

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

Nastaran Ghahhari

University of Geneva

Magdalena Sznurkowska

University of Basel and University Hospital Basel

Nicolas Hulo

Institute of Genetics and Genomics University of Geneva Switzerland

Lilia Bernasconi

University of Geneva

Nicola Aceto

ETH Zurich

Didier Picard (✉ didier.picard@unige.ch)

Université de Genève <https://orcid.org/0000-0001-8816-9668>

Article

Keywords: epithelial to mesenchymal transition (EMT), breast cancer, estrogen receptor α (ER α)

Posted Date: May 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-493200/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

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

Nastaran Mohammadi Ghahhari¹, Magdalena K. Sznurkowska², Nicolas
Hulo³, Lilia Bernasconi¹, Nicola Aceto^{2,4} & Didier Picard^{1,*}

¹Département de Biologie Cellulaire, Université de Genève, Sciences III, 1211

Genève 4, Switzerland

²Cancer Metastasis Laboratory, Department of Biomedicine, University of Basel and
University Hospital Basel, 4058 Basel, Switzerland

³Institute of Genetics and Genomics of Geneva, Université de Genève, 1211 Genève
4, Switzerland

⁴Current address: Department of Biology, Institute of Molecular Health Sciences, ETH
Zurich, 8093 Zürich, Switzerland

* Correspondence: didier.picard@unige.ch

Abstract

The epithelial to mesenchymal transition (EMT) has been proposed to contribute to the metastatic spread of breast cancer cells. EMT-promoting transcription factors determine a continuum of different EMT states. In contrast, estrogen receptor α (ER α) helps to maintain the epithelial phenotype of breast cancer cells and its expression is crucial for effective endocrine therapies. Determining whether and how EMT-associated transcription factors such as ZEB1 modulate ER α signaling during early stages of EMT could promote the discovery of novel therapeutic approaches to suppress metastasis. We have discovered that, shortly after induction of EMT and while cells are still epithelial, ZEB1 modulates ER α -mediated transcription induced by estrogen or cAMP signaling in breast cancer cells. Based on these findings and our *ex vivo* and xenograft results, we suggest that the functional interaction between ZEB1 and ER α may alter the tissue tropism of metastatic breast cancer cells towards bone.

Introduction

ER α is a nuclear hormone receptor that mediates the transcriptional regulation of specific target genes during normal mammary development and breast tumorigenesis^{1,2}. Because ER α drives two thirds of breast cancers, it has been recognized as an important prognostic marker and a therapeutic target. Using potent ER α antagonists such as tamoxifen and fulvestrant (ICI), ER α -positive (ER α^+) breast tumors are targeted with antiestrogen therapy. However, more than a quarter of all breast cancer patients develop antiestrogen resistance, which remains a major hurdle in managing their clinical outcome³⁻⁵. Among a plethora of mechanisms that have been found to contribute to endocrine resistance, there are distinct changes in the tumor microenvironment, which stimulate cancer cell proliferation and induce invasiveness^{6,7}. In this context, EMT of non-invasive breast cancer cells has been proposed to play a key role in their progression to high-grade metastatic tumors and differential responses to endocrine therapy^{5,8,9}.

EMT is orchestrated through the action of a number of transcription factors (EMT-TFs), which shape the malignant transformation of carcinoma cells by modifying gene expression¹⁰⁻¹². ZEB1/2, SNAIL1/2, and TWIST1/2 are core EMT-TFs, which regulate the transitions among different EMT stages in an interdependent fashion^{6,8,13}. A substantial set of genes involved in the maintenance of the epithelial state (for example the E-cadherin gene *CDH1*) are repressed upon activation of EMT-TFs. In parallel, with the activation of genes associated with the mesenchymal state, a partial mesenchymal phenotype is acquired^{14,15}. However, recent studies suggest that a partial and reversible EMT phenotype or an intermediate hybrid state

of breast cancer cells is associated with metastasis, chemoresistance, and poor prognosis for the patients^{10,16,17}.

ZEB1 is a key factor for cell fate determination, tumor initiation, cancer cell plasticity, and metastatic dissemination^{13,18,19}. ZEB1 is generally considered to be a transcriptional repressor, but it can also act as a transcriptional activator^{20,21}.

Compared to luminal breast cancer subtypes, ZEB1 is highly expressed in triple-negative breast cancers, which express neither ER α nor progesterone receptor (PR), which is encoded by an ER α target gene^{22,23}. Comprehensive analyses of samples from breast cancer patients support the co-existence of epithelial cells with low levels of ER α with mesenchymal cells expressing high levels of ER α in the same tumor microenvironment¹⁷; the latter are reminiscent of cells in an EMT hybrid state. Although loss of ER α function promotes an EMT-associated phenotype in breast cancer cells²⁴⁻²⁶, ER α activation can also induce EMT in other hormone-inducible cancers^{11,27}.

The transcriptional activity of ER α can be switched on by both cognate ligand and ligand-independent pathways to regulate cell functions in the mammary epithelium²⁸. Binding of 17 β -estradiol (E2) to its hormone binding domain (HBD) triggers the binding of an ER α homodimer to specific DNA sequences containing estrogen response elements (EREs), often dependent on the prior binding of pioneer factors such as FOXA1, GATA3, and AP2 γ to chromatin^{4,29-31}. cAMP-activated protein kinase A (PKA) activates ER α primarily indirectly by promoting the phosphorylation of ER α coregulators including CARM1, LSD1, CREB1, and their interactions with ER α ³²⁻³⁵.

These findings motivated us to investigate the effects of the EMT-inducer ZEB1 on both liganded and unliganded ER α transcriptional responses. Indeed, it was unknown whether the two factors cooperate to modulate EMT programs in breast cancer and to transform non-metastatic into more invasive cancer cells. The discovery of such mechanisms may reveal novel molecular targets that could lead to more effective therapeutic strategies to prevent breast cancer progression. By analyzing the ZEB1-ER α interdependent transcriptional activities, we have been able to reveal new mechanisms by which ZEB1 drives tumor progression and invasion of ER α ⁺ breast cancer cells.

Results

ZEB1 enhances ER α transcriptional activity during early EMT stages. We used the luminal breast cancer cell line MCF7 and its variant MCF7-V, which displays more robust ER α responses, and the luminal breast cancer cell line T-47D to establish cells stably expressing ZEB1 from a doxycycline (DOX)-inducible (Tet-on) lentiviral vector. We monitored the expression of EMT-associated markers in the absence (–DOX) or presence (+DOX) of DOX after short-term (1-2 weeks) and long-term (8-12 weeks) expression of ZEB1 to achieve partial and complete EMT, respectively. A complete EMT indicated by the expression of mesenchymal markers vimentin and N-cadherin was detected after long-term expression of ZEB1 (Fig. 1a and Supplementary Fig. 1a). ZEB1 downregulated ER α expression, in agreement with previous reports^{36,37}. Surprisingly, one to two weeks after induction of ZEB1 expression, ER α expression was still maintained (Fig. 1a). Because ZEB1 did not

affect ER α levels in the short-term, we wondered whether it affects ER α transcriptional activity. This was explored with luciferase reporter assays with various cell lines. ZEB1 significantly enhanced an ERE-containing luciferase reporter activity upon activation of ER α by E2 or forskolin/3-isobutyl-1-methylxanthine (FI) (Fig. 1b, c and Supplementary Fig. 1b). Other EMT-TFs, TWIST1 and PRRX1³⁸, repressed the ER α transcriptional response; however, expression of ZEB1 reversed this effect and increased ER α activity (Fig. 1d). Because ZEB1 expression also correlates with the presence of other hormone receptors³⁹, we used a construct containing the progesterone response element (PRE-Luc) to measure the PR activity. In contrast to ER α , PR activity was strongly repressed by ZEB1 (Fig. 1e and Supplementary Fig. 1c), indicating a specific ER α -dependent response to ZEB1.

ZEB1 repressed the E-cadherin promoter and stimulated the transcription from the vascular endothelial growth factor A (VEGFA) promoter as expected^{40,41}, the latter being more strongly activated in the presence of ER α (Fig. 1f, g). Transforming growth factor β (TGF β) induces the EMT⁸, and SMAD4 is a TGF β -mediating transcriptional co-repressor for ER α in breast cancer⁴². Although TGF β alone repressed ER α activity, ZEB1 significantly reversed this effect (Fig. 1h). ZEB1 did not affect the activation of a SMAD reporter (Fig. 1i), confirming that the ERE is essential for ZEB1-mediated enhanced ER α activity. To determine if ZEB1 affects the expression of ER α target genes, we examined their mRNA levels. ZEB1 increased expression levels of all assessed genes (Fig. 1j and Supplementary Fig. 1d, e). Conversely, long-term expression of ZEB1, resulting in a mesenchymal phenotype, reduced the expression of ER α targets (Supplementary Fig. 1f). Overall, these findings indicate that ZEB1 potentiates ER α -mediated transcription in ligand-

dependent and -independent manners, possibly with functional relevance to the early stages of EMT.

Increased ZEB1-induced invasion ability of breast cancer cells is associated

with ER α . Activation of ER α by E2 increases breast cancer cell invasiveness^{43,44},

but the effect of activating ER α with cAMP on invasion is not clear. We used a native-like three-dimensional (3D) tumor microenvironment model⁴⁵. Notably in the presence of E2, tumor spheroids expressing ZEB1 (+DOX) grew to a larger size within 96 hours (h) of embedding in collagen compared to control spheroids (-DOX), and displayed a significantly increased dissemination from the main spheroid body into the surrounding matrix (Fig. 1k, l). cAMP/PKA signaling was induced using FI. In accordance with recent findings that PKA activation reverses the EMT and induces a mesenchymal-epithelial transition (MET)⁴⁶, FI suppressed the invasion of cells. Moreover, ZEB1-expressing cells preceded the -DOX cells in invasion, but the antiestrogen ICI suppressed invasion (Fig. 1k, l). These observations suggest that, in response to E2, ZEB1 enhances ER α -mediated cell invasion. 3D invasion assays with T47-D cells showed similar results (Supplementary Fig. 1g). We used the Gene expression-based Outcome for Breast cancer Online (GOBO) tool to correlate ZEB1 expression with outcome in ER α ⁺ and ER α ⁻ breast cancer patients. Interestingly, higher levels of ZEB1 improves overall survival (OS) and distant metastasis-free survival (DMSF) of ER α ⁺ patients (Supplementary Fig. 1h, i). In contrast, ER α ⁻ patients showed no correlation with OS, but higher ZEB1 expression adversely affected the DMSF (Supplementary Fig. 1j, k). This indicates that with relatively high levels of ZEB1, the ER α status determines outcome.

