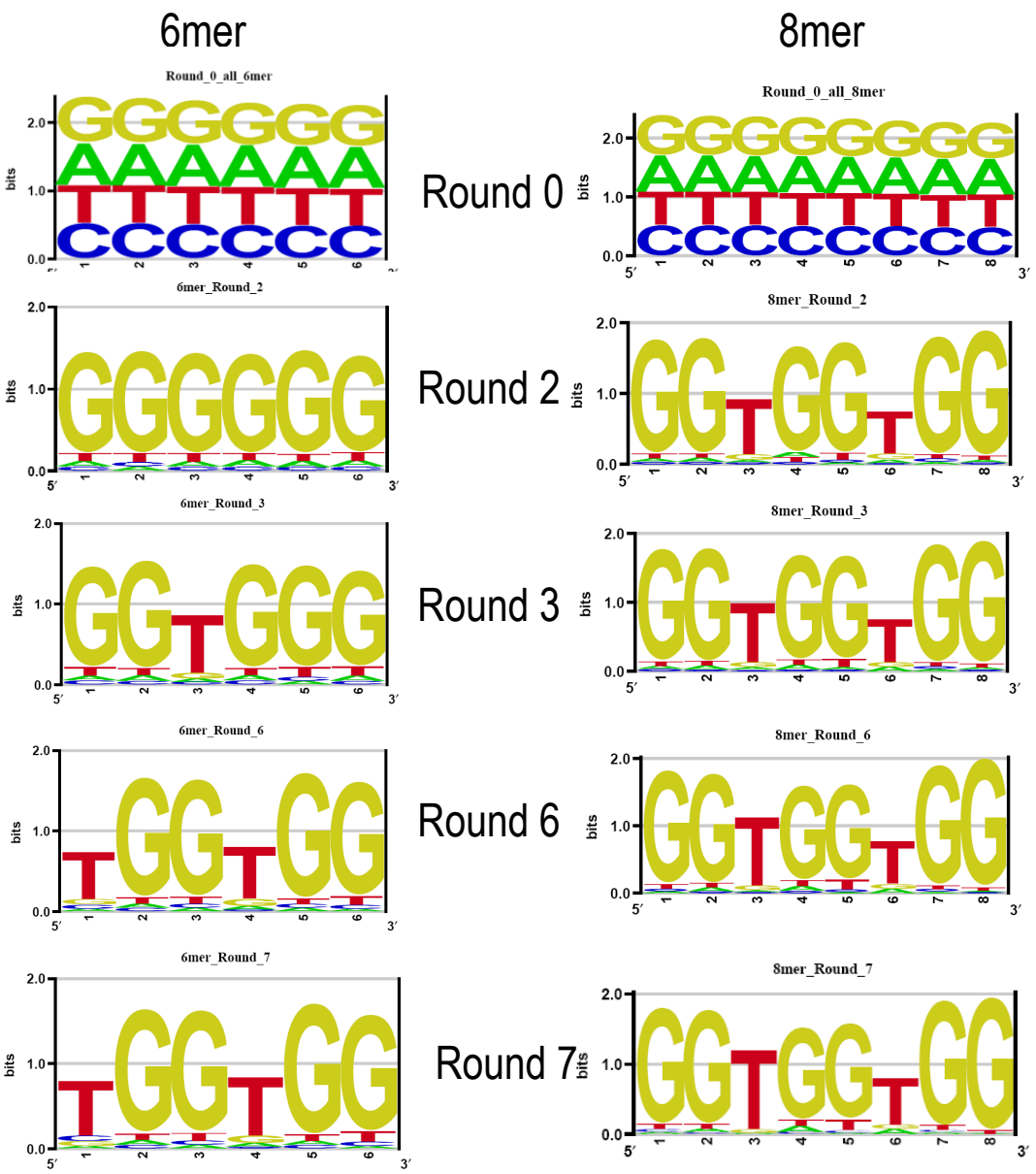
A



B

Top 12 6-mers by round				
Round 0	Round 2	Round 3	Round 6	Round 7
GGGTCA	GGGGGG	GGTGGG	TGGTGG	TGGTGG
GGTCAA	GGGGGT	TGGTGG	GGTGGG	GGTGGG
GGTCAG	GTGGGG	GTGGGG	GTGGTG	GTGGTG
AGGTCA	GGTGGG	GGGGGT	GGTGGT	GTGGGG
GGTCAT	TGGGGG	GTGGTG	GTGGGG	GTGTGG
TGGTCA	GGGGTG	GGGGTG	TGTGGG	GGTGTG
GGTAAA	GGGTGG	GGTGGT	GTGGGT	TGTGGG
GGTACA	TGGTGG	GGGGGG	GGGGGT	GGTGGT
GGTACG	GGGAGG	GGTGTG	GGGATC	GTGGGT
ACGGTA	GGTGGT	TGGGGG	GTGTGG	TGGGGT
GGAAAA	GGGGAG	GTGTGG	GGTGTG	GGGGGT
AAGGTC	GTGTGG	GGGTGG	TGGGGG	TGGGGG

Top 12 8-mers by round				
Round 0	Round 2	Round 3	Round 6	Round 7
GGTCAAGG	GGTGGTGG	GGTGGTGG	GGTGGTGG	GGTGGTGG
GGGTCAAG	GTGGTGGG	GTGGTGGG	GTGGTGGG	GTGGTGGG
GGTCAAAG	GGTGGGGG	GGTGTGGG	GGTGTGGG	GGTGTGGG
GGGTCAGG	GGTGTGGG	GGGTGTGG	GGGTGTGG	GGGTGTGG
GGTCAGGG	GGGGGGGG	GGTGGGGG	TGGTGGTG	TGGTGGTG
GGTCAATG	GGGTGTGG	GGTGGAGG	GGTGGGGG	GGTGTTGG
GGGTCAAA	GTGGGGGG	TGGTGGGG	TGGTGGGG	TGGTGGGG
AGGGTCAA	GGGGGTGG	TGGTGGTG	GGTGTTGG	GGTGGGGG
GGGGTCAA	GGGGGGGT	GGGGGTGG	GGTGGAGG	GTGTGGGG
AGGTCAAG	GGGGTGGG	GTGTGGGG	GTGTGGGG	GGTGGAGG
GGGTCAAT	TGGTGGGG	GGGTTGGG	GGGAGTGG	TGGTGGGT
GGTCATGG	GGGTGGGG	GGTGTTGG	TGGTGGGT	GGGAGTGG

Figure S1. SELEX-Seq Motif Analysis. (A) Position weight matrix (PWM) plots for the enrichment of 6-mer and 8-mer motifs at rounds 0, 2, 3, 6, and 7 of SELEX-Seq. (B) List of the top 12 enriched 6-mer and 8-mer motifs at each round. Motifs in red were previously identified by bioinformatics analysis.

