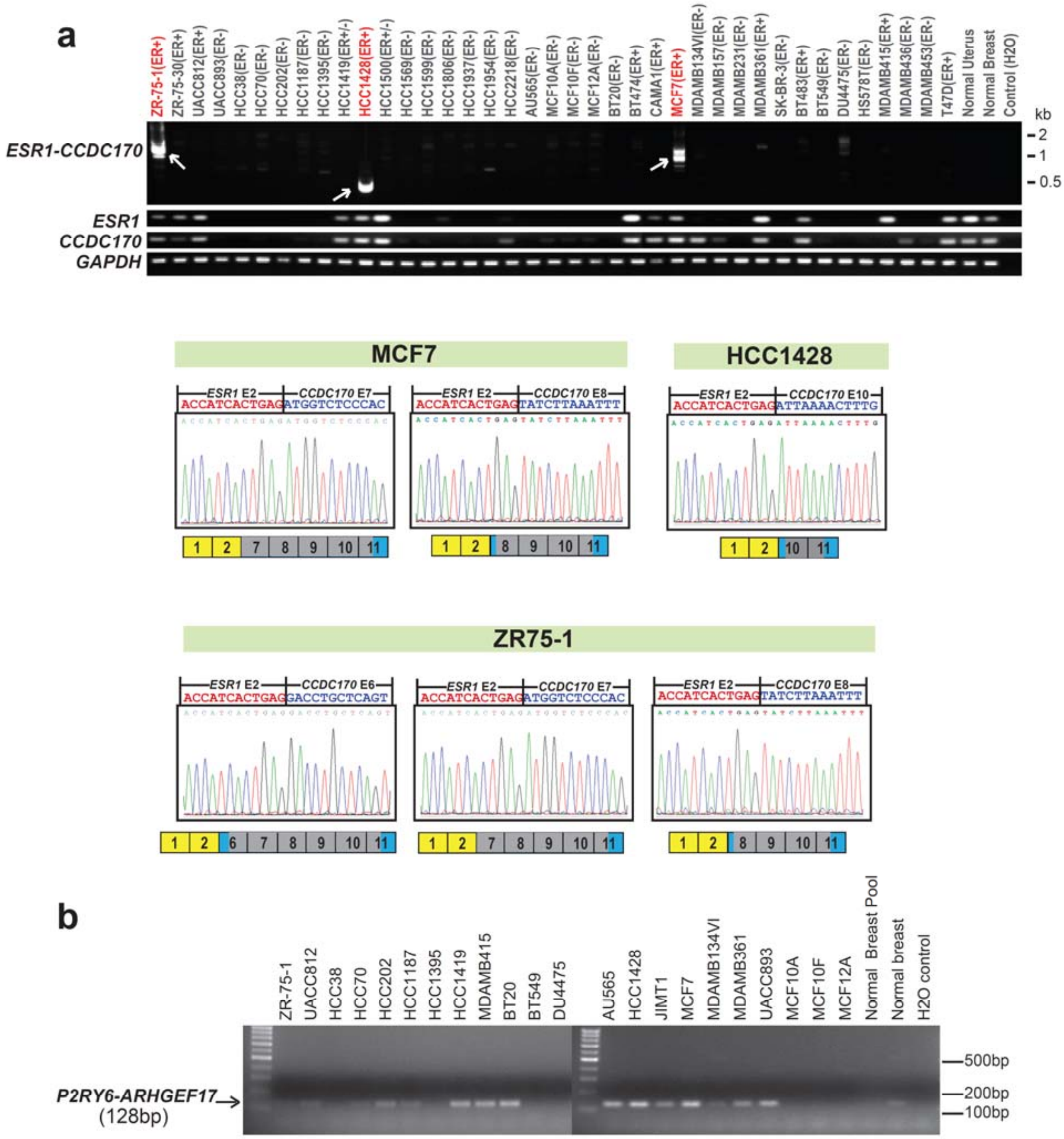
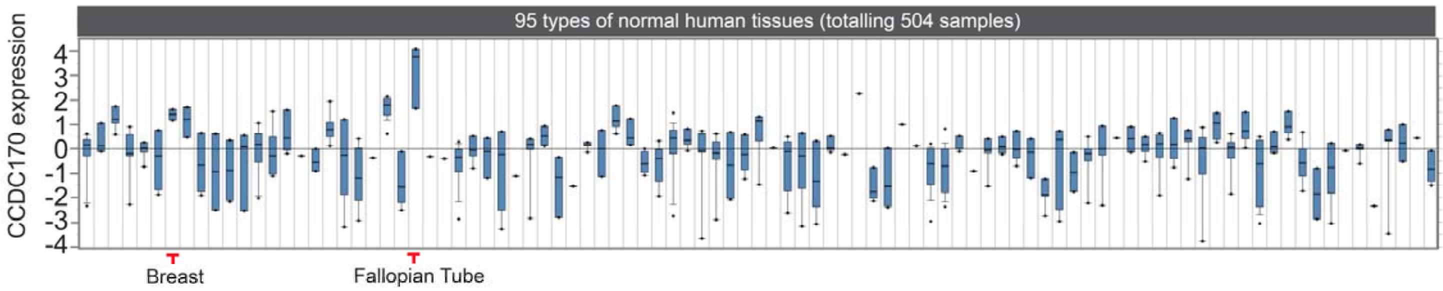
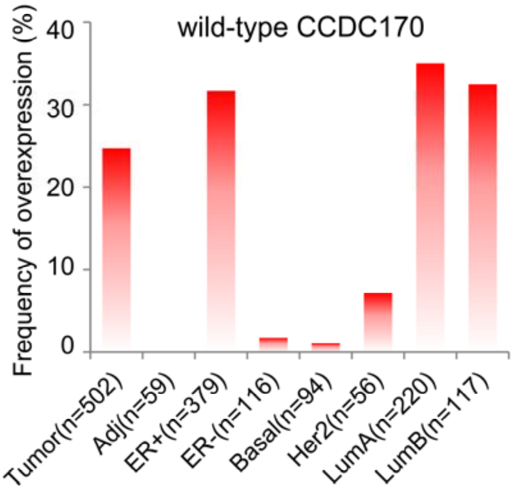
SUPPLEMENTARY INFORMATION