ZEB1 induces transitional sensitivity to ER α antagonist. We determined with the GOBO tool whether outcome in breast cancer patients treated with the ER α antagonist tamoxifen correlates with ZEB1 expression levels. Indeed, improved relapse-free survival (RFS) and DMSF in tamoxifen-treated patients with higher expression of ZEB1 suggested that an increased ER α activation due to ZEB1 could induce a transitional sensitivity to 4-OHT during early EMT stages (Supplementary Fig. 1l, m). We performed a cell cycle assay with increasing concentrations of the active tamoxifen metabolite 4-hydroxytamoxifen (4-OHT). ZEB1 sensitized the cells to 4-OHT as indicated by the fact that ~75% of the cells were arrested in the G0/G1 cell cycle phase compared to control cells with ~60% G0/G1 arrest (Supplementary Fig. 1n).

ZEB1 and ER α form a transcriptional complex. Our data suggest that, during early/hybrid EMT states, ZEB1 can functionally modulate ER α responses. To determine the underlying mechanisms, we tested whether ZEB1 affects the recruitment of ER α to its chromatin binding sites. ZEB1 increased ER α recruitment to the known enhancers of the ER α target genes *GREB1* and *TFF1*. Without activation of ER α , ZEB1 could not increase ER α recruitment (Fig. 2a and Supplementary Fig. 2a). We performed ER α chromatin immunoprecipitation-sequencing (ChIP-seq) with MCF7-V-ZEB1 cells to reveal possible global changes induced by ZEB1. Comparing these data to our previously published ER α ChIP-seq data for the MCF7-V parent cells³⁵, we uncovered 3,149 novel ER α binding sites (ERBSs) induced by E2 and 2,156 ERBSs induced by FI, all unlocked by ZEB1 (Figure 2b, c, and Supplementary Data 1). The Genomic Regions Enrichment of Annotations Tool (GREAT) for the functional annotation of ZEB1-induced ERBSs revealed biological functions

predominantly related to EMT, migration and activation of WNT signaling (Fig. 2d, e). GREAT predicted several phenotypes associated with abnormal bone morphogenesis, being in line with the proposed functions of ZEB1 during bone development and osteoblast differentiation (Supplementary Fig. 2b, c)²¹. Interestingly, following *de novo* motif analysis we found the ZEB1 motif enriched in hundreds of ERBSs (Fig. 2f and Supplementary Data 2). Genome browser views of some top novel ERBSs associated with the *TBX2*, *ANXA3*, *CEP89*, and *SLC25A24* genes show a significantly increased ER α recruitment upon ZEB1 expression (Supplementary Fig. 2d and Supplementary Data 1), which we verified by ChIP-qPCR (Fig. 2g, h). We wondered whether ZEB1 might also be recruited to ERBSs. This appears to be the case at least for some since ZEB1 significantly bound to the *GREB1* +5 kb and *TFF1* +0.5 kb ERBSs, but not to other more remote *TFF1* regions (Fig. 2i).

We extended our findings by comparing ZEB1 binding sites and ERBSs on a genome-wide scale. We performed ER α ChIP-seq with wild-type MCF7-V cells (lacking the construct for inducible ZEB1 expression). After peak-calling and applying stringent cutoffs, we acquired 22,271 regions induced by E2 and 13,497 ERBSs with FI. We overlapped the ER α peaks with 32,907 ZEB1-binding sites from MDA-MB-231 and 13,514 sites from Hs578T breast cancer cells⁴⁷. Remarkably, E2-induced ER α shares 2,747 binding sites with ZEB1 and 1,323 regions overlap with ZEB1 for FI-activated ER α in MDA-MB-231 cells. We observed similar percentages of ERBSs shared with ZEB1 in Hs578T cells (Fig. 2j-l). With a re-ChIP of candidate sites from the intersection of ZEB1 and ER α shared binding sites we showed that the two TFs can be simultaneously present at the same chromatin location (Fig. 2m). Overall, we

show that ZEB1 stimulates the ligand-dependent and -independent recruitment of ER α and confirm that the two TFs share several cis-regulatory regions as part of joint transcriptional complexes.

ZEB1 interacts with ER α and AP2 γ is required for the ZEB1-induced ER α

activity. Our findings suggested that ZEB1 and ER α could be present in the same complex of TFs, interacting directly or indirectly. Co-immunoprecipitations (co-IPs) of HA-tagged ZEB1 or ER α confirmed that ZEB1 and ER α are present in the same protein complexes, notably upon activation of ER α with E2 (Fig. 3a). We found that the zinc finger cluster 1 of ZEB1 and the F-domain of ER α may play key roles in mediating the stimulation of ER α activity by ZEB1. However, we could not detect a direct physical interaction of these two domains. This suggests that a direct interaction either depends on other domains or that the interaction is indirect (Supplementary Results and Supplementary Fig. 3).

We performed a motif analysis for the regions present at the intersection of ZEB1 and ERBSs. As expected, the binding sites of ER α and ZEB1 were significantly enriched motifs (Supplementary Fig. 4a). Remarkably, we found the binding sites of the AP2 family, including TFAP2A/B/C, as some of the most highly enriched sequences, ranking higher than motifs for forkhead factors such as the ER α pioneer factor FOXA1. We also found the enrichment of AP2 motifs in the ZEB1-induced differentially-bound ERBSs (Fig. 2). AP2 factors play key roles in regulating differentiation, with the *TFAP2C* gene being expressed in adult mammary myoepithelial cells. AP2 γ is involved in breast cancer proliferation and metastasis. Moreover, AP2 γ is recruited to ERBSs to regulate transcription⁴⁸⁻⁵⁰. AP2 γ , FOXA1

and ER α co-target genes of the luminal phenotype during breast cancer progression^{51,52}. We decided to characterize the possible involvement of AP2 γ and FOXA1 in the ZEB1-ER α transcriptional complex. Intersecting published ChIP-seq for AP2 γ ⁴⁸ and ZEB1⁴⁷ with our own data for ER α , we found > 1,500 sites shared among the three factors (Fig. 3b). Overlapping ZEB1 and ERBSs with FOXA1 peaks²⁹ resulted in only 85 shared binding sites (Fig. 3c), suggesting that FOXA1 is not a defining factor for the ZEB1-ER α complex.

We compared averaged ChIP-seq signal intensities for ZEB1, ER α , AP2 γ , and FOXA1 with those of known factors and chromatin marks of estrogen-regulated active enhancers including GATA3, P300, H3K27ac, H3K4me1, and H3K9me3²⁹. Compared to FOXA1, AP2 γ shows a stronger signal around the center of the common binding sites with ZEB1 and ER α (Fig. 3d) and that ZEB1 and AP2 γ can be a part of ER α TF complexes at sites of open chromatin associated with the histone marks H3K27ac and H3K4me1 (Supplementary Fig. 4b-d). Furthermore, ZEB1, ER α , AP2 γ , and FOXA1 seem to co-localize to certain binding sites (Fig. 3e). We looked at the effects of reducing AP2 γ and FOXA1 levels for the ZEB1-stimulated ER α activity. We knocked down AP2 γ or FOXA1 expression (Supplementary Fig. 4e) and found that the activation of the ERE-Luc reporter is reduced upon depletion of AP2 γ ; this reduction is even more prominent in the presence of ZEB1. The knockdown of FOXA1 did not affect ER α activity by itself in this particular experimental setup. However, it compromised ER α activity in the presence of ZEB1 even more strongly than without it, and the same could be observed with the knockdown of AP2 γ (Fig. 3f).

AP2 γ and FOXA1 interact with ER α as part of the same protein complexes²⁹, and we could confirm by co-IPs that AP2 γ and FOXA1 also form complexes with ZEB1 (Supplementary Fig. 4f). We explored whether AP2 γ and FOXA1 are required for the physical association of ZEB1 and ER α . For either ZEB1 or ER α , reducing AP2 γ and FOXA1 levels (Supplementary Fig. 4g) resulted in the complete loss of the interaction with the other factor (Fig. 3g). We assessed the consequences of the absence of AP2 γ and FOXA1 on the recruitment of ER α . Upon FOXA1 knockdown we observed a similar loss of ER α recruitment to *GREB1* and *TFF1* enhancers independently of ZEB1 expression (Fig. 3h). Remarkably, in the presence of ZEB1, but not in its absence, the loss of AP2 γ caused a very significant decrease in ER α recruitment (Fig. 3i). This mirrors the larger impact of the AP2 γ knockdown on the ZEB1-stimulated ER α activity in our reporter assays. Knowing that ZEB1 could also increase the *TFAP2C* mRNA levels (Supplementary Fig. 4h), we assumed that ER α binding to the *TFAP2C* enhancer should be also enhanced by ZEB1. Our ChIP-seq data in $-/+$ DOX cells demonstrated that ZEB1 enhances ER α recruitment to the 5' UTR of *TFAP2C* (Fig. 3j), which we could confirm by ChIP-qPCR (Fig. 3k). Enhanced binding of ER α to the *TFAP2C* promoter increased the expression levels of the AP2 γ protein (Fig. 3l). Overall, we show that while ZEB1 activates a distinct ER α response during early-EMT, AP2 γ binding is necessary for effective and functional ligand-dependent and -independent activation of ZEB1-ER α -bound enhancers.

ZEB1 reprograms the ER α transcriptome towards a metastatic profile. To investigate the impact of ZEB1 on ER α -regulated gene expression, we performed RNA-seq of $-/+$ DOX cells treated with either vehicle, E2 or FI for 6 h (Fig. 4 and

Supplementary Fig. 5). With E2 a total number of 7,107 genes in –DOX cells (Supplementary Data 3) and 6,748 genes in +DOX (Supplementary Data 4) were affected. FI up/down-regulated the expression of 11,063 genes without ZEB1 (Supplementary Data 3) and 10,570 genes upon ZEB1 expression (Supplementary Data 4) (Fig. 4a and Supplementary Fig. 5a). 70-80% of these differentially expressed genes were downregulated ≥ 1.5 -fold upon ER α activation (Supplementary Data 3 and 4). ZEB1 caused an overall increase in the expression of ER α target genes (Supplementary Data 4). We performed gene set enrichment analysis (GSEA) followed by clustering the output list of enriched GO terms and generated annotated enrichment maps (Supplementary Fig. 5b-e). ZEB1 shifted the cells towards a more mesenchymal and invasive phenotype, with enrichment of genes associated with cell morphogenesis, neuronal differentiation, increased cell motility, and extensive cytoskeletal changes (Supplementary Fig. 5d, e). We used the ClusterProfiler package in R⁵³ to classify gene sets with the GO term collections "Biological Process", "Cellular Component", and "Molecular Function". Several functions and processes associated with cell-cell junction, adhesion, and cellular anatomical entity characteristic of the epithelial cell phenotype were inhibited by ZEB1-ER α transcriptional activity (Supplementary Fig. 5f-h).