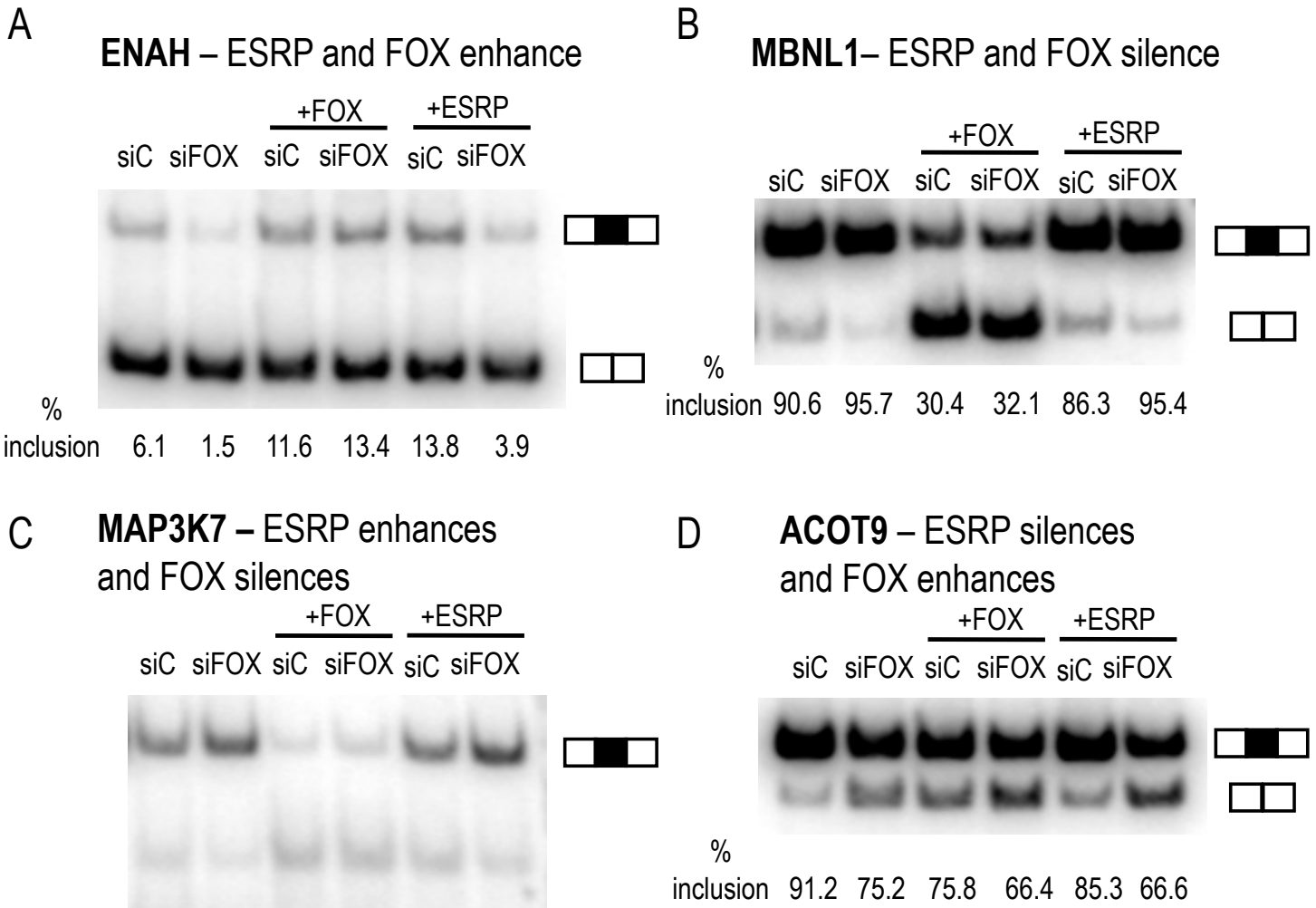


Figure S2. ESRP and Fox2 exhibit combinatorial regulation of the alternative splicing of cassette exons in MDA-MB-231 cells. RT-PCR analysis of ESRP and Fox2 co-regulated exons was performed in MDA-MB-231 cells with knockdown of Fox2 (siFOX) versus cells with control siRNAs (siC). Additional these events were analyzed in cells where knockdown of Fox2 was coupled with overexpression of Fox2 (+FOX) or ESRP1 (+ESRP). (A) *ENAH* is cooperatively enhanced by ESRP and Fox2, (B) *MBNL1* is cooperatively silenced by ESRP and Fox2, (C) *MAP3K7* is enhanced by ESRP and silenced by Fox2 (D) *ACOT9* is silenced by ESRP and enhanced by Fox2. (E) A Western blot demonstrates the knockdown of Fox2 and overexpression of Fox2 and ESRP. Note that due to an absence of ESRP expression in MDA-MB-231 and higher Fox2 levels than in PNT2 cells the analysis is complicated by different baseline exon inclusion levels compared to PNT2 cells.