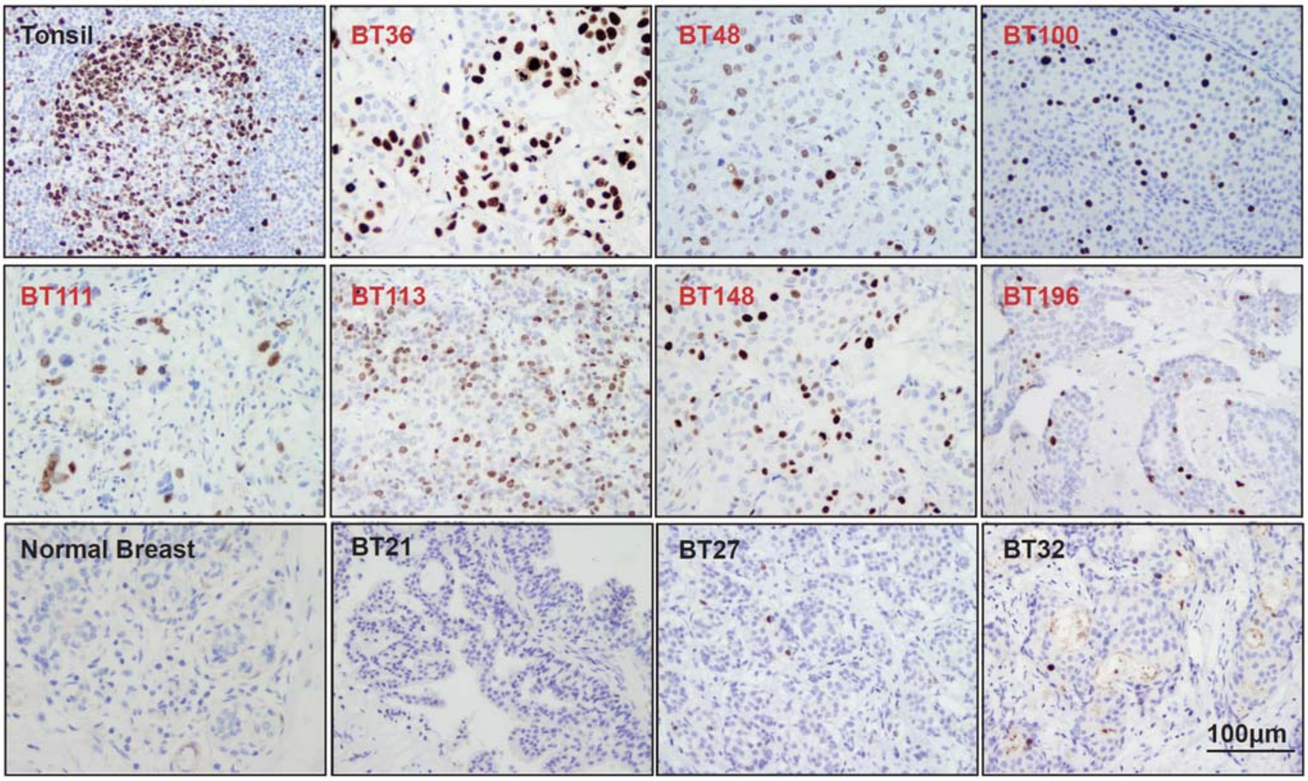
Supplementary Figure 1. Detecting *ESR1-CCDC170* and *P2RY6-ARHGEF17* chimerical transcripts in breast cancer cell lines by RT-PCR. (**a**) RT-PCR results using a forward primer (AACAGGCTCGAAAGGTCCATGC) in the first exon of *ESR1* and a reverse primer (GCCTAGCATCTGCGACACCACC) in the last exon of *CCDC170* (upper), and representative chromatograms of the fusion junction of each *ESR1-CCDC170* variant (lower). RT-PCR analysis of wt*ESR1* and wt*CCDC170* was performed using a forward primer from the second exon and a reverse primer from the third exon of the respective genes. The exon numbers are based on reference sequence NM_001122742 for *ESR1*, NM_025059 for *CCDC170.* (**b**) RT-PCR validation of the *P2RY6-ARHGEF17* fusion in a panel of breast cell lines and normal breast controls. The expected size of the *P2RY6-ARHGEF17* PCR product is 128bp.



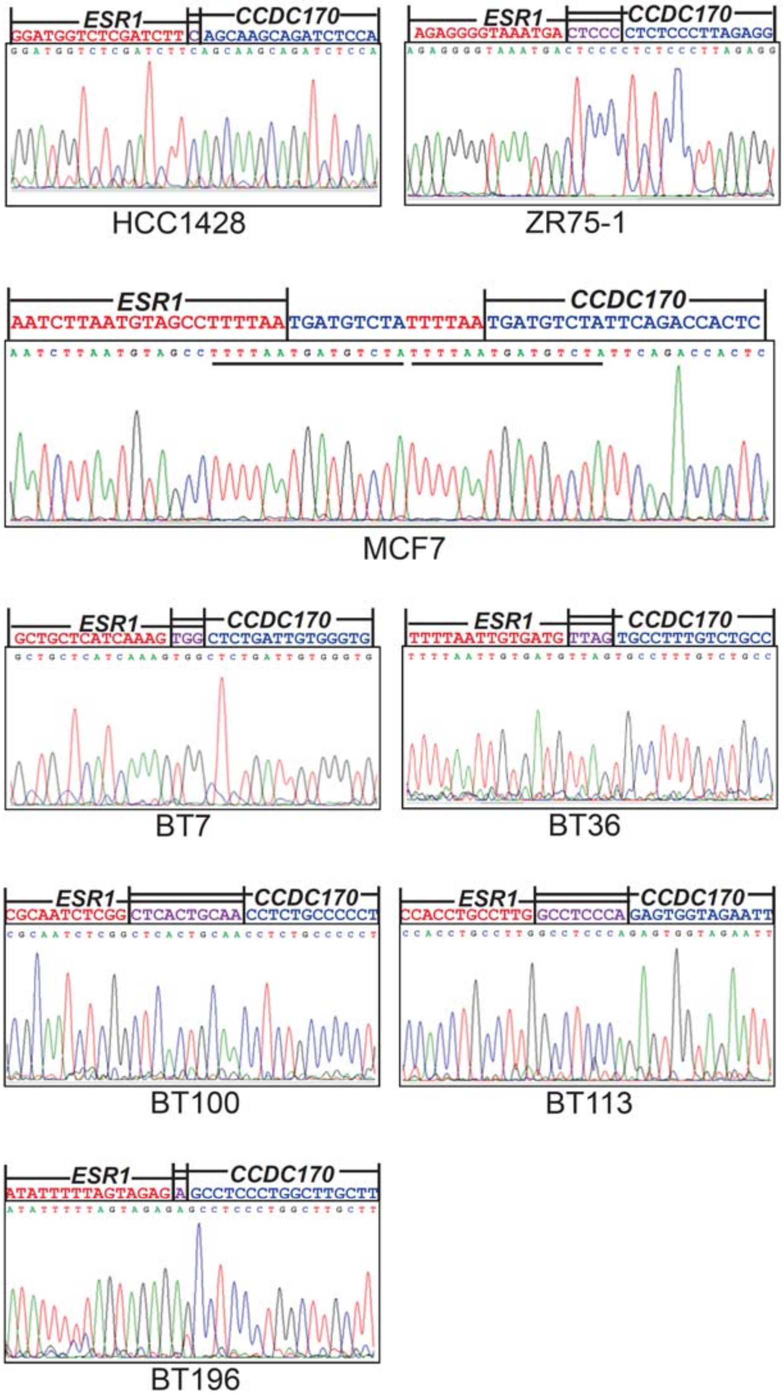
Supplementary Figure 2. The expression of wild-type *CCDC170* in 95 types of normal human tissues. Affymetrix U133 plus 2.0 gene expression data for normal human tissues are from the Human Body Index dataset (GSE7307) and are analyzed using Oncomine (www.oncomine.org).



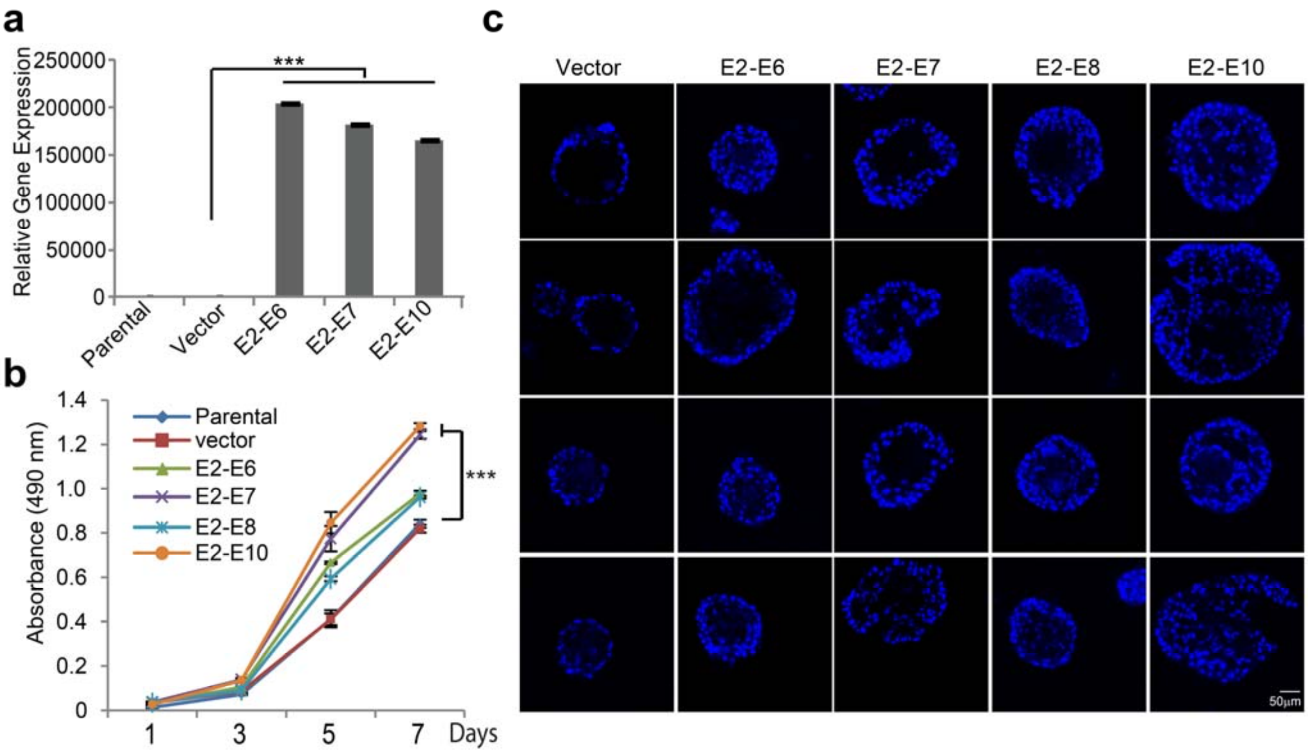
Supplementary Figure 3. The frequency of wild-type *CCDC170* overexpression in different breast cancer subtypes. Overexpression of wild-type *CCDC170* is almost exclusive to ER+ breast tumors with no significant difference observed between the luminal A and B tumors.