Most remarkably, we found that ZEB1 unlocked > 1,000 previously undescribed ER α target genes (Fig. 4b, c and Supplementary Data 5). The unique GSEA terms from each group showed that these genes are most likely involved in EMT-related phenotypes, stem cell differentiation, bone morphogenesis, and ossification (Fig. 4d-g for unique terms; full GSEA list in Supplementary Fig. 6a-d). To confirm that the novel genes unlocked by ZEB1 are ER α targets, we used ICI as ER α antagonist and

analyzed mRNA levels from several top hits (Fig. 5 and Supplementary Fig. 7a, b). We selected *DIO2*, *MUC16*, *DSCAM*, and *ESR2*, on the one hand, and *MUC2*, *P2RX7*, *HSPB8*, and *SCG2*, on the other, which are upregulated by ZEB1 in the presence of E2 and FI, respectively (Fig. 5a). *MUC16*, *DSCAM*, and *ESR2* could be confirmed to be E2-induced ER α targets upregulated by ZEB1 (Fig. 5b). Among the FI-induced ER α -dependent targets, *P2RX7* and *HSPB8* were highly upregulated in the presence of ZEB1. The expression profile of *MUC2* was highly unusual in that it was massively induced by ZEB1, but only in the presence of ICI as if ER α repressed this induction under very specific conditions (Fig. 5c).

Using 3D spheroid invasion assays, we found that the knock down of *MUC16*, *DSCAM*, *ESR2*, and *P2RX7* (Supplementary Fig. 7c) inhibited the invasion of ZEB1-expressing cells, while the invasion capacity of wild-type cells remained unaffected (Fig. 5d). *DIO2* depletion had no significant impact on the ability of spheroids to invade in the presence of ZEB1 and E2, corroborating our conclusion that *DIO2* is not a genuine ER α target (Supplementary Fig. 7d). Loss of *MUC2*, *HSPB8*, and *SCG2* completely disrupted the formation of tumor spheroids in cells expressing ZEB1. Cumulatively, these data confirm the pleiotropic role of ZEB1 in modulating the ER α transcriptome and for the acquisition of an invasive cell phenotype during the early stages of EMT.

ZEB1 induces different EMT transition states. We monitored the progressive loss of EpCAM during EMT at different time points upon induction of ZEB1 expression in MCF7-V-ZEB1 cells until a mesenchymal phenotype was obvious by visual inspection (Supplementary Fig. 8a). To determine whether ZEB1 promotes cellular

heterogeneity by inducing different intermediate EMT states we performed droplet-mediated single cell RNA sequencing (scRNA-seq) of FACS-isolated EpCAM^{high} and EpCAM^{low} cells (Supplementary Fig. 8b). We identified the highly variable features in both populations including the metastasis-associated genes *MALAT1*, *TFF1* and *IGFBP5* (Supplementary Fig. 8c). We identified 11 distinct subpopulations of cells, representing five clusters of the EpCAM^{high} and six of the EpCAM^{low} subpopulations (Fig. 6a). EpCAM^{high} clusters presented higher expression of *ESR1*, supporting the notion that the progressive loss of the epithelial state is associated with reduced levels of ER α (Fig. 6b-e). We confirmed the expression of several epithelial and mesenchymal markers at the single-cell level pertaining to epithelial, mesenchymal, and hybrid states (Fig. 6f-h and Supplementary Fig. 8d). This included genes such as *BRIPI*, *MKI67*, and *CLDN7* as epithelial cell markers⁵⁴⁻⁵⁷ (Fig. 6f). Interestingly, we found genes such as *ANXA2*, *KRT8*, *HSPB1*, and *MDK* expressed in both EpCAM^{high} or EpCAM^{low} clusters, suggesting that these genes could participate in the establishment of a hybrid EMT state (Fig. 6g). Genes such as *LGALS1*, *S100A6*, and *LRRC75A* facilitate invasion and were mostly expressed in the EpCAM^{low} clusters (Fig. 6h).

We examined the expression of integrin markers in our dataset and could not detect the expression of the surface markers CD61 (*ITGB3*) and CD106 (*VCAM1*), whereas the surface marker CD51 (*ITGAV*) was expressed in some cells from the EpCAM^{high} clusters (Supplementary Fig. 8e). This suggests the existence of a triple negative subpopulation (CD61⁻/CD51⁻/CD106⁻) in the EpCAM^{low} population, which is a characteristic of early/hybrid EMT stages¹⁶. We also observed homogeneous expression of CD29 (*ITGB1*) and CD59 in all subpopulations (Supplementary Fig.

8f), but the expression of the tetraspanin CD151 was primarily enriched in the EpCAM^{low} population (Fig. 6i, j and Supplementary Fig. 8g). The tetraspanin CD151 is a transmembrane integrin involved in metastasis to bone⁵⁸⁻⁶¹. High expression of CD151 supports tumor growth, and this dependency is associated with ZEB1/2. CD151 dependency and higher expression of ZEB1 are positively correlated with EGFR inhibitor resistance⁶². Therefore, we investigated whether CD151 could be a potential therapeutic target during early EMT stages when ZEB1 is expressed, and how this could modulate the response to EGFR inhibitors. We knocked down CD151 (Supplementary Fig. 8h) and measured the viability of cells exposed to the EGFR inhibitors gefitinib and erlotinib. Depletion of CD151 in the absence of ZEB1 inhibited cell growth. Remarkably, cells expressing ZEB1 showed a higher sensitivity to both inhibitors, and the loss of CD151 further accentuated this effect (Fig. 6k, l). We used the GOBO database and extracted the data for ER α ⁺ and ER α ⁻ breast cancer patients in relation to ZEB1 and CD151 expression levels. The OS and DMSF of ER α ⁺ patients with higher expression of ZEB1 and CD151 were improved (Fig. 6m), whereas the correlation was inversed for ER α ⁻ breast cancer patients (Fig. 6n). Overall, these data suggest that CD151 dependency is associated with the ZEB1 and ER α status, inducing a transitional sensitivity to EGFR antagonists during early EMT.

ZEB1 reprograms breast cancer cells to induce metastasis *in vivo*. High levels of ZEB1 in invasive ER α ⁻ tumors are associated with the expression of genes suggested to be involved in breast cancer bone metastasis⁶³. Our data indicate that ZEB1 modulates gene signatures related to bone development and abnormal phenotypes in ER α ⁺ breast cancer cells (Fig. 2 and Fig. 4). This led us to ask

whether ZEB1 modifies the organ tropism of ER α ⁺ breast tumor metastases. The conventional MCF7 xenograft models frequently develop metastatic lesions in lungs, brain, liver, and spleen, but not in bones⁶⁴⁻⁶⁶. We therefore compared the metastatic potential of control ZsGreen-expressing and ZEB1-expressing wild-type MCF7 cells (+DOX) in xenograft experiments with immunocompromised mice. Since the cells were also marked with luciferase for *in vivo* detection, we used their bioluminescence as an approximation of their relative sizes. Surprisingly, compared to control ZsGreen-expressing cells, ZEB1-expressing cells induced the formation of smaller primary tumors (Fig. 7a). The bioluminescence imaging of organs from the same sacrificed individuals revealed overt metastatic lesions in bones and lungs (Fig. 7b, c, and Supplementary Fig. 8i), but not in liver and brain in either groups (Supplementary Fig. 8j, k). Although ZEB1 expression led to smaller primary tumors, these animals had a similar or even increased metastatic burden as indicated by increased metastatic indices (ratio of total bioluminescence counts of the metastatic organ over those of the primary tumor) (Fig. 7b, c). 3 out of 3 mice from the ZEB1 group developed overt bone metastases, whereas in the control group only 2 out of 5 mice showed significant metastatic lesions in bones, as judged by the bioluminescence and considering only those with values above 10² (Fig. 7a-d). To further support the tentative conclusion that ZEB1 may redirect some metastasis formation to bones, we used two different orthogonal approaches. With an *ex vivo* bone invasion assay using murine femoral bones, we observed that induction of ZEB1 expression greatly induced the invasion of the ZsGreen-marked MCF7-V cells into the bone when compared to the control cells without ZEB1 (Fig. 7e). In a standard transwell assay, we also readily detected the migration of the ZEB1-

expressing cells through the transwell membrane towards bone in the lower chamber, but not to muscle or decellularized bone (Fig. 7f).

Discussion

The induction of EMT-TFs and the acquisition of mesenchymal characteristics are associated with loss of ER α and resistance to antiestrogen therapies^{12,36,37,67}, but the impact of early/hybrid states of EMT on ER α signaling had not been investigated. Here, we describe key roles for ZEB1 during early EMT stages in enhancing ER α responses and suggest that the functional ZEB1-ER α interaction may modulate the tissue tropism of breast cancer metastases. Our results reveal that ZEB1 interacts with ER α at shared binding sites at the enhancers of genes involved in EMT, invasion, and bone morphogenesis (Fig. 8). We demonstrate that the interaction between ZEB1 and ER α not only confers augmented transcriptional activation of liganded ER α at the genomic regions that they co-occupy, but it also alters the regulation of gene expression by cAMP/PKA-activated unliganded ER α .

Whereas ZEB1 has been known as a repressor for some time, its transcriptional activation function has only been described more recently^{20,21}. In line with a previous report³⁷, we observed that ZEB1 suppresses ER α expression once a mesenchymal phenotype is achieved. ZEB1 is known to activate the transcription of TGF β /BMP pathway genes to support osteoblast differentiation²¹. In ER α ⁻ breast cancers ZEB1 activates the transcription of YAP target genes²⁰. We show that ER α protein levels in luminal breast cancer cells remain unchanged up to 1-2 weeks after the induction of ZEB1 expression while cells maintain epithelial features. Moreover, we could not

detect any changes in the number of *ESR1* transcripts by RNA-seq. This suggests that ER α expression could be dynamically modulated at different stages of EMT and that ZEB1 may not have any impact on ER α levels at early/hybrid stages of EMT. Binding of ZEB1 to the two E-box elements within the upstream CpG-rich region of the *ESR1* promoter represses ER α expression³⁷. In contrast, our ER α reporter construct contains an ERE and its ER α -mediated expression is significantly boosted by ZEB1. We also demonstrate that a *VEGFA* promoter is more strongly activated by both ER α and ZEB1, and that ZEB1 reverses the inhibitory effect of TGF β on ER α activity. Therefore, we speculate that the existence of an ERE is indispensable for ZEB1 to enhance ER α transcriptional activity. Unlike for ER α ⁻ tumors^{20,22,37}, the presence of ZEB1 in ER α ⁺ tumors may improve disease outcome, perhaps because enhanced ER α activity favors the maintenance of epithelial features and the response to antiestrogens such as tamoxifen.

