Supplementary Information

Cooperative interaction between ERα and the EMT-inducer ZEB1 reprograms breast cancer cells for metastasis

Ghahhari et al.

Supplementary Information includes Supplementary Results, Supplementary Methods, eight Supplementary Figures with legends, and four Supplementary Tables

Supplementary Results

Zinc finger cluster 1 of ZEB1 and F-domain of ERα are required for augmented ERα activity. ZEB1 is composed of two zinc finger clusters (ZF1 and ZF2) with a central homeodomain. ZEB1 uses its ZF1/2 clusters to bind DNA through specific E-boxes[1,2]. To determine which domain(s) of ZEB1 are involved in its interaction with ERα, we generated several truncation mutants (Supplementary Fig. 3a). Full-length ZEB1 induced ERα activity, while the deletion of ZF1 resulted in the loss of ZEB1-induced ERα regulation. In contrast, the truncation of ZF2 did not affect the ZEB1-mediated activation of the ERE-Luc reporter. Double-deletion mutants of the ZF1/2 significantly reduced the impact on ERα activity and the homeodomain of ZEB1 appeared to be dispensable for ERα-augmented activity (Supplementary Fig. 3b). These findings point to the importance of ZF1 of ZEB1 for increasing ERα activity. To further explore the possible involvement of various ZEB1 domains in the physical interaction with ERα, we performed co-IPs and compared the interaction of ZEB1 mutants with full-length ERα. Mirroring our observations with the luciferase assays, ZF1 was required for the ZEB1-ERα interaction. Excision of neither ZF2 nor the homeodomain affected the ERα interaction with ZEB1 (Supplementary Fig. 3c).

Next, we sought to identify the ERα domain(s) that mediate the activation by ZEB1. We used Gal4-ERα fusion proteins, which consist of the Gal4 DNA binding domain fused to the activation function 1 (AF1), the DNA binding domain (DBD), the hormone binding domain (HBD) alone including the AF2, or the HBD with the C-terminal F-domain of ERα (Supplementary Fig. 3d). In response to E2, only the activity of the HBD chimera with the F-domain proved to be stimulated by ZEB1, albeit only modestly, suggesting that the F-domain of ERα plays a key role for the enhancement of ERα activity by ZEB1 (Supplementary Fig. 3e). Indeed, the F-domain modifies the HBD function, and in its absence ERα fails to activate transcription through SP1[3]. To explore whether the ZEB1 ZF1 and the F-domain of ERα are sufficient to mediate a direct physical interaction, we did a GST pulldown with the purified recombinant proteins GST-F-domain and His-tagged ZF1 (Supplementary Fig. 3f). We could not observe a direct interaction of these ZEB1 and ERα domains, suggesting that either a direct interaction requires other domains as well or that the ZEB1-ERα interaction is indirect through another protein(s) in the same complex.

Supplementary Methods

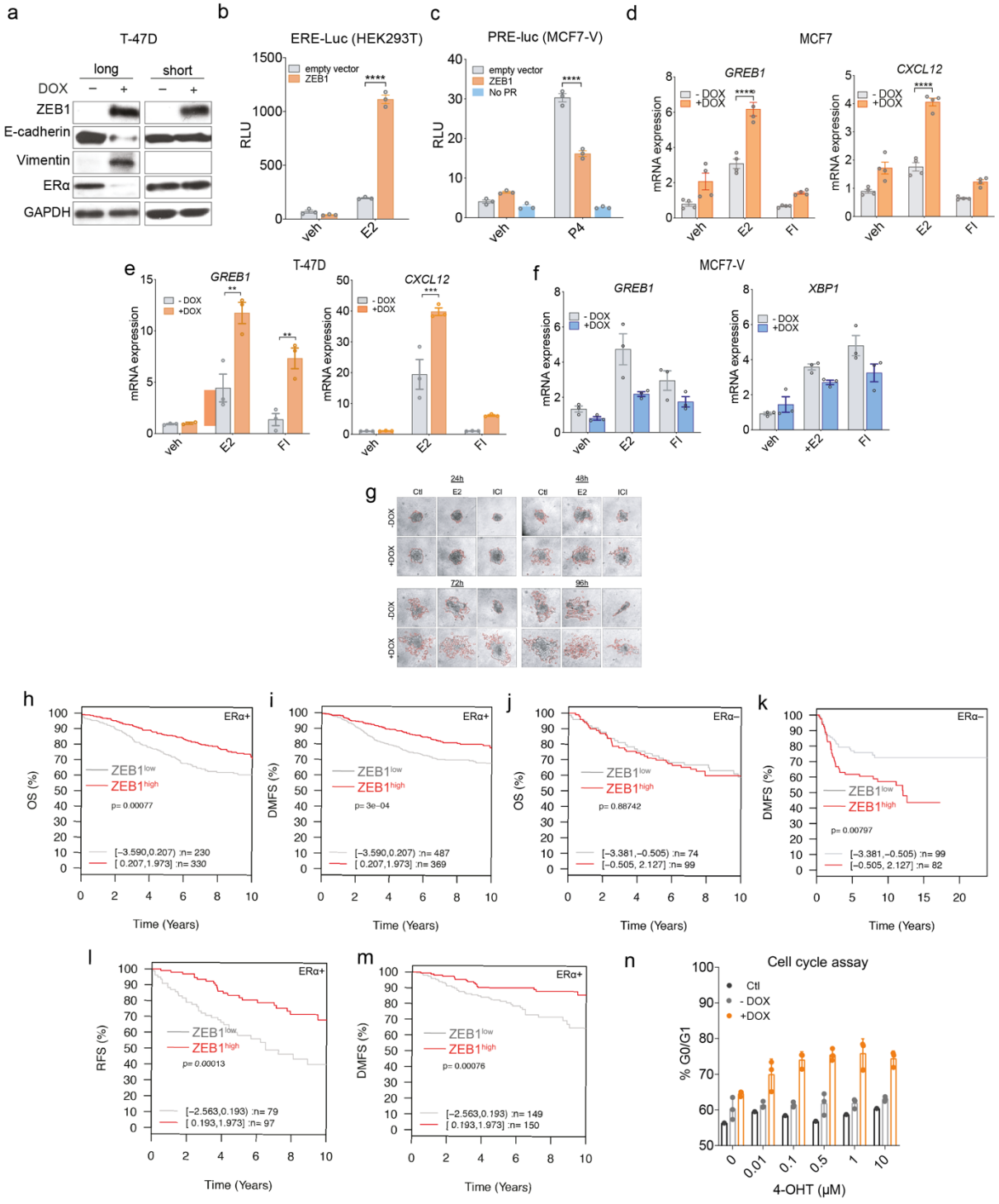
Cell lines. HEK293T and MCF7 cells were purchased from the American Type Culture Collection (ATCC). T-47D cells of the European Collection of Authenticated Cell Cultures (ECACC) were purchased from Sigma Aldrich). MCF7-V cells were a gift from Dr. Wilbert Zwart (Netherlands Cancer Institute, Amsterdam); their highly polymorphic short tandem repeat loci (STRs) were profiled using commercial services (ATCC and Microsynth) and found to be closely (88%) related to wild-type MCF7 cells; specifically, their eight core STR markers were as follows (alleles are indicated in parenthesis): D5S818 (11, 12), D13S317 (11), D7S820 (8, 9), D16S539 (11, 12), vWA (14, 15), TH01 (6), TPOX (9), and CSF1PO (10, 11).