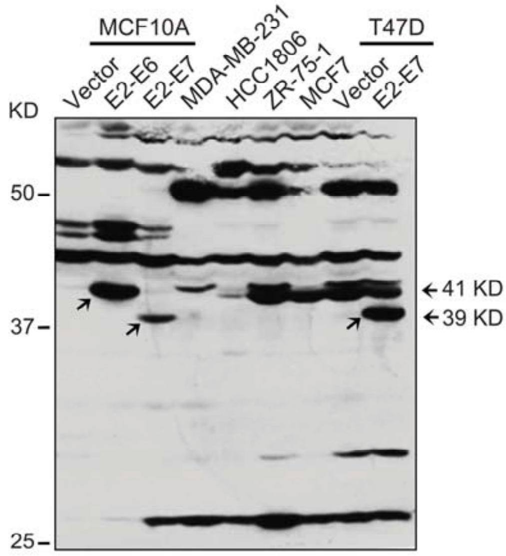
Supplementary Figure 4. Ki67 staining of all evaluable fusion-positive and representative fusion-negative breast cancer tissues. The fusion-positive samples are highlighted in red. Normal human tonsil tissue and normal breast tissue are used as positive and negative controls respectively.



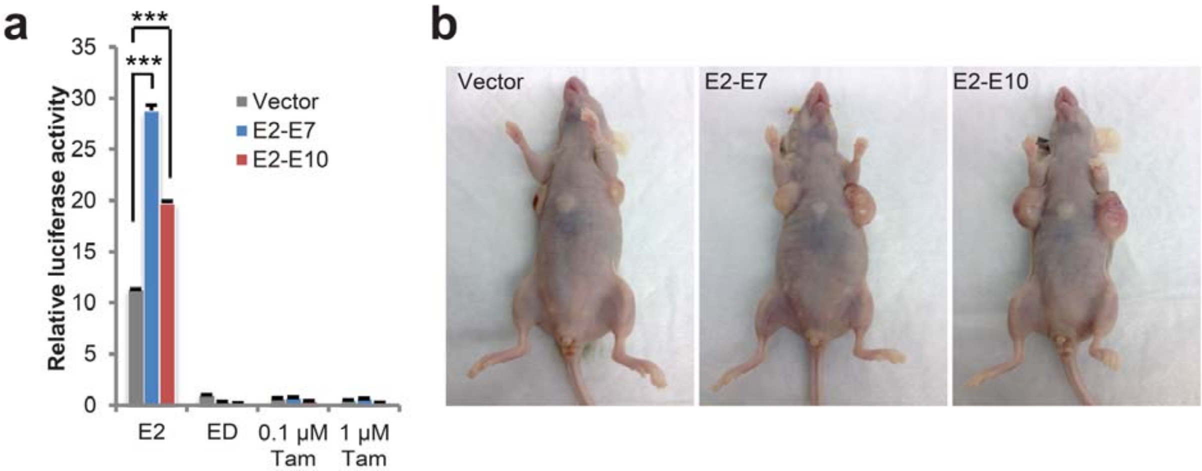
Supplementary Figure 5. ***ESR1-CCDC170*** **genomic fusion points in fusion positive breast cancer cell lines and tissues.** Representative capillary sequencing chromatograms of the *ESR1-CCDC170* genomic fusion junctions in the three positive breast cancer cell lines and five positive breast tumor tissues are shown.



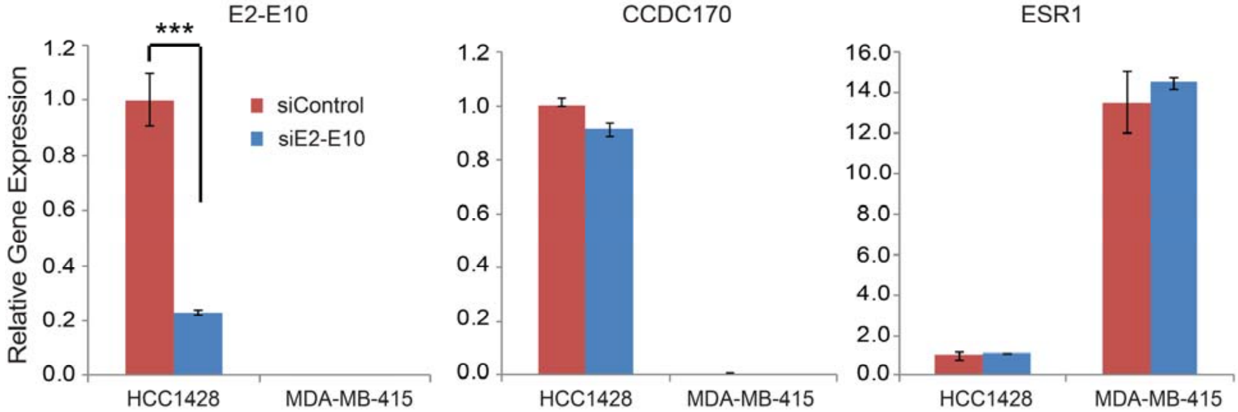
Supplementary Figure 6. The impact of ΔCCDC170 expression on the proliferation and acini formation of MCF10A cells. (**a**) Quantitative RT-PCR validation of the expression of ΔCCDC170 variants in transduced MCF10A cells. ***p<0.001 (t-test). Error bars represent the standard deviation of two replicate measurements per condition. (**b**) The effect of ectopic expression of ΔCCDC170 variants on the proliferation of MCF10A cells. ***p<0.001 (t-test based on Day-7 data). Error bars represent the standard deviation of three replicate measurements per condition. (**c**) Representative images of MCF10A cells expressing the vector control or different ΔCCDC170 fusion variants cultured on matrigel for 60 days. Nuclei were stained with DAPI and visualized by confocal microscopy with x40 magnification. Scale bar represents 50μm.



