Major Classification: Biological Sciences

Minor Classification: Medical Sciences

Title: The Embryonic Transcription Factor SOX9 Drives Human Breast Cancer Endocrine Resistance

Short Title: SOX9 Drives Breast Cancer Endocrine Resistance

Rinath Jeselsohn[1,2,3] MacIntosh Cornwell[1], Matthew Pun[1], Gilles Buchwalter[1,2], Mai Nguyen[1], Clyde Bango[1], Ying Huang[1], Yanan Kuang[4], Cloud Paweletz[4], Xiaoyong Fu[5,6,7], Agostina Nardone[5,6,7], Carmine De Angelis[5,6,7], Simone Detre[8], Andrew Dodson[8], Hisham Mohammed[9], Jason S Carroll[9], Michaela Bowden[1], Prakash Rao[1], Henry Long[1], Fugen Li[1], Mitchell Dowsett[8,10], Rachel Schiff[5,6,7], Myles Brown[1,2,3]

1. Department of Medical Oncology, Dana Farber Cancer Institute, 450 Brookline Ave, Boston MA 02215

2. Center for Functional Cancer Epigenetics, Dana Farber Cancer Institute, 450 Brookline Ave, Boston MA 02215

3. Breast Oncology Center, Dana Farber Cancer Institute, 450 Brookline Ave, Boston MA 02215

4. Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, 360 Longwood Ave, Boston, MA 02215

5. Lester and Sue Smith Breast Center, Baylor College of Medicine, One Baylor Plaza, Houston TX 77030

6. Dan L. Duncan Cancer Center, One Baylor Plaza, Houston TX 77030

7. Department of Molecular and Cellular Biology, One Baylor Plaza, Houston TX 77030

8. Ralph Lauren Centre for Breast Cancer Research, Royal Marsden Hospital, Fulham Rd London UK SW3 6JB

9. Nuclear Transcription Factor Laboratory, Cancer Research UK, Cambridge Institute, Cambridge University, Li Ka Shing Centre Robinson Way, Cambridge, UK CB2 0RE

10. The Breast Cancer Now Toby Robin's Research Centre, Institute of Cancer Research, 123 Old Brompton Rd, London UK SW7 3RP

Corresponding Authors:
Myles Brown, M.D. email: myles_brown@dfci.harvard.edu
Rinath Jeselsohn M.D. email rinath_jeseloshn@dfci.harvard.edu
Dana Farber Cancer Institute , 450 Brookline Ave, D730
Boston, MA 02215 Phone: 617-632-3948 Fax: 617-582-8501

Key words: estrogen receptor, breast cancer, cistrome, endocrine resistance, RUNX2, SOX9

Abstract:

The estrogen receptor (ER) drives the growth of most luminal breast cancers and is the primary target of endocrine therapy. Although ER blockade with drugs such as tamoxifen is very effective, a major clinical limitation is the development of endocrine resistance especially in the setting of metastatic disease. Pre-clinical and clinical observations suggest that even following the development of endocrine resistance, estrogen receptor signaling continues to exert a pivotal role in tumor progression in the majority of cases. Through the analysis of the ER cistrome in tamoxifen resistant breast cancer cells we have uncovered a role for a novel RUNX2-ER complex that stimulates the transcription of a set of genes including most notably the stem cell factor SOX9 that promote proliferation and a metastatic phenotype. We show that upregulation of SOX9 is sufficient to cause relative endocrine resistance. The gain of SOX9 as an ER regulated gene associated with tamoxifen resistance was validated in a unique set of clinical samples supporting the need for the development of improved ER antagonists.

Significance Statement:

Resistance to endocrine treatment remains a significant clinical obstacle. *ESR1* mutations were found to be the mechanism of endocrine resistance in a substantial number of patients with metastatic ER positive breast. However, these mutations are primarily linked to aromatase inhibitor resistance and are not strongly associated with tamoxifen resistance. Herein, we show that tamoxifen treatment promotes a RUNX2-ER complex, which

mediates an altered ER cistrome that facilitates the upregulation of SOX9. We show that upregulation of SOX9, an embryonic transcription factor with key roles in metastases, is a driver of endocrine resistance in the setting of tamoxifen treatment. Our data provides novel putative targets for the development of new strategies to treat tamoxifen resistant breast cancer.

/body

Introduction:

Approximately 70% of breast cancers are hormone receptor positive (HR+) and express estrogen receptor (ERα), progesterone receptor or both. ERα is a nuclear receptor that is a key driver of tumor development and progression and is the most important therapeutic target in HR+ breast cancers. Therapies targeting ERα signaling include aromatase inhibitors that inhibit ERα by blocking the synthesis of estrogen in peripheral tissues and are effective in post-menopausal women or in combination with ovarian suppression in pre-menopausal women. Another type of endocrine therapy is the use of pure anti-estrogens, also called selective estrogen receptor degraders (SERDs), which do not have agonistic activity and cause ERα degradation. A third class of endocrine therapy consists of the selective estrogen receptor modulators (SERMs), such as tamoxifen, which bind to ERα and competitively inhibit estrogen binding in the breast. The efficacy of tamoxifen in the treatment of breast cancer patients was first confirmed in clinical trials conducted over thirty years ago (1-4) and tamoxifen remains an important drug in the adjuvant and metastatic setting of HR+ disease.

Despite the known efficacy of endocrine treatments, endocrine resistance remains an important clinical challenge. A significant number of patients with early stage disease will develop disease recurrence after adjuvant endocrine treatment and in metastatic disease the majority of patients will eventually develop resistance (5). Loss of ERα expression is seen in 10-15% of patients who develop resistance to endocrine treatment. Thus, in the majority of these cases ERα continues to be expressed (6-8). Moreover, ligand independent ERα activity is a key feature underlying the mechanism of endocrine resistance in multiple pre-clinical studies. These mechanisms include increased ERα co-activator interactions(9) and crosstalk between ERα and growth factor receptor pathways, such as the human epidermal growth factor receptor 2 (HER2) and insulin-like growth factor receptor 1 (IGF1R)(10, 11). This crosstalk leads to the activation of the PI3K signaling pathway, which facilitates ligand independent activation of ERα. In addition, we and other groups have detected *ESR*1 ligand binding domain (LBD) mutations in approximately 20% of patients with metastatic HR+ disease. These mutations confer constitutive ligand independent activity and resistance to estrogen deprivation (12). Also, recently reported is a *YAP1-ESR1* translocation, whereby the ER LBD is lost leading to ligand independent growth and endocrine resistance (13). More recently upregulation of FOXA1, a pioneer transcription factor for ERα in breast cancer, has been implicated as a mechanism of endocrine resistance (14). Taken together, these pre-clinical studies suggest that the ERα transcriptional activity remains an important factor in the majority of cases exhibiting endocrine resistance.

A number of clinical trials also support the notion that ERα is a pivotal driver of endocrine resistant breast cancer and continues to be an important therapeutic target. In the metastatic setting, approximately 30% of patients who progress on an aromatase inhibitor respond to fulvestrant (15, 16). Additionally, increasing the dose of fulvestrant resulted in improved disease free survival and overall survival(17). Furthermore, there is evidence that the combination of fulvestrant with an aromatase inhibitor is superior to an aromatase inhibitor alone in first line treatment for metastatic disease(18). In early stage disease it has been demonstrated that prolonging endocrine treatment or enhancing endocrine blockade with the combination of ovarian suppression with an aromatase inhibitor or tamoxifen in pre-menopausal women can improve clinical outcomes in certain patients(19, 20) suggesting that more effective inhibition of ERα signaling may also prevent the development of endocrine resistance.

Upon ERα activation, ERα is recruited to thousands of sites across the genome, defining its cistrome. This process is highly organized through epigenetic events that restrict the recruitment of the receptor to a subset of its potential binding sites in a cell-type specific manner(21-24). In addition to the lineage specific transcriptional program, the ERα cistrome is also dictated by the specific stimuli, as an example the transcriptional response to growth factor stimulation is different from that of estrogen and regulates genes that are overexpressed in HR+ breast cancers that overexpress ERBB2, which may explain endocrine resistance in this setting (25). Moreover, the ERα cistrome is heterogeneous in HR+ breast tumors and distinct binding sites are associated with clinical outcomes(26). Collectively, based on the evidence that the ERα transcriptional activity is a key driver in

endocrine resistance, the ERα cistrome is cell type and stimuli specific and the ERα cistrome linked to clinical outcomes, we hypothesized that altered ERα binding to the genome and the resulting changes of its transcriptional program, constitute a fundamental mechanism of endocrine resistance. To test this hypothesis, we studied the alterations in the ERα cistrome in breast cancer cell line models of endocrine resistance. Given the recent clinical data supporting the continuation of tamoxifen treatment up to ten years and the combination of ovarian suppression with tamoxifen in subsets of pre-menopausal patients (19, 27, 28) our study focused on a model of tamoxifen and estrogen deprivation resistance.

In the present study we demonstrate that with the acquisition of tamoxifen resistance, the transcription factor RUNX2 is upregulated and in complex with ERα induces a distinct ERα cistrome, which regulates the transcription of a set of genes that promote a metastatic phenotype. We show that the distinct ERα–RUNX2 binding sites are also increased in metastatic HR+ patient samples when compared to early stage tumors. Furthermore, the stem cell transcription factor SOX9 is induced by the ERα-RUNX2 complex and is sufficient for the development of resistance to estrogen deprivation and tamoxifen.

Results

Tamoxifen resistant cell growth is ERα dependent

Using an established tamoxifen resistant (TAMR) cell line model derived from MCF7 cells that were grown in long term estrogen deprived conditions and tamoxifen (TAM) (14, 29), we assessed the contribution of ERα to cell proliferation. These studies were performed in the presence of TAM since removal of TAM in this model abrogates cell growth (Figure in

S1). The TAM growth dependency in TAMR cell line models has been demonstrated previously in pre-clinical studies (30) and in the clinic there is evidence of tumor regression with tamoxifen withdrawal in a subset of patients with metastatic disease (31). We first confirmed that ERα is expressed in TAMR cells and, in fact, detected higher levels of ERα when compared to the parental MCF7 cells (Figure 1A). ERα reduction by siRNA inhibited cell proliferation in parental and TAMR cells (Figure 1A). In addition, fulvestrant treatment resulted in a dose dependent reduction in ERα levels and growth inhibition (Figure 1B). Thus, cell proliferation in TAMR remains dependent on ERα.

We further characterized the TAMR cells and observed that these cells acquired morphological changes consisting of cellular elongation with a mesenchymal-like appearance and a more dispersed growth pattern compared to the parental cells (Figure 1C). The TAMR cells also develop filopodia, which are plasma membrane protrusions known to be important for cell migration (32,33). In line with the morphological changes, the TAMR cells exhibit a significantly increased migratory capacity compared to the parental cells as demonstrated by the radius assay and a Boyden chamber (Figure 1C). Furthermore, differential gene expression by RNA-seq analysis revealed 1092 genes up-regulated and 855 genes down-regulated in TAMR compared to parental cells (using DEseq (32) with a log2FC > 1 or < -1). Functional annotation of the up-regulated genes with DAVID (33) showed an enrichment of Gene Ontology (GO) terms that included *response to oxidative stress*, *positive regulation of cell migration, negative regulation of apoptosis, response to hormone stimulus, regulation of programmed cell death and filopodium assembly* ($P < 0.001$). These functions are consistent with the morphological changes and indicative of a metastatic phenotype.

