## **Supplemental Information**

### Bisulfite converted DNA was amplified by PCR using the following primers:

### **ChIP** primers

GGATGAGGGGGAGTCGAATG
CAAAAAGTCCCAGGTAGGCG
TCTTCCAGGAGTCTGACCC
AGTGTGGGGCGGAGAGGTG
TGGACAGGTGTGAGTGTGTG
GTCGCCCTCGGTGATCTC
GCCCTCCGCTCCCGCC
GGTTTCCACCTCCAGAAGG

#### Microarray-based mRNA analysis

Gene expression profiling was performed by the core facility of the German Cancer Research Center using a whole-genome Human Sentrix-6 v2 BeadChip (Illumina) containing 48701 transcript probes consisting of genes and known alternative splice variants from the RefSeq database release 17 and UniGene build 188. RNA from patient samples was extracted using a miRNeasy Kit (Qiagen). The quality of total RNA was checked by gel analysis using the total RNA Nanochip assay on an Agilent 2100 Bioanalyzer (Agilent Technologies GmbH). Biotin-labeled cRNA samples for hybridization were prepared according to Illumina's recommended sample labeling procedure based on a modified Eberwine protocol (1). Hybridization was performed at 58°C for 20 h. Spike-in controls for low, medium and highly abundant RNAs were added, as well as mismatch control and biotinylation control oligonucleotides. Microarray scanning was done using a Beadstation array scanner, the setting adjusted to a scaling factor of 1 and PMT settings at 430. Data extraction was done for all beads individually, and outliers were removed when > 2.5 MAD (median absolute deviation). All remaining data points were used for the calculation of the mean average signal for a given probe, and standard deviation for each probe was calculated. Data analysis was performed by normalization of the signals using the quantile normalization algorithm without background subtraction.

#### **Plasmids and oligonucleotides**

Segments of the 3' untranslated region (3' UTR) of RASD1 (600 bp) containing target sites for miR-375 were amplified from genomic DNA with the following primer pairs:

RASD1 forward5'-CACACTAGTGGACCTTTTTGTTAAGTCAAATCC-3'RASD1 reverse5'-CACAAGCTTTTAAAGACCTCAGACACACAGA-3'

The amplicon was cloned into the multiple cloning site of a pMIR-REPORT Luciferase vector (Ambion) located at the 3'UTR of the firefly luciferase gene.

## **Oligonucleotides used for transfection:**

RASD1 siRNA GeneSolution (1027416, Qiagen): pool of siRADS1 2 (CGGTGTCAACATGGACACGAA), siRASD1 4 (CACCAAGTCTTGCCTCAAGAA), siRASD1 5 (CCGCAAGGTCGGTGCAGTA), siRASD1 6 (GAGGGTGGATTTATCTTCTCA); silencer negative control No. 1 (AM4611, Ambion); pre-miR-375 (PM10327, Ambion); pre-miR negative control No. 1 (AM17110, Ambion); anti-miR-375 (AM10327, Ambion); anti-miR negative control No.1 (AM 17010, Ambion). Overexpression assays for RASD1 were performed using RASD1 ORF sequences cloned into the expression vector pT-REX-DEST 30 (12301016, Invitrogen).

## Supplemental Reference

1. Eberwine J, Yeh H, Miyashiro K, et al. Analysis of gene expression in single live neurons. Proc Natl Acad Sci U S A 1992;89:3010-4.

## Supplementary figure legends

**Supplementary Figure S1.** Overview of the genomic region encompassing miR-375 (chromosome 2:219865848-219868241; UCSC Genome Browser Feb. 2009 assembly). The pre-miR-375 hairpin sequence is shown underlined and in capital letters. The two CpG-rich regions are marked by blue lettering. In yellow, potential CTCF-sites identified by the MatInspector tool are indicated (CpG-site no. 18 is shown in capital letters and underlined at position 219866125). Gray boxes mark potential ZEB1 binding sites (E- and Z-boxes). Underlined and in red are the sequence homologous to the mouse miR-375 promoter region identified by Avnit-Sagi et al. is highlighted.

**Supplementary Figure S2.** Validation of knock-down and overexpression assays. qRT-PCR analysis of miR-375 levels in (A) MCF-7 cells following transfection by anti-miRs and (B) HEK293T cells upon transfection with Pre-miRs. (C) qRT-PCR analysis of RASD1 in MCF-7 cells transfected with ORF vectors. (D) qRT-PCR analysis of RASD1 in MCF-7 cells transfected with siRNAs. (E) qRT-PCR analysis of *ERa* mRNA in MCF-7 cells transfected with siRNAs. (F) Immunoblots for ERa in MCF-7 cells transfected with siRNAs.