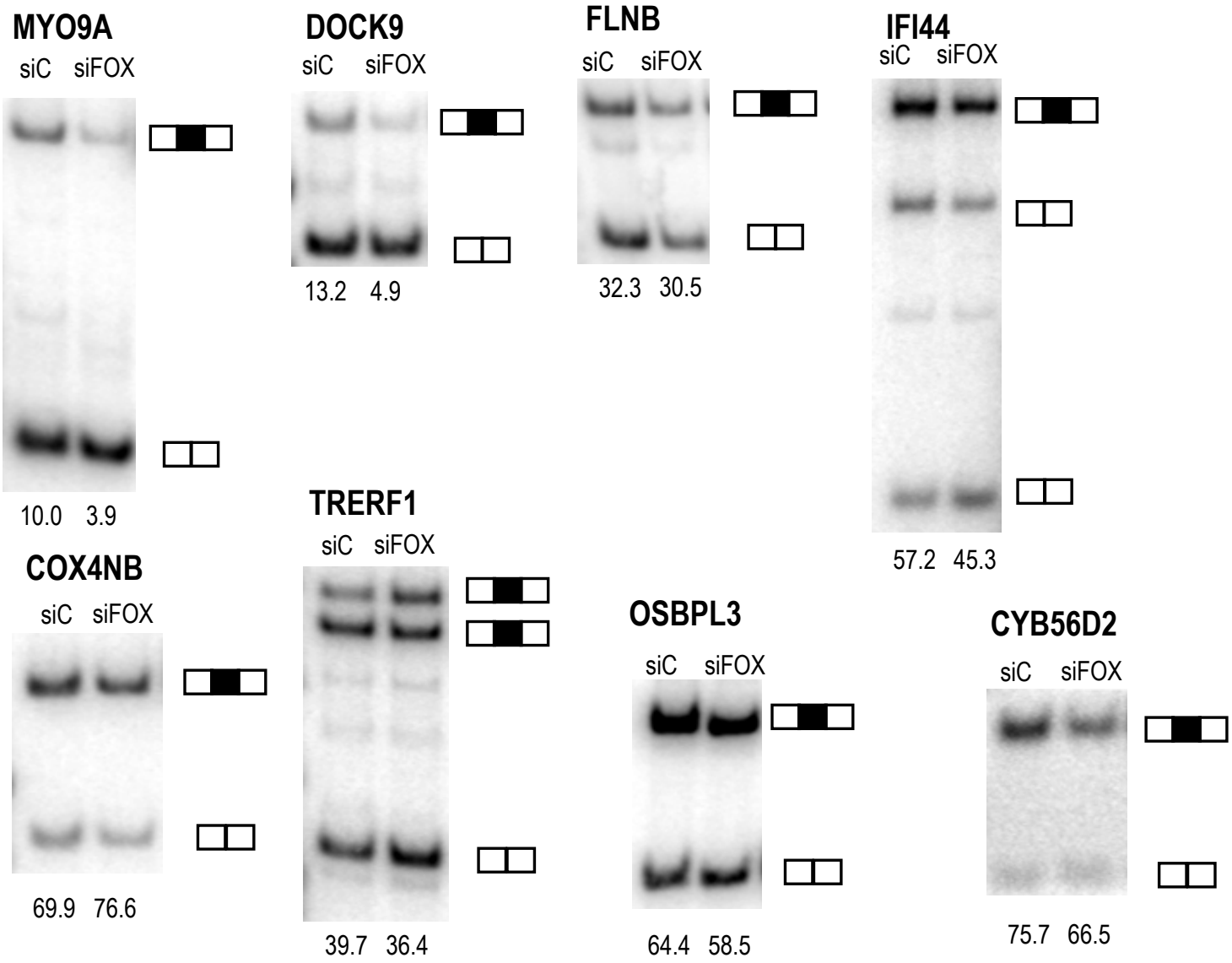
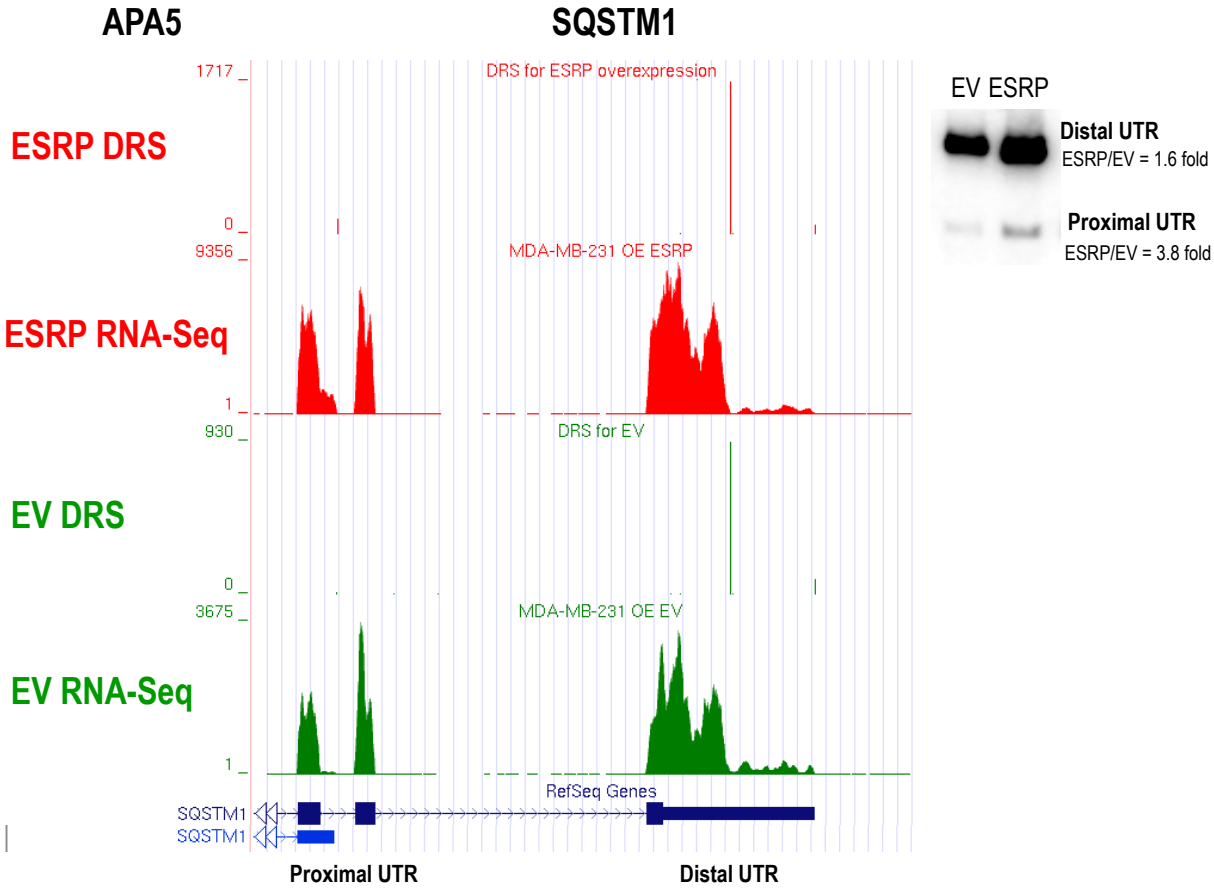
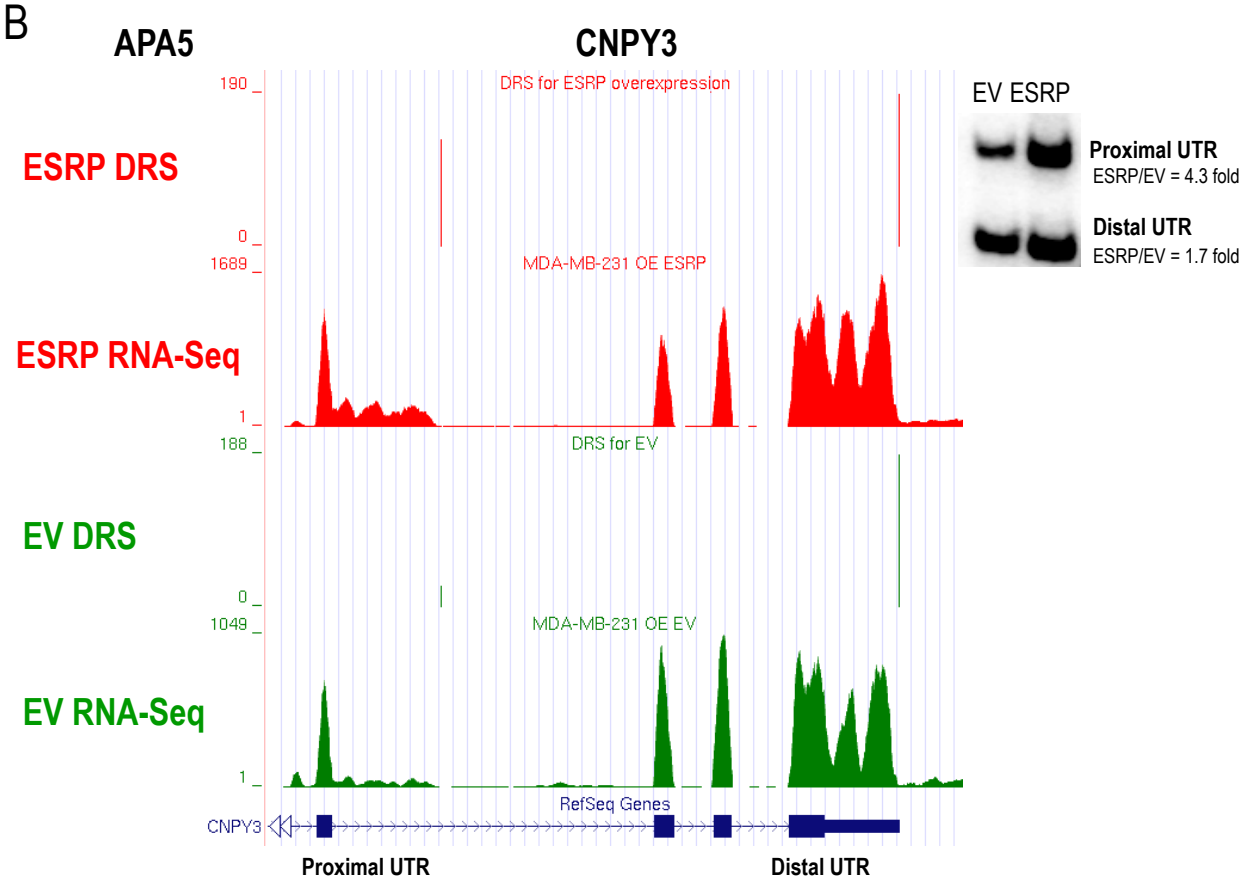