Our findings support the notion that the pioneer factor AP2 γ promotes the formation of a ZEB1-ER α TF complex. Although AP2 γ maintains the mammary epithelial state in ER α ⁺ breast cancers^{68,69}, it can stimulate both EMT and MET by inducing open chromatin states⁷⁰⁻⁷². Moreover, we discovered that ZEB1 significantly enhances the ER α recruitment to the AP2 γ gene *TFAP2C*, resulting in the upregulation of AP2 γ . This suggests a positive feedforward loop where increased AP2 γ levels further stimulate the transcriptional activity of the ZEB1–ER α complex. Interestingly, in addition to the indirect binding of ZEB1 to ERBSs through factors such as AP2 γ (Fig. 8), we found ZEB1 motifs in a small subset of ERBSs, suggesting that ZEB1 might directly bind to ERBSs of specific genes during early EMT states. We uncovered

several other motifs, which are associated with ZEB1–ERBSs; this suggests that yet other TFs could contribute to shaping ZEB1-ER α TF complexes at early EMT states.

Clinical data suggest that cancer cells may start spreading very early in the course of tumorigenesis⁷³. Besides, a complete EMT is an extremely rare event in human carcinomas⁹. Tumor cells evolve to different hybrid EMT states, each characterized by different metastatic capabilities^{16,74}. Strikingly, the circulating tumor cells from breast tumors at different stages show the characteristics of hybrid EMT states and predominantly retain E-cadherin expression⁷⁵⁻⁷⁷. It is thus highly possible that each of these states is heterogeneous, expresses a unique set of markers and different levels of ER α . The ZEB1–ER α cooperative interaction stimulates the transcription of some target genes, which are normally expressed during EMT of embryonic development. We also discovered that the induction of ZEB1 expression leads to the coexistence of subpopulations of epithelial, mesenchymal and hybrid cell states expressing factors associated with cancer cell migration and invasion. Several of these factors are expressed in both epithelial EpCAM^{high} and mesenchymal-like EpCAM^{low} cells. For example, annexin A2 (gene *ANXA2*) plays a key role in EMT and metastasis to bone⁷⁸; keratin 8 (*KRT8*) is a keratin expressed in hybrid EMT states⁷; the small heat shock protein Hsp27 (*HSPB1*) is a regulator of EMT and determining factor in breast cancer stem cells⁷⁹; midkine (*MDK*) is involved in regulating the estrogen-mediated EMT at the early stages of breast tumorigenesis⁸⁰. Some of these markers could be gatekeepers of early/hybrid EMT states. It remains to be investigated whether they are involved in switching tumor-initiating cells between different EMT states, and whether they could be validated as therapeutic targets.

We discovered and validated the tetraspanin CD151 in the EpCAM^{low} cell population as a potential target of early EMT stages. In a large cohort of breast cancer patients, elevated CD151 levels were significantly correlated with ER α levels, tumor stage, metastatic potential, and patient survival⁸¹. We observed that breast cancer cell growth is enhanced by short-term ZEB1 overexpression. Intriguingly, these breast cancer cells become more sensitive to EGFR inhibitors, which is also observed upon separate or combined depletion of CD151. We predict that screening epithelial tumors with partial EMT, and potentially moderate levels of ZEB1 and low levels of CD151 could help detecting patients with higher sensitivity to EGFR inhibitors. While the impact of low levels of CD151 is similar in other types of cancer, these cancers have been reported to be more sensitive to EGFR inhibitors when ZEB1 levels are low^{62,82,83}.

Functional annotation of differentially-bound ERBSs unlocked by ZEB1 revealed functions associated with EMT, invasion, activation of WNT pathway, and notably bone morphogenesis. This is in agreement with the established action of ZEB1 during osteoblast differentiation and skeletal morphogenesis²¹. *In vitro* studies suggest that ZEB1 not only initiates invasion, but that conditioned medium from these invasive ER α ⁻ breast cancer cells also promotes the maturation of osteoclasts while repressing osteoblast differentiation⁸⁴. In breast cancer patients, the regulation of the BMP pathway by ZEB1 is predicted to correlate with the incidence of metastases in bones, but not in brain or lungs⁶³. We demonstrate with *ex vivo* invasion assays and with xenograft experiments in mice that ZEB1 may enhance the invasive and metastatic capacity of cells, and modify the organ tropism of

disseminating ER α ⁺ breast tumor cells towards bone tissue. However, the clinical relevance of the ZEB1-ER α target genes in different subtypes of invasive and non-metastatic breast tumors for growth, the formation of distant metastasis, and therapeutic resistance need to be further investigated.

In conclusion, the present work highlights a mechanism by which ER α signaling is pushed towards activating novel targets, which shape a phenotype specific to early stages of EMT and metastasis in breast cancer. The exact components of the TF complexes during the course of EMT and the factors responsible for the gradual loss of ER α remain elusive. A small-molecule screening of cells at early stages of EMT could provide a gateway to developing therapeutic agents targeting the metastatic dissemination at the very early stages.

Methods

Antibodies and other reagents. The anti-ER α rabbit polyclonal antiserum (C1355) for ChIP (5 μ g per IP) was from Millipore (Billerica, MA). Rabbit polyclonal antisera against ER α (A300-498A) for immunoblots and co-IPs (1:1000 for immunoblots; 1 μ g/mg of protein extract), against ZEB1 (A301-921A) for ChIP (10 μ g per IP), 1:250 for immunoblots, and co-IPs (1 μ g/mg of protein extract), and against vimentin (A301-620A; 1:500 for immunoblots) were from Bethyl Laboratories. Mouse monoclonal antibody against N-cadherin (13A9) (1:1000 for immunoblots) was from Cell Signaling Technology (Beverly, USA). Mouse monoclonal anti-GAPDH (6C5, ab8245; 1:30,000 for immunoblots) and goat polyclonal antiserum against FOXA1 (1:1000 for immunoblots and 1 μ g/mg of proteins for co-IPs) were from Abcam.

Mouse monoclonals against AP2 γ (6E4/4) and CD151 (H-8) were from Santa Cruz Biotechnology (Santa Cruz, CA, USA) (1:500 for immunoblots). Mouse monoclonal anti-E-cadherin (C36) (1:8000 for immunoblots) and the BV421 mouse anti-CD326 (EpCAM) and IgG1 k isotype control (used for FACS at 1 μ g per 3 x 10⁵ cells) was purchased from BD Transduction Laboratories. Small interfering RNAs (siRNAs) specific for *TFAP2C* (EHU019581) and *FOXA1* (EHU155811), universal negative control siRNA (SIC007), and the X-tremeGENE siRNA Transfection Reagent were obtained from Sigma-Aldrich (St Louis, USA).

Doxycycline hyclate (Sigma-Aldrich) was used at a concentration of 2 μ g/ml for all tissue culture experiments. 17 β -estradiol (E2), progesterone (P2), and ICI 182780 were from Sigma-Aldrich. Forskolin, 3-isobutyl-1-methylxanthine (IBMX) and recombinant human TGF β 1 derived from HEK293T cells were from PeproTech (London, UK). Collagen I from rat tail was from Enzo Life Science, UK. Dynabeads-Protein G (10009D, Thermo Scientific) were used for the ChIP experiments (100 μ l) and co-IPs (50 μ l). Protease inhibitor cocktail (A32965, Thermo Scientific) was used for preparing all cell lysates. Purified rabbit IgG was used as a reference antibody in co-IPs (Sigma-Aldrich, St Louis, USA). HRP-conjugated anti-mouse and anti-rabbit secondary antibodies for immunoblotting were from Agilent Dako (1:8000). The EGFR inhibitors gefitinib and erlotinib, and 4-hydroxytamoxifen (4-OHT) were from Sigma-Aldrich. AlamarBlue Cell Viability Reagent was obtained from Invitrogen (Thermo Scientific). PEI MAX 40K (Polysciences) and jetOPTIMUS (Polyplus Transfection) were used for transient transfections. Puromycin was from Cayman Chemical.

Plasmids. For ZEB1 expression, the lentiviral doxycycline-inducible construct pTRIPz-puro-HA-ZEB1 was used (a gift from Alain Puisieux's laboratory). Plasmid pTRIPz-puro-HA-ZsGreen was used as a control for pTRIPz-puro-HA-ZEB1. To construct the former plasmid, ZsGreen coding sequences were inserted in place of those for HA-ZEB1 in plasmid pTRIPz-puro-HA-ZEB1. Plasmid pBABE-puro-mTWIST⁸⁵ was from Addgene (ID #1783), and pCMV6-PRRX1 (#RC213276) was purchased from Origene. The plasmid pHAGE-fullEF1a-lZsGreen (plasmid ID 233 from the DNA Resource Core at the Harvard Medical School, Boston) was used to label cells with constitutive ZsGreen expression. We used plasmid HEG0 to express the full-length human ER α ⁸⁶ and plasmid pSG5-hPR for the expression of human PR⁸⁷.

The following luciferase reporters were used: GK1⁸⁸ for Gal4 fusion proteins, EREtkLuc (XETL) for ER α , PRE-TATA-Luc (a gift from D. McDonnell) for PR, pGL4.10-VEGFprom-Luc (-1000 to -1) for VEGFA (Addgene #66128)⁸⁹, proE-cad670-Luc for E-cadherin (Addgene #42083)⁹⁰, SBE4-Luc for SMAD (a gift from Bert Vogelstein; Addgene #16495)⁹¹, and the renilla luciferase transfection control reporter pRL-CMV from Promega (E2261).

For knockdowns of *TFAP2C*, *FOXA1*, *MUC16*, *DIO2*, *ESR2*, *P2RX7*, *MMP10*, *SCG2*, *MUC2*, *DSCAM*, and *CD151* the shRNA constructs were generated using the pLKO.1 vector (Open Biosystems) and the target sequences listed in Supplementary Table 1 To produce lentiviruses, the plasmids pMD2G and psPAX2 were used (gifts from Didier Trono's laboratory). The lentiviral UBC-GFP-T2A-Luciferase dual

reporter for *in vivo* imaging and the pMDLg and pRSV-Rev packaging plasmids were from BioCat GmbH.

Cell culture. The human breast carcinoma cell lines MCF7 and its variant MCF7-V (see Supplementary Methods for more details), and human embryonic kidney HEK293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) complemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. Human T-47D ductal carcinoma cells were cultured in RPMI-1640 medium supplemented with 0.2 units/ml bovine insulin, FBS to a final concentration of 10% and 1% penicillin/streptomycin. To deprive the cells from steroids, they were cultured for at least five days in DMEM without phenol red complemented with 5% charcoal-stripped FBS, 2 mM L-glutamine and 1% penicillin/streptomycin. Cells were split with 0.05% (w/v) trypsin in phosphate-buffered saline (PBS), containing 0.02% (w/v) EDTA at least 2x per week. All cells were maintained in 5% CO₂ in a humidified incubator at 37°C. Cells were regularly checked for mycoplasma contamination.