Since the TAMR cells grow under the selective pressure of estrogen deprivation and tamoxifen treatment and remain ER dependent, we tested the cells for the presence of the most common ligand binding *ESR1* mutations (Y537S, Y537N, Y537C, D538G and E380Q). Using droplet digital PCR we did not detect these mutations (Figure in S2), confirming that these mutations do not contribute to endocrine resistance in these cells and suggesting that the mechanism of resistance is due to other alterations in the ER transcriptional axis. In addition we did not detect other mutations or splice variations in *ESR1* in the TAMR cells by RNA-seq.

Tamoxifen resistance leads to the redistribution of ERα-chromatin binding

To determine whether there is differential ERα recruitment after the acquisition of resistance to tamoxifen and estrogen deprivation, we compared the ERα cistromes of parental cells after estradiol stimulation with TAMR and long-term estrogen deprived (LTED) cells (also derived from MCF7 cells) in the absence of ligand stimulation (figure 2A). While the LTED cells without ligand stimulation had a relatively small number of binding sites (Q value of <0.001) and 95% of these peaks overlapped with the parental cistrome, the TAMR cells had a higher number of estradiol (E2) independent binding sites (Q value <0.001) and 37% of these binding sites were unique to the TAMR cells when compared to the parental cells. In addition, over 60% of the parental ligand stimulated binding sites were lost in the TAMR cells without E2 stimulation. Of interest, GATA motifs were enriched in the ERα binding events depleted in the TAMR cells (-log10 (p-value)=686) (Figure 2B and 2C). This finding was shown in previous work (26) and explained here by the loss of GATA3 protein expression with the acquisition of tamoxifen resistance (figure 2E). GATA3 is a transcription factor that promotes mammary luminal

differentiation and a key determinant of luminal type breast cancers (34). A link between GATA3 and ERα that promotes HR+ tumor development is well established; GATA3 and ERα participate in a positive feedback loop in which each transcription factor stimulates the expression of the other and GATA3 is also a pioneer factor and required for ERα binding at sites that lack active histone modifications (35, 36). On the other hand, GATA3 was shown to suppress epithelial-mesenchymal transition (EMT) and metastases and is associated with a favorable outcome in HR+ breast cancers (34, 37). Consistent with the latter role of GATA3, we detected loss of GATA3 expression and consequently loss of ER binding at sites enriched in GATA motifs in the TAMR cells, which have an increased migratory capacity and EMT phenotype.

Motif analysis of the sites that were gained following the acquisition of tamoxifen resistance revealed that these sites were enriched for several motifs with the RUNX motif being highly significant (-log10 p-value=293) (Figure 2B and 2C). The RUNX transcription factors consist of three family members that are master regulators of differentiation in distinct tissues (Runx1-hematopoeisis, Runx2-bone, Runx3-neuronal/gastrointestinal). The RUNX transcription factors also function as tumor suppressors or oncogenes in a cell context–dependent manner(38). Because the three RUNX transcription factors share a common motif (TGTGGT consensus), we looked at the expression level of the three transcription factors in the parental and TAMR cells. While RUNX3 levels were down regulated in TAMR compared to the parental cells, RUNX1 and RUNX2 expression levels were increased in TAMR cells (Figure 2D). We next performed RNA-seq to evaluate the transcript levels of RUNX1, RUNX2 in TAMR compared to matched parental cells in two other cell line models, MDAMB415 and 600MPE. RUNX2

levels were upregulated in both TAMR cell lines compared to the matched parental cells. In contrast, RUNX1 levels were decreased in the MDAMB415 TAMR and 600MPE TAMR cells compared to the parental cells (Figure S3A). In addition, we were able to confirm up-regulation of RUNX2 and not RUNX1 in TAMR cells at the protein level (Figures 2F and S3B). Collectively, the upregulation of RUNX2 in a number of cell lines, the established role of RUNX2 in bone metastases and EMT (39-42) and emerging data on the association between RUNX2 and poor outcomes in HR+ breast cancer (43, 44) led us to focus our initial studies on RUNX2.

The ERα-RUNX regulated genes are enriched in genes associated with metastases and poor outcomes in breast cancer

To determine which genes are uniquely up-regulated by ERα transcriptional activity specifically at sites of the RUNX motif in TAMR cells, we used *binding and expression target analysis* (BETA) to integrate the TAMR unique ERα-RUNX binding sites with the genes differentially expressed between parental and TAMR cell as determined by RNA-seq (45). We identified 461 genes that are up-regulated by ERα-RUNX after the acquisition of tamoxifen resistance with a rank product of <0.01. Ranked gene set enrichment analysis (GSEA) (46) identified signatures up-regulated in MCF7 cells after the expression of MEK (NES=2.5, q-val=0.015), ERBB2 (NES=2.3, q-val=0.091) and EGFR (NES=2.3,q-val=0.071) as the top scoring gene sets enriched in the 461 ERα-RUNX2 up-regulated genes in TAMR. These results suggest that the ERα-RUNX transcriptional activity is involved in the crosstalk between ERα and the receptor tyrosine kinase signaling pathways, which is a key mechanism of ligand independent activation of ERα and endocrine resistance. This finding is supported by previous studies that have shown that RUNX2 is a

downstream mediator of the PI3K/AKT pathway(47, 48). In addition, other gene sets that were highly scored in the ERα-RUNX2 up-regulated genes include signatures of TNFα signaling, IL2-STAT5 signaling and EMT, which are important signatures of the metastatic phenotype (Figure 3A).

To address the relevance of the TAMR ERα-RUNX transcriptional activity in breast tumors, we established the correlation between the TAMR ERα-RUNX regulated genes and breast tumor expression signatures using Oncomine Concepts Map analysis (49). Significant correlations were revealed between the TAMR ERα-RUNX2 regulated genes and gene expression signatures of poor outcome including death and metastatic disease (Figure 3A). We also determined the ERα cistrome of ER+ positive cancer cells isolated from a metastatic pleural effusion from a patient that had progressive metastatic disease after tamoxifen treatment. A significant proportion of the ERα binding sites in this tumor contained a RUNX motif and there was a significant overlap between the ERα-RUNX cistrome detected in the TAMR cells and that from this clinical case of tamoxifen resistant metastatic breast cancer ($p < 3e\text{-}7$, random permutation) (figure S4A). In addition, we compared the ERα-RUNX binding sites that were gained in the TAMR cells to published data mapping ER binding sites in ER positive (ER+) primary tumors of good and poor outcomes and in metastatic ER+ tumors (26). We found that the overlap between the TAMR ERα-RUNX2 binding sites and ER binding sites in the metastatic tumors was significantly higher than the overlap between the TAMR ERα-RUNX binding sites and ERα binding sites in good ($p=0.01$) and poor outcome primary tumors ($p=0.04$) (Figure S4B). Thus, the tamoxifen resistant unique ERα cistrome associated with RUNX motifs supports transcriptional changes that promote metastases and is associated with breast

tumors of poor outcome and metastases.

A RUNX2-ER interaction occurs with the acquisition of tamoxifen resistance and regulates genes that mediate EMT and metastases.

Because we detected the enrichment of RUNX motifs in the TAMR unique ERα cistrome we performed a protein immunoprecipitation study to examine whether there is a direct interaction between ERα and RUNX2. Indeed, when using a RUNX2 antibody for the immunoprecipitation and blotting for ER we detected a RUNX2-ERα interaction. This interaction was detectable in the parental MCF7 cells, mildly increased after 3 days of estrogen deprivation and was greatest with the acquisition of tamoxifen and estrogen deprivation resistance (Figure 3B). This was not merely due to the increase in ERα levels as the ratio between the amount of ER that was co-precipitated with RUNX2 comparing the TAMR to parental cells was higher than the ratio of the level of ERα in TAMR to parental cells in the input determined by the normalized relative density (The ratio is 5.8 in the immune-precipitated samples and 2 in the input samples). To validate this interaction and to identify other proteins that interact with RUNX2 in the context of chromatin binding, we used the RIME approach (rapid immunoprecipitation mass spectrometry of endogenous proteins) (50). As expected, RUNX2 and CBFβ, a known RUNX2 heterodimer partner, were among the most significantly enriched proteins after removing the non-specific proteins also found in the IgG control (Figure 3C, supplemental table 1 in S5). Other interacting proteins identified included ER and known ER co-regulators such as AIB1 (NCOA3), P300, CARM1(51-53) and the ER pioneer factors; FOXA1 and AP2γ (21, 54). In addition, GRHL2, which was recently found to bind to FOXA1 was also detected (55).

Furthermore, the RIME assay for RUNX2 failed to identify RUNX2 or significant interacting proteins in the parental cells, which is in keeping with the low RUNX2 levels observed in these cells.

Since we confirmed an interaction between ERα and RUNX2 in TAMR cells, we hypothesized that with the acquisition of tamoxifen resistance, RUNX2 is upregulated and interacts with ER and the ER-RUNX2 complex is in part responsible for the reprogramming of the ER cistrome and transcriptional activity in tamoxifen resistance. To test this hypothesis we generated MCF7 cells stably expressing doxycycline (DOX) inducible HA-tagged RUNX2 (Figure 3D). Globally, RNA-seq analysis revealed a large number of differentially expressed transcripts when comparing cells without and with DOX treatment (FDR<0.01 and log2FC > 1 or < -1). Ranked GSEA revealed that the genes upregulated by the induction of RUNX2 expression are most significantly enriched in the Hallmark gene set of epithelial to mesenchymal transition (NES=2.06, q-val=0.002), consistent with the metastatic phenotype we observed in the TAMR cells (Figure 3E). Furthermore, we saw that overexpression of RUNX2 led to a significant increase in invasion and a trend towards increased migration (Figure 3F). Next, we performed RUNX2 ChIP-seq in the MCF7 cells after RUNX2 induction using an HA-antibody (Figure 4A). We compared the RUNX2 cistrome to the TAMR unique ER-RUNX2 binding sites and found that 30% of the TAMR unique binding sites overlapped with RUNX2 binding sites. This overlap is statistically significant ($p < 3e-7$, random permutation). Moreover, we found that 44% of the 461 genes upregulated by the RUNX2- ERα transcriptional activity overlapped with the RUNX2 specific upregulated genes defined by integrating the RUNX2 ChIP-seq and the differential gene expression induced by RUNX2 expression in the DOX-

inducible cell line applying BETA (46). To test if the overlapping RUNX2-ERα up-regulated genes are co-regulated by both transcription factors, we looked at the expression levels of three of the overlapping genes that are known to have key functions in metastases (SOX9, EDN1, JAG1) after knock-down of ERα or RUNX2. By RT-PCR we confirmed that all three genes are regulated both by ERα and RUNX2. In contrast, MMP9 and MMP13(39) (40), which are published RUNX2 target genes and are within our list of RUNX2 upregulated genes that do not overlap with the RUNX2-ERα upregulated genes, are down-regulated after RUNX2 knock-down but not ER knock-down (Figure 4B). Lastly, we demonstrated that down regulation of RUNX2 abrogated proliferation (Figure 4C). Since there is a recently published small molecule (56) that inhibits the RUNX-CBFβ binding and we showed that RUNX1 is expressed the MCF7-TAMR cells, we tested the effect of RUNX1 down-regulation by siRUNX1 and showed that down-regulation of RUNX1 also had a growth inhibitory effect, suggesting that a non-selective RUNX inhibitor could be effective in this setting. Additionally, we confirmed that siRUNX1 was specific to RUNX1 (Figure 4E). Taken together, these results imply that with the acquisition of tamoxifen resistance ERα interacts with RUNX2 and these two transcription factors co-regulate a unique set of genes that mediate proliferation and a metastatic phenotype.