Reagents. The mouse monoclonal antibody against GST (B-14, sc-138, 1:1000) was from Santa Cruz Biotechnology. The mouse monoclonal antibody against the His6 tag (anti-6-His (H1029), 1:5000) was from Sigma Aldrich. Glutathione-coupled magnetic agarose beads (GoldBio cat No. G-250) were used for GST-pulldowns.

Plasmids. The constructs pTRIPz-ΔZF1-ZEB1, pTRIPz-ΔZF2-ZEB1, pTRIPz-ΔZF1/2-ZEB1, and pTRIPz-ΔHD-ZEB1 for the expression of the corresponding ZEB1 truncation mutants were generated by Gibson assembly (In-Fusion HD Cloning Plus, Takara Bio) with the primers listed in Supplementary Table 4. Briefly, for pTRIPz-ΔZF1-ZEB1, the sequences encoding the zinc finger (ZF) cluster 1 domain were deleted with the assembly of two PCR amplicons corresponding to sequences for aa 1-149 and aa 272-1061 from the full-length ZEB1. HA-ZEB1 fragment was excised by restriction digestion to generate the empty lentiviral vector pTRIPz. To construct pTRIPz-ΔZF2-ZEB1, two PCR fragments corresponding to coding sequences for aa 1-881 and aa 960-1061 were assembled into the empty vector pTRIPz. For plasmid pTRIPz-ΔZF1/2-ZEB1, three fragments were PCR amplified from the full-length ZEB1 (corresponding to sequences for aa 1-149, aa 272-881 and aa 960-1061) and were ligated into the empty vector. Finally, to engineer the construct pTRIPz-ΔHD-ZEB1 for the homeodomain mutant, two PCR amplicons equivalent to the coding sequences for aa 1-560 and aa 619-1061 from full-length ZEB1 were assembled into the vector pTRIPz. pET32-ZEB1-ZF1 was generated by PCR amplification of the sequences for ZF1 of ZEB1 from plasmid pTRIPz-puro-HA-ZEB1 and inserted into the bacterial expression vector pET-32Ek/LIC (Novagen). To generate an expression vector for GST fused to the ERα F-domain, the respective sequences were PCR amplified from plasmid HEG0 and inserted into the bacterial expression vector pGEX-2T87. The following constructs were used for the Gal4 fusions: plasmid pSCTEV gal93 for the expression of the following fusions with the Gal4 DNA binding domain (DBD, aa 1-93)[4]: Gal93.ER(G) (also referred to as Gal4.ERα) with the ERα HBD (aa 282-595), Gal4.ER(AF1) with AF1 (aa 82-152)[5], Gal4.ER(AF2) with AF2 (aa 282-595), Gal93.ERΔF with the HBD without the F-domain (aa 282-553)[6], and Gal93.ER-DBD with the ERα DBD (aa 180-262)[7].

Protein purifications. His-ZF1 was expressed in the bacterial strain Rosetta (DE3) by induction with 1mM IPTG at 18°C for 4 h. Bacterial pellets were resuspended in the lysis buffer (50 mM Na_2HPO_4, 2 mM CHAPS, 500 mM NaCl, 5 mM 2-mercaptoethanol, 10 mM imidazole pH 8) with 0.1 mg/ml lysozyme and 0.1 mg/ml DNase. Resuspended bacterial pellets were then lysed with an Emulsiflex (4 passages), and cell debris were discarded after centrifugation at 13000 x g for 30 minutes at 4°C. The soluble fraction was affinity-purified on a His-Trap FF column (GE Healthcare Life Sciences) with an imidazole gradient of 10-400 mM with the elution buffer (50 mM NaH_2PO_4, 100 mM NaCl, 5 mM 2-Mercaptoethanol, 400 mM Imidazole pH 8) followed by size exclusion chromatography on a Superdex 200/ 16.6. GST-F-domain was expressed in the same bacterial strain by induction with 1 mM IPTG at 25°C for 3 h. Bacterial pellets were lysed in PBS supplemented with 1 x EDTA-free protease inhibitor cocktail (A32965, Thermo Scientific), DNase 0.1 mg/ml, and 0.1 mg/ml lysozyme. Extracts were processed as above, incubated with glutathione-agarose beads (GoldBio cat No. G-250) in PBS; the bound material was eluted with 10 mM reduced L-glutathione in 50 mM Tris-HCl pH 8.