Figure S3. Fox2 regulation of validated ESRP-regulated cassette exons containing UGCAUG motifs. Eight validated ESRP-regulated cassette exons containing UGCAUG motifs, but that were not previously known Fox2 targets, were tested for Fox2 regulation in MDA-MB-231 control cells (siC) versus Fox2 knockdown cells (siFOX). Exon inclusion percentage for each sample is shown beneath the gel images. Exon coordinates and location of UGCAUG motifs are shown in Supplemental Table 6. The MYO9A, DOCK9, OSBPL3, and CYB56D2 were predicted to be Fox enhanced and all showed an expected decrease of at least 5% exon inclusion with Fox2 knockdown. Of 4 predicted Fox2 silenced exons in COX4NB, TRERF1, FLNB and IFI44, only COX4NB showed an increase in exon inclusion with Fox2 knockdown.

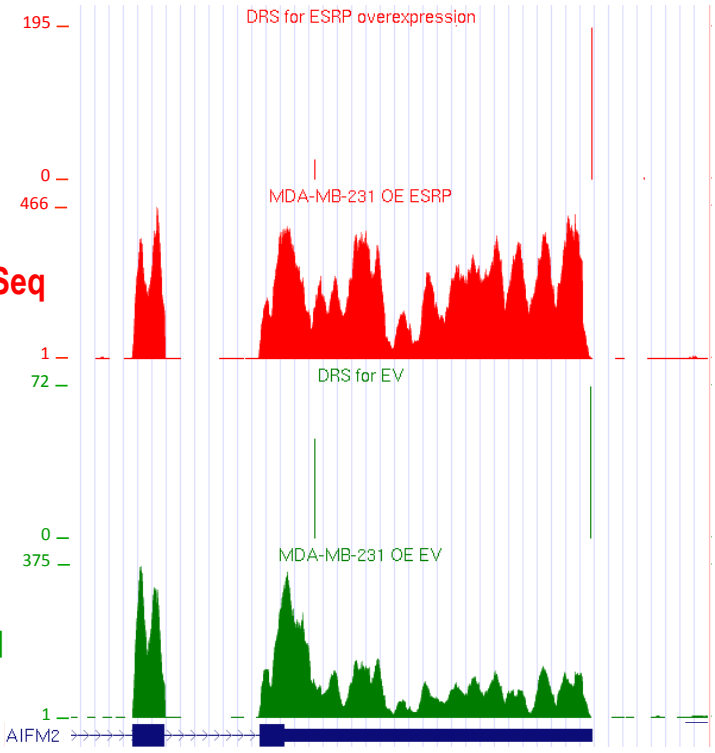
A

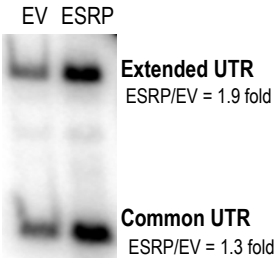

APA5
SQSTM1
ESRP DRS
ESRP RNA-Seq
EV DRS
EV RNA-Seq
DRS for ESRP overexpression
MDA-MB-231 OE ESRP
DRS for EV
MDA-MB-231 OE EV
RefSeq Genes
SQSTM1
SQSTM1
Proximal UTR
Distal UTR
EV ESRP
Distal UTR
ESRP/EV = 1.6 fold
Proximal UTR
ESRP/EV = 3.8 fold