SOX9 is a ERα-RUNX2 target gene and is sufficient for tamoxifen resistance

SOX9 is a transcription factor that regulates stem and progenitor cells in adult tissues. It is upregulated in basal cell breast cancers and confers a tumor initiating stem/progenitor cell and metastatic phenotype in breast cancer cells (57, 58). Since we found that SOX9 is one of the target genes of the ERα-RUNX2 complex, we next turned our attention to the role of

SOX9 in tamoxifen resistance.

We first confirmed the upregulation of SOX9 in TAMR cells compared to parental cells and after the induction of RUNX2 expression at the protein level (Figure 5A). We also tested other ER+ cell line models of tamoxifen resistance and found that similar to the MCF7 cells, in TAMR cell lines in which ERα expression is maintained (MDAMB415 and 600MPE) RUNX2 and SOX9 are upregulated after the development of tamoxifen resistance. In contrast, in the T47D cell line, ERα is suppressed after the development of tamoxifen resistance and RUNX2 and SOX9 are both down regulated (Figure 5B). We next determined that TAMR cell growth is dependent on SOX9 by knock-down of SOX9 in TAMR cells (Figure 5C). Subsequently we showed that overexpression SOX9 in parental cells, which does not affect ER levels, leads to a growth advantage in full medium and estrogen-deprived conditions compared to cells expressing an empty vector (EV) (Figure 5D). Furthermore, dose response studies showed that overexpression of SOX9 leads to relative resistance to tamoxifen with a 2.3 -fold increase in the tamoxifen IC50 in the SOX9 overexpressing cells compared to the EV expressing cells with a p value close to significant (empty vector TAM IC50 = $6X10^{-10}$ M, SOX9 overexpression TAM IC50=$1.4X10^{-9}$M, p=0.054) (Figure 5D). We analyzed the transcriptional changes after SOX9 overexpression in PAR MCF7 cells and identified 277 genes that were upregaulated compared to PAR control (EV) (Supplemental Figure 4A) (FDR<0.01). The top ranked gene sets enriched in the PAR cells overexpressing SOX9 were genes down regulated during apoptosis in breast cancer cell lines (NES =1.94), and genes associated with acquired endocrine resistance in breast tumors expressing ER and ERBB2 (NES=1.55). In

addition, Kegg Pathways analysis with GAGE (general applicable gene set enrichment for pathway analysis)(59) revealed that the top ranked pathway enriched in the SOX9 overexpressing PAR cells was the JAK-STAT pathway (Figure S6). In summary, we show that SOX9 is upregulated in TAMR by the ERα-RUNX2 complex, TAMR cell proliferation is dependent on SOX9 and SOX9 upregulation is sufficient to cause resistance to estrogen deprivation and decreased sensitivity to tamoxifen treatment.

In order to validate our pre-clinical findings we examined a unique set of clinical samples from patients that developed local recurrences during or after adjuvant treatment with tamoxifen. This sample set included 42 pairs of primary ERα+ breast cancers and the matched recurrent lesions. Comprehensive clinical data was available for 39 of the patients (table in S7). The majority of the recurrent lesions remained ERα+ (67%) and the majority of the patients developed the recurrence while on tamoxifen treatment (80%). For the 20% of patients that developed the recurrence after tamoxifen treatment, the average time between completion of tamoxifen and disease recurrence was 35 months. The average duration of tamoxifen treatment was 34 months (ranging between 6-83 months). We were able to examine the expression of SOX9 by IHC in the 42 tumor pairs (Figure 6A). SOX9 showed strong nuclear staining of breast cancer epithelial cells in the ERα+ primary and recurrent tumors. Taking into account the staining intensity and percentage of positive epithelial cells, there was variation in the level of SOX9 expression between tumors. Nonetheless, there was a significant increase in the expression of SOX9 after the development of tamoxifen resistance (expression in treatment naïve tumors, mean±se: 0.268±0.028, expression in tamoxifen resistant recurrent tumors, mean±se: 0.342±0.031, p value of 0.033 [t=-2.207]) (Figure 6B). Furthermore, we performed a subset analyses and

looked at the pairs in which the recurrent disease remained ERα+ (N=28) and found that in this sub-group SOX9 expression was significantly up-regulated after the development of tamoxifen resistance (p value of 0.004), whereas in the recurrent tumors with loss of ERα expression (N=14), SOX9 expression was not up-regulated (Figure 6C). These clinical results show a correlation between tamoxifen resistance and SOX9 expression in ERα+ tumors and support the notion that SOX9 is an ERα transcriptional target gene in the setting of tamoxifen resistance.

Discussion

Our results demonstrate that differential recruitment of ERα resulting in an aberrant transcriptional network is a fundamental mechanism of resistance to tamoxifen treatment. Furthermore, as endocrine resistance is commonly associated with the development of distant recurrences (60), our results indicate that the aberrant ERα transcriptional axis also has a key role in driving an EMT and pro-metastatic program. More specifically, we found that after long-term exposure to tamoxifen and estrogen deprivation, ERα-chromatin binding is enriched in RUNX motifs mediated by a direct interaction between ERα and RUNX2. Our RIME data suggest that in addition to RUNX2 and known ERα co-regulators the SWI/SNF chromatin-remodeling complex, which we have previously shown to be important for ERα activity and is emerging as important cancer driver in (61), may be involved in RUNX2-ERα chromatin binding.

In this study we found that other transcription factor binding motifs of potential interest were enriched in the TAMR ERα DNA-binding sites. These include the FOXM1 motif, which was enriched in the TAMR ERα unique binding sites consistent with published

studies showing that FOXM1 interacts with ER and can also promote tamoxifen resistance(62, 63). Furthermore, we demonstrated the loss of transcription factors essential for the ER transcription axis in luminal type breast cancers, such as GATA3. Thus, multiple alterations in the ERα cistrome orchestrate resistance to tamoxifen and disease progression underscoring the therapeutic implications of better targeting of ERα. Our study also depicts the complexity of endocrine resistance and the fact that there are multiple mechanisms of resistance. As an example, in the TAMR cells derived from T47D cells, ERα is down regulated and RUNX2 and SOX9 are not upregulated, implying that the mechanism of resistance is different in this model and likely due to the loss of ERα expression. Likewise, in the clinic, there are multiple mechanisms of endocrine resistance and studies to identify biomarkers of the specific drivers of resistance will be important.

A RUNX1-ERα interaction has been described in the literature in a triple negative breast cancer cell line engineered to ectopically express ERα (64). RUNX1 has been shown to function as both a tumor suppressor and an oncogene in breast cancer and recently point mutations in RUNX1 and CBFB have been detected in ER+ breast cancers by next generation sequencing studies(65-67). Similar to the cell and context dependent activity of RUNX1, the RUNX2-ERα complex appears to be distinct in a cell and context dependent manner. A RUNX2–ERα interaction has been described previously in COS7 cells transfected with an expression vector encoding ERα (68). While in the COS7 cell type, ERα repressed RUNX2 transcriptional activity in a ligand dependent manner, here we show an interaction between endogenous RUNX2 and ERα proteins in an ERα+ endocrine resistant breast cancer cell line, and these two transcription factors in concert mediate transcription of genes that promote EMT and metastases. In addition, we also show that

down regulation of RUNX1 has a moderate inhibitory effect on the TAMR cells, suggesting that RUNX1 likely also has a role in endocrine resistance in this model and future studies are needed for further evaluation.

A number have studies have shown that RUNX2 can increase invasiveness, but these studies were in triple negative breast cancer models (69-71). Likewise, in clinical samples of primary breast cancers RUNX2 is expressed and found to be prognostic mainly in triple negative breast cancers (44, 72). However, in a more recent study, RUNX2 level in metastatic ER+ breast cancers was associated with poor outcomes (73). In line with these clinical results, our RUNX2-ER gene signature correlated with gene signatures of metastases and poor outcomes. Our findings have important clinical implications since a small molecule inhibitor of the RUNX-CBFβ complex was recently identified (56). One of the genes regulated by the ERα –RUNX2 complex of particular significance is SOX9. SOX9 is a master regulator of embryonic stem cells and more recently was shown to have key roles in the development of metastases in triple negative and HER2 positive breast cancer. Here, we show that SOX9 is expressed in ERα positive breast cancers, increases with progression and plays an important role in endocrine resistance.

In this study we were also able to extend our findings to endocrine resistant clinical samples. We found a significant overlap between the ERα–RUNX2 DNA binding sites unique to the TAMR cell line model and the ERα binding sites in ER+ breast cancer clinical samples. Furthermore, we found that the overlap between the TAMR unique ERα–RUNX2 DNA binding sites and metastatic ERα binding sites was higher compared to the overlap with the ER binding sites in primary treatment naïve tumor samples. We also

detected an association between the ERα–RUNX2 co-regulated genes that are upregulated in TAMR and gene sets of poor prognosis in breast cancer. Lastly, we show that upregulation of SOX9 expression is linked to tamoxifen resistance in clinical samples. Taken together, these results support the notion that in a subset of breast cancers with acquired tamoxifen resistance, ERα activity is maintained but reprogrammed leading to an aberrant ERα cistrome. In part, the aberrant chromatin binding is due to the non-canonical activation of ERα in complex with RUNX2. This complex results in the activation of a set of genes that include SOX9 and promote endocrine resistance and metastases. These results highlight the importance of developing improved ERα antagonists as well as agents that target other key proteins in complex with ER and key ER transcriptional targets, such as RUNX2 and SOX9, respectively.