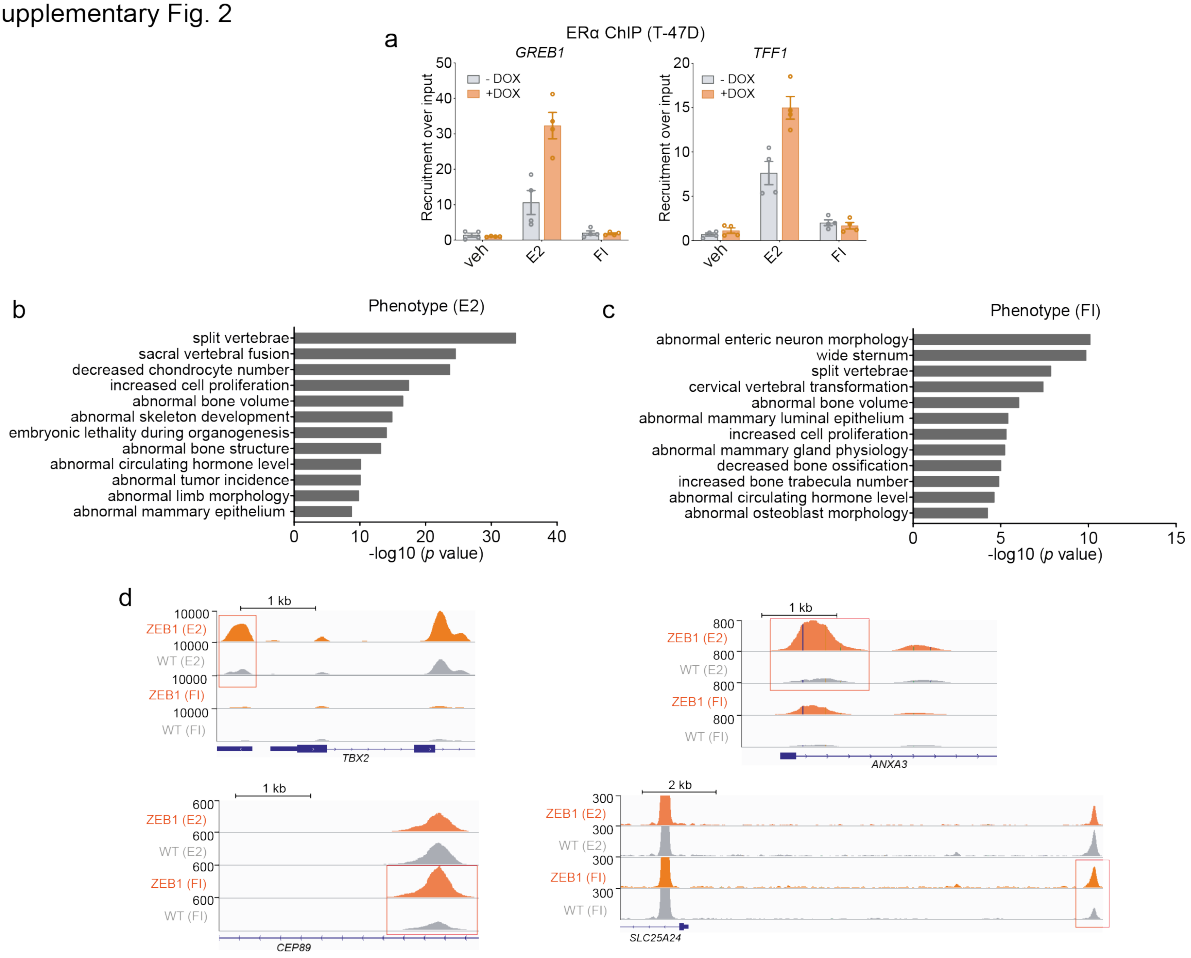
Supplementary Fig. 1



Supplementary Fig. 1 ZEB1 enhances ERα activity and improves breast cancer patient outcome. a T-47D-ZEB1 cells express ZEB1 in the presence of DOX. Immunoblots show E-cadherin and vimentin levels in cells after long-term (8-12 weeks) and short-term (1-2 weeks) expression of ZEB1. **b, c** Luciferase reporter assays in HEK293T cells transiently transfected with ERE-Luc (**b**) or with PRE-Luc in MCF7-V cells (**c**). **d, e** mRNA levels of ERα target genes in MCF7 (**d**) and T-47D (**e**) cells after induction of ZEB1 expression for one week and following 6 h of treatments

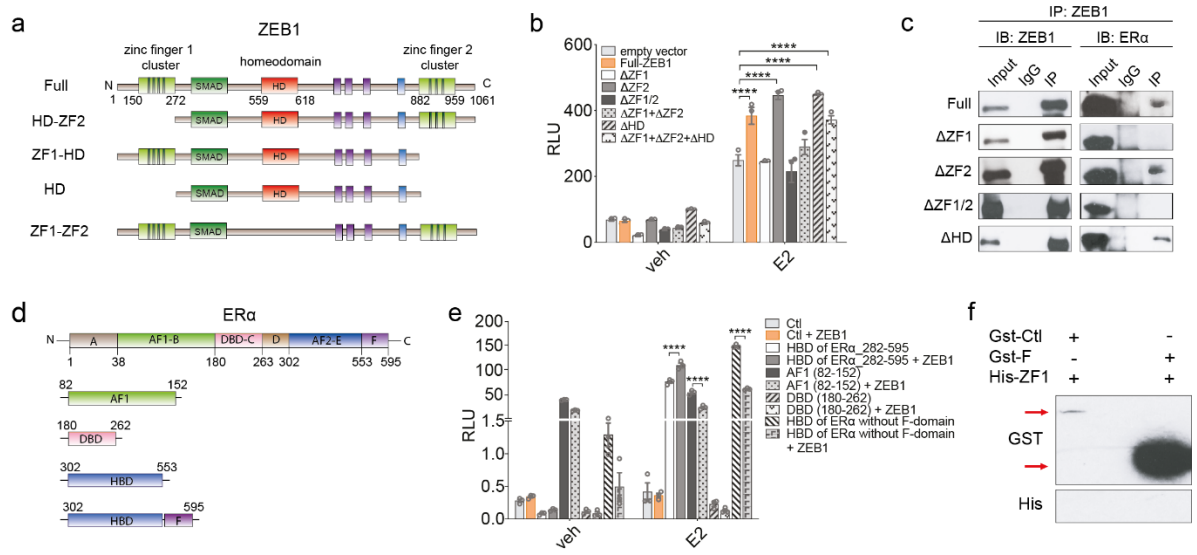
as indicated. **f** mRNA levels of ERα target genes in MCF7-V-ZEB1 cells upon long-term expression of ZEB1 (+DOX). RT-qPCRs were performed following 6 h of indicated treatments. **g** Representative three-dimensional (3D) tumor invasion assays with T-47D-ZEB1 cells treated as indicated (n = 2 biologically independent experiments). **h-k** Kaplan-Meier plots for overall survival (OS) and distant metastasis-free survival (DMFS) of $ERα^+$ (**h, i**) and $ERα^-$ (**j, k**) breast cancer patients classified as tumors expressing high levels (red line) and low levels (black line) of ZEB1. **l, m** Kaplan-Meier plots for relapse free survival (RFS) (**l**) and DMFS (**m**) of $ERα^+$ breast cancer patients treated with tamoxifen classified as tumors expressing high levels (red line) and low levels (gray line) of ZEB1. **n** Cell cycle assay with MCF7-V-ZEB1 cells after 72 h of treatment with increasing concentration of 4-OHT. Ctl, veh, E2, FI, and ICI stand for control, vehicle, 17β-estradiol, forskolin + IBMX, and fulvestrant, respectively. Panels **b-f** and panel **n** represent n = 3 biologically independent experiments. All error bars represent standard errors of the means (mean ± SEM). Statistical significance was determined with a two-way ANOVA. Asterisks indicate significant differences compared to the –DOX (veh) (** $P = 0.001$ and *** $P = 0.0002$, **** $P < 0.0001$). Source data are provided as a Source Data file.

Supplementary Fig. 2



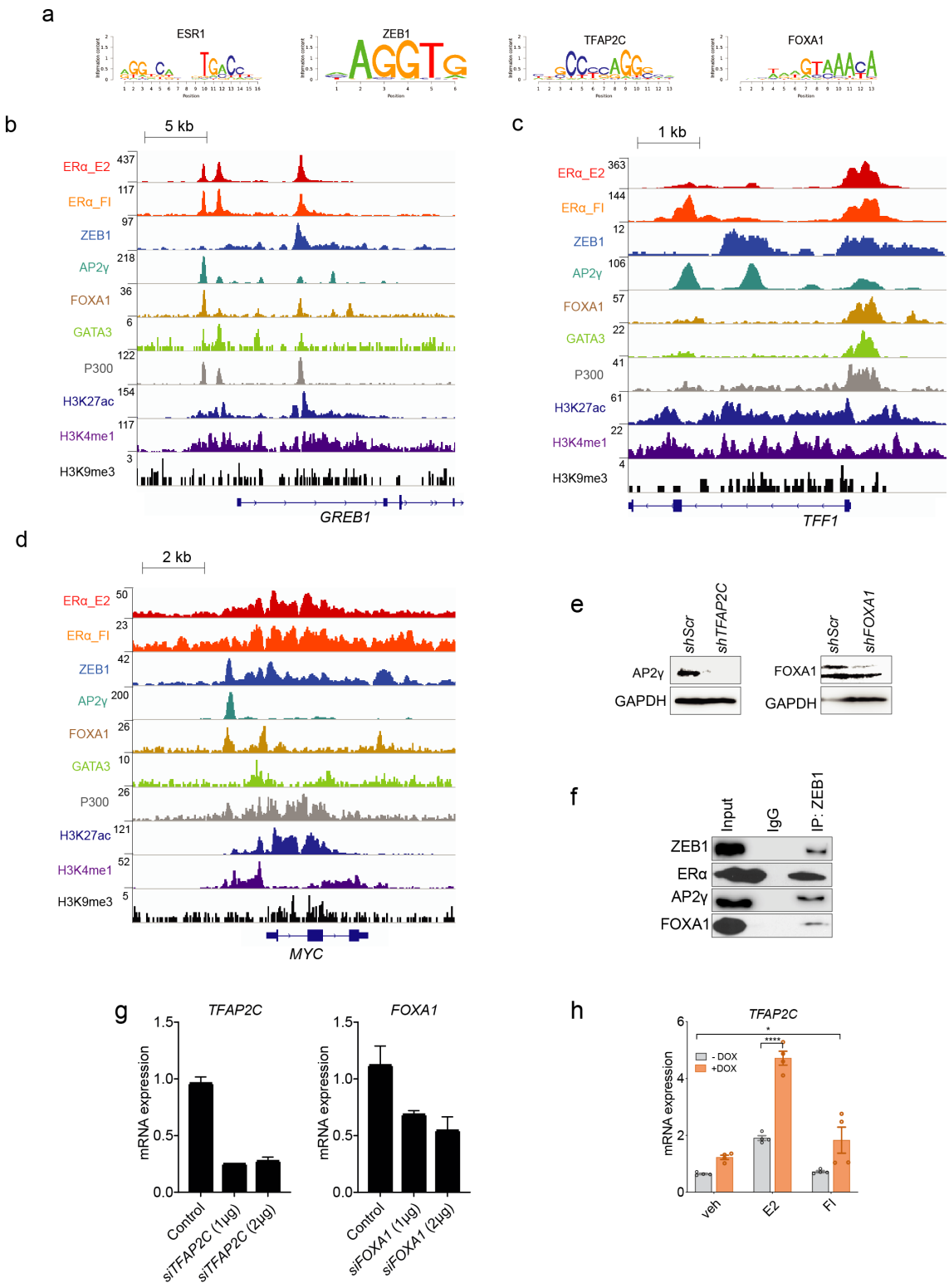
Supplementary Fig. 2 ZEB1 potentiates ERα binding on novel binding sites. a ChIP-qPCR of ERα on the proximal binding sites of the *GREB1* and *TFF1* genes in T-47D-ZEB1 cells. All error bars represent standard errors of the means (mean ± SEM) (n = 4 biologically independent experiments). **b, c** Functional annotations for the biological phenotypes of E2- (**b**) or FI-only (**c**) ERα binding sites, using GREAT. **d** Genome bowser views of ERα binding sites adjacent to the *TBX2*, *ANXA3*, *CEP89*, and *SLC25A24* genes. Orange boxes indicate the sites with altered ERα binding in presence of indicated treatments. Veh, E2, and FI stand for vehicle, 17β-estradiol, and forskolin + IBMX, respectively. Source data are provided in the Source Data file and Supplementary Table 1.