B

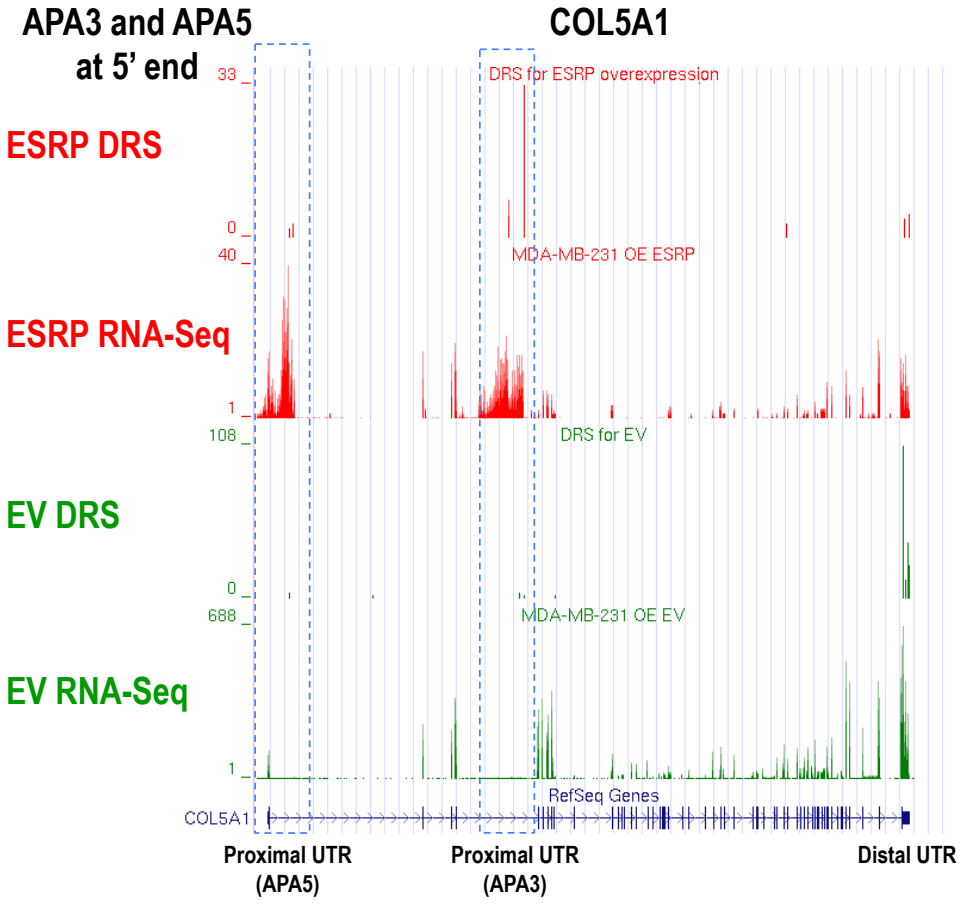

APA5
CNPY3
ESRP DRS
ESRP RNA-Seq
EV DRS
EV RNA-Seq
DRS for ESRP overexpression
MDA-MB-231 OE ESRP
DRS for EV
MDA-MB-231 OE EV
RefSeq Genes
CNPY3
Proximal UTR
Distal UTR
EV ESRP
Proximal UTR
ESRP/EV = 4.3 fold
Distal UTR
ESRP/EV = 1.7 fold

C


APA
AIFM2
DRS for ESRP overexpression
195
ESRP DRS
0
466
MDA-MB-231 OE ESRP
ESRP RNA-Seq
1
72
DRS for EV
EV DRS
0
375
MDA-MB-231 OE EV
EV RNA-Seq
1
AIFM2
Common UTR
Extended UTR


EV ESRP
Extended UTR
ESRP/EV = 1.9 fold
Common UTR
ESRP/EV = 1.3 fold

D


APA3 and APA5
at 5' end
COL5A1
33
DRS for ESRP overexpression
ESRP DRS
0
40
MDA-MB-231 OE ESRP
ESRP RNA-Seq
1
108
DRS for EV
EV DRS
0
688
MDA-MB-231 OE EV
EV RNA-Seq
1
RefSeq Genes
COL5A1
Proximal UTR
(APA5)
Proximal UTR
(APA3)
Distal UTR