Virus production and transduction. HEK293T cells were seeded to a density of 3×10^7 in a 150 mm dish in standard medium 24 hours (h) before transfection. Lentiviral constructs were co-transfected with plasmids pMD2G and psPAX2. All transfections were performed using the calcium phosphate transfection method. 16 h later, the medium was replaced by fresh one and lentivirus/retrovirus-containing supernatants were collected every 24 h during the next 3 days. Supernatants were filtered and mixed with a 40% sterile polyethylene glycol 8000 (PEG 8000; Sigma-Aldrich) solution by rotating at 4°C for at least 2 h. The mixes were then centrifuged at 4000 x g at 4°C for 30 minutes to pellet the viral particles. Each pellet was then

gently dissolved in 1 ml of medium to yield concentrated viral stocks. Cells were infected with the concentrated viruses. 24-48 h later, infected cells were selected with puromycin (2 µg/ml for MCF7 and MCF7-V and 3 µg/ml for T-47D cells) for 24 h. Doxycycline-inducible cells were maintained in 2 µg/ml of DOX for the specified duration mentioned in the text.

Luciferase reporter assays. 24 h prior to transfections, 6×10^4 cells for HEK293T and 4×10^4 cells for MCF7, MCF7-V or T-47D cells were seeded in complete or steroid-deprived medium in each well of a 24-well plate. Cells were transfected with the indicated plasmids and corresponding firefly luciferase reporters and pRL-CMV for renilla luciferase expression using PEI MAX for HEK293T and jetOPTIMUS for other cells. After 8 h, medium was changed and 24 h after transfection specific treatments were added for 18-24 h including vehicle, E2 (10 nM), FI (10 µM forskolin + 100 µM IBMX), P4 (10 nM), and TGFβ1 (10 ng/ml). Luciferase activity was measured with the dual-luciferase reporter assay (Promega). The firefly and renilla luciferase activities were measured with a bioluminescence plate reader. Normalization to the renilla luciferase internal control was performed to quantify the activity.

Cell cycle assay. One day before the treatments cells were seeded at a density of 3×10^5 per well of 6-well plates in complete medium. The next day, cells were treated with different concentrations of 4-OHT (Sigma-Aldrich) for 72 h. To perform the assay, cells were harvested with trypsin-EDTA and washed with PBS. Cells were fixed in cold 80% ethanol by adding it drop wise to the pellet while vortexing at low speed, followed by incubation on ice for 30 minutes. After the centrifugation for 10

minutes at 700 x rpm, the pellets were washed 2x with cold PBS. Cells were then treated with 100 μ l of RNase (100 μ g/ml) for 10 minutes. 300 μ l of propidium iodide (50 μ g/ml stock) was added to cells to stain the DNA. A Gallios Flow Cytometer (Beckman Coulter) was used to measure the forward scatter and side scatter. Using the FlowJo software, cell debris were gated out, doublets were excluded and the PI histogram plot was applied.

Protein extraction, co-immunoprecipitations and immunoblots. Cells were washed with tris-buffered saline (TBS), detached with trypsin-EDTA, and harvested by adding complete medium and centrifuging at 1000 x rpm for 5 minutes. The pellets were washed with PBS once and then lysed in ice-cold lysis buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 10% glycerol, 10 mM Na-molybdate, and 1x protease inhibitor cocktail). Cell suspensions were sonicated for 20 cycles of 20 seconds at high power with the Bioruptor sonicator (Diagenode). After centrifugation at the maximum speed for 5 minutes, supernatants were collected and protein amounts measured with the Bradford assay. For immunoprecipitations, 2 mg of protein extracts were mixed with a specific antibody or control IgG of the same species and incubated overnight at 4°C on a rotating wheel. On the next day, 50 μ l of washed Dynabeads-Protein G were added and incubated for 2 h at 4°C. Using a magnetic stand, beads were harvested and washed 5x with the lysis buffer supplemented with 0.1% Triton X-100. After the last wash, proteins were eluted from the beads with NuPAGE LDS Sample Buffer (Thermo Scientific) and 10 mM DTT in boiling water for 5 minutes. To obtain cell extracts for immunoblotting without immunoprecipitation, cell pellets were lysed with lysis buffer supplemented with 0.1% Triton X-100 and protein extracts were mixed with the sample buffer complemented

with 10 mM DTT and heated in boiling water for 5 minutes. Immunoprecipitates and input protein extracts were loaded and separated by SDS-PAGE and transferred to a nitrocellulose membrane. After blocking the membranes with 5% fat-free milk powder in TBS with 0.2% Tween-20 (TBS-T) for 20 minutes, specific primary antibodies were added and incubated overnight at 4°C. Membranes were washed 3x with TBS-T and incubated with a secondary antibody coupled to horse radish peroxidase (Agilent Dako) for 1 h at room temperature. After several washes of the membranes with TBS-T, protein bands were developed and visualized with an ECL kit (Enhanced ChemiLuminescence, Advansta).

Chromatin immunoprecipitation (ChIP). ChIP experiments were performed using a previously described protocol⁹². A minimum of 2×10^7 cells were plated in 150 mm dishes in specified medium. 100 µl of magnetic beads were washed with 1 ml of blocking solution (0.5% BSA (w/v) in PBS) and incubated with specific antibodies rotating overnight at 4°C. For ERα ChIPs, cells were treated with either vehicle, E2, or FI for 90 minutes. DNA-protein complexes were then crosslinked with 1% formaldehyde for 10 minutes with gentle swirling. The formaldehyde was quenched by adding 125 mM L-glycine. Cells were rinsed 3x with ice-cold PBS and harvested in lysis buffer 1 (50 mM Hepes-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 or IGEPAL CA-630, 0.25% Triton X-100) and rocked at 4°C for 10 minutes and pelleted at $2000 \times g$ for 5 minutes at 4°C. The pellets were resuspended in lysis buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) and rocked gently at 4°C for 5 minutes. Nuclei were pelleted by spinning at $2000 \times g$ for 5 minutes at 4°C. Each pellet was resuspended in lysis buffer 3 (10 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1%

Na-deoxycholate, 0.5% N-lauroylsarcosine, and 1x protease inhibitor cocktail) and subjected to sonication with 30 cycles each 30 seconds at high power. After adding 0.1% Triton X-100, cell debris were discarded by centrifugation at 16,000 x g for 5 minutes to collect the nuclear extracts.

We used micrococcal nuclease digestion for the preparation of samples for the high-resolution MNase ChIP-seq with wild-type MCF7-V cells using a previously published protocol⁹³. After the step with the lysis buffer 1, nuclear pellets were resuspended in NPS buffer (0.5 mM spermidine, 0.075% IGEPAL CA-630, 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 5 mM MgCl₂ and 1 mM CaCl₂) and centrifuged at 2,000 x g for 5 minutes. Pellets were resuspended in 1 ml NPS buffer, 3 units MNase for 4 mg of nuclear protein were added, and samples were incubated at 37°C for 10 minutes. The reaction was stopped by adding 15 mM EDTA and 3 mM EGTA and the buffer L was added to samples (50 mM Hepes-KOH, pH 8.0, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100 and 0.1% sodium deoxycholate; supplemented with protease inhibitors). Samples were centrifuged at 16,000 x g for 10 minutes and nuclear extracts were collected.

For both types of ChIP experiments, 10 µl of cell lysate was saved as input DNA and stored at -20°C. The antibody/magnetic bead mix was added to the extracts for each IP and incubated overnight at 4°C on a rocker. Dynabeads were collected with a magnetic stand and washed 10 x with RIPA buffer (50 mM Hepes-KOH pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40 or IGEPAL CA-630, 0.7% Na-deoxycholate) with the last wash being done with TBS. Inputs and IPs were reverse-crosslinked with the elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA and 1% SDS) at 65°C overnight

in a shaker at 700 x rpm. Dynabeads were discarded and elutions were diluted in TE buffer and incubated with 25 μ g/ml RNase for 1 h at 37°C followed by an incubation with 200 μ g/ml proteinase K for 2 h at 57°C. DNA was isolated by extraction with phenol–chloroform–isoamyl alcohol (25:24:1), and the phases were separated with 2 ml Phase Lock Gel Light tubes (5 PRIME) with centrifugation at 10,000 x g for 10 minutes. Aqueous layers were collected in new tubes with 10 μ g glycogen (VWR) and 200 mM NaCl. 100% ethanol was added and samples were incubated for 1 h at -80°C and then centrifuged at 16,000 x g for 30 minutes at 4°C. Pellets were washed once with 80% ethanol and dried at room temperature. Pellets were resuspended in nuclease-free water. qPCR was performed with the primers in Supplementary Table 2 to measure the fold enrichment over input.

Re-ChIP. For the Re-ChIP experiments, an ER α ChIP was first performed as described in the previous section, with the exception that after the last wash with RIPA buffer, 25 μ l of 10 mM DTT was used to elute the bound chromatin from the beads by incubation at 37°C with shaking for 30 minutes⁹⁴. The supernatant was removed and diluted at least 20 x with the re-ChIP dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.1). The second ChIP was performed with the anti-ZEB1 antibody or control IgG followed by the standard ChIP procedure. All re-ChIP values are relative to the IgG control.

ChIP-seq and bioinformatic analysis. For the ChIP DNA library preparation of wild-type MCF7-V and MCF7-ZEB1 cells, the Illumina TruSeq protocol was applied for each replicate and DNA was sequenced using a HiSeq 4000 machine to produce 100 bp paired-end reads. For the ChIP-seq of MCF7-V cells sequences were aligned

to the Human Reference Genome (assembly hg19, NCBI build 37, February 2009) with BWA-MEM (Version 0.7.17)⁹⁵. Peak calling was carried out by using the MACS2 tool (Version 2.1.0) of the Galaxy tool suite (<https://usegalaxy.org>)⁹⁶. We predicted the fragment sizes generated during the fragmentation step of the library preparation from the alignment results and called the peaks with the input file as control. Only the statistically significant binding sites were kept by R (version 3.6.2) based on the confidence level ($-10 \times \log_{10} P$ value) of the peak center. Motif analysis was performed with the SeqPos tool using the JASPAR motif matrix. The generation of aggregation plots was done using the cistrome platform (<http://cistrome.org/ap/root>)⁹⁷. Venn diagrams were produced with the VennDiagram package in R.