Supplementary Fig. 3



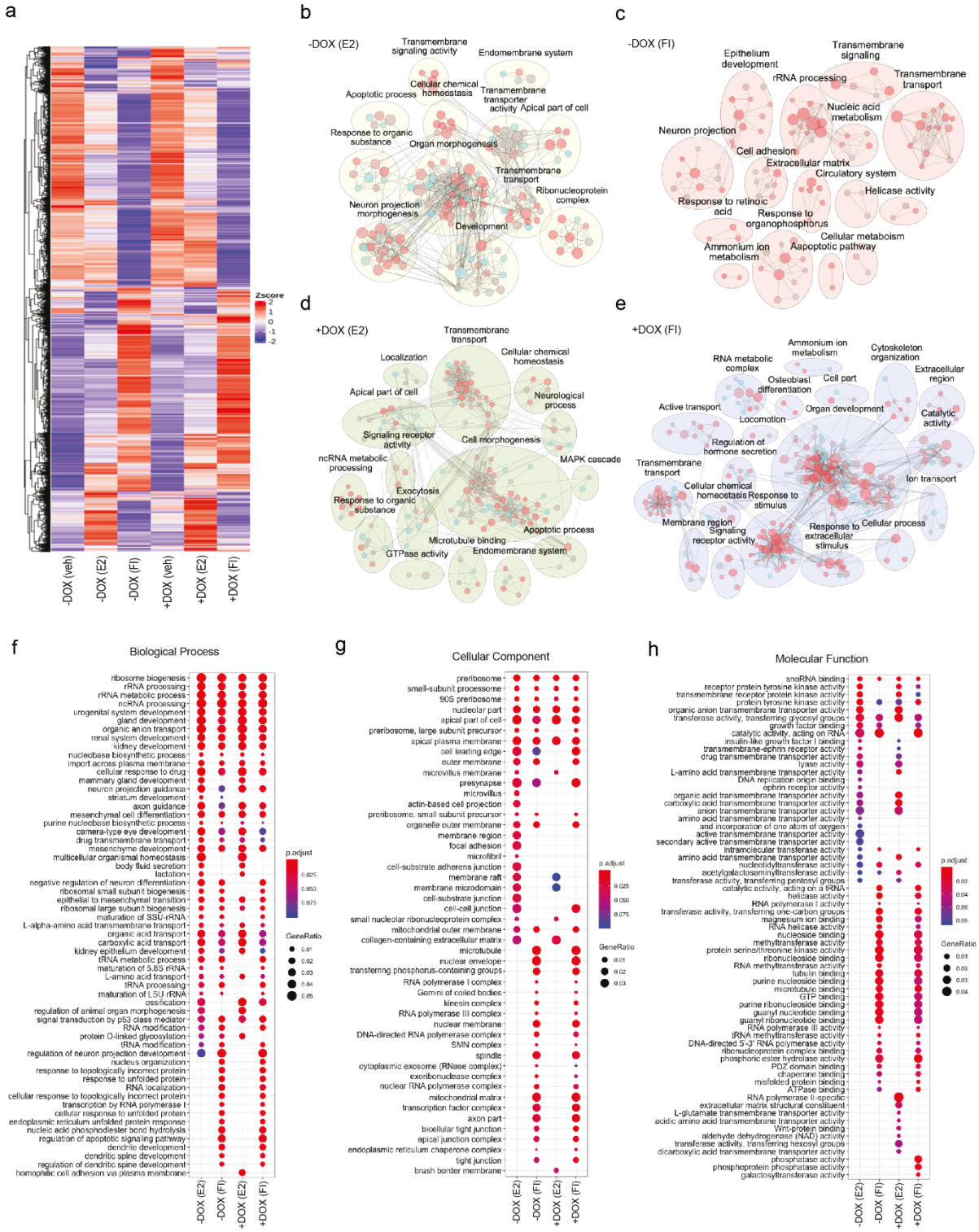
Supplementary Fig. 3 Analyses of ZEB1 and ERα domains in potentiating the ERα activity. a Schematic representation of ZEB1 truncation mutants. **b** ERE-luciferase reporter assays were performed upon transient coexpression of ERα with full-length or truncated ZEB1 mutants in HEK293T cells with the indicated treatments. **c** IP experiment assessing the interaction of ZEB1 truncations with ERα. Full-length ZEB1 and mutants were expressed with ERα in HEK293T cells. **d** Schematic overview of different ERα domains fused to the Gal4 DNA binding domain. **e** ERE-luciferase reporter assays in HEK293T cells were transiently coexpressing ZEB1 and Gal4-ERα domain fusion proteins as indicated. Ctl (control) is the plasmid pSCTEV gal93 for expression of the Gal4-DBD by itself. The luciferase activities in presence of indicated treatments are represented as fold activation over control vector values. **f** Immunoblots of a GST-pulldown experiment with the purified recombinant proteins GST-F-domain and His-tagged ZF1. An unrelated GST-tagged protein was used as negative control. ZF, HD, AF1, DBD, HBD, veh, and E2 stand for zinc finger, homeodomain, activation function 1, DNA binding domain, hormone binding domain, vehicle, and 17β-estradiol, respectively. For all luciferase reporter assays, statistical significance was determined with a two-way ANOVA (n = 3 biologically independent experiments). All graphs with error bars represent the means ±SEM. Asterisks indicate significant differences (**** $P < 0.0001$). Source data are provided as a Source Data file.