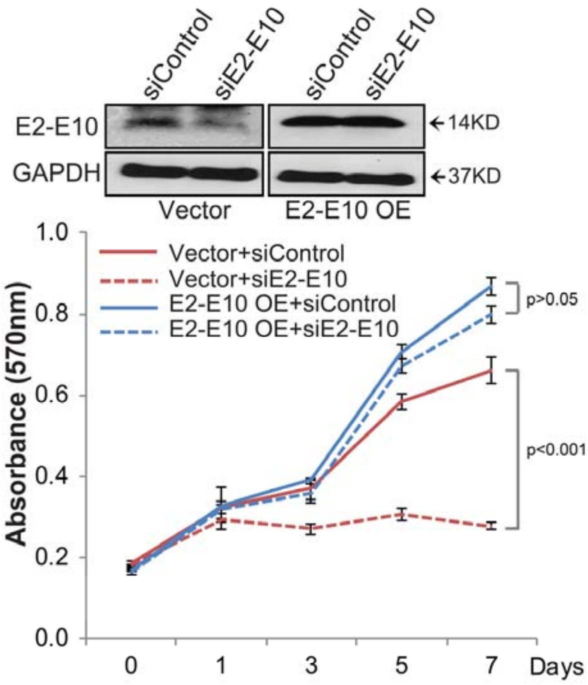
Supplementary Figure 7. Western blot detecting the ΔCCDC170 protein products expressed in the *ESR1-CCDC170* positive breast cancer cell lines, ZR-75-1 and MCF-7. The engineered MCF10A and T47D cells overexpressing the indicated fusion variants and fusion-negative breast cancer cell lines, MDAMB231 and HCC1806 are used as positive or negative controls respectively. The MCF10A and T47D cells transduced with pLenti7.3 vector containing an YFP ORF is also used as negative controls (Vector). The predicted size of the protein product encoded by the E2-E6 fusion variant expressed in ZR-75-1 cells is 41KD. While the blot did reveal a 41KD band in ZR-75-1 cells, this band is also detected in the cell lines that are negative for this fusion variant suggesting the possible presence of cross-reacting proteins. We did not detect the predicted 39KD band encoded by the E2-E7 fusion variant in MCF7 cells. Presumably this could be attributable to the relative low level expression of the E2-E7 fusion in this cell line (Figure 1e).



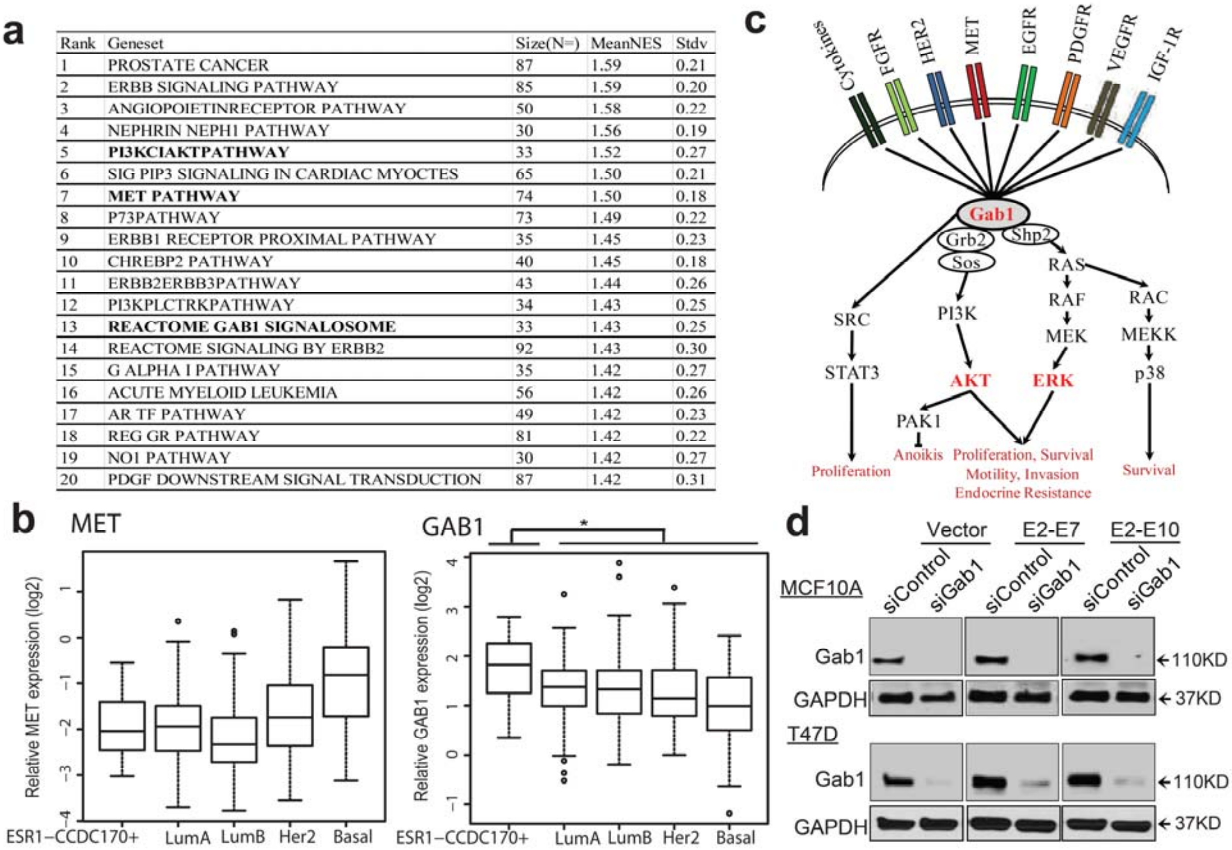
Supplementary Figure 8. The impact of ΔCCDC170 expression in ER+ T47D breast cancer cells on ER transcriptional activity *in vitro* and tumor formation *in vivo*. (**a**) The impact of ΔCCDC170 expression on ER transcriptional activity in T47D cells assayed by estrogen-responsive element luciferase reporter assay following estrogen (E2), estrogen deprivation (ED), or 4-OH tamoxifen treatment (Tam). Error bars represent the standard deviation of two replicate measurements per condition. ***p<0.001 (t-test). (**b**) Representative images of xenograft tumors engrafted from T47D cells expressing the vector, E2-E7, or E2-E10 ΔCCDC170 variants in athymic nude mice.



Supplementary Figure 9. Quantitative RT-PCR validation of E2-E10 fusion, wt*CCDC170* and wt*ESR1* expression in fusion-positive HCC1428 and fusion-negative MDA-MB-415 cell lines transfected with E2-E10 siRNA. Error bars represent the standard deviation of two replicate measurements per condition. *****P<0.001 (t-test).

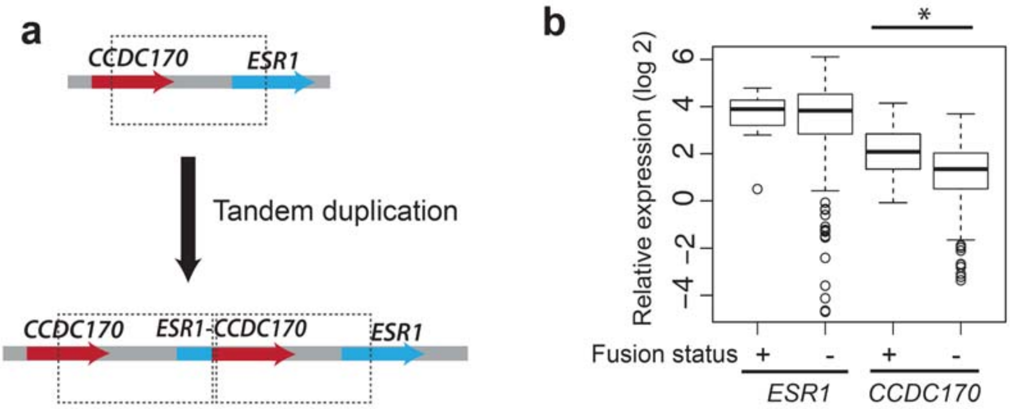


Supplementary Figure 10. The rescue effect of forced expression of the E2-E10 ORF on the proliferation of HCC1428 cells treated with siE2-E10. The inhibitory effect on the proliferation of HCC1428 cells resulting from E2-E10 knockdown was rescued by ectopic E2-E10 expression (MTT assay). Error bars represent the standard deviation of three replicate measurements per condition. P-value was determined using t-test based on Day 7 data.