Methods

Cell culture and proliferation assays: The TAMR and LTED cells were derived from MCF7 cells (MCF7L originally from Dr. Marc Lippman's lab) using methods previously reported(29). Other TAMR cells were derived from 600MPE (originally from Dr. Joe Gray's lab) and MDAMB415 (purchased from ATCC). All the cells were authenticated and regularly tested for mycoplasma contamination. The MCF7 cells were maintained in RPMI/1640 or supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S). The endocrine resistant cells were kept in phenol-red free medium supplemented with 10% heat-inactivated charcoal-stripped (CS)-FBS and 1% PS. The TAMR cells were grown with the addition of 100 nM 4-OH-tamoxifen (H7904, Sigma). All cells were incubated at 37°C in 5% CO_2.

For proliferation assays the breast cancer cells were plated in 24 well plates ($2.5x10^4$/well).

At indicated time points, the cells were trypsinized and collected. The number of viable cells was determined by Trypan blue exclusion staining and directly assessed with a hemacytometer using independent triplicates.

Cell migration-invasion assay: The Radius assay and Boyden Chamber assays were performed following the manufacturer's instructions (Cell Biolabs).

Western-blotting and protein immunoprecipitation: For Western blot analyis, cells were lysed in 50mM Tris-HCl pH 7.5, 150mM NaCl, 1mM EDTA, 1mM EGTA, 0.5% NP-40, 1% Triton X-100 supplemented with protease inhibitors and subjected to SDS-PAGE. Antibodies used were: ERα (sc-543, Santa Cruz), GATA3 (607102, Biolegend) RUNX2 (D130-3, MBL) SOX9 (Ab5535, EMD Millipore) Beta-Actin (Sigma), GAPDH (Santa Cruz), HA (Ab9110, Abcam). Protein immunoprecipitation was carried out as previously described (74).

Chromatin Immunoprecipitation (ChIP)-Sequencing: ChIP experiments were conducted as described previously(75). ChIP-seq reads were aligned to the hg19 genome assembly using Bowtie(76) and ChIP-seq peaks were called using MACS 2.0(77, 78). Regions of enrichment comparing ChIP and input control signal exceeding $q < 0.01$ were called as peaks. Read densities were calculated for each peak in reads per million per nucleotide (RPM), which were used for comparison of cistromes across samples. We used SeqPos available at www.cistrome.org for motif analysis. Correlation between RNA-seq differential gene expression and transcription factor binding based on ChIP-seq was performed with the Binding and Expression Target Analysis (BETA) basic algorithm (45).

All the ChIP-seq data have been deposited in the Gene Expression Omnibus database (GSE86538).

RNA sequencing: Total RNA was isolated using an RNeasy Mini Kit (Qiagen). RNA-seq libraries were made using the TruSeq RNA Sample Preparation Kit (Illumina) adapted for use on the Sciclone (Perkin-Elmer) liquid handler. Samples were sequenced on an Illumina Nextseq500 and the 75-bp-long reads were aligned to hg19 with STAR aligner. Cufflinks was used to generate expression value (RPKM) for each gene (79) and the differential expression analysis was performed using the DEseq method. For gene ontology we used the DAVID web site. Gene set enrichment analysis was performed using the online tool from the Broad Institute. All the RNA-seq data have been deposited in the Gene Expression Omnibus database (GSE86538).

Small interfering RNA (siRNA) transfection: siRNAs were chemically synthesized by Dharmacon Inc. A non-specific siRNA duplex was used as negative control. For transfection, cells were seeded in complete medium without antibiotics the day before the experiment. After 24 h, cells were transfected with 50 nM of siRNA, using Lipofectanmine 2000 transfection reagent (Invitrogen), according to the manufacturer's protocol.

Lentiviral Infection:

For the Dox-inducible RUNX2 MCF7 cells, RUNX2-HA tag cDNA (GeneCopoeia) was transferred to the pInducer 20 destination vector (5) using the Gateway system (Invitrogen). Lentivirus was produced in 293T cells to infect cells in media containing polybrene (8μg/mL). Cells were selected after the infection with G418.

RIME: The RIME experiment using the TAMR and parental MCF7 cells was conducted

as published previously(80). The RUNX2 antibody (D130-3, MBL) was used for the experiments and as controls an IgG antibody (SC-2025, Santa Cruz) was used. Mass spectrometry (MS) was performed using an LTQ Velos-Orbitrap MS (Thermo Scientific) coupled to an Ultimate RSLCnano-LC system (Dionex). Raw MS data files were processed using Proteome Discoverer v.1.3 (Thermo Scientific). Processed files were searched against the SwissProt human database using the Mascot search engine version 2.3.0.

Patient Tissue Analysis: A total of 42 matched pairs of ER^+ breast cancer tumors from patients treated with tamoxifen, including treatment naïve primary tumors and local recurrent tumors that developed tamoxifen treatment. The tissue samples were evaluated on two TMA's obtained from consented patients at the Royal Marsden Hospital, London UK. More details of these tissue samples were previously published. Immunohistochemical staining of SOX9 was performed on 4-µm sections of the TMA; using the Bond Refine Detection System following the manufacturer's protocols on the Leica Bond III automated immunostainer. The sections were automatically deparaffinized, antigen retrieval was done with EDTA buffer (pH 9.0) and processed for 20 min. The slides were incubated with the antibody against SOX9 (8G5, Mouse monoclonal, D130-3, MBL) at a dilution of 1:500. The sections were then treated according to the streptavidin–biotin–peroxidase complex method (Bond Polymer Refine Detection, Leica Microsystems) with diaminobenzidine (DAB) as a chromogen and counterstained with hematoxylin. Incubation with the biotinylated universal secondary antibody was then performed. Visualization was performed with 3,3'-diaminobenzidine (DAB) as the chromogen substrate. Testis tissue was used as positive controls for SOX9. Omission of the primary antibody was utilized use as a negative control. Once stained, the TMA slides were scanned on the Olympus BX-51

W1 microscope using Vectra 2.0.8 software (Perkin Elmer). Cores that were disrupted or contained insufficient tissue were eliminated from the analysis. Following the standard bright-field TMA scanning protocol, a chromogenic spectral library was composed using the spectra of both the counterstain (haematoxylin) and the immunostain (DAB). Tissue segmentation algorithms were subsequently constructed using inForm V2.0.2 (Perkin Elmer). Initially a training set comprising three classes of tissue was created (i.e. tumor, stroma and other). Representative regions of interest (ROI) for each of these classes were marked on 15-20 images from the TMAs. The software was trained on these areas and tested to determine how accurately it could differentiate between the tissue classes. Cell segmentation algorithms were then constructed for the cell nucleus. Cell segmentation algorithms identified nuclei as pixels above the minimum signal value. The spectral library and algorithms were then run on all samples. Poorly segmented cores were manually corrected via touchscreen editing following pathology review. Data generated from every cell per core, where expression of SOX9 was observed, was utilized to calculate a core mean intensity value for subsequent statistical analysis.

Statistical analysis Statistical analyses were performed using unpaired Student's t tests, and P values less than 0.05 were considered statistically significant. Error bars in figures represent SEM. For the patient tissue sample analysis a two sided (paired student t-test) was performed.

Study Approval: The pleural effusion was collected with patient consent and DFCI/HCC

IRB approval (protocol 12-259). The tissue samples for the generation of the TMAs were obtained from consented patients at the Royal Marsden Hospital, London UK.

Acknowledgements: This study was supported in part by grants from Susan G. Komen for the Cure (to MB), Breast Cancer SPORE Career Development Award (to RJ), Royal Marsden NIHR Biomedical Research Centre (to MD).

References:

1. Lerner HJ, Band PR, Israel L, & Leung BS (1976) Phase II study of tamoxifen: report of 74 patients with stage IV breast cancer. *Cancer treatment reports* 60(10):1431-1435.
2. Morgan LR, Jr.*, et al.* (1976) Therapeutic use of tamoxifen in advanced breast cancer: correlation with biochemical parameters. *Cancer treatment reports* 60(10):1437-1443.
3. Ingle JN*, et al.* (1981) Randomized clinical trial of diethylstilbestrol versus tamoxifen in postmenopausal women with advanced breast cancer. *The New England journal of medicine* 304(1):16-21.
4. Ingle JN*, et al.* (1986) Randomized trial of bilateral oophorectomy versus tamoxifen in premenopausal women with metastatic breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 4(2):178-185.
5. Early Breast Cancer Trialists' Collaborative (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 365(9472):1687-1717.
6. Lindstrom LS*, et al.* (2012) Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 30(21):2601-2608.
7. Curigliano G*, et al.* (2011) Should liver metastases of breast cancer be biopsied to improve treatment choice? *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 22(10):2227-2233.
8. Gong Y, Han EY, Guo M, Pusztai L, & Sneige N (2011) Stability of estrogen receptor status in breast carcinoma: a comparison between primary and metastatic tumors with regard to disease course and intervening systemic therapy. *Cancer* 117(4):705-713.
9. Torres-Arzayus MI, Zhao J, Bronson R, & Brown M (2010) Estrogen-dependent and estrogen-independent mechanisms contribute to AIB1-mediated tumor formation. *Cancer research* 70(10):4102-4111.
10. Massarweh S*, et al.* (2008) Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. *Cancer research* 68(3):826-833.
11. Miller TW*, et al.* (2010) Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *The Journal of clinical investigation* 120(7):2406-2413.
12. Jeselsohn R, Buchwalter G, De Angelis C, Brown M, & Schiff R (2015) ESR1 mutations-a mechanism for acquired endocrine resistance in breast cancer. *Nature reviews. Clinical oncology* 12(10):573-583.
13. Li S*, et al.* (2013) Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell reports* 4(6):1116-1130.
14. Fu X*, et al.* (2016) FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer.