Supplementary Fig. 4



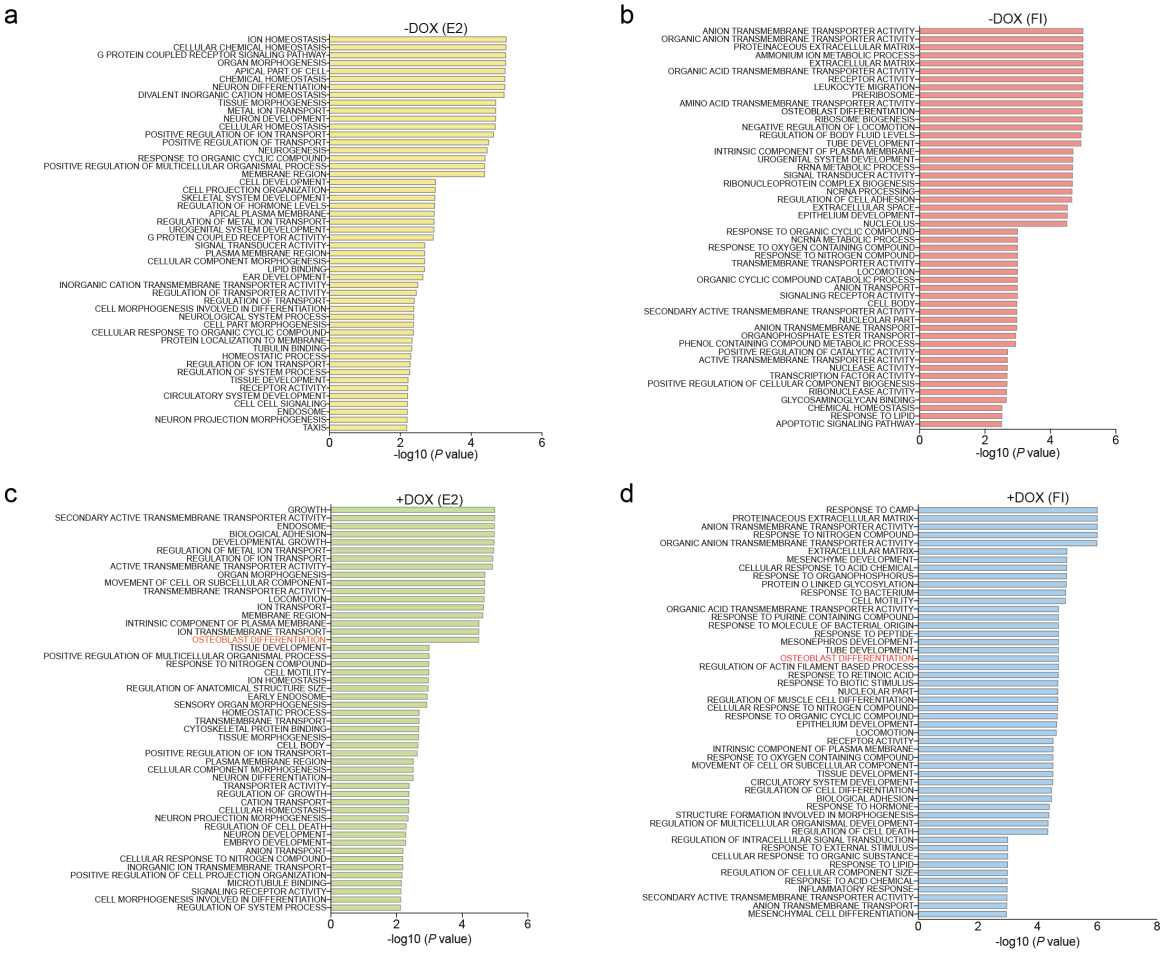
Supplementary Fig. 4 ZEB1 is a part of the ERα TF-complex during early EMT stages and AP2γ determines the interaction between ZEB1 and ERα. a Motif enrichment analysis for the regions at the intersections of ZEB1 and ERα binding sites. Motifs were determined by SeqPos analysis using known or *de novo* motif search with the JASPAR motif matrix. **b-d** Examples of genome browser views for the *GREB1* (**b**), *TFF1* (**c**) and *MYC* (**d**) genes, of the binding sites of ZEB1, ERα, AP2γ, FOXA1, GATA3, and P300, and the open chromatin histone marks H3K27ac, H3K4me1, and H3K9me3. Except for ERα, ChIP-seq data were from publicly available datasets: GSE89206 (ZEB1), GSE21234 (TFAP2C), GSE25315 (FOXA1), GSE60270 (GATA3, P300, H3K27ac, H3K4me1, and H3K9me3). **e** Immunoblots presenting the knockdown efficiencies in cells infected with lentiviral constructs for shRNAs targeting *TFAP2C* and *FOXA1* mRNAs. Scrambled shRNA (shScr) was used as negative control. **f** Immunoblots of ZEB1, ERα, AP2γ, and FOXA1 in ZEB1 IPs. An IP with an IgG antibody was used as a negative control. **g** RT-qPCR analyses in cells transiently transfected with a mixture of siRNAs targeting *TFAP2C* (left panel) or *FOXA1* (right panel) in HEK293T cells. A universal negative control siRNA was used as a control. **h** *TFAP2C* mRNA expression levels were analyzed by RT-qPCR following 6 h of treatments as indicated. Veh, E2, and FI stand for vehicle, 17β-estradiol, and forskolin + IBMX, respectively (n = 4 biologically independent experiments). All error bars represent standard errors of the means (mean ± SEM). Statistical significance was determined with a two-way ANOVA (* $P = 0.01$, **** $P < 0.0001$). Source data are provided as a Source Data file.