Rank	Geneset	Size(N=)	MeanNES	Stdv
1	PROSTATE CANCER	87	1.59	0.21
2	ERBB SIGNALING PATHWAY	85	1.59	0.20
3	ANGIOPOIETINRECEPTOR PATHWAY	50	1.58	0.22
4	NEPHRIN NEPH1 PATHWAY	30	1.56	0.19
5	**PI3KCIAKTPATHWAY**	33	1.52	0.27
6	SIG PIP3 SIGNALING IN CARDIAC MYOCTES	65	1.50	0.21
7	**MET PATHWAY**	74	1.50	0.18
8	P73PATHWAY	73	1.49	0.22
9	ERBB1 RECEPTOR PROXIMAL PATHWAY	35	1.45	0.23
10	CHREBP2 PATHWAY	40	1.45	0.18
11	ERBB2ERBB3PATHWAY	43	1.44	0.26
12	PI3KPLCTRKPATHWAY	34	1.43	0.25
13	**REACTOME GAB1 SIGNALOSOME**	33	1.43	0.25
14	REACTOME SIGNALING BY ERBB2	92	1.43	0.30
15	G ALPHA I PATHWAY	35	1.42	0.27
16	ACUTE MYELOID LEUKEMIA	56	1.42	0.26
17	AR TF PATHWAY	49	1.42	0.23
18	REG GR PATHWAY	81	1.42	0.22
19	NO1 PATHWAY	30	1.42	0.27
20	PDGF DOWNSTREAM SIGNAL TRANSDUCTION	87	1.42	0.31

Supplementary Figure 11. Gab1 is a crucial component of the signaling pathways activated in ΔCCDC170 expressing ER+ breast cancer cells. (**a**) Top 20 pathways upregulated in *ESR1-CCDC170* positive breast tumors, as revealed by Gene Set Enrichment Analysis (GSEA). The pathways are ranked by their mean normalized enrichment scores (mean NES). (**b**) The expression of c-Met and Gab1 in *ESR1-CCDC170*-positive breast tumors as opposed to the fusion-negative breast tumors of different clinical subtypes. * $p<0.05$ (t-test). The gene expression data analyses in **a-b** are based on 12 *ESR1-CCDC170*+ cases that have available Agilent gene expression data from TCGA. The 2 cases supported by 2 fusion reads (see Supplementary Table 1) are included to increase the number of cases in the analysis. Error bars represent the standard deviations (*ESR1-CCDC170*+, n=12; LumA, n=229; LumB, n=120; Her2, n=56; Basal, n=96). (**c**) Schematic of the Gab1 pathway and key signaling intermediates (based on *Birchmeier et al*, Molecular cell biology 2003 4:915, and *Raghav et al*, Transl Lung Cancer Res 2012 1:179). (**d**) Western blot validation of Gab1 knockdown in MCF10A and T47D cells expressing ΔCCDC170 variants.



Supplementary Figure 12. Tandem duplication as a potential genetic mechanism generating *ESR1-CCDC170* fusions. (**a**) The schematic of the tandem duplication potentially generating the *ESR1-CCDC170* fusions. **(b)** Relative expression of wild-type *ESR1* and *CCDC170* in 12 fusion-positive and 380 fusion-negative ER+ breast cancer cases with available Agilent gene expression data. Gene expression data shown here are from TCGA. Error bars represent the standard deviation (+, n=12; -, n=380). * p<0.05 (t-test).

Supplementary Table 1. The clinicopathological data for the *ESR1-CCDC170* positive tumors revealed by BWA alignment of reconstructed chimerical sequences.

TCGA sample ID	BWA fusion reads (n=)			CNG	PAM50 subtype		Histology	Receptor Status			AJCC classifications				MP stat
	Total	Encompass	Junction		Gene Exp	RNAseq		ER	PR	Her2(IHC)	T	N	M	Stage	
TCGA-A1-A0SM-01	5	2	3	Yes	LumB	LumA	IDC	+	-	+	T2	N0 (i-)	M0	IIA	NA
TCGA-A1-A0SN-01	21	17	4	Yes		LumB	IDC	+	+	+	T1c	N1	MX	IIA	Post
TCGA-A2-A0CT-01	39	25	14	Yes	LumB	LumA	IDC	+	-	+/-	T2	N0 (i-)	M0	IIA	Post
TCGA-A2-A0YG-01	10	5	5	No	LumB	LumB	IDC	+	+	+	T2	N3a	M0	IIIC	Post
TCGA-A2-A25B-01	7	4	3	No		LumB	IDC	+	+	+/-	T2	N1	M0	IIB	Pre
TCGA-A2-A4S3-01	23	12	11	No			IDC	+	+	-	T2	N1a	M0	IIB	Post
TCGA-A7-A4SC-01	10	6	4	NA			ILC	+	-	+/-	T3	N0	MX	IIB	Post
TCGA-A8-A08P-01	6	2	4	No	LumB	LumB	IDC	+	+	+	T2	N2a	M0	IIIA	Post
TCGA-A8-A09Q-01	4	3	1	No	LumB	LumB	IDC	+	+	-	T4b	N2a	M0	IIIB	Post
TCGA-AC-A3OD-01	4	3	1	No		LumA	ILC	+	+	-	T2	N1mi	MX	IIB	Post
TCGA-AN-A0AK-01	3	0	3	No	LumB	LumB	IDC	+	-	+	T2	N0	M0	IIA	Post
TCGA-AN-A0FW-01*	2	0	2	No	LumA	LumA	IDC	+	+/-	-	T2	N2	M0	IIIA	Post
TCGA-AR-A24R-01	69	31	38	Yes		LumB	IDC	+	+	-	T1	N2	M0	IIIA	Pre
TCGA-BH-A0B1-01	5	3	2	No	LumA	LumA	IDC	+	+	-	T2	N1a	M0	IIB	Post
TCGA-BH-A0C0-01	36	23	13	Yes	LumB	LumB	IDC	+	+	+/-	T1c	N1a	M0	IIA	Post
TCGA-BH-A0C7-01*	2	1	1	No	LumB	LumB	IDC	+	-		T2	N1mi	M0	IIB	Peri
TCGA-BH-A0HY-01	5	3	2	Yes	Her2	LumB	IDC	+	-	+	T1c	N0	M0	I	Post
TCGA-BH-A18R-01	87	47	40	Yes	Her2	Her2	IDC	+/-	-	+	T2b	N1	M0	IIA	NA
TCGA-BH-A1EY-01	3	0	3	No		LumA	IDC	+	+	-	T2	N0	M0	IIA	Post
TCGA-BH-A1FD-01	39	21	18	Yes		LumB	IDC	+	+	-	T1c	N0	M0	I	Post
TCGA-D8-A1JP-01	41	25	16	Yes		LumA	IDC	+	+	-	T1c	N0	M0	IA	Post
TCGA-D8-A1XA-01	6	2	4	Yes		LumA	IDC	+	+	+/-	T1c	N0	M0	IA	Post
TCGA-D8-A27N-01	48	25	23	Yes		LumB	IDC	+	+	+	T2	N2a	M0	IIIA	Pre

Note: Gene-expression based PAM50 subtypes are derived from TCGA publication (*Nature* 2012); RNAseq based PAM50 subtypes are derived from UCSC Cancer Genome Browser (https://genome-cancer.ucsc.edu/). CNG, copy-number gains between the *CCDC170* and *ESR1* genomic loci that affect 3' *CCDC170* and 5' *ESR1*. MP stat, Menopausal status. *These two cases are marginal cases supported by two fusion reads, which are not included in the assessment of fusion-positive rates.