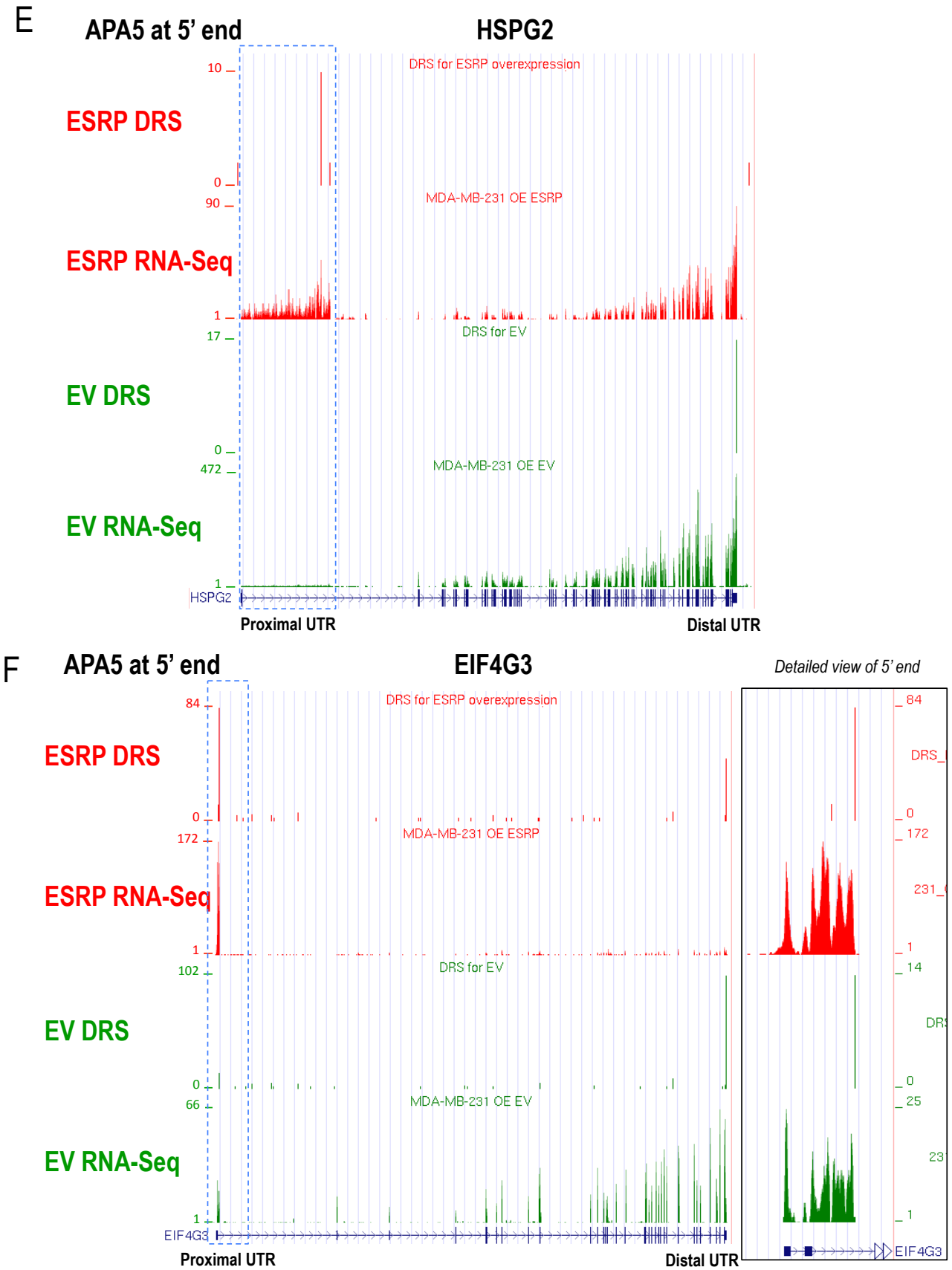


Figure S4. Additional examples of ESRP-regulated alternative 3' ends. APA5 events inSQSTM1 (A) and CNPY3 (B) and an APA event in AIFM2 (C) are shown with UCSC browser views of RNA-Seq data from MDA-MB-231 control cells (EV, green) or ESRP1 expressing cells (ESRP, red) with RT-PCR validations.)(D-F) Examples of APA5 and APA3 events near the 5' end that result in decreased overall expression for these genes in ESRP1 expressing cells.

Table S1 – Summary of RNA-Seq data

	Samples	Total # reads	Genomic Reads (hg19)		Splice Junction Reads		
			Overall	Unique	Overall	Unique	Filtered*
MDA-MB-231	ESRP	136.2M	111.0M (81.5%)	91.3M (67.0%)	17.3M (12.7%)	17.1M (12.6%)	**11.0M (8.1%)**
	EV	120.1M	93.1M (77.5%)	76.0M (63.3%)	15.3M (12.7%)	15.1M (12.6%)	**10.1M (8.4%)**
	Total	256.3M	204.1M (79.6%)	167.3M (65.3%)	32.6M (12.7%)	32.2M (12.6%)	**21.1M (8.2%)**
PNT2	siGFP	59.1M	48.6M (82.3%)	41.7M (70.6%)	7.1M (12.0%)	7.0M (11.8%)	**6.2M (10.4%)**
	siESRP1/2	74.5M	59.6M (80%)	50.7M (68.1%)	8.2M (11.0%)	8.1M (10.8%)	**7.1M (9.5%)**
	Total	133.6M	108.2M (81%)	92.4M (69.2%)	15.3M (11.5%)	15.1M (11.3%)	**13.2M (9.9%)**

Bowtie, first 50 bp, allowing up to 3 bp mismatches
Junction: all possible junctions from Ensembl transcripts, 84 bp long
****Filtered: Uniquely mappable reads after combining genome and junction annotations together***

Table S1. Summary of RNA-Seq read data. The total number of reads obtained from each experimental sample are shown and well as the combined number of reads for each cell system. Allowing 3 mismatches, reads were mapped to hg19. The number of reads that mapped overall and uniquely are shown, with the percentage of total reads shown in parentheses. Reads were also mapped to a custom splice junction database of all exon-exon junctions from Ensembl transcripts. The number of splice junction reads that map overall and uniquely are shown.

Table S11. Details of the subnetwork of functional interactions of ESRP-target networks. The 276 genes containing RT-PCR validated ESRP regulated exons were mapped to the extended FI network. A total of 250 genes (90.6%) were in the network. The edge-betweenness algorithm was used to find 19 network modules, 13 of which were greater than size 5. Table S11A 1 shows a list of the modules and the nodes within them. Table S11B shows a list of the pathways significantly enriched within these modules