Supplementary Fig. 5



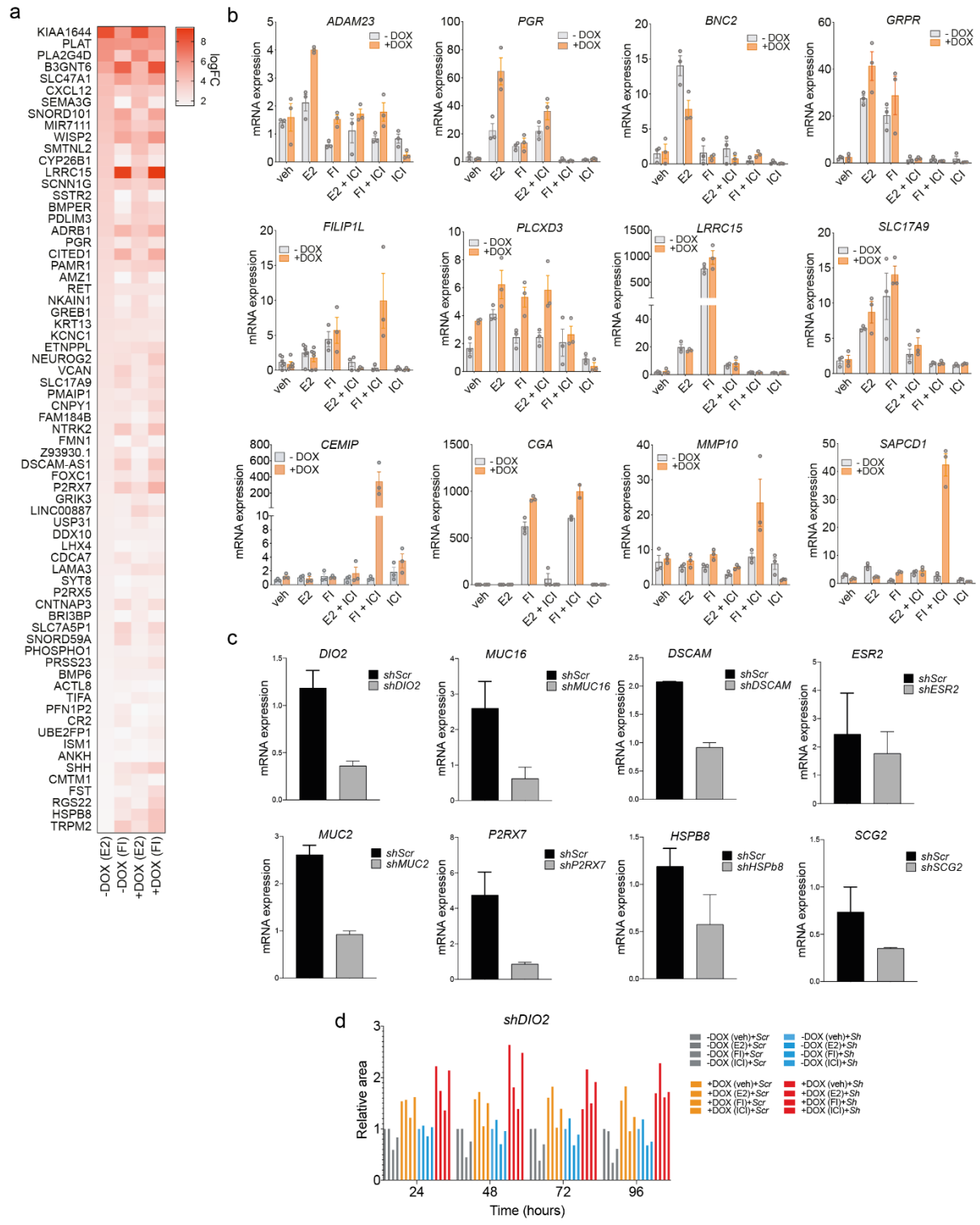
Supplementary Fig. 5 GO and GSEA analyses of the RNA-seq data. a Clustering heat map of the RNA-seq data (–DOX (vehicle), –DOX (E2), –DOX (FI), +DOX (vehicle), +DOX (E2), and +DOX (FI)), based on merged replicates from two independent biological replicates. **b-e** Cytoscape enrichment map visualization of a GSEA of the RNA-seq data comparing the functions of gene set clusters in each of the following groups: –DOX (E2) (**b**), +DOX (E2) (**c**), –DOX (FI) (**d**), and +DOX (FI). (**e**). Red node color indicates positive gene set enrichment and augmentation of the function, and blue nodes are indicative of suppression of that function. Node size represents the number of genes enriched in each function (FDR < 0.05). **f-h** Bubble plots of top GO terms compared in the four groups: –DOX (E2), –DOX (FI), +DOX (E2), and +DOX (FI). The GO classes were separated into three major categories: biological process (**f**), cellular component (**g**) and molecular function (**h**). Veh, E2, and FI stand for vehicle, 17β-estradiol, and forskolin + IBMX, respectively. P values are shown by different colors and the sizes of the bubbles indicate the gene count of each pathway. Source data are provided in Supplementary Data 3-5.

Supplementary Fig. 6



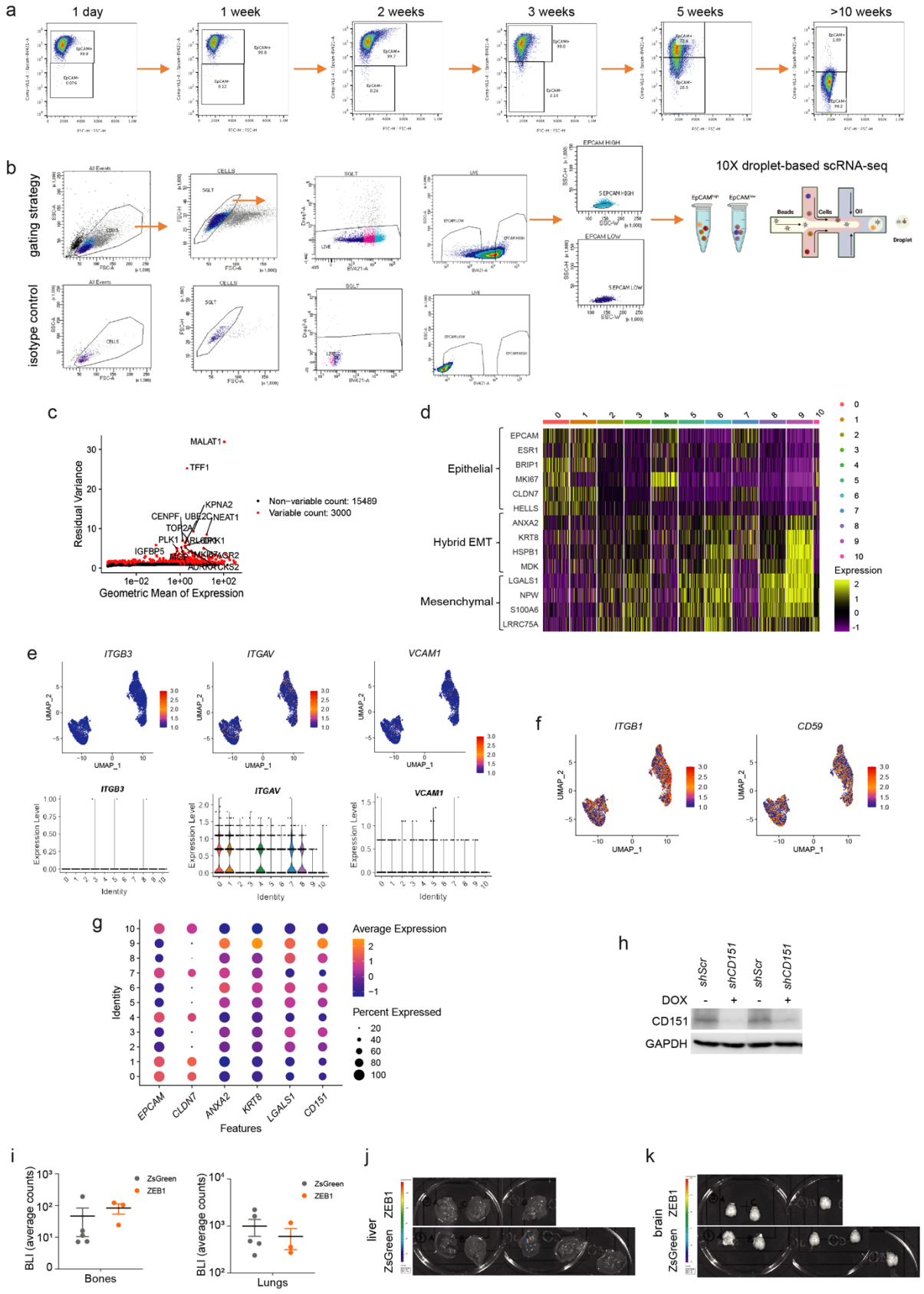
Supplementary Fig. 6 GO analysis of the RNA-seq data. a-d The top 50 GO terms generated from the RNA-seq data by GSEA of differentially expressed ERα target genes in MCF7-V-ZEB1 cells with (+DOX) or without (–DOX) ZEB1 and stimulated with E2 or FI. Bar plots show the -log10(p-value) of GO terms associated with genes whose expression significantly changes (-log10(p-value) > 1.5) with E2 and FI treatments with or without ZEB1 expression. E2 and FI stand for 17β-estradiol and forskolin + IBMX, respectively. Source data are provided as a Source Data file.