Supplementary Table 2. Summary of the clinicopathological data for the breast tumor samples analyzed in this study

Sample ID	AGE	Sex	Race	RT-PCR	Capillary Sequencing	gPCR	MP stat	ER	PR	HER-2	Grade	AJCC TNM	AJCC Stage	Ki67 Index (%)
BT1	75	F	Caucasian	-			Post	+	+	-	1	T2NXM0	IIA	9.73
BT2	68	F	Caucasian	-			Post	+	-	-	2	T1cNXM0	I	12.98
BT3	64	F	Caucasian	-			Post	+	-	-	2	T1NXM0	I	0.40
BT4	37	F	Caucasian	-			NA	+	-	NA	1	T2N1MX	IIB	21.15
BT5	49	F	Caucasian	-			Pre	+	+	-	1	T1cNXM0	I	
BT6	70	F	Caucasian	-			Post	+	+	-	2	T2N2M0	IIIA	10.97
BT7	63	F	Caucasian	+	Yes	+	Post	+	+	-	2	T2N0M0	IIA	
BT8	51	F	Caucasian	-			NA	+	-	-	2	T1N0M0	I	0.40
BT9	NA	F	Caucasian	-			NA	+	+	-	1	T1N0M0	I	8.79
BT10	74	F	Caucasian	-			Post	+	+	-	1	T2NXM0	IIA	6.00
BT11	46	F	Caucasian	-			Pre	+	+	-	1	T2N2M0	IIIA	13.06
BT12	NA	F	Caucasian	Weak	Yes		Post	+	-	-	1	T1cN1MX	IIA	7.00
BT13	47	F	Caucasian	-			Pre	+	+	-	1	T2N1M0	IIB	22.74
BT14	53	F	Caucasian	-			Post	+	+	-	1	T4N2M0	IIIB	56.00
BT15	32	F	Caucasian	-			Pre	+	-	+	1	T2NXM0	IIA	22.39
BT16	51	F	Caucasian	-			Pre	+	-	-	1	T1N0M0	I	5.00
BT17	69	F	Caucasian	-			Post	+	+	-	2	T2N0M0	IIA	16.60
BT18	53	F	Caucasian	-			Pre	+	+	-	1	T2N0M0	IIA	2.06
BT19	53	F	Caucasian	-			Pre	+	+	NA	1	T1cN1MX	IIA	28.00
BT20	56	F	Caucasian	-			Post	+	+	-	1	T1NXMX	I	21.25
BT21	69	F	Caucasian	Weak	Yes		Post	+	-	-	3	T4N1M0	IIIB	2.67
BT22	40	F	Caucasian	Weak	Yes		NA	+	+	-	1	T1cN1M0	IIA	25.13
BT23	58	F	Caucasian	Weak	Yes		Post	+	+	-	1	T2N1M0	IIB	23.00
BT24	74	F	Caucasian	Weak	Yes		Post	+	+	-	1	T2NXMX	IIA	14.34
BT25	48	F	Caucasian	Weak	Yes		Pre	+	+	-	1	T2NXM0	IIA	27.20
BT26	59	F	Caucasian	-			Post	+	+	-	1	T2N1M0	IIB	44.56
BT27	74	F	Caucasian	-			Post	+	-	-	2	T1N1M0	IIA	3.81
BT28	53	F	Caucasian	-			Post	+	-	-	1	T1N1M0	IIA	3.00
BT29	52	F	Caucasian	-			Pre	+	+	-	1	T2NXM0	IIA	6.18
BT30	73	F	Caucasian	-			Post	+	-	-	2	T2NXM0	IIA	12.62
BT31	55	F	Caucasian	-			Post	+	+	-	1	T2NXM0	IIA	18.00
BT32	39	F	Caucasian	-			Pre	+	+	-	1	T2NXM0	IIA	4.00
BT33	36	F	Caucasian	-			Pre	+	-	-	1	T1NXM0	I	6.12

BT34	51	F	Caucasian	-			Pre	+	+	-	1	T2NXM0	II	6.00
BT35	52	F	Caucasian	-			Pre	+	-	-	1	T2N2M0	IIIA	38.74
BT36	62	F	Caucasian	+	Yes	+	Post	+	-	NA	1	T2N2M0	IIIA	64.07
BT37	NA	F	Caucasian	-			Pre	+	+	+	2	T2N2M0	IIIA	
BT38	62	F	Caucasian	-			Post	+	-	NA	1	T1cNXM0	I	13.44
BT39	44	F	Caucasian	-			Pre	+	+	-	1	T1N1M0	IIA	32.50
BT40	68	F	Caucasian	-			Post	+	+	-	1	T4NXM0	IIIB	5.70
BT41	41	F	Caucasian	-			Pre	+	+	-	1	T2NXM0	IIA	20.70
BT42	45	F	Caucasian	-			Pre	+	+	-	2	T2N1M0	IIB	30.00
BT43	74	F	Caucasian	-			Post	+	+	NA	1	T2N1M0	IIB	22.00
BT44	71	F	Caucasian	-			Post	+	-	-	1	T2NXM0	IIA	41.00
BT45	61	F	Caucasian	-			Post	+	+	-	1	T1cNXM0	I	38.00
BT46	70	F	Caucasian	Weak	Yes		Post	+	-	NA	1	T2N1M0	IIB	17.00
BT47	70	F	Caucasian	-			Post	+	+	-	1	T1cNXM0	I	11.50
BT48	56	F	Caucasian	+	Yes		NA	+	-	-	1	T1NXM0	I	28.96
BT49	47	F	Caucasian	-			Post	+	+	-	1	T2N2M0	IIIA	32.00
BT50	77	F	Caucasian	-			Post	+	+	-	1	T2N0M0	IIA	22.00
BT51	47	F	Caucasian	Weak	Yes		Pre	+	+	-	1	T2N0M0	IIA	27.88
BT52	58	F	Caucasian	-			Post	+	+	-	1	T1cNXM0	I	2.92
BT53	36	F	Asian	-			Pre	NA	NA	NA	1			
BT54	42	F	Caucasian	-			Pre	+	-	-	1	T1N0M0	I	2.12
BT55	59	F	Caucasian	-			NA	+	+	-	2	T1N1M0	IIA	6.50
BT56	52	F	Caucasian	-			Post	+	-	-	2	T1N1M0	IIA	42.00
BT57	43	F	Caucasian	-			Pre	+	+	NA	1	T2N2M0	IIIA	15.68
BT58	65	F	Caucasian	Weak	Yes		Post	+	+	-	2	T2NXMX	IIA	8.00
BT59	53	F	Caucasian	-			Pre	+	+	-	1	T2N1M0	IIB	3.29
BT60	75	F	Caucasian	-			Post	+	-	-	1	T2N1M0	IIB	7.32
BT61	79	F	Caucasian	-			Post	+	+	-	2	T4N2M0	IIIB	19.00
BT62	45	F	Caucasian	-			Pre	+	+	-	2	T2NXMX	IIA	6.20
BT63	79	F	Caucasian	-			Post	+	+	-	2	T2NXM0	IIA	18.73
BT64	46	F	Caucasian	-			Pre	+	+	-	1	T1cNXM0	I	12.57
BT65	59	F	Caucasian	-			Post	+	+	-	1	T1NXM0	I	11.40
BT66	43	F	Caucasian	-			Pre	+	+	-	1	T2N0M0	IIA	0.60
BT67	52	F	African American	-			Post	+	+	-	1	T2N2M0	IIIA	13.11
BT68	68	F	Caucasian	-			Post	+	-	-	1	T3N1M0	IIIA	0.99