A

Module	Nodes in Module	Node List
0	44	ACAA1,ACSL3,AKR1A1,AKR1B1,AKT1,CREBBP,ENO2,HIF1A,HK2,HMBS,IKBKG, JUN,LASS2,MAPK14,MAPK3,MEF2D,MLPH,MST1R,MST4,NCOA1,NCOR2,NDN, PDIA3,PFKM,PPP1R7,PRMT2,RELA,RIPK2,RPH3A,RXRA,SOD2,SP1,STAT3, TCF7L2,TCIRG1,TLE1,TNFRSF10A,TNIP2,TP53,TRERF1,TSC2,TUFT1,UBE2K, YWHAH
1	38	ACTC1,ACTG1,ACTN1,ATP2B4,BSG,CAV1,COL16A1,CSDA,CTNNA1,CTNNB1, CTNND1,CTTN,DLG1,DNM2,FKBP1A,FLNA,FNBP1,GIT2,GLMN,GUSB,ITGA6, ITGAV,MACF1,MAGI1,MET,MLLT4,MYH10,MYO6,PIP4K2C,PPFIBP1,PTPRF ,PTPRM,SLC7A6,TFAP2A,TGFBR1,TJP1,YAP1,ZFYVE9
2	33	ARHGAP1,ARHGAP17,ARHGEF11,ARHGEF12,ATG16L2,BAIAP2,CACNG6,CIT, EHMT2,EPHB4,EXOC1,EXOC7,FAM13B1,GRB2,HRAS,IL7R,ITSN1,KLHL20, MINK1,MYO9A,MYO9B,NCK1,PDGFA,PIK3R1,RAC1,RAC2,RAPGEF1,RAPGEF2, RASGRF2,RHOA,RHOQ,SPTAN1,STARD8
3	31	ACTR1A,ADD3,AKAP11,CCNK,CDC2,CDCA3,CDK2,CEP164,CEP78,CSNK1E, DVL3,EIF2AK4,FGFR1OP,GSK3B,HSP90AA1,LRRFIP2,MAST2,NAE1,NGLY1, PLK1,PRC1,PRKACB,PRKAR2B,PUM2,RAD23B,RPS27A,SIRT3,SKP2,STX2, STXBP2,WEE1
4	19	ARFGAP2,ARRB1,CLTC,DCBLD1,F2,GABARAP,GNAS,GNG7,HGFAC,MPP3, OPRK1,PAQR5,PLA2G1B,PLAA,PXDN,SPINT2,THBS1,TNFAIP6,TTC3
5	12	ATP13A3,CCAR1,FASTK,GOLGA4,HNRNPA3,HNRNPM,KTN1,PTBP1,RBM4, SNRPG,SYNCRIP,TIA1
6	12	CTSD,EIF2AK2,LRP8,MAP2K4,MAP2K7,MAP3K7,MAP3K9,PSAP,SMPD2,SPAG9, TARBP2,TNFRSF1A
7	11	CHM,CHMP2A,MITD1,RAB11A,RAB18,RAB5A,RAB7A,VPS24,VPS39,YIF1A,YIPF5
8	10	ARF1,ARFIP1,CASP3,GGA1,HTT,KIF13A,M6PR,PACSIN1,PACSIN2,SLK
9	8	APBB2,APH1A,APLP2,JAG2,NOTCH1,NUMB,PCSK4,POFUT1
10	6	PEX1,PEX19,PEX26,PEX3,PEX6,PMP22
11	6	CDS2,DGKQ,IMPA1,PLCB1,RTN1,RTN2
12	5	IDH3A,IDH3B,IMPDH2,MTHFD1,RRM1
13	3	EMP3,P2RX7,RYR3
14	3	PLOD1,PLOD2,PLOD3
15	3	NF2,ZNF267,ZNF675
16	2	RPSA,VARS
17	2	ETV7,PBX2
18	2	TIMM17A,TIMM44

B

Module	Pathway	Ratio	Number Of Protein In Pathway	Protein From Module	FDR	Nodes
2	Signaling by Rho GTPases(R)	0.0168	126	9	<1.00e-03	MYO9B,ARHGAP17, ARHGEF12, ITSN1,FAM13B1, MYO9A,ARHGEF11, ARHGAP1,STARD8
9	NOTCH(C)	0.0088	66	4	<1.00e-03	APH1A, NUMB,JAG2, POFUT1
9	Notch signaling pathway(K)	0.0062	47	3	5.00E-04	APH1A,NUMB,JAG2
1	Nectin adhesion pathway(N)	0.014	105	5	1.00E-03	MYO6,CTNND1,CTTN,MLLT4,DNM2
3	G2/M Transition(R)	0.011	83	4	1.00E-03	CEP78,CEP164,CSNK1E,FGFR1OP
1	E-cadherin signaling in the nascent adherens junction(N)	0.0061	46	4	2.00E-03	MYO6,CTNND1,CTTN,MLLT4
9	Signaling by Notch(R)	0.0021	16	2	3.00E-03	APH1A,JAG2
1	Tight junction(K)	0.0177	133	5	3.33E-03	MAGI1,CSDA,CTTN,MLLT4,MYH10
9	Alzheimer disease-presenilin pathway(P)	0.0173	130	3	4.00E-03	APH1A,APBB2,PCSK4
9	Notch signaling pathway(P)	0.0029	22	2	4.00E-03	APH1A,NUMB
4	how progesterone initiates the oocyte maturation(B)	0.0029	22	2	8.00E-03	GNAS,PAQR5
9	Alzheimer disease-amyloid secretase pathway(P)	0.0084	63	2	2.23E-02	APH1A,PCSK4
16	Valine, leucine and isoleucine biosynthesis(K)	0.0015	11	1	4.60E-02	VARS

Supplementary Table Legends

Table S2. Candidate ESRP-regulated cassette exons identified by RNA-Seq analysis. Candidate cassette exons with FDR <0.05 are listed in order of increasing p-value. Separate tabs show the data from ESRP overexpression in MDA-MB-231 (MDA-MB-231 OE) and ESRP knockdown in PNT2 (PNT2 KD). Columns show the transcript information, exon chromosomal coordinates, upstream junction read count (UJC), downstream junction read count (DJC), skipping junction read count (SJC), exon inclusion level (Ψ) for each sample, change in exon inclusion between samples (delta Ψ), probability, p-value and FDR.

Table S3. Validations of ESRP-regulated cassette exons. RT-PCR validated cassette exons identified in either MDA-MB-231 OE or PNT2 KD experiment are shown in separate tabs ranked by the magnitude of the splicing change. Columns indicate the chromosomal coordinates, RT-PCR measured % exon inclusion level for each sample and % change in exon inclusion between samples. Exons that have not been previously identified as alternatively spliced cassette exons are shown in bold. Exons that did not show at least a 5% change in splicing or showed a change opposite to that predicted by RNA-Seq are shown in gray. The additional validations tab contains validations for events that did not meet our statistical cutoffs but were still tested by RT-PCR. The final tab shows the complete list of RT-PCR validated ESRP-regulated cassette exons (i.e. those identified by RNA-Seq, HJAY, and Exon Array analysis) that have >10% change in at least one experimental system ranked by change in inclusion level (delta Ψ). Columns show the gene, chromosomal coordinates, exon size and the experiment and cell line in which each cassette exon was identified as a candidate ESRP target well as the function of ESRP (enhanced or silenced). Events listed as "additional validations" represent exons in genes of interest that did not meet the statistical criteria of previous analyses, but were also tested and found to show significant changes by RT-PCR.