Proceedings of the National Academy of Sciences of the United States of America 113(43):E6600-E6609.
15. Perey L, *et al.* (2007) Clinical benefit of fulvestrant in postmenopausal women with advanced breast cancer and primary or acquired resistance to aromatase inhibitors: final results of phase II Swiss Group for Clinical Cancer Research Trial (SAKK 21/00). *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 18(1):64-69.
16. Ingle JN, *et al.* (2006) Fulvestrant in women with advanced breast cancer after progression on prior aromatase inhibitor therapy: North Central Cancer Treatment Group Trial N0032. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 24(7):1052-1056.
17. Di Leo A, *et al.* (2010) Results of the CONFIRM phase III trial comparing fulvestrant 250 mg with fulvestrant 500 mg in postmenopausal women with estrogen receptor-positive advanced breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 28(30):4594-4600.
18. Mehta RS, *et al.* (2012) Combination anastrozole and fulvestrant in metastatic breast cancer. *The New England journal of medicine* 367(5):435-444.
19. Davies C, *et al.* (2013) Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* 381(9869):805-816.
20. Goss PE, *et al.* (2003) A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *The New England journal of medicine* 349(19):1793-1802.
21. Carroll JS, *et al.* (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* 122(1):33-43.
22. Carroll JS, *et al.* (2006) Genome-wide analysis of estrogen receptor binding sites. *Nature genetics* 38(11):1289-1297.
23. Eeckhoute J, Carroll JS, Geistlinger TR, Torres-Arzayus MI, & Brown M (2006) A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. *Genes & development* 20(18):2513-2526.
24. Krum SA, *et al.* (2008) Unique ERalpha cistromes control cell type-specific gene regulation. *Mol Endocrinol* 22(11):2393-2406.
25. Lupien M, *et al.* (2010) Growth factor stimulation induces a distinct ER(alpha) cistrome underlying breast cancer endocrine resistance. *Genes & development* 24(19):2219-2227.
26. Ross-Innes CS, *et al.* (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481(7381):389-393.
27. Francis PA, Regan MM, & Fleming GF (2015) Adjuvant ovarian suppression in premenopausal breast cancer. *The New England journal of medicine* 372(17):1673.
28. Burstein HJ, *et al.* (2016) Adjuvant Endocrine Therapy for Women With Hormone Receptor-Positive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update on Ovarian Suppression. *Journal of*

clinical oncology : official journal of the American Society of Clinical Oncology 34(14):1689-1701.
29. Morrison G*, et al.* (2014) Therapeutic potential of the dual EGFR/HER2 inhibitor AZD8931 in circumventing endocrine resistance. *Breast cancer research and treatment* 144(2):263-272.
30. Gottardis MM, Jiang SY, Jeng MH, & Jordan VC (1989) Inhibition of tamoxifen-stimulated growth of an MCF-7 tumor variant in athymic mice by novel steroidal antiestrogens. *Cancer research* 49(15):4090-4093.
31. Agrawal A, Robertson JF, & Cheung KL (2011) Clinical relevance of "withdrawal therapy" as a form of hormonal manipulation for breast cancer. *World journal of surgical oncology* 9:101.
32. Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome biology* 11(10):R106.
33. Huang da W*, et al.* (2008) DAVID gene ID conversion tool. *Bioinformation* 2(10):428-430.
34. Asselin-Labat ML*, et al.* (2007) Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nature cell biology* 9(2):201-209.
35. Eeckhoute J*, et al.* (2007) Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer research* 67(13):6477-6483.
36. Theodorou V, Stark R, Menon S, & Carroll JS (2013) GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome research* 23(1):12-22.
37. Kouros-Mehr H*, et al.* (2008) GATA-3 links tumor differentiation and dissemination in a luminal breast cancer model. *Cancer cell* 13(2):141-152.
38. Blyth K, Cameron ER, & Neil JC (2005) The RUNX genes: gain or loss of function in cancer. *Nature reviews. Cancer* 5(5):376-387.
39. Pratap J*, et al.* (2005) The Runx2 osteogenic transcription factor regulates matrix metalloproteinase 9 in bone metastatic cancer cells and controls cell invasion. *Molecular and cellular biology* 25(19):8581-8591.
40. Selvamurugan N, Kwok S, & Partridge NC (2004) Smad3 interacts with JunB and Cbfa1/Runx2 for transforming growth factor-beta1-stimulated collagenase-3 expression in human breast cancer cells. *The Journal of biological chemistry* 279(26):27764-27773.
41. Li XQ*, et al.* (2015) ITGBL1 Is a Runx2 Transcriptional Target and Promotes Breast Cancer Bone Metastasis by Activating the TGFbeta Signaling Pathway. *Cancer research* 75(16):3302-3313.
42. Sharp JA, Waltham M, Williams ED, Henderson MA, & Thompson EW (2004) Transfection of MDA-MB-231 human breast carcinoma cells with bone sialoprotein (BSP) stimulates migration and invasion in vitro and growth of primary and secondary tumors in nude mice. *Clinical & experimental metastasis* 21(1):19-29.
43. El-Gendi SM & Mostafa MF (2015) Runx2 Expression as a Potential Prognostic Marker in Invasive Ductal Breast Carcinoma. *Pathology oncology research : POR*.

44. Yang Z, Zhang B, Liu B, Xie Y, & Cao X (2015) Combined Runx2 and Snail overexpression is associated with a poor prognosis in breast cancer. *Tumour Biol* 36(6):4565-4573.
45. Wang S, *et al.* (2013) Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nature protocols* 8(12):2502-2515.
46. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102(43):15545-15550.
47. Pande S, *et al.* (2013) Oncogenic cooperation between PI3K/Akt signaling and transcription factor Runx2 promotes the invasive properties of metastatic breast cancer cells. *Journal of cellular physiology* 228(8):1784-1792.
48. Cohen-Solal KA, Boregowda RK, & Lasfar A (2015) RUNX2 and the PI3K/AKT axis reciprocal activation as a driving force for tumor progression. *Molecular cancer* 14:137.
49. Rhodes DR, *et al.* (2007) Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 9(2):166-180.
50. Mohammed H, *et al.* (2016) Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc* 11(2):316-326.
51. Anzick SL, *et al.* (1997) AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* 277(5328):965-968.
52. Hanstein B, *et al.* (1996) p300 is a component of an estrogen receptor coactivator complex. *Proceedings of the National Academy of Sciences of the United States of America* 93(21):11540-11545.
53. Shang Y, Hu X, DiRenzo J, Lazar MA, & Brown M (2000) Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* 103(6):843-852.
54. Tan SK, *et al.* (2011) AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *The EMBO journal* 30(13):2569-2581.
55. Jozwik KM, Chernukhin I, Serandour AA, Nagarajan S, & Carroll JS (2016) FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. *Cell reports* 17(10):2715-2723.
56. Illendula A, *et al.* (2016) Small Molecule Inhibitor of CBFbeta-RUNX Binding for RUNX Transcription Factor Driven Cancers. *EBioMedicine* 8:117-131.
57. Guo W, *et al.* (2012) Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* 148(5):1015-1028.
58. Malladi S, *et al.* (2016) Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* 165(1):45-60.
59. Luo W, Friedman MS, Shedden K, Hankenson KD, & Woolf PJ (2009) GAGE: generally applicable gene set enrichment for pathway analysis. *BMC bioinformatics* 10:161.
60. Metzger-Filho O, *et al.* (2013) Patterns of Recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: results from international breast cancer study group trials VIII and IX. *Journal of clinical*

oncology : official journal of the American Society of Clinical Oncology 31(25):3083-3090.
61. Hohmann AF & Vakoc CR (2014) A rationale to target the SWI/SNF complex for cancer therapy. *Trends in genetics : TIG* 30(8):356-363.
62. Bergamaschi A, *et al.* (2014) The forkhead transcription factor FOXM1 promotes endocrine resistance and invasiveness in estrogen receptor-positive breast cancer by expansion of stem-like cancer cells. *Breast cancer research : BCR* 16(5):436.
63. Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, & Balasubramanian S (2013) Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. *Genome biology* 14(1):R6.
64. Stender JD, *et al.* (2010) Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. *Molecular and cellular biology* 30(16):3943-3955.
65. Browne G, *et al.* (2015) Runx1 is associated with breast cancer progression in MMTV-PyMT transgenic mice and its depletion in vitro inhibits migration and invasion. *J Cell Physiol* 230(10):2522-2532.
66. Ellis MJ, *et al.* (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 486(7403):353-360.
67. Janes KA (2011) RUNX1 and its understudied role in breast cancer. *Cell Cycle* 10(20):3461-3465.
68. Khalid O, *et al.* (2008) Modulation of Runx2 activity by estrogen receptor-alpha: implications for osteoporosis and breast cancer. *Endocrinology* 149(12):5984-5995.
69. Owens TW, *et al.* (2014) Runx2 is a novel regulator of mammary epithelial cell fate in development and breast cancer. *Cancer research* 74(18):5277-5286.
70. Taipaleenmaki H, *et al.* (2015) Targeting of Runx2 by miR-135 and miR-203 Impairs Progression of Breast Cancer and Metastatic Bone Disease. *Cancer research* 75(7):1433-1444.
71. Tandon M, Chen Z, & Pratap J (2014) Runx2 activates PI3K/Akt signaling via mTORC2 regulation in invasive breast cancer cells. *Breast cancer research : BCR* 16(1):R16.
72. McDonald L, *et al.* (2014) RUNX2 correlates with subtype-specific breast cancer in a human tissue microarray, and ectopic expression of Runx2 perturbs differentiation in the mouse mammary gland. *Dis Model Mech* 7(5):525-534.
73. El-Gendi SM & Mostafa MF (2016) Runx2 Expression as a Potential Prognostic Marker in Invasive Ductal Breast Carcinoma. *Pathol Oncol Res* 22(3):461-470.
74. Ni M, *et al.* (2011) Targeting androgen receptor in estrogen receptor-negative breast cancer. *Cancer cell* 20(1):119-131.
75. Wang Q, *et al.* (2009) Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell* 138(2):245-256.
76. Langmead B, Trapnell C, Pop M, & Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 10(3):R25.

77. Zhang Y, *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology* 9(9):R137.
78. Feng J, Liu T, Qin B, Zhang Y, & Liu XS (2012) Identifying ChIP-seq enrichment using MACS. *Nature protocols* 7(9):1728-1740.
79. Trapnell C, *et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* 7(3):562-578.
80. Mohammed H, *et al.* (2013) Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell reports* 3(2):342-349.
81. Luo W & Brouwer C (2013) Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29(14):1830-1831.

Figure Legends:

Figure 1. TAMR cell proliferation is ERα dependent and increased migratory capacity with the acquisition of tamoxifen resistance: A. Cell proliferation analyzed by cell counting on days 1, 3 and 5 after down regulation of ERα with siRNA in TAMR (tamoxifen resistant) and PAR (parental) cells. As controls TAMR and PAR cells were transfected with a siCON. ERα down-regulation was confirmed by western blot. Western blot also confirms ERα expression in TAMR. B-actin was used as a loading control. B. Western blot for ERα and cell proliferation curves after treatment of TAMR cells with fulvestrant (FULV) 10^{-7}M and 10^{-6}M or vehicle (Veh) control. C. Brightfield microscopy picture of 2D culture of PAR and TAMR cells. Magnification 20X. Radius assay and Boyden chamber assay testing PAR cell migration in FM (full medium) and TAMR in hormone depleted white medium. * p-value= 0.02 ** p-value=0.003, error bars denote SEM figure represents the results of three replicates . D. Volcano plot of the RNAseq differential gene expression between PAR and TAMR cells using DE-seq2, log 2 fold change >1 or <-1, p-val <0.05. The red dots are the genes that are significantly up-regulated in TAMR cells and blue dots are the genes down regulated. Functional annotation of genes up-regulated in TAMR cells. Top overrepresented gene categories from gene ontology (GO) using DAVID are shown here. P-values represented by red line and % of genes by blue bars.

Figure 2. Redistribution of ERα chromatin binding with the acquisition of tamoxifen resistance: A. Venn diagram showing the overlap of ERα binding in TAMR (tamoxifen resistant) and LTED (long term estrogen deprived) cells without E2 stimulation and PAR

(parental cells) with E2 stimulation. B. Top motifs enriched and lost in the TAMR cells compared to parental cells. C. Heatmap showing clustered ERα binding signals enriched in RUNX and GATA in TAMR and parental cells. The windows represent ±1-kb regions from the center of the ERα binding events. The color scale shows relative enrichment based on raw signal. D. Relative mRNA levels of the RUNX transcription factors determined by RT-PCR calculated by ΔΔCT in parental, LTED and TAMR cells. E. Immunoblots for RUNX2 and GATA3 using cell lysates of PAR, LTED and TAMR cells.