BT69	45	F	Caucasian	-			NA	+	-	-	1	T2N0M0	IIA	12.98
BT70	53	F	Caucasian	-			Post	+	+	NA	1	T2N3M0	IIIC	13.15
BT71	53	F	Caucasian	-			Pre	+	+	-	1	T2N0MX	IIA	10.62
BT72	44	F	Caucasian	Weak	Yes		Pre	+	-	+	1	T2N2M0	IIIA	20.63
BT73	50	F	Caucasian	-			Pre	+	+	-	1	T2N3M0	IIIC	10.17
BT74	37	F	Caucasian	-			Pre	+	+	-	1	T3N1M0	IIIA	17.67
BT75	51	F	Caucasian	-			Pre	+	+	-	1	T2N1M0	IIB	27.97
BT76	55	F	Caucasian	-			Post	+	+	-	1	T2N0M0	IIA	10.10
BT77	56	F	Caucasian	-			NA	+	+	-	1	T2N1M0	IIB	6.76
BT78	65	F	Caucasian	-			Post	+	+	-	2	T1bN0M0	I	8.46
BT79	47	F	Caucasian	Weak	Yes		NA	+	+	-	1	T1N1M0	IIA	44.00
BT80	64	F	Caucasian	-			Post	+	+	-	1	T2N1M0	IIB	7.33
BT81	67	F	Caucasian	-			Post	+	+	-	2	T2N0M0	IIA	0.20
BT82	48	F	Caucasian	-			NA	+	+	NA	1	T2N2M0	IIIA	33.25
BT83	47	F	Caucasian	-			NA	+	-	+	1	T1cN0M0	I	6.43
BT84	NA	F	Caucasian	-			NA	+	+	NA	1	T1N1M0	IIA	9.91
BT85	NA	F	Caucasian	-			NA	+	-	NA	1	T2NXM0	IIA	11.05
BT86	81	F	Caucasian	-			Post	+	+	-	1	T2N1aM0	IIB	0.20
BT87	53	F	Caucasian	-			Post	+	+	-	1	T1NXM0	I	18.34
BT88	47	F	Caucasian	-			Pre	+	+	NA	1	T2N0M0	IIA	27.48
BT89	79	F	Caucasian	-			Post	+	-	NA	1	T4N1M0	IIIB	31.73
BT90	88	F	Caucasian	-			Post	+	+	-	1	T1N0M0	I	10.99
BT91	45	F	Caucasian	-			Pre	+	+	-	1	T2N2M0	IIIA	10.66
BT92	69	F	Caucasian	-			Post	+	-	NA	1	T1N2M0	IIIA	24.07
BT93	66	F	Caucasian	Weak	Yes		Post	+	+	-	1	T1N0M0	I	0.39
BT94	54	F	Caucasian	-			Pre	+	+	-	3	T4N3M0	IIIC	7.14
BT95	41	F	Caucasian	Weak	Yes		Pre	+	+	NA	1	T2N1M0	IIB	26.34
BT96	56	F	Caucasian	-			Post	+	+	-	2	T2N2M0	IIIA	17.57
BT97	51	F	Caucasian	-			Post	+	-	-	1	T1NXM0	I	18.00
BT98	51	F	Caucasian	-			Post	+	+	-	1	T1NXM0	I	9.07
BT99	39	F	Caucasian	-			Pre	+	+	NA	1	T2NXM0	IIA	23.87
BT100	70	F	Caucasian	+	Yes	+	Post	+	+	-	1	T2N0M0	IIA	17.41
BT101	78	F	Caucasian	-			Post	+	+	-	1	T2NXM0	IIA	11.01
BT102	60	F	Caucasian	-			Post	+	-	-	1	T2N1M0	IIB	70.97
BT103	74	F	Caucasian	-			Post	+	-	-	1	T4N1M0	IIIB	13.12
BT104	66	F	Caucasian	-			Post	+	-	-	2	T2N0M0	IIA	11.38

BT105	58	F	Caucasian	-			Pre	+	+	-	1	T1cN1M0	IIA	2.00
BT106	32	F	Caucasian	-			Pre	+	-	+	1	T2NXM0	IIA	
BT107	45	F	Caucasian	-			Pre	+	-	-	1	T4N1M0	IIIB	8.06
BT108	46	F	Caucasian	-			Pre	+	+	-	1	T3N1M0	IIIA	36.28
BT109	55	F	Caucasian	-			Post	+	-	-	1	T2NXMX	IIA	28.22
BT110	67	F	Caucasian	Weak	Yes		Post	+	+	-	1	T2N0M0	IIA	0.39
BT111	65	F	Caucasian	+	Yes		Post	+	-	-	1	T1N0M0	I	39.37
BT112	41	F	Caucasian	-			Pre	+	+	-	2	T2N0M0	IIA	0.99
BT113	65	F	Caucasian	+	Yes	+	Post	+	-	-	1	T1cN2M0	IIIA	36.13
BT114	NA	F	Caucasian	-			Post	+	-	+	1	T2N0M0	IIA	0.59
BT115	61	F	Caucasian	-			Post	+	+	-	1	T2N0MX	IIA	0.60
BT116	32	F	Caucasian	-			Pre	+	+	+	1	T2N0M0	IIA	13.66
BT117	59	F	Caucasian	-			Post	+	+	-	2	T2N0M0	IIA	12.00
BT118	43	F	Caucasian	-			Post	+	+	-	1	T4N1M0	IIIB	9.26
BT119	42	F	Caucasian	-			Pre	+	+	-	2	T2N0M0	IIA	24.02
BT120	55	F	Caucasian	-			Post	+	+	-	1	T2N2M0	IIIA	6.68
BT121	72	F	Caucasian	-			Post	+	+	-	1	T2NXM0	IIA	10.97
BT122	53	F	Caucasian	-			Post	+	+	-	2	T2N0M0	IIA	
BT123	69	F	Caucasian	-			Pre	+	+	-	1	T3N0M0	IIB	7.55
BT124	50	F	Caucasian	-			Post	+	+	-	1	T2N0M0	IIA	17.25
BT125	66	F	Caucasian	Weak	Yes		Post	+	-	NA	1	T2N1M0	IIB	12.98
BT126	51	F	Caucasian	-			Post	+	-	NA	1	T2N1MX	IIB	28.48
BT127	59	F	Caucasian	Weak	Yes		Post	+	-	NA	1	T2N1MX	IIB	5.00
BT128	50	F	Caucasian	-			Post	+	+	NA	2	T1N1M0	IIA	12.00
BT129	50	F	Caucasian	-			Post	+	-	+	2	T2N0M0	IIA	19.00
BT130	NA	F	Caucasian	-			NA	+	+	NA	1	T1N2M0	IIIA	13.17
BT131	NA	F	Caucasian	-			NA	+	+	NA	1	T2N1M0	IIB	6.00
BT132	54	F	Caucasian	-			Post	+	+	NA	2	T1N1M0	IIA	24.69
BT133	51	F	Caucasian	-			Post	+	+	+	1	T1N0M0	I	28.46
BT134	75	F	Caucasian	-			Post	+	+	NA	1	T2NXMX	IIA	0.60
BT135	57	F	Caucasian	Weak	Yes		Post	+	+	-	1	T2NXM0	IIA	43.43
BT136	77	F	Caucasian	-			Post	+	+	NA	1	T1NXM0	I	6.74
BT137	58	F	Caucasian	Weak	Yes		Post	+	-	NA	1	T1cNXM0	I	51.92
BT138	68	F	Caucasian	-			Post	+	+	NA	1	T2N2M0	IIIA	38.32
BT139	60	F	Caucasian	-			Post	+	+	NA	1	T1NXM0	I	36.59
BT140	69	F	Caucasian	-			Post	+	+	NA	1	T2NXM0	IIA	53.42