For the ChIP-seq of MCF7-V-ZEB1 cells, we used four biologically independent replicates for each of the treatment groups. We used the ER α ChIP-seq data from our previously published dataset in MCF7-V cells ([GSE109103](#))³⁵ to perform a differential binding analysis³⁵. FASTQ reads were aligned to the human genome hg19 using BWA-MEM (Version 0.7.17) with standard settings. The quality of the ChIP-seq data was assessed as describe in the encode project (<https://www.encodeproject.org/data-standards/terms/>). MACS2 (Version 2.1.0) with default parameters was used to call peaks on each replicate. These peaks were then used to build a reference dataset of binding regions. For each treatment (E2, FI), peaks were added in the reference dataset when they were found in at least two replicates. To define standard binding regions each peak summit was extended 50 nucleotides on both sides. This procedure identified 40,720 reference ERBSs for E2 and 26,739 for FI. FeatureCounts (version 2.0.0) was then used to count reads per

binding regions and generated a count table for E2 and FI treatments. The count tables were then analyzed in R with the edgeR package. Binding regions with very low number of reads were filtered out (mean of all replicates CPM < 5). The count tables were then normalized, the common dispersion and the tagwise dispersion were estimated with the estimateDisp function. After fitting to a binomial model, the differentially binding sites that were statistically significant were identified with the exact test. For each of the differentially binding regions, sequences of 100 bp surrounding the peak summit were retrieved with the samtools program (version 1.10-3) from the hg19 genome. Identification of specific binding motifs was done with FIMO (version 5.0.5)⁹⁸ by using the HOCOMOCO v10 collection of transcription factor binding models for human⁹⁹.

The following publicly available ChIP-seq GEO datasets were used for comparisons: GSE109103 (ER α)³⁵, GSE89206 (ZEB1)⁴⁷, GSE21234 (TFAP2C)⁴⁸, GSE25315 (FOXA1)¹⁰⁰, and GSE60270 (GATA3, P300, H3K27ac, H3K4me1, and H3K9me3)²⁹. Sequence Read Archive (SRA) files from each dataset were transferred from the NCBI server and after converting them into FASTQ files, the same procedure as for ChIP-seq analyses mentioned above was followed. The Integrative Genomics Viewer (IGV version 2.8.0) was used to browse and illustrate the binding sites.

RNA extraction, reverse-transcription and quantitative PCR. Cells were seeded in steroid-deprived medium for at least five days prior to treatments. RNA was extracted using the guanidium-acid-phenol method from 5 x 10⁵ cells per well of 6-well plates. Briefly, cells were lysed with the TRI reagent (4 M guanidium thiocyanate, 25 mM sodium citrate, 0.5% N-lauroylsarcosine, 0.1 M 2-

mercaptoethanol, pH 7). 2 M Na-acetate pH 4, aquaphenol and chloroform:isoamyl alcohol (49:1) were added to the cell lysates and mixed vigorously. After centrifugation at $10,000 \times g$ for 20 minutes at 4°C , the top phases were collected and RNA was precipitated by adding absolute ethanol and centrifugation at $16000 \times g$ for 30 minutes at 4°C . RNA pellets were washed twice with 70% ethanol and the pellets were dried at room temperature and resuspended in nuclease-free water.

To prepare the samples for RNA-sequencing, the RNeasy Mini Kit and columns (QIAGEN) were used to extract and purify high-quality RNA. A NanoDrop (Thermo Scientific) was used to measure RNA concentrations. Total RNA was reverse transcribed using random primers (Promega) and the GoScript Reverse Transcription System according to the manufacturer's instructions (Promega). qPCR analyses were conducted in 10 μl reaction mixtures including the GoTaq master mix (Promega), cDNA and specific primer pairs (Supplementary Table 3) with a Biorad CFX96 thermocycler. RNA levels were standardized with the *GAPDH* mRNA as the internal standard control and relative gene expression levels were calculated by the $\Delta\Delta\text{Ct}$ method.

RNA-seq. Before sequencing, the quality of RNA was evaluated with a Qubit 4 Fluorometer (Thermo Scientific). We obtained two biologically independent replicates for each treatment group. The samples were sequenced on a HiSeq 4000 (Illumina). Approximately 50 million paired-end reads of 100 bp per sample were obtained and the quality of reads was checked with the fastqc tool. Sequences were mapped against the Human Reference Genome (assembly hg38, UCSC, August 2015) with the STAR (Version 2.7.0) and count tables were produced with the featureCounts

function (Version 2.0.0) in R, containing the number of mapped reads per gene (Supplementary Data 3-5). Differential Expression (DE analysis) was performed and fold changes were calculated with the edgeR Bioconductor package in R¹⁰¹. Genes with a fold change of expression ≥ 1.5 were considered upregulated and those with a fold change of expression ≤ 1.5 were considered downregulated. Only the genes that had a FDR < 0.05 were included in the analyses. All volcano plots, Venn diagrams, heat maps and GO terms were generated in R (version 3.6.2). For the gene set enrichment analysis (GSEA)¹⁰², normalized fold changes from each gene were used from each replicate to generate a table, which was converted to the gct format and used with the GSEA (v4.1.0) software along with the gene ontology gene sets from ontology gene sets of the MSigDB collection, all from the Broad Institute. GSEA tables were then visualized and analyzed with Enrichment Map from Cytoscape version 3.8.2¹⁰³ to cluster and annotate the GO interactome.

FACS and scRNA-seq. For sorting, cells were harvested with trypsin-EDTA and washed in the FACS solution composed of 1x PBS, 2% FBS and 1mM EDTA. 6×10^6 cells were stained with the BV421-conjugated anti-human CD326 (EpCAM) or the IgG1/k isotype control antibodies for 30 minutes at 4°C protected from light. Cells were washed twice with the FACS solution and resuspended in 1 ml of the solution for sorting. The BD FACS Aria III Cell Sorter (BD Biosciences) was used to select living single cells based on the forward and side scatter to separate the doublets and DRAQ7 dye exclusion to exclude dead cells. To perform scRNA-seq of FACS-isolated EpCAM^{high} and EpCAM^{low} cells, a minimum of >6,000 cells for each group was used. The Chromium Next GEM Single Cell 3' v3.1 workflow was followed for

the library preparation and sequencing was performed with the Chromium Controller system from 10x Genomics.

Bioinformatic analysis of scRNA-seq data. Cell Ranger 4.0

(<http://10xgenomics.com>) was used to process Chromium single cell 3' RNA-seq output and generate the count table. FASTQ reads were aligned to the reference genome GRCh38-2020-A downloaded from 10X genomics website. All analyses were then carried out in R with the Seurat 3.0 package¹⁰⁴. QC metrics per cells were calculated with the function PercentageFeatureSet. Only cells with less than 25% mitochondrial RNA and feature counts between 200 and 7000 were kept. EpCAM^{high} and EpCAM^{low} cells were merged with the function FindIntegrationAnchors, which takes a list of Seurat objects as input and uses these anchors to integrate the two datasets together with IntegrateData. Data were normalized and scaled with the Sctransform function. The percentage of mitochondrial RNA was used as a variable to regress out in a second non-regularized linear regression. Dimensional reduction was performed by PCA and UMAP and selection of markers specific to clusters of interest were done with the FindConservedMarkers function.

Three-dimensional invasion assays. ZsGreen-labeled cells were seeded in 200 μ l of steroid-deprived medium at a density of 1,000 cells per each well of 96-well plates coated with 1.5% of agarose, and incubated for 3 days to allow the formation of compact tumor spheroids⁴⁵. For the assay, the plates were placed on ice and 150 μ l of medium was gently removed. Using ice-cold tips, 50 μ l of neutralized collagen I (3 mg/ml, at 2x the final concentration) was dispensed to the bottom of each well with 6 replicates/condition. The plates were centrifuged at 300 x g for 3 minutes to ensure

that the single tumor spheroid of each well is located in the center. After 30 minutes of incubation at 37°C to allow the collagen I to solidify, 100 µl per well of medium with 2x concentrations of E2, FI, ICI, and 4-OHT was added. Images were acquired using an automated ImageXpress Micro XL confocal microscope (Molecular Devices) every day for 96 h. Images were analyzed with the MetaXpress high-content image acquisition and analysis software.

Cell viability assay. To measure cell viability in the context of erlotinib and gefitinib treatments, cells were seeded at a density of 3×10^3 cells per well of 96-well plates ($n = 8$) overnight and treated with 10 µM erlotinib or 20 µM gefitinib for 72 h. Cell viability was determined using AlamarBlue (Invitrogen) according to manufacturer's protocol. Values were measured using a fluorescence-based plate reader with excitation at 560 nm and emission at 590 nm, which was normalized to the control treated with the corresponding concentration of vehicle.

Bone invasion assay and multi-photon microscopy. For the following experiments, we used control MCF7-V (Ctl) and MCF7-V-ZEB1 cells transformed with the plasmid pHAGE-CMV-fullIEF1a-IRES-ZsGreen for constitutive expression of ZsGreen. The femur bones were removed from sacrificed mice and cleaned in sterile saline solution. The upper part of the femoral bones was sliced into ~5 mm pieces. 20 µl collagen I was pipetted into center of each well of 24-well transwell plates to stabilize the bone. Bone pieces were placed directly into the collagen and 1 ml of DMEM supplemented with 10% FBS, 2% penicillin/streptomycin and 1% fungizone was gently added on top of the bones. Bones were incubated for 24 h in the incubator to normalize. Medium was carefully replaced with fresh medium containing

2% FBS. 2×10^5 ZsGreen-labeled control (Ctl) or MCF7-V-ZEB1 cells were seeded into transwell inserts with a $8 \mu\text{m}$ pore size (Corning) and coated with 20% Matrigel (Invitrogen), in medium containing $2 \mu\text{g/ml}$ DOX and 2% FBS. The inserts were incubated for 30 minutes to allow the cells to settle at the bottom and were placed on top of the bones. The negative control contained decellularized bone (bone boiled for 10 minutes). After one week, chambers were rinsed with PBS and cells inside the transwell inserts were removed using cotton swabs, and the cells on the lower surface of the membrane were fixed in 1% PFA for 10 minutes and stained with 0.5% crystal violet. After washing the stained inserts twice with PBS, the invaded cells were imaged under a light microscope.

To assess the invasion of cells into the bone, pieces of bone were washed 2x in PBS and fixed in 4% PFA for a minimum of 3 days at 4°C ; PFA was changed daily. Bones were washed and incubated in 2 ml of 0.05% Alizarin Red solution containing 1% KOH and 1 mM HEPES pH = 7.5 gently shaking for 20 minutes. Pieces of bone were fixed on 35 mm glass bottom dishes being soaked in $100 \mu\text{l}$ of 1% low-melting agarose. A SP8 DIVE upright multiphoton confocal microscope (Leica Microsystems) was used to image the invasion of GFP-labeled cells into bone surface area from 0- $200 \mu\text{m}$ in depth. The Imaris 9.6 software was used to build the 2D and 3D images.