# Supplementary Figure S1

## chr2:

219865848	-	gcctactacatcgcctgggtttgaatcccaccccttaaaccagcagcata
219865898	-	tttggcaacttggaatcttttagccacacttggaacgtccgtaacattga
219865948	-	aacaagaatgccccatagtgtagtggatgagggggggggtcgaatgggaatct
219865998	-	ccgtagagctcgcggcacaatgtaccgcacggacaggcgagcga
219866048	-	gcgcgggcactgcggccgccctcatgctagcccccccccc
219866098	-	cccggccttcgccaacctgcaggagg <u>CG</u> ctgctggagaacatgatccaga
219866148	-	acatcctggtggaggcgagccgcgggggggggggggtggtactcacctcgcggcca
219866198	-	cgcgtcatcgccctgccgccgttctgcgtgcccgggtaagggcgccggga
219866248	-	ccagcggccgcgcagaccgcgcgagaggggccaccgggggcagccaggctc
219866298	-	tgcccgaagcctcgcccccacgcctacctgggactttttgtcttccagga
219866348	-	gtctgaccccggacacgctgctgccgacgcagcaggggtactccac
219866398	-	${\tt ccggtggtgccacttcctaccgaccttccgtaaatgcccgcgcccagcct}$
219866448	-	ct <mark>ccgacatgccgctaggggtcacgcctg</mark> gccccctctccaccgccagta
219866498	-	a a agc at ctagt ctttttctgt at ctaccttg a cttt at ttccccg acg
219866548	-	agggggaacgcgcgcgcgagctctgagagggaaggatggcagggagcagc
219866598	-	${\tt ccgggcggactgcagaggcgcggggaacttagggcccgggaggatgca}$
219866648	-	<pre>cctctccgcccacactgcagctggactgagacctggccgctagctgccgg</pre>
219866698	-	ccgctggacc <mark>gggtctggccagtggagggcgctccgg</mark> acggggcctggct
219866748	-	ggcctttgccccagcgcaccttggtgcacagtaactgagtgccaggcctg
219866798	-	gccagcgaag <b>caggtg</b> tccactgggcagggcggccacccctcacctttcc
219866848	-	gcgccgcgcctgctcggaggcttagcacccccagggccagccccttagca
219866898	-	tttgccgagctccgagtgcccaagaggctgggggaggaagaact <mark>gaggcac</mark>
219866948	-	gccgcggagggccaggaactaggctggggcgggtcgaggtcaccactgga
219866998	-	tttcgggccacagaccctgctaagcgactcccactgtgaatcagtcctca
219867048	-	gacgaggatggatacccacctctcggttcctcctaacccccgcccg
219867098	-	cttctgtcctctgcttctcggctcctccccagagggcaggtgga
219867148	-	aaagtgacatctggtgttgttccagaggcgcccagggtccccagcaccct
219867198	-	cccctccgcccccgccaccaaggcctcggagaagctccggtctcagagcc
219867248	-	cgggcggggggggggggggggggggggggggggggggg
219867298	-	gagcaaacgggcgcctggacaggtgtgagtgtgtgtgtgt
219867348	-	$actgaacaggcagta {\tt taagagcacacggagcccgggctccgggac} aagct$
219867398	-	ccaaggcgtgggctggggacgaagccaagctaggctgggccgggggg
219867448	-	ggccaggaaagccgggcggaggg <mark>tggctgggaaaggagggggggggc</mark> cagg
219867498	-	tgcgcacctgcggtggccgccgcggccgccgacgtgtcagccgcagatgc
219867548	-	gttcaggtgaggcggaggctagcgggggcgctgtgcagcactgagctcgc
219867598	-	ggaagaccaggaccaggagatcaccgagggcgaccgccaggccccgg <mark>gcc</mark>
219867648	-	ctccgctcccgCCCCGCGACGAGCCCCTCGCACAAACCGGACCTGAGCGT
219867698	-	<u>TTTGTTCGTTCGGCTCGCGTGAGGC</u> aggggggggcgtctctcagcaccagccc
219867748	-	ggggggccggcctgatcgccacgcagg <mark>cacctg</mark> ccgccgccaccgccaccg
219867798	-	ccatctcaaccgtacgggtgggagaggctgtgcgccgctccaggggagat
219867848	-	ccggctcccatccggccccacccgccctgccttgccctgcccgcagcttc
219867898	-	tgggctgccaggctccattctgcgtcgtaatacgggctgagccggctcca
219867948	-	gcgggtggcgaagggggccgtggcccagaacagaaacccttctggaggtg
219867998	-	gaaacccggattctagtccgcctctgtccccgcgacttctatggccttgg
219868048	-	gcaagtcatttcccggcctcttggttttcccatttgctcagtgaacggtt
219868098	-	ggtctggcctcccgcaggggctaggccctccccagcctcctgtcccggtt
219868148	-	cctgggctcctggtcatgctggggaagatgcaggccctcacggagttgga
219868198	-	gccaaggctgaggtgtggaaaatctgggcgggggaagctgaggg

## **Supplementary Figure S2**



#	Cell line	Origin	Tumor type	Tumorigenic	ERa+
1	MCF-7	Pleural effusion	Invasive ductal carcinoma	Yes	Yes
2	T47D	Pleural effusion	Invasive ductal carcinoma	Yes	Yes
3	SK-BR3	Pleural effusion	Adenocarcinoma	Yes	No
4	MDA-MB-231	Pleural effusion	Adenocarcinoma	Yes	No
5	MDA-MB-435	Pleural effusion	Invasive ductal carcinoma	Yes	No
6	HDQ-P1	Primary breast	Ductal infiltrating carcinoma	Yes	No
7	MCF-10A	Primary breast	Fibrocystic disease	No	No
8	MCF-12A	Primary breast	Fibrocystic disease	No	No
9	BT474	Primary breast	Invasive ductal carcinoma	Yes	Yes
10	ZR7530	Ascites fluid	Invasive ductal carcinoma	Yes	Yes

## Supplementary Table S1. Mammary cell lines used in this study

Cell lines 1-8 were used for miRNA profiling by microarrays.

#	Age	Histological subtype	ERα	Grade	Lymph node metastasis
1	45	IDC	Р	III	Yes
2	49	IDC	N	III	Yes
3	51	IDC	Р	II	Yes
4	46	AMC	N	III	Yes
5	74	IDC	Р	III	Yes
6	45	IDC	Ν	III	Yes
7	63	FD	Р	N/A	N/A
8	42	IDC	Р	III	Yes
9	53	IDC	Р	III	Yes

Supplementary Table S2. Clinical information of breast cancer specimens

IDC: Invasive ductal carcinoma; AMC: Atypical medullary carcinoma; FD: Fibrocystic disease; P: positive; N: Negative; N/A: Not applicable