Supplementary Fig. 7



Supplementary Fig. 7 Expression and functional analyses of ERα target genes. **a** Heatmap showing the log2 fold changes (log2FC) of top differentially expressed genes, based on the RNA-seq analysis of MCF7-V cells. **b** Bar graphs showing mRNA levels of selected genes, based on an RT-qPCR analysis (n = 3 biologically independent experiments). All error bars represent standard errors of the means (mean ± SEM). **c** Knockdown efficiencies of selected genes assessed by RT-qPCR. **d** Bar graphs representing the changes in the invasion ability of spheroids with (+DOX) or without (–DOX) ZEB1 expression in cells infected with scrambled RNA (scrRNA) or shRNAs targeting the *DIO2* gene. Veh, E2, FI, and ICI stand for vehicle, 17β-estradiol, forskolin + IBMX, and fulvestrant, respectively. The mean of area values (n ≥ 3 biologically independent experiments) was normalized to the mean of the values of the –DOX (shScr) condition treated with veh. Source data are provided as a Source Data file.

Supplementary Fig. 8



Supplementary Fig. 8 FACS analysis strategy, scRNA-seq and additional results of xenograft experiments. a Representative dot plots of EpCAM quantitation by FACS. Time-course of cells gated into $EpCAM^{high}$ and $EpCAM^{low}$ cell populations following DOX-induced ZEB1 expression in MCF7-V-ZEB1 cells. **b** Gating strategy for the FACS-sorting of EpCAM-stained cells for the scRNA-seq analysis of $EpCAM^{high}$ and $EpCAM^{low}$ population of cells. An IgG1 isotype control was used to exclude non-specific antibody bindings. The scheme on the right was created with BioRender.com. **c** Variable feature plot demonstrates the 15 most highly variable genes from the normalized scRNA-seq data of the combined scRNA-seq datasets of the $EpCAM^{high}$ and $EpCAM^{low}$ populations. **d** Heatmap representing the heterogenous expression of specified genes from the epithelial, hybrid EMT, and mesenchymal states distributed across the 11 clusters of the scRNA-seq. **e** Feature plots and violin plots showing the distribution of single cell transcripts for the *ITGB3* (CD61), *ITGAV* (CD51) and *VCAM1* (CD106) genes. **f** Feature plots presenting the distribution of *ITGB1* (CD29) and *CD59* gene expression levels from single cells. **g** Dot plot showing examples of gene expression levels from epithelial (*EpCAM*; *CLDN7*), hybrid (*ANXA3*; *KRT8*) and mesenchymal (*LGALS2*; *CD151*) states. **h** Control immunoblots showin the efficiency of knockdown of CD151 protein levels in cells targeted with a pool of two different shRNAs. **i** Bar graphs showing the quantitative analyses of BLI counts of the indicated organs in the xenograft experiments with ZsGreen- or ZEB1-expressing wild-type MCF7 cells injected into the mammary fat pad ($n \geq 3$ animals); the corresponding BLI are shown in Fig. 7b-c. The BLI of the other two organs were not quantitated because of very low values, consistent with the marked tissue preferences of metastases. Note the log10 scale of the Y-axes ($n \geq 3$ animals, error bars = mean ± SEM). **j**, **k** BLI images of liver and brain, respectively, showing no metastatic lesions. Source data are provided as a Source Data file.

Supplementary Table 1. Details of the shRNA target sequences used for the knockdown experiments.

shRNA	Target sequence 5'-3'	Reference / TRCN*/manual design**
shScr	CCTAAGGTTAAGTCGCCCTCG	[8]
sh1*TFAP2C*	GCTGAGCTATCTCCTAACTTT	TRCN0000019744
sh2*TFAP2C*	CCTCAGCTCTACGTCTAAATA	TRCN0000019746
sh1*FOXA1*	GCGAAGTTTAATGATCCACAA	TRCN0000014878
sh2*FOXA1*	GAACACCTACATGACCATGAA	[9]
sh*MUC16*	GCCAAGACTTTGGCTTCAGAA	TRCN0000180917
sh*DIO2*	GTATTGCCTTGGCTCTATTTG	TRCN0000294313
sh1*ESR2*	CGCCAGTTATCACATCTGTAT	[10]
sh2*ESR2*	AGGCCATGATCCTGCTCAATT	[10]
sh1*P2RX7*	CCTGGCTACAACTTCAGATAC	[11]
sh2*P2RX7*	CCTGGCTACAACTTCAGATAC	Designed manually
sh1*DSCAM*	AAAGAGTTTAGCTGAAATGCT	[12]
sh2*DSCAM*	CCTCCCGAAATTGAGATCAAA	Designed manually
sh1*MUC2*	CGGAGTTTACATCGACAACTA	TRCN0000073548
sh2*MUC2*	CGACTACAAGATACGTGTCAA	TRCN0000073552
sh1*CD151*	ACCTGCTGTTTACCTACAATT	TRCN0000057646
sh2*CD151*	GCAGTCACCACCACCCGAAAT	TRCN0000381613

* TRCN: The RNA consortium number (https://www.broadinstitute.org/rnai/trc)
** shRNAs were designed manually with an online tool at http://sirna.wi.mit.edu using the gene of interest accession number.

Supplementary Table 2. Primer sequences used in the ChIP and re-ChIP qPCR experiments.

ChIP and re-ChIP	Forward primer 5'-3'	Reverse primer 5'-3'
GREB1 -2 kb	GAAGGGCAGAGCTGATAACG	GACCCAGTTGCCACACTTTT
GREB1 +5 kb	GTACTGTGGCTCCAGTCCAAGTA	CAAATGCCACCGTTTCGTGTCT
TFF1 +0.5 kb	CACCCCGTGAGCCACTGT	CTGCAGAAGTGATTCATAGTGAGA
TFF1 +2 kb	GCAATGGGTTTCCACCTCCT	CACCAGGTGAAGATGGAGCC
TFF1 -9 kb	AGCAAGACCAGGGTTGCAT	ACGAGCAGATGGAGGTGGGT
XBP1 -9 kb	ATACTTGGCAGCCTGTGACC	GGTCCACAAAGCAGGAAAAA
CDH1 +8 kb	TCACCCTGCCTGCTTCTGTGT	ACCCTGAGCACCCTGCACA
CDH1 TSS	GGCCGGCAGGTGAACCCTCA	GGGCTGGAGTCTGAACTGA
RARA +17 kb	TTGTTCTGCTACAGCCAGGGT	CTGACCCGTAGTGCACACAGAAT
MYC TSS	TGAGTATAAAAGCCGGTTTTC	TCCCTTCCCAGGACGCCCGCA
c-MYC intron	GCCAGTCCAACCGGCTTATG	GGTTCTCCCAAGCAGGAGCA
TBX2	CGGAGCTGGGTCGGAGG	GCAAAGCATTGACACGACCC
ANXA3	CTGACGGGTGGCTTGGG	ACTGAGGTCCCTCCAGCTC
SLC25A24	AGTGAGATTGAGTGACTTTCCCA	AGGTGGTTGGAACCAGACAAA
CEP89	ACTCCCTGATGCATTGAGCA	GCCCCTGGACATTTGACTGA

Supplementary Table 3. Primer sequences used for RT-qPCR.