Xenograft experiments. Lentiviral particles for labelling wild-type MCF7 cells were generated by co-transfection of the UBC-GFP-T2A-Luciferase dual reporter and the packaging plasmids pMDLg, pMD2G and pRSV-Rev. ZsGreen as negative control or ZEB1 were stably expressed in labeled cells with the DOX-inducible lentiviral vectors. Cells were cultured with $2 \mu\text{g/ml}$ of doxycycline for 2 passages before

injections. NOD/scid GAMMA (NSG) mice, aged 10-12 weeks, were injected with 1×10^6 cells orthotopically into the mammary gland and oestrogen pellet (E2-M -17 β -estradiol, 60 days, Belma technologies) was implanted in the back. Mice were provided with fresh DOX-containing water (2 mg/ml DOX and 5% sucrose) protected from light. The experiment was terminated after 10 weeks. All mouse samples were collected over 3 days at the same time during the day for each comparative group. All xenograft experiments were carried out in compliance with institutional and cantonal guidelines (approved mouse protocol #2781, cantonal veterinary office of Basel-City). NSG mice were purchased from Jackson Laboratory and kept in pathogen-free conditions specified by the University of Basel and cantonal veterinary office of Basel-City.

For the bioluminescence (BLI) imaging and quantification mice bearing cells with GFP/Luc were injected intraperitoneally with 3 mg of D-Firefly-Luciferin (Gold Bio, LUCK- 5G). After 8 minutes, bioluminescent images of the full mouse were taken at IVIS Lumina LT (Perkin Elmer). After euthanasia primary tumor and metastatic organs were imaged separately. The same exposure of 1 minute was set for all organs, except for primary tumors and whole-mouse imaging where both 1 minute and auto-exposure settings were applied. Bioluminescence signal analysis was carried out with Living Image, and average BLI count was computed. Metastatic index was calculated as the ratio of the total BLI counts of the metastatic organ over those of the primary tumor.

Statistical analyses. All experiments were performed at least in three replicate experiments, unless mentioned otherwise. All statistical analyses were performed

with GraphPad Prism (version 8.3.0) or R. The mean comparison tests were calculated using the unpaired and two-tailed Student's *t*-test or the two-way analysis of variance. P values < 0.05 were considered statistically significant. All error bars represent standard error of the mean.

Data availability

All data supporting the findings of this study are available from the corresponding authors upon reasonable request. The ChIP-seq, RNA-seq, and scRNA-seq data are available through the GEO accession [GSE173562](#) (using the editor/reviewer token: qngtycyohxyjtsl). Source data of uncropped immunoblot images and the individual data points are provided with this paper.

References

1. Lin, C.-Y. et al. Discovery of estrogen receptor α target genes and response elements in breast tumor cells. *Genome Biol.* **5**, R66 (2004).
2. Siersbaek, R., Kumar, S. & Carroll, J. S. Signaling pathways and steroid receptors modulating estrogen receptor α function in breast cancer. *Genes Dev.* **32**, 1141-1154 (2018).
3. Musgrove, E. A. & Sutherland, R. L. Biological determinants of endocrine resistance in breast cancer. *Nat. Rev. Cancer* **9**, 631-643 (2009).
4. Farcas, A. M., Nagarajan, S., Cosulich, S. & Carroll, J. S. Genome-wide estrogen receptor activity in breast cancer. *Endocrinology* **162**, bqaa224 (2021).

5. Gupta, P. B., Pastushenko, I., Skibinski, A., Blanpain, C. & Kuperwasser, C. Phenotypic plasticity: Driver of cancer initiation, progression, and therapy resistance. *Cell Stem Cell* **24**, 65-78 (2019).
6. Yang, J. et al. Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell. Biol.* (2020).
7. Kroger, C. et al. Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proc. Natl. Acad. Sci. USA* **116**, 7353-7362 (2019).
8. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell. Biol.* **20**, 69-84 (2019).
9. Shibue, T. & Weinberg, R. A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611-629 (2017).
10. Aiello, N. M. et al. EMT subtype influences epithelial plasticity and mode of cell migration. *Dev. Cell* **45**, 681-695 (2018).
11. Park, S. H., Cheung, L. W., Wong, A. S. & Leung, P. C. Estrogen regulates Snail and Slug in the down-regulation of E-cadherin and induces metastatic potential of ovarian cancer cells through estrogen receptor α . *Mol. Endocrinol.* **22**, 2085-2098 (2008).
12. Puisieux, A., Brabletz, T. & Caramel, J. Oncogenic roles of EMT-inducing transcription factors. *Nat. Cell Biol.* **16**, 488-494 (2014).
13. Krebs, A. M. et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat. Cell Biol.* **19**, 518-529 (2017).

14. Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256-260 (2015).
15. Pastushenko, I. & Blanpain, C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* **29**, 212-226 (2019).
16. Pastushenko, I. et al. Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463-468 (2018).
17. Jackson, H. W. et al. The single-cell pathology landscape of breast cancer. *Nature* **578**, 615-620 (2020).
18. Caramel, J., Ligier, M. & Puisieux, A. Pleiotropic roles for ZEB1 in cancer. *Cancer Res.* **78**, 30-35 (2018).
19. Hanrahan, K. et al. The role of epithelial-mesenchymal transition drivers ZEB1 and ZEB2 in mediating docetaxel-resistant prostate cancer. *Mol. Oncol.* **11**, 251-265 (2017).
20. Lehmann, W. et al. ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types. *Nat. Commun.* **7**, 10498 (2016).
21. Postigo, A. A. Opposing functions of ZEB proteins in the regulation of the TGF β /BMP signaling pathway. *EMBO J.* **22**, 2443-2452 (2003).
22. Wu, H.-T. et al. Oncogenic functions of the EMT-related transcription factor ZEB1 in breast cancer. *J. Transl. Med.* **18**, 51 (2020).
23. Karihtala, P. et al. Vimentin, zeb1 and Sip1 are up-regulated in triple-negative and basal-like breast cancers: association with an aggressive tumour phenotype. *Breast Cancer Res. Treat.* **138**, 81-90 (2013).
24. Liu, B.-W. et al. Estrogen receptor- α -miR-1271-SNAI2 feedback loop regulates transforming growth factor- β -induced breast cancer progression. *J. Exp. Clin. Cancer Res.* **38**, 109 (2019).

25. Al Saleh, S., Al Mulla, F. & Luqmani, Y. A. Estrogen Receptor silencing induces epithelial to mesenchymal transition in human breast cancer cells. *PLoS ONE* **6**, e20610 (2011).
26. Bouris, P. et al. Estrogen receptor α mediates epithelial to mesenchymal transition, expression of specific matrix effectors and functional properties of breast cancer cells. *Matrix Biol.* **43**, 42-60 (2015).
27. Di Zazzo, E. et al. Estrogen receptors in epithelial-mesenchymal transition of prostate cancer. *Cancers* **11**, 1418 (2019).
28. Bennesch, M. A. & Picard, D. Tipping the balance: ligand-independent activation of steroid receptors. *Mol. Endocrinol.* **29**, 349-363 (2015).
29. Liu, Z. et al. Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell* **159**, 358-373 (2014).
30. Jozwik, K. M. & Carroll, J. S. Pioneer factors in hormone-dependent cancers. *Nat. Rev. Cancer* **12**, 381-385 (2012).
31. Droog, M., Mensink, M. & Zwart, W. The estrogen receptor α -cistrome beyond breast cancer. *Mol. Endocrinol.* **30**, 1046-1058 (2016).
32. Acconcia, F. & Kumar, R. Signaling regulation of genomic and nongenomic functions of estrogen receptors. *Cancer Lett.* **238**, 1-14 (2006).
33. Bennesch, M. A., Segala, G., Wider, D. & Picard, D. LSD1 engages a corepressor complex for the activation of the estrogen receptor α by estrogen and cAMP. *Nucleic Acids Res.* **44**, 8655-8670 (2016).
34. 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).

35. Berto, M., Jean, V., Zwart, W. & Picard, D. ER α activity depends on interaction and target site corecruitment with phosphorylated CREB1. *Life Sci. Alliance* **1**, e201800055 (2018).
36. Vesuna, F. et al. Twist contributes to hormone resistance in breast cancer by downregulating estrogen receptor- α . *Oncogene* **31**, 3223-3234 (2012).
37. Zhang, J. et al. ZEB1 induces ER- α promoter hypermethylation and confers antiestrogen resistance in breast cancer. *Cell Death Dis.* **8**, e2732 (2017).
38. Brabletz, T. EMT and MET in metastasis: where are the cancer stem cells? *Cancer Cell* **22**, 699-701 (2012).
39. Soini, Y. et al. Transcription factors zeb1, twist and snai1 in breast carcinoma. *BMC Cancer* **11**, 73 (2011).
40. Liu, L. et al. ZEB1 upregulates VEGF expression and stimulates angiogenesis in breast cancer. *PLoS ONE* **11**, e0148774 (2016).
41. Sánchez-Tilló, E. et al. ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. *Oncogene* **29**, 3490-3500 (2010).
42. Wu, L. et al. Smad4 as a transcription corepressor for estrogen receptor α . *J. Biol. Chem.* **278**, 15192-15200 (2003).
43. Han, R. et al. Estrogen promotes progression of hormone-dependent breast cancer through CCL2-CCR2 axis by upregulation of Twist via PI3K/AKT/NF- κ B signaling. *Sci. Rep.* **8**, 9575 (2018).
44. Sanchez, A. M. et al. Estrogen Receptor- α promotes breast cancer cell motility and invasion via focal adhesion kinase and N-WASP. *Mol. Endocrinol.* **24**, 2114-2125 (2010).

45. Vinci, M., Box, C. & eccles, s. Three-dimensional (3D) tumor spheroid invasion assay. *J. Vis. Exp.* **2015**, e52686 (2015).
46. Pattabiraman, D. R. et al. Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, aad3680 (2016).
47. Katsura, A. et al. ZEB1-regulated inflammatory phenotype in breast cancer cells. *Mol. Oncol.* **11**, 1241-1262 (2017).
48. Woodfield, G. W., Chen, Y., Bair, T. B., Domann, F. E. & Weigel, R. J. Identification of primary gene targets of TFAP2C in hormone responsive breast carcinoma cells. *Genes Chromosomes Cancer* **49**, 948-962 (2010).
49. Wu, V. T. et al. A TFAP2C gene signature is predictive of outcome in HER2-positive breast cancer. *Mol. Cancer Res.* **18**, 46-56 (2020).
50. Franke, C. M. et al. TFAP2C regulates carbonic anhydrase XII in human breast cancer. *Oncogene* **39**, 1290-1301 (2020).
51. Jeselsohn, R. et al. TransCONFIRM: Identification of a genetic signature of response to fulvestrant in advanced hormone receptor–positive breast cancer. *Clin. Cancer Res.* **22**, 5755-5764 (2016).
52. Fu, X. et al. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proc. Natl. Acad. Sci. USA* **113**, E6600-E6609 (2016).
53. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**, 284-287 (2012).
54. Li, L. T., Jiang, G., Chen, Q. & Zheng, J. N. Ki67 is a promising molecular target in the diagnosis of cancer. *Mol. Med. Rep.* **11**, 1566-1572 (2015).