Figure 3. The ERα binding sites unique to TAMR cells and enriched in RUNX2 motifs are associated with poor outcomes in breast cancer and endogenous RUNX2 and ERα interact to mediate transcriptional changes that promote epithelial to mesenchymal transition: A. Ranked gene set enrichment analysis (GSEA) of the genes determined to be up-regulated by ERα–RUNX2 in TAMR cells based on BETA analysis. NES= normalized enrichment score, $q<0.25$. Oncomine Concepts Map analysis (Compendia Bioscience) was used to compare the ER-RUNX2 gene signature in TAMR cells against published gene signatures from primary breast cancers. This analysis showed significant correlations between the ER-RUNX2 induced genes and gene expression signatures of poor outcome. The association between the molecular concepts of different gene signatures is represented as a graph using Cytoscape (www.cytoscape.org). A node represents a gene set and significantly associated ($q<0.2$) sets were connected by an edge. The node of the ER-RUNX2 gene set is in red and the nodes of poor outcomes are in blue. The size of the nodes is proportional to the number of overlapping genes between the corresponding gene

set and the ERα-RUNX2 gene set and the thickness of the edges that connect between the nodes is proportional to the rank of the association significance. B. Immunoprecipitation of endogenous RUNX2 using nuclear extracts and immunoblotting for endogenous ERα with IgG and input controls. PAR=parental, TAMR=tamoxifen resistant, FM=full medium, WM= white medium. Normalized protein quantification was done using ImageJ (imageJ.nih.gov). C. RUNX2 RIME results in TAMR cells depicted in a word cloud. D. Western blot for RUNX2 and HA confirming stable DOX-inducible expression of HA-tagged RUNX2 in MCF7 parental cells after treatment with DOX. EV=empty vector. E. Heat map of a K-means 2 clustering of the top 1,000 genes differentially expressed between the DOX-inducible RUNX2 MCF7 cells with DOX treatment (pInd-RUNX2_DOX) or no DOX treatment (pInd-RUNX2). R1-R3 denotes replicates 1-3. On the left is an enrichment plot from GSEA showing the Hallmark epithelial-mesenchymal transition gene set, which is the gene set most significantly enriched after the induction of RUNX2 expression. F. Results of a Boyden chamber assay comparing migration and invasion in MCF7 cells with and without DOX-induction of RUNX2. * denotes <0.05. Error bars represent the SEM of three replicates.

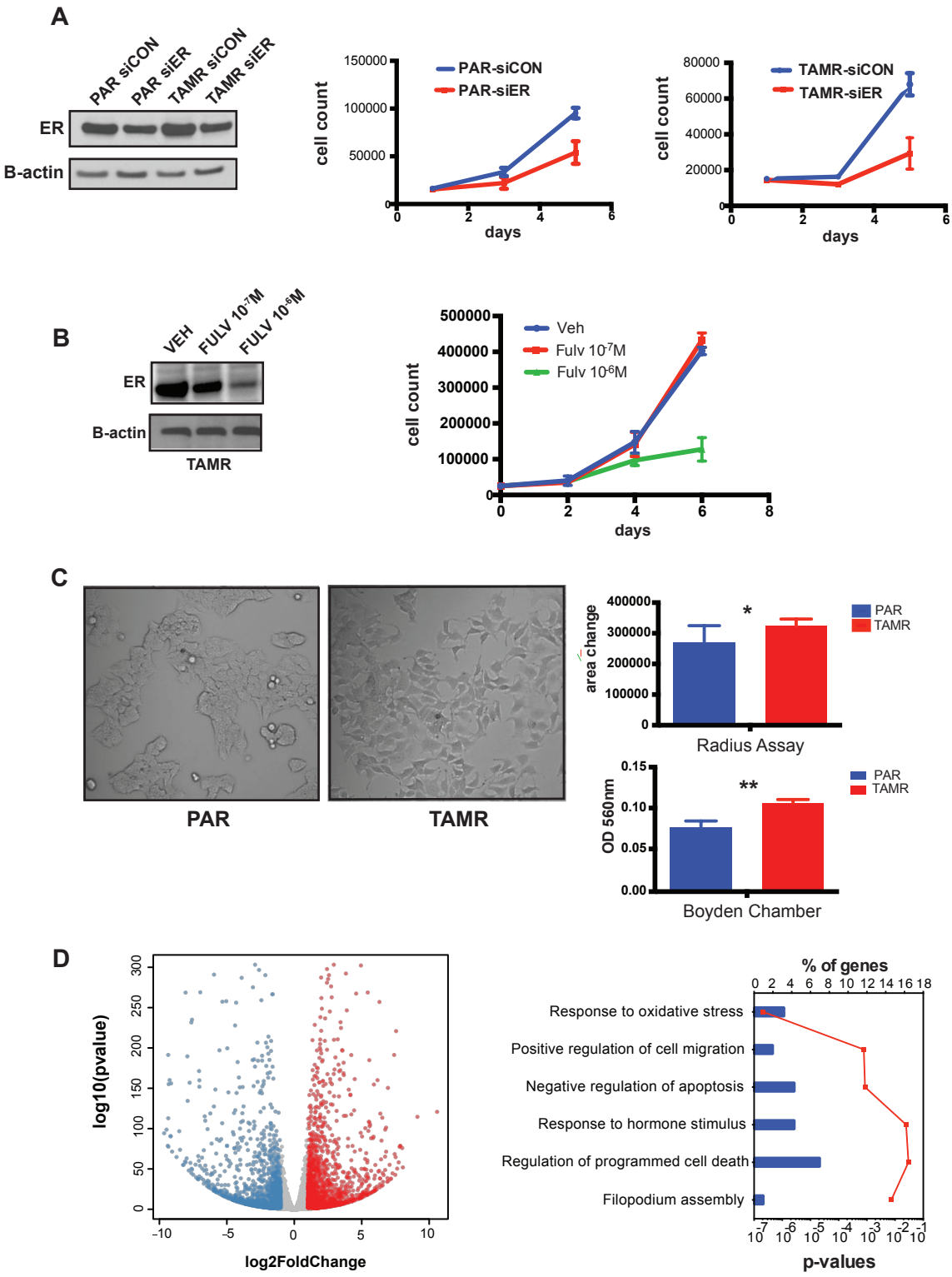
Fig 4. The RUNX2 cistrome highly overlaps with the ER-RUNX2 cistrome: A. Venn diagram (left) depicts the overlap between the ERα binding sites with a RUNX2 motif in TAMR cells and RUNX2 binding sites in MCF7 cells with DOX-inducible expression of RUNX2. Venn diagram (right) showing the overlap between the genes up-regulated by the ERα-RUNX2 complex in TAMR cells and genes upregulated by RUNX2 in MCF7. Both gene sets were determined by integrating ChIPseq and RNAseq applying

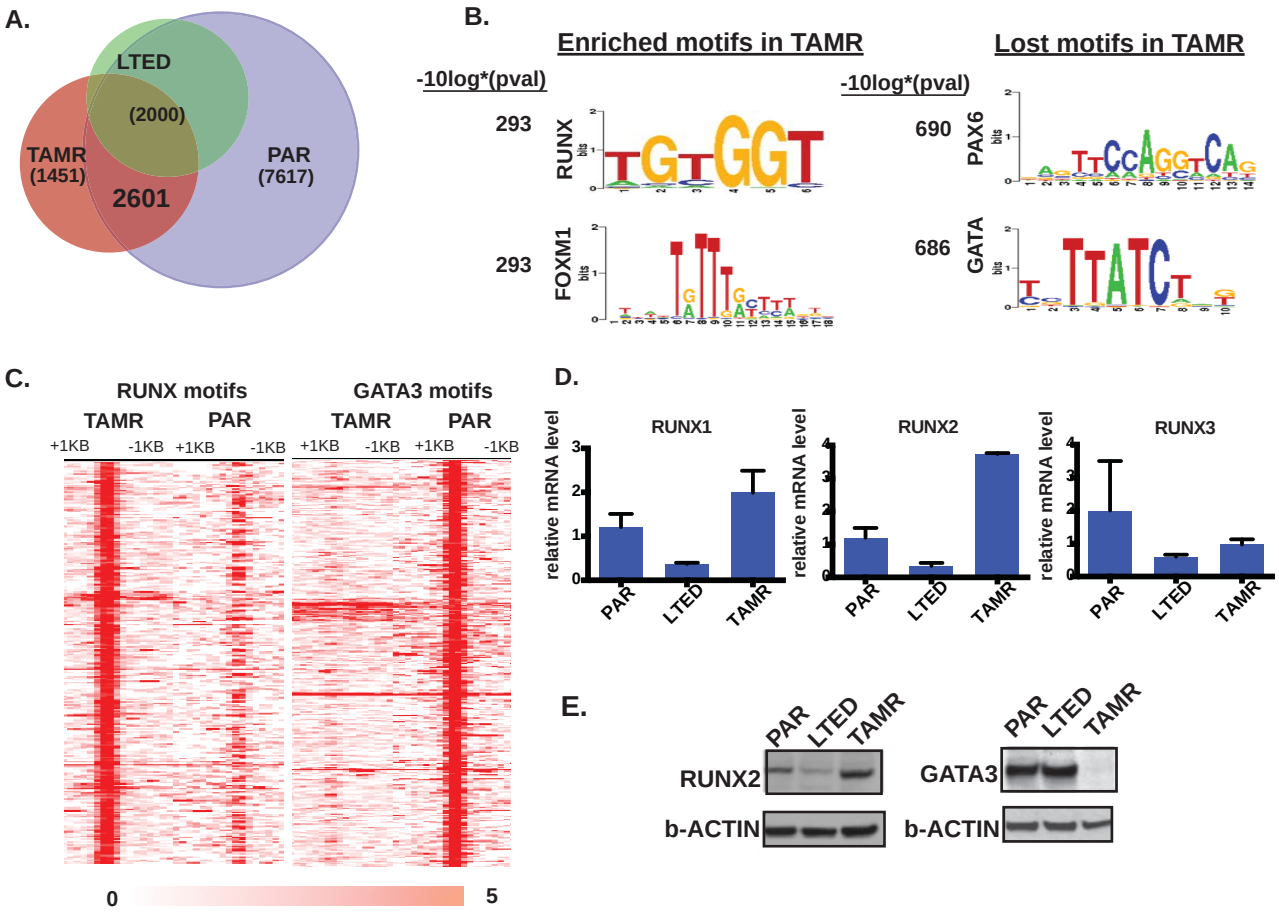
BETA[45]. B. Relative mRNA levels of genes regulated by the ERα- RUNX2 complex after transfection of TAMR cells with either siERα, siRUNX2 or siControl (siCON) and extraction of RNA on day 3 after transfection. Shown here are the relative mRNA levels after siERα knock down or siRUNX2 knock down compared to siControl. C. Cell proliferation analyzed by cell counting on days 1, 3 and 5 after down regulation of RUNX2 with siRNA. D. Cell proliferation analyzed by cell counting on days 1, 3 and 5 after down regulation RUNX1 with siRUNX1 in TAMR cells. As controls TAMR cells were transfected with a siCON. E. Relative mRNA levels of RUNX2 (left) and RUNX1 (right) in control MCF7 cells and after down regulation of RUNX1 and RUNX2.