Table S4. ESRP binding motif analysis from SELEX-Seq data. All possible n-mers (from 1mer to 8mers) were enumerated and the frequency of each n-mer in the 20 nt random sequence region from each round of SELEX, defined as the total occurrence of an n-mer divided by the total number of unique n-mer positions within all 20 nt random sequence regions from a given round, was calculated. Tabs are shown for each n-mer from 1 through 8. In each tab the motif density from round 0 up to round 7 is shown sorted according to the frequency in the 7th round.

Table S5. P-values for the difference in ESRP motif score in the set of validated ESRP enhanced and silenced cassette exons at each nucleotide position. On separate tabs for ESRP enhanced exons and silenced exons the p-value is shown at each position within each region shown in Fig. 4. Nucleotide positions associated with p-values <0.001 are highlighted in yellow. Note that for regions 1 and 3 negative values for nucleotide position are exonic and positive values indicate intronic sequences downstream of the exon. In regions 2 and 4 negative values are associated with the upstream intron and positive values the downstream exon.

Table S6. Validated Fox2 and ESRP regulated cassette exons and identification of the Fox motif within validated ESRP regulated cassette exons. Cassette exons

validated as Fox2 and ESRP targets by RT-PCR are shown in the first tab with chromosomal coordinates, % change by Fox and ESRP and function of Fox and ESRP (enhanced or silenced). Events in which Fox2 and ESRP induce the same changes in splicing and those in which they induce the opposite changes in splicing are shown. Data for Fox regulation was obtained from Venables, et. al 2009. Location of the UGCAUG Fox binding motifs within or in the 250 nt flanking the 276 validated ESRP-regulated cassette exons. Columns indicate the chromosomal coordinates of regulated cassette exons, as well as the chromosomal coordinates of upstream, exonic, and downstream UGCAUG sites. N indicates no UGCAUG sites in that region. Data is shown on separate tabs for ESRP enhanced versus silenced cassette exons. The summary tab shows the total number of UGCAUG-containing ESRP regulated exons and flanking introns compared to a control set with the percentage of total shown in parentheses. Exons in bold were tested for Fox2 regulation by RT-PCR as shown in Fig. 2.

Table S7. Comparison of ESRP-regulated cassette exons with cassette exons with validated changes in breast cancer subtypes. Cassette exons with differential splicing between Basal B (Claudin-low) vs. Luminal Breast Cancer cell lines are predominantly regulated by the ESRPs. We used the Affymetrix Human Junction Array (HJAY) data to analyze the differential alternative splicing between two types of breast cancer cell lines: luminal cell lines and basal B (also referred to as claudin-low) cell lines. The HJAY data was extracted from a study of the alternative splicing in breast cancer cell lines. The luminal cell lines we analyzed included 600MPE, BT474, CAMA1, MCF7, MDAMB361, MDAMB415, MDAMB453, SKBR3, SUM185PE, SUM52PE, T47D, ZR751 and ZR75B. The basalB cell lines included BT549, HBL100, HCC38, MDAMB157, MDAMB231, MDAMB435 and SUM159PT. Based on the HJAY data of luminal cell lines and basalB cell lines, the exon-level differential alternative splicing P values were calculated by MADS+ software. The computed P values for differential alternative splicing were transformed to false discovery rate (FDR) using the Benjamini-Hochberg algorithm. HJAY analysis of ESRP regulated exons from previous MDA-MB-231 and PNT2 data including p-value and direction of change showed that the exons with the most statistical difference in splicing in Basal B vs. luminal cells mostly showed the same direction of change in response to ESRP expression in which there was similarly high statistical evidence of ESRP regulation. Exons were ranked based on p-value for difference between Basal B and luminal cell types. For direction change "+" indicates that the exon inclusion is greater in luminal than Basal B and in the ESRP columns "+" indicates that ESRP expression increases exon inclusion. Also shown are the top 20 events for which validation of ESRP regulation was carried out along with the percent change in exon inclusion.

Table S8. Candidate ESRP-regulated alternative 3' and 5' splice sites identified by RNA-Seq analysis. Candidate exons with alternative 3' and 5' splice sites with predicted ESRP-regulated splicing changes >5% are listed in order of increasing p-value. Separate tabs show the data from ESRP overexpression in MDA-MB-231 (MDA-MB-231 OE) and ESRP knockdown in PNT2 (PNT2 KD). Columns show the transcript information, chromosomal coordinates for long and short exons and flanking exons, long exon junction read count (LJC), short exon junction read count (SJC), exon inclusion level

(PSI) for the long form of the exon in each sample, change in long exon inclusion between samples (deltaPSI), probability, p-value and FDR.

Table S9. Candidate ESRP-regulated alternative 3' ends identified by DRS and RNA-Seq analysis. Candidate exons with ESRP-regulated alternative 3' ends (APA, APA3 and APA5) are shown on three separate tabs. The first tab shows the complete list of 355 events identified by DRS. The second tab shows the list of 160 events identified by DRS and validated by RNA-Seq (p-value <0.01). The third tab shows the list of 108 events used to draw the RNA map, with percentages of polyA usage change greater than 0.1 and duplicate polyA sites removed. Columns show chromosomal coordinates, gene symbol, polyA read counts from DRS, p-value, FDR, % polyA_1 (DRS), % polyA_2, (DRS), % change, normalized % change (corrected for total number of mapped DRS reads), read counts from RNA-Seq within 300nt upstream of polyA sites, % polyA_1 (RNA-Seq), % polyA_2 (RNA-Seq), on-sided Fisher exact text and p-values for polyA_1 greater in EV and ESRP, genomic region containing two polyAs, type and novelty. In all cases, site 1 represents the more 5' or proximal polyA site and site 2 represents the more 3' or distal polyA site. In the third tab, RPKM values are shown for genes where ESRP promotes the proximal polyA site, as well as log 2 fold change, p-value and FDR for overall expression change.

Table S10. P-values for the difference in ESRP motif score in the set of 160 ESRP regulated alternative polyadenylation sites at each nucleotide position within 250 nucleotides upstream and downstream. The p-value is shown at each position within the 250 nucleotide regions upstream and downstream of the polyadenylation sites shown in Fig. 7D. Values are listed separately for ESRP promoted and ESRP inhibited sites. Nucleotide positions associated with p-values <0.001 are highlighted in yellow.