BT141	64	F	Caucasian	Weak	Yes	Post	+	+	NA	1	T2N1M0	IIB	29.34
BT142	64	F	Caucasian	Weak	Yes	Post	+	+	NA	3	T2NXMX	IIA	8.13
BT143	68	F	Caucasian	-		Post	+	+	-	3	T2NXM0	IIA	19.21
BT144	59	F	Caucasian	-		Post	+	+	NA	1	T2NXM0	IIA	13.37
BT145	53	F	Caucasian	-		Post	+	+	NA	1	T1cNXMX	IC	27.85
BT146	65	F	Caucasian	-		Post	+	+	NA	2	T2N1M0	IIB	6.24
BT147	43	F	Caucasian	-		Post	+	-	NA	1	T4N2MX	IIIB	3.43
BT148	70	F	Caucasian	+	Yes	Post	+	+	NA	1	T4N2MX	IIIB	30.75
BT149	37	F	Caucasian	-		Pre	+	+	NA	1	T1N0MX	I	6.70
BT150	67	F	Caucasian	-		Post	+	+	NA	1	T2N1M0	IIB	4.44
BT151	56	F	Caucasian	-		Post	+	+	NA	1	T2NXM0	IIA	7.73
BT152	86	F	Caucasian	-		Post	+	-	-	1	T2NXM0	IIA	0.40
BT153	72	F	Caucasian	-		Post	+	+	-	3	T1N0M0	I	11.84
BT154	37	F	Caucasian	-		Pre	+	+	NA	1	T2NXM0	IIA	34.64
BT155	68	F	Caucasian	-		Post	+	-	+	2	T1cN0M0	I	2.95
BT156	30	F	Caucasian	-		Pre	+	+	NA	1	T1N0M0	I	8.42
BT157	48	F	Caucasian	-		Post	+	+	NA	1	T1cN0M0	I	10.30
BT158	59	F	Caucasian	-		Post	+	+	NA	1	T1cNXM0	I	13.22
BT159	41	F	Caucasian	-		Pre	+	-	-	1	T2N1MX	IIB	0.79
BT160	58	F	Caucasian	Weak	Yes	Post	+	+	-	1	T1N0M0	I	32.25
BT161	54	F	Caucasian	-		Post	+	+	-	2	T2N1M0	IIB	20.62
BT162	47	F	Caucasian	-		NA	+	+	NA	1	T1cN0M0	I	1.38
BT163	71	F	Caucasian	-		Post	+	+	NA	1	T2NXMX	IIA	13.29
BT164	51	F	Caucasian	-		NA	+	+	NA	1	T3N0MX	IIB	19.75
BT165	58	F	Caucasian	Weak	Yes	Post	+	+	NA	1	T1cN1M0	IIB	8.75
BT166	69	F	Caucasian	-		Post	+	+	NA	1	T1cN1M0	IIA	8.38
BT167	47	F	Caucasian	-		NA	+	+	NA	1	T1cN2M0	IIIA	0.40
BT168	53	F	Caucasian	-		Post	+	-	NA	1	T1cN0MX	I	9.78
BT169	NA	F	Caucasian	-		Post	+	+	NA	3	T1N0M0	I	12.42
BT170	70	F	Caucasian	-		Post	+	+	-	2	T2N3M0	IIIC	2.96
BT171	71	F	Caucasian	-		Post	+	-	NA	1	T1cN2M0	IIIA	2.93
BT172	44	F	Caucasian	-		Pre	+	+	-	1	T2N0M0	IIA	0.60
BT173	63	F	Caucasian	-		Post	+	+	+	1	T1cN0M0	I	13.27
BT174	45	F	Caucasian	-		NA	+	-	-	3	T1cN0M0	I	0.99
BT175	37	F	Caucasian	-		Pre	+	+	NA	1	T1cN1M0	IIA	27.07
BT176	42	F	Caucasian	-		Pre	+	+	NA	1	T1cNXM0	I	56.76

BT177	70	F	Caucasian	-			Post	+	+	NA	1	T2N0M0	IIA	15.68
BT178	56	F	Caucasian	-			Post	+	+	-	1	T1NXM0	I	38.34
BT179	50	F	Caucasian	-			Pre	+	+	NA	1	T1cN1M0	IIA	22.13
BT180	44	F	Asian	-			Pre	+	-	+	1	T2NXMX	IIA	47.36
BT181	81	F	Caucasian	Weak	Yes		Post	+	+	NA	1	T2N0MX	IIA	15.50
BT182	39	F	Caucasian	-			Pre	+	+	-	1	T2NXM0	IIA	4.14
BT183	52	F	Asian	-			Post	+	+	NA	1	T2N2M0	IIIA	12.75
BT184	58	F	Caucasian	-			Post	+	+	NA	1	T2NXM0	IIA	17.16
BT185	43	F	Caucasian	-			Pre	+	+	NA	1	T2NXM0	IIA	11.13
BT186	56	F	Caucasian	-			Post	+	+	NA	1	T3N0M0	IIB	20.97
BT187	67	F	Caucasian	-			Post	+	-	+	2	T1cN0M0	I	4.28
BT188	54	F	Caucasian	-			Post	+	+	NA	2	T2N0M0	IIA	2.12
BT189	48	F	Caucasian	-			Pre	+	+	NA	1	T2N1M0	IIB	13.18
BT190	51	F	Caucasian	-			Post	+	+	+	1	T2N2M0	IIIA	17.59
BT191	63	F	Caucasian	Weak	Yes		Post	+	-	-	1	T3N2M0	IIIA	0.00
BT192	66	F	Caucasian	-			Post	+	-	-	2	T4NXM0	IIIB	12.94
BT193	52	F	Caucasian	-			Post	+	-	NA	1	T2N0M0	IIA	59.92
BT194	59	F	Caucasian	Weak	Yes		Post	+	+	-	1	T2N0M0	IIA	20.56
BT195	60	F	Asian	Weak	Yes		Post	+	+	NA	1	T3N1M0	IIIA	2.47
BT196	49	F	Caucasian	+	Yes	+	Post	+	+	-	2	T2N0M0	IIA	7.98
BT197	64	F	Caucasian	-			Post	+	+	-	2	T1cN0M0	I	
BT198	85	F	Caucasian	-			Post	+	+	-	1	T2N2M0	IIIA	27.28
BT199	53	F	Caucasian	-			Post	+	+	+	1	T4N1M0	IIIB	0.96
BT200	66	F	Caucasian	-			Post	+	+	-	1	T1N0MX	I	13.14

Note: MP stat, Menopausal status.

Supplementary Table 3. Primer sequences for RT-PCR and genomic PCR.

ESR1-CCDC170 RT PCR primers	
ESR1-CCDC170_F1	CTGCGGTACCAAATATCAGCAC
ESR1-CCDC170_R1	CTTCTCCAGTTGGTCTCTGGAT
ESR1-CCDC170_F2	AACAGGCTCGAAAGGTCCATGC
ESR1-CCDC170_R2	GCCTAGCATCTGCGACACCACC
P2RY6-ARHGEF17 RT PCR primers	
P2RY6-ARHGEF17_F1	GCTCCACGAGTGGGAATTTG
P2RY6-ARHGEF17_R1	AGCGACTCCACATACGACTG
ESR1 RT PCR primers	
ESR1_F1	CGGGACTGCGGTACCAAATA
ESR1_R1	CAGGGCAGAAGGCTCAGAAA
CCDC170 RT PCR primers	
CCDC170_F1	CGGAATGTGGCTCAAAATGCT
CCDC170_R1	AGAAGTGGAAAGTGCTGCTGA
GAPDH RT PCR primers	
GAPDH_F1	CCCACTCCTCCACCTTTGAC
GAPDH_R1	TCCTCTTGTGCTCTTGCTGG
ESR1-CCDC170 Genomic PCR primers	
ZR-75-1_F1	GCTGGGACTGAAAGCAGAATT
ZR-75-1_R1	AGAAGAGGCTAAAAATTAGGGGACA
MCF7_F1	CATTTGGTTCCTCAGTCCCAGA
MCF7_R1	ATGGCCGGAAATTGGCTAACTA
HCC1428_F1	AGTGGAAGGCTCAACTGCAT
HCC1428_R1	AGAGCTGAGGCAAACACAGG
BT7_F1	GTTGGCCTGACACATAATAGCA
BT7_R1	AACCTACAGACAAGCTACTAAAGAC
BT36_F1	ATTTTTCCCCCATTCTGTAGGT
BT36_R1	ACACCAGCTATATGCTTGATTTCAG
BT100_F1	TGAGGCTCCCAGGATTATTG
BT100_R1	ACTGGGTTATCAGTGCCATTGT
BT113_F1	TCTTACCCCGTCAATAATTTCTCCT
BT113_R1	AGGTTGCCTCTGAAGAGAACTG
BT196_F1	ATGGGAAGAAAAAGTGATGGTG
BT196_R1	AGCAACCACTTATGGGTACAGG