Figure 5. SOX9 is upregulated in TAMR and mediates resistance to estrogen deprivation and tamoxifen. A. Western blot for SOX9 in parental (PAR) and tamoxifen resistant (TAMR) MCF7 cells (left) and in MCF7 cells infected with an empty vector control and a DOX-inducible RUNX2 expression vector with and without doxycycline treatment. B. Heat maps of mRNA levels of *ESR1*, *RUNX2* and *SOX9* in parental (PAR) and tamoxifen resistant (TAMR) cell lines derived from 600MPE, MDAMB415 and T47D cell lines. C. Cell proliferation analyzed by cell counting on days 1, 3 and 5 in TAMR cells after SOX9 silencing with 2 different siRNAs or with a siRNA control (si-CON). Western blot for SOX9 in TAMR cells showing successful knock-down of SOX9. D. Western blot of SOX9 in parental MCF7 cells after stable expression of SOX9 and an empty vector (EV) (top). Western blot shows stable ERα expression after expression of SOX9 in MCF7 parental cells. Cell proliferation analyzed by

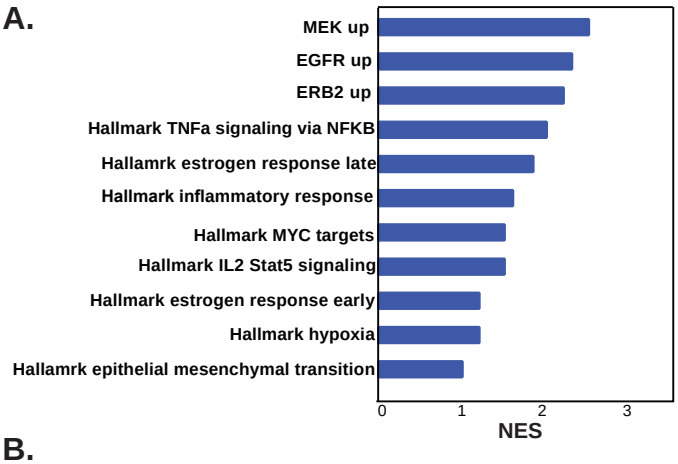
cell counting on days 1, 3 and 5 in MCF7 parental cells after SOX9 expression or an empty vector (EV) control in full medium (top left) or white medium (top right). Dose response curves for tamoxifen treatment in MCF7 cells expressing SOX9 and an empty vector (EV) as control. Tamoxifen IC50 in SOX9 overexpressing cells is $1.4X10^{-9}$ M and in control cells expressing an EV $6X10^{-10}$ M.

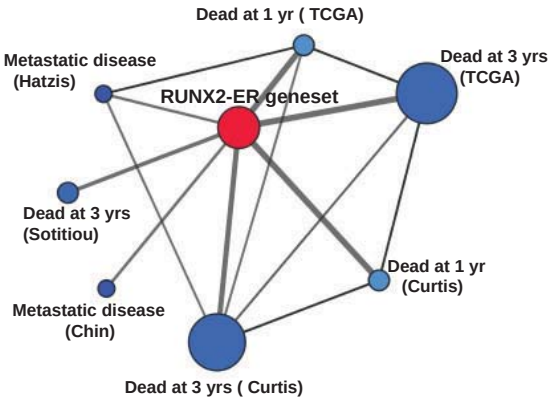
Figure 6. Tamoxifen resistance in clinical samples is associated with SOX9 expression: A. Representative figures of immunohistochemistry staining for SOX9 showing increase in nuclear staining in the recurrent tamoxifen-resistant samples. Magnification 40X. B. SOX9 expression in ER+ primary and recurrent breast cancer samples. The scatter plot shows that in a number of cases there was a decrease in SOX9 expression after the development of tamoxifen resistance, but in the majority of the cases (67%) there was an increase and overall there was a significant increase C. SOX9 expression levels in the recurrent tumors that remained ER positive and the recurrent tumors that changed to ER negative and their matched primary tumors. D. Scheme of the model of the ERα-RUNX2 model in tamoxifen resistance.


A
PAR siCON
PAR siER
TAMR siCON
TAMR siER
ER
B-actin
cell count
days
PAR-siCON
PAR-siER
TAMR-siCON
TAMR-siER
B
VEH
FULV 10-7M
FULV 10-6M
TAMR
Veh
Fulv 10-7M
Fulv 10-6M
C
PAR
TAMR
area change
Radius Assay
OD 560nm
Boyden Chamber
D
log10(pvalue)
log2FoldChange
% of genes
Response to oxidative stress
Positive regulation of cell migration
Negative regulation of apoptosis
Response to hormone stimulus
Regulation of programmed cell death
Filopodium assembly
p-values

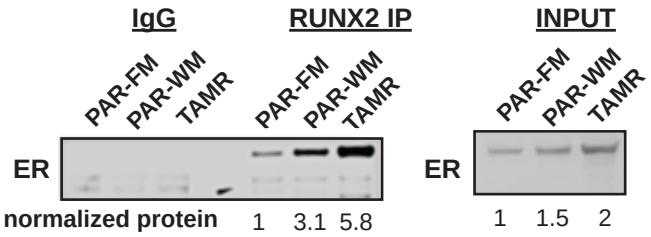

A.
LTED
(2000)
TAMR
(1451)
PAR
(7617)
2601
B.
Enriched motifs in TAMR
Lost motifs in TAMR
-10log*(pval)
293
RUNX
293
FOXM1
-10log*(pval)
690
PAX6
686
GATA
C.
RUNX motifs
GATA3 motifs
TAMR
PAR
+1KB
-1KB
0
5
D.
RUNX1
RUNX2
RUNX3
relative mRNA level
PAR
LTED
TAMR
E.
PAR
LTED
TAMR
RUNX2
b-ACTIN
GATA3
b-ACTIN

A.



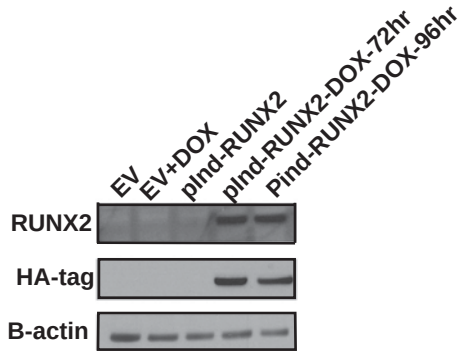


B.

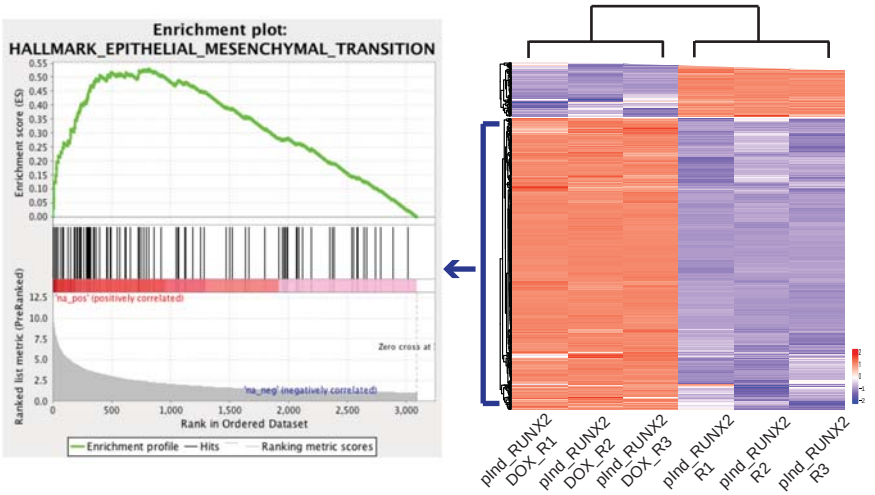


C.

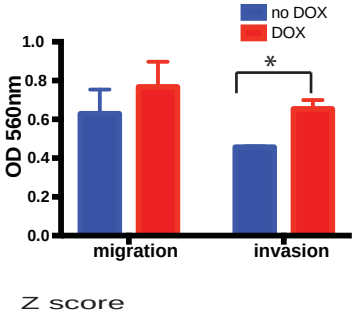
D.



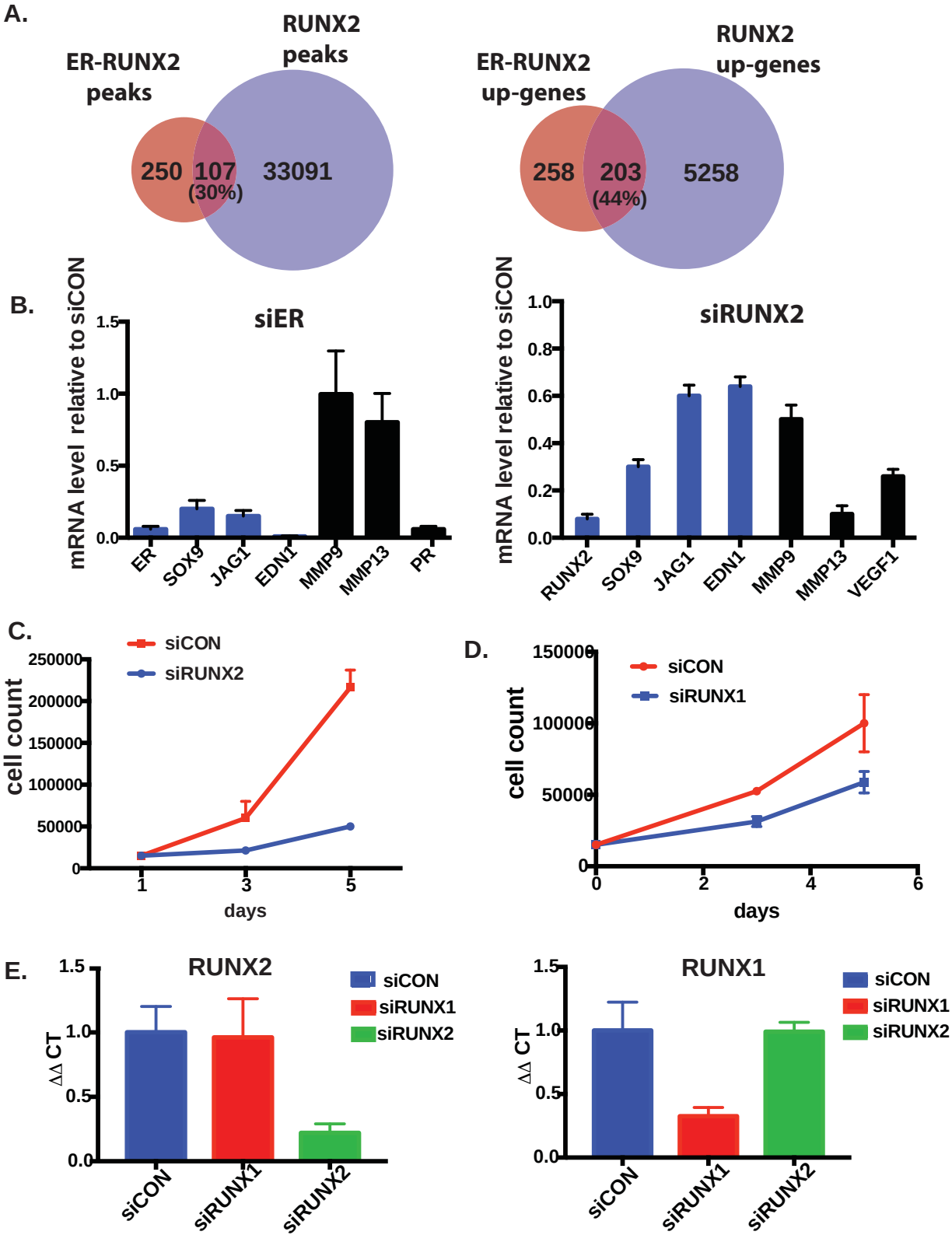
E.