mRNA	Forward primer 5'-3'	Reverse primer 5'-3'
GREB1	GGCAGGACCAGCTTCTGA	CTGTTCCCACCACCTTGG
CXCL12	AACACTCCAAACTGTGCCCT	AGTGGGTCTAGCGGAAAGTC
XBP1	CCCTCCAGAACATCTCCCCAT	ACATGACTGGGTCCAAGTTGT
CCDC110	CTGTCCTTCCTCGACCAAACA	AACACACTGCTCAGTAACGGA
ADRB1	CCGGGAACAGGAACACAC	GAAAGCAAAAGGAAATATGTC
MYC	CACCTTGTAGCACGTCCTG	GACTCCCCAAGATGTGGTGG
MUC16	CCTTAACGGTTACAATGAACCTGG	GTGTGAGGGTCTTCAGGTGG
ESR2	AGAGTCCCTGGTGTGAAGCAA	GACAGCGCAGAAGTGAGCATC
P2RX7	ACGTTTGCTTTGCTCTGGTG	ACCTTGGTGTGCACAGAACT
HSPB8	GTGTGTGAATGTGCACAGCTT	TCATGTTTGCCAGACACCTCC
SAPCD1	TTTACACTCAGAGCCTGGTCGC	CTAGCCTTGTTCAGTGGACTTGG
CEMIP	ACCGAGCACATTCCAACTACCG	GGCAGAGATGATTGAGAGGAACG
DIO2	TCGATGCCTACAAACAGGTGAA	CTGGGTACCATTGCCACTGT
MMP10	GACAGAAGATGCATCAGGCAC	AGCTTCAGTGTTGGCTGAGT
ADAM23	GACCCAAGGGTCCTAGTGC	GATGGGGCCTTGCTGAGTAG
SCG2	CGGAGAACGGGGAGGAATATG	GCCATGTTTGAAAGATTTCCTCTT
MUC2	CTGCTATGTCGAGGACACCC	GAGTTGGTACACACGCAGGA
FILIP1L	GAACTGCAGGCTCGAGATGA	TGAGCTTCCAGCAAAGCCAG
PLCXD3	CTCGTCTCAGGGGAAAAACGA	ATCATGAGACCCTGGAATGGC
CGA	TGTGCAGGATTGCCCAGAAT	CTGAAGTATTGGGGCACCCG
SLC17A9	TCTCCAGCGTCTTTGCTCTG	GCCACTGTGGTTGAAGGTCT
PGR	AGCCAGAGCCCACAATACAG	AGTTGTGCTGCCCTTCCATT
BMPER	TGTGTTTGAGGGTGTGCAGT	ATTGTGTCCTGCCTCCAGTG
GRPR	CGCTCTCGGCAGACAGATAC	GGTCTGGTTGGTGCTTTCCT
DSCAM	ATCGGCCAGGTGGTCCAG	TCCGGCTTTTCTGAGGTTCC
BNC2	AAATCAGAGGACAGGCTTAGTGA	TTGAGATGTATCAACCCCACAAC
GAPDH	GCACAAGAGGAAGAGAGAGACC	AGGGGAGATTCAGTGTGGTG

Supplementary Table 4. Primer sequences used for the Gibson assemblies related to the generation of plasmids for expression of ZEB1 truncation mutants.

ZEB1 mutants	Primers
Insert 1 ΔZF1	F: 5' GAACCGTCAGATCGCACCGGCCGGTGAGGACCATGGCATATC 3' R: 5' TCAAGCTAATCAACTGGGAAAATGCATCTGGTG 3'
Insert 2 ΔZF1	F: 5' TTCCCAGTTGATTAGCTTGATGCCTGTGAATGG 3' R: 5' TCAGGATCCGGTTACCTCGATCGAGCCTTCTGTTGGGTTAACC 3'
Insert 1 ΔZF2	F: 5' GAACCGTCAGATCGCACCGGCCGGTGAGGACCATGGCATATC 3' R: 5' AGGAGTAGCGCATTCCATTCTCTGTCTTCCGAGT 3'
Insert 2 ΔZF2	F: 5' GAATGGAATGCGCTACTCCTACTGCAAGAGAGG 3' R: 5' TCAGGATCCGGTTACCTCGATCGAGCCTTCTGTTGGGTTAACC 3'
Insert 1 ΔZF1/2	F: 5' GAACCGTCAGATCGCACCGGCCGGTGAGGACCATGGCATATC 3' R: 5' TCAAGCTAATCAACTGGGAAAATGCATCTGGTG 3'
Insert 2 ΔZF1/2	F: 5' TTCCCAGTTGATTAGCTTGATGCCTGTGAATGG 3' R: 5' AGGAGTAGCGCATTCCATTCTCTGTCTTCCGAGT 3'
Insert 3 ΔZF1/2	F: 5' GAATGGAATGCGCTACTCCTACTGCAAGAGAGG 3' R: 5' TCAGGATCCGGTTACCTCGATCGAGCCTTCTGTTGGGTTAACC 3'
Insert 1 ΔHD	F: 5' GAACCGTCAGATCGCACCGGTGAGGACCATGGCATATCC 3' R: 5' GGTCAGGAGATCCGTTTCCAGCTGATGAAGC 3'
Insert 2 ΔHD	F: 5' TGGAAACGGATCTCCTGACCCCCCTTCTCC 3' R: 5' TCAGGATCCGGTTACCTCGAGCCTTCTGTTGGGTTAACCTGAAG 3'

References for the Supplementary Information

1. Zhang, P., Sun, Y. & Ma, L. ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. *Cell Cycle* **14**, 481-487 (2015).
2. Vandewalle, C., Van Roy, F. & Berx, G. The role of the ZEB family of transcription factors in development and disease. *Cell. Mol. Life Sci.* **66**, 773-787 (2009).
3. Koide, A. et al. Identification of regions within the F domain of the human estrogen receptor α that are important for modulating transactivation and protein-protein interactions. *Mol. Endocrinol.* **21**, 829-842 (2007).
4. Seipel, K., Georgiev, O. & Schaffner, W. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11**, 4961-4968 (1992).
5. Gburcik, V., Bot, N., Maggiolini, M. & Picard, D. SPBP is a phosphoserine-specific repressor of estrogen receptor α. *Mol. Cell. Biol.* **25**, 3421-3430 (2005).
6. Maggiolini, M. et al. Estrogen receptor α mediates the proliferative but not the cytotoxic dose-dependent effects of two major phytoestrogens on human breast cancer cells. *Mol. Pharmacol.* **60**, 595-602 (2001).
7. Carascossa, S., Dudek, P., Cenni, B., Briand, P. A. & Picard, D. CARM1 mediates the ligand-independent and tamoxifen-resistant activation of the estrogen receptor α by cAMP. *Genes Dev.* **24**, 708-719 (2010).
8. Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the Rictor-mTOR complex. *Science* **307**, 1098-1101 (2005).
9. Ni, M. et al. Amplitude modulation of androgen signaling by c-MYC. *Genes Dev.* **27**, 734-748 (2013).
10. Dondi, D. et al. Estrogen receptor β and the progression of prostate cancer: role of 5α-androstane-3β,17β-diol. *Endocr. Relat. Cancer* **17**, 731-742 (2010).
11. Gui, Y. et al. Uridine adenosine tetraphosphate induces contraction of airway smooth muscle. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **301**, L789-794 (2011).
12. Purohit, A. A. et al. Down syndrome cell adhesion molecule (DSCAM) associates with uncoordinated-5C (UNC5C) in netrin-1-mediated growth cone collapse. *J. Biol. Chem.* **287**, 27126-27138 (2012).