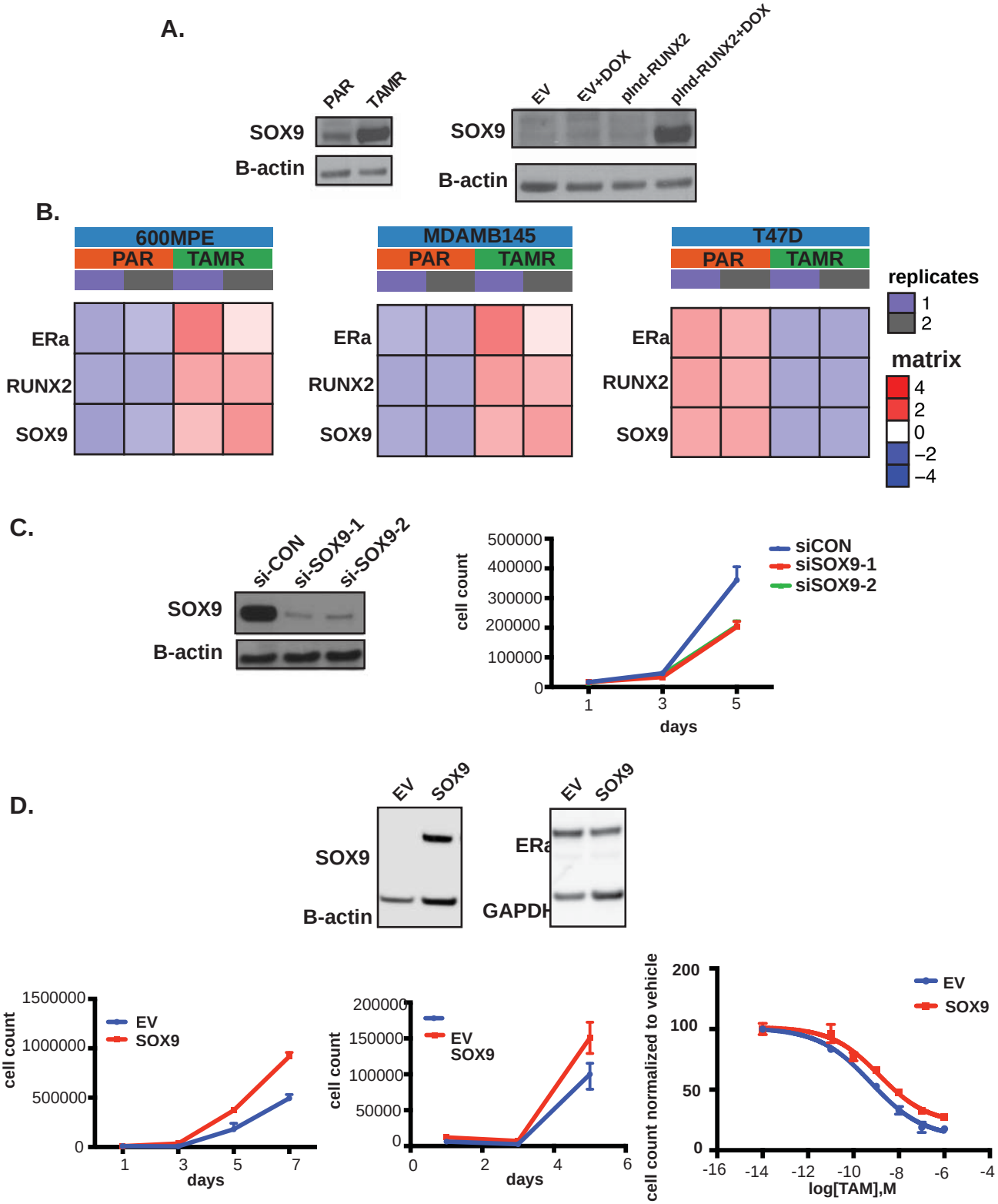


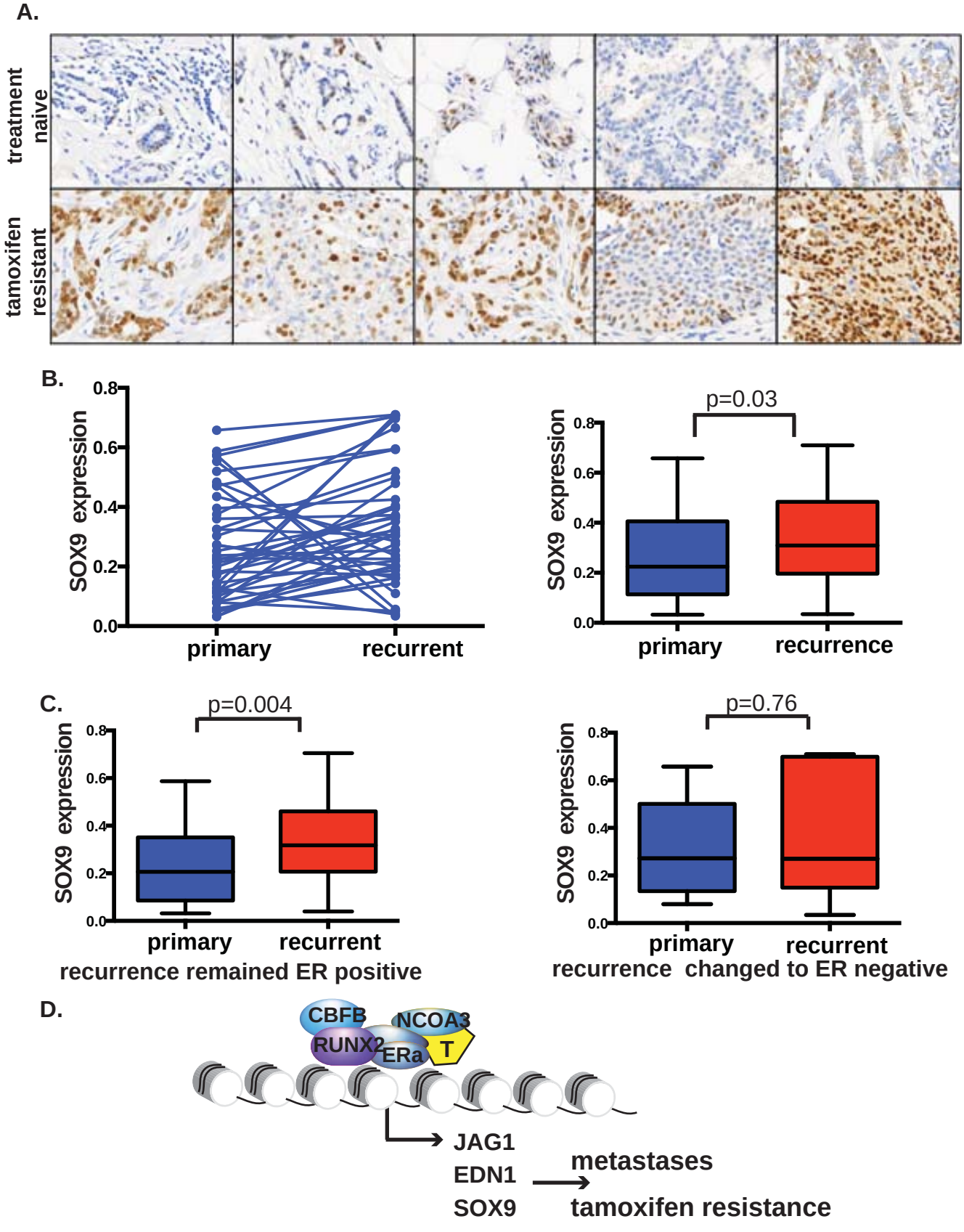
F.



Z score


A.
ER-RUNX2 peaks
RUNX2 peaks
250
107
(30%)
33091
ER-RUNX2 up-genes
RUNX2 up-genes
258
203
(44%)
5258
B.
siER
mRNA level relative to siCON
ER
SOX9
JAG1
EDN1
MMP9
MMP13
PR
siRUNX2
RUNX2
VEGF1
C.
siCON
siRUNX2
cell count
days
D.
siRUNX1
E.
RUNX2
RUNX1
ΔΔ CT


A.
PAR
TAMR
SOX9
B-actin
EV
EV+DOX
pInd-RUNX2
pInd-RUNX2+DOX
SOX9
B-actin
B.
600MPE
PAR
TAMR
MDAMB145
PAR
TAMR
T47D
PAR
TAMR
ERa
RUNX2
SOX9
replicates
1
2
matrix
4
2
0
−2
−4
C.
si-CON
si-SOX9-1
si-SOX9-2
SOX9
B-actin
cell count
500000
400000
300000
200000
100000
0
1
3
5
days
siCON
siSOX9-1
siSOX9-2
D.
EV
SOX9
SOX9
B-actin
ERa
GAPDH
cell count
1500000
1000000
500000
0
1
3
5
7
days
EV
SOX9
cell count
200000
150000
100000
50000
0
0
2
4
6
days
EV
SOX9
cell count normalized to vehicle
200
100
50
0
-16
-14
-12
-10
-8
-6
-4
log[TAM],M
EV
SOX9


A.
treatment naive
tamoxifen resistant
B.
SOX9 expression
0.0
0.2
0.4
0.6
0.8
primary
recurrent
p=0.03
primary
recurrence
C.
p=0.004
recurrence remained ER positive
p=0.76
recurrence changed to ER negative
D.
CBFB
NCOA3
RUNX2
ERa
T
JAG1
EDN1
SOX9
metastases
tamoxifen resistance