55. Kominsky, S. L. et al. Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast. *Oncogene* **22**, 2021-2033 (2003).
56. Rafnar, T. et al. Mutations in BRIP1 confer high risk of ovarian cancer. *Nat. Genet.* **43**, 1104-U1191 (2011).
57. Blackman, B., Russell, T., Nordeen, S. K., Medina, D. & Neville, M. C. Claudin 7 expression and localization in the normal murine mammary gland and murine mammary tumors. *Breast Cancer Res.* **7**, R248-R255 (2005).
58. Palmer, T. D., Ashby, W. J., Lewis, J. D. & Zijlstra, A. Targeting tumor cell motility to prevent metastasis. *Adv. Drug Deliv. Rev.* **63**, 568-581 (2011).
59. Zhao, S. J. et al. CD151 promotes breast cancer metastasis by activating TGF- β 1/Smad signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* **22**, 7314-7322 (2018).
60. Sadej, R., Grudowska, A., Turczyk, L., Kordek, R. & Romanska, H. M. CD151 in cancer progression and metastasis: a complex scenario. *Lab. Invest.* **94**, 41-51 (2014).
61. Yoo, S. H., Lee, K., Chae, J. Y. & Moon, K. C. CD151 expression can predict cancer progression in clear cell renal cell carcinoma. *Histopathology* **58**, 191-197 (2011).
62. Medrano, M. et al. Interrogation of functional cell-surface markers identifies CD151 dependency in high-grade serous ovarian cancer. *Cell Rep.* **18**, 2343-2358 (2017).
63. Mock, K. et al. The EMT-activator ZEB1 induces bone metastasis associated genes including BMP-inhibitors. *Oncotarget* **6**, 14399-14412 (2015).

64. Shafie, S. M. & Liotta, L. A. Formation of metastasis by human-breast carcinoma-cells (MCF-7) in nude-mice. *Cancer Lett.* **11**, 81-87 (1980).
65. Comsa, S., Cimpean, A. M. & Raica, M. The story of MCF-7 breast cancer cell line: 40 years of experience in research. *Anticancer Res.* **35**, 3147-3154 (2015).
66. Ziegler, E., Hansen, M. T., Haase, M., Emons, G. & Grundker, C. Generation of MCF-7 cells with aggressive metastatic potential in vitro and in vivo. *Breast Cancer Res. Treat.* **148**, 269-277 (2014).
67. Jehanno, C. et al. Nuclear translocation of MRTFA in MCF7 breast cancer cells shifts ER α nuclear/genomic to extra-nuclear/non genomic actions. *Mol. Cell. Endocrinol.* **530**, 111282 (2021).
68. Xiao, T. et al. Estrogen-regulated feedback loop limits the efficacy of estrogen receptor–targeted breast cancer therapy. *Proc. Natl. Acad. Sci. USA* **115**, 7869-7878 (2018).
69. Cyr, A. R. et al. TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis. *Oncogene* **34**, 436-444 (2015).
70. Kim, W. et al. TFAP2C-mediated upregulation of TGFBR1 promotes lung tumorigenesis and epithelial-mesenchymal transition. *Exp. Mol. Med.* **48**, e273 (2016).
71. Hardy, K. et al. Identification of chromatin accessibility domains in human breast cancer stem cells. *Nucleus* **7**, 50-67 (2016).
72. Wang, Y. et al. TFAP2C facilitates somatic cell reprogramming by inhibiting c-Myc-dependent apoptosis and promoting mesenchymal-to-epithelial transition. *Cell Death Dis.* **11**, 482 (2020).

73. Husemann, Y. et al. Systemic spread is an early step in breast cancer. *Cancer Cell* **13**, 58-68 (2008).
74. Godin, L. et al. A novel approach for quantifying cancer cells showing hybrid epithelial/mesenchymal states in large series of tissue samples: towards a new prognostic marker. *Cancers* **12**, 906 (2020).
75. Yu, M. et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339**, 580-584 (2013).
76. Na, T. Y., Schecterson, L., Mendonsa, A. M. & Gumbiner, B. M. The functional activity of E-cadherin controls tumor cell metastasis at multiple steps. *Proc. Natl. Acad. Sci. USA* **117**, 5931-5937 (2020).
77. Aceto, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110-1122 (2014).
78. Wang, T., Wang, Z. Y., Niu, R. F. & Wang, L. Crucial role of Anxa2 in cancer progression: highlights on its novel regulatory mechanism. *Cancer Biol. Med.* **16**, 671-687 (2019).
79. Wei, L. et al. Hsp27 participates in the maintenance of breast cancer stem cells through regulation of epithelial-mesenchymal transition and nuclear factor- κ B. *Breast Cancer Res.* **13** (2011).
80. Filippou, P. S., Karagiannis, G. S. & Constantinidou, A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene* **39**, 2040-2054 (2020).
81. Kwon, M. J. et al. Clinical significance of CD151 overexpression in subtypes of invasive breast cancer. *Brit. J. Cancer* **106**, 923-930 (2012).
82. Byers, L. A. et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a

- therapeutic target for overcoming EGFR inhibitor resistance. *Clin. Cancer Res.* **19**, 279-290 (2013).
83. Shien, K. et al. Acquired resistance to EGFR inhibitors is associated with a manifestation of stem cell-like properties in cancer cells. *Cancer Res.* **73**, 3051-3061 (2013).
 84. Hu, F. et al. δ EF1 promotes osteolytic metastasis of MDA-MB-231 breast cancer cells by regulating MMP-1 expression. *Biochim. Biophys. Acta* **1809**, 200-210 (2011).
 85. Yang, J. et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927-939 (2004).
 86. Tora, L. et al. The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. *EMBO J.* **8**, 1981-1986 (1989).
 87. Scheidegger, K. J., Cenni, B., Picard, D. & Delafontaine, P. Estradiol decreases IGF-1 and IGF-1 receptor expression in rat aortic smooth muscle cells: mechanisms for its atheroprotective effects. *J. Biol. Chem.* **275**, 38921-38928 (2000).
 88. Webb, P. et al. Estrogen receptor activation function 1 works by binding p160 coactivator proteins. *Mol. Endocrinol.* **12**, 1605-1618 (1998).
 89. Wood, L. W. et al. Thyroid transcription factor 1 reprograms angiogenic activities of secretome. *Sci. Rep.* **6**, 19857 (2016).
 90. Mazda, M., Nishi, K., Naito, Y. & Ui-Tei, K. E-cadherin is transcriptionally activated via suppression of ZEB1 transcriptional repressor by small RNA-mediated gene silencing. *PLoS ONE* **6**, e28688 (2011).
 91. Zawel, L. et al. Human Smad3 and Smad4 are sequence-specific transcription activators. *Mol. Cell* **1**, 611-617 (1998).

92. Schmidt, D. et al. ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods* **48**, 240-248 (2009).
93. Infante, J. J., Law, G. L. & Young, E. T. Analysis of nucleosome positioning using a nucleosome-scanning assay. *Methods Mol. Biol.* **833**, 63-87 (2012).
94. Métivier, R. et al. Estrogen receptor- α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751-763 (2003).
95. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
96. Afgan, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537-W544 (2018).
97. Liu, T. et al. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* **12**, R83 (2011).
98. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011).
99. Kulakovskiy, I. V. et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252-D259 (2018).
100. Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* **43**, 27-33 (2011).
101. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2009).

102. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545-15550 (2005).
103. Isserlin, R., Merico, D., Voisin, V. & Bader, G. D. Enrichment Map - a Cytoscape app to visualize and explore OMICs pathway enrichment results. *F1000Res* **3**, 141 (2014).
104. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902 (2019).

Acknowledgements

We are grateful to the core facilities of the Faculties of Medicine and Science of the University of Geneva: the Flow Cytometry facility, the Genomics Platform and the Bioimaging Center of iGE3, and the high throughput screening facility (ACCESS Geneva) for their excellent technical support. We thank Dr. Jean-Pierre Aubry and Grégory Schneider for their generous technical assistance for the FACS procedures. Research in the Aceto lab is supported by the European Research Council (678834 and 840636), the European Union (801159-B2B), the Swiss National Science Foundation (PP0P3_163938, PP00P3_190077, IZLIZ3_182962), the Swiss Cancer League (KFS-3811-02-2016, KLS-4222-08-2017, KLS-4834-08-2019), the Basel Cancer League (KLbB-4173-03-2017, KLbB-4763-02-2019), the two Cantons of Basel through the ETH Zürich (PMB-01-16), the University of Basel and the ETH Zurich, Switzerland. The work of the Picard lab for this study was supported by the Medic Foundation and the Canton de Genève.

Author contributions

N.M.G. conceived the study, designed and performed most of the experiments, analyzed most of the data, prepared the figures, and wrote the manuscript. M.K.S. conducted the xenograft experiments and recorded the data. N.H. extensively contributed to the bioinformatics analyses of the ChIP-seq, RNA-seq, and sc-RNA-seq data. L.B. contributed to experiments with recombinant proteins. N.A. contributed to designing and supervised the xenograft experiments. D.P. supervised and conceived the study, contributed to the design of the experiments, and wrote and critically edited the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

Fig. 1 Expression of ZEB1 potentiates ER α activity and invasion of breast

cancer cells. a MCF7-V-ZEB1 cells in the presence of doxycycline (+DOX) express ZEB1. Immunoblots show the levels of EMT markers and ER α after long-term (8-12 weeks) and short-term (1-2 weeks) expression of ZEB1. **b-i** Luciferase reporter assays with transiently transfected cells as indicated. Except for the experiments of panels **(b)** and **(c)**, all assays were done with HEK293T cells. The activities of ER α and PR, and the E-cadherin, VEGF, and TGF β -responsive promoter activities were monitored with the reporter plasmids ERE-Luc, PRE-Luc, and proE-cad670-Luc, VEGFprom-Luc, and SBE4-Luc, respectively. The luciferase activities (RLU) are expressed relative to the activities of the internal transfection standard, Renilla luciferase. Graphs are based on $n \geq 4$ biologically independent experiments. **j** Expression of ER α target genes in MCF7-V-ZEB1 cells; mRNA levels were analyzed by RT-qPCR following 6 h of treatments as indicated; $n = 4$ biologically independent experiments. **k** Representative images of a three-dimensional (3D) tumor invasion assay with MCF7-V-ZEB1 cells. **(l)** Invasion kinetics based on the area of 4-6 independent spheroids as shown in panel **k** (note that most error bars are too small to be visible). Ctl, veh, E2, FI, and ICI stand for control, vehicle, 17 β -estradiol, forskolin + IBMX, and fulvestrant, respectively. All error bars represent standard errors of the means (mean \pm SEM). Asterisks indicate significant differences compared to the control expressing empty vector (* $P < 0.05$, ** $P = 0.001$ and **** $P < 0.0001$). Statistical significance was determined with unpaired and two-sided Student's t-tests for panels **f**, **j**, and **i** and a two-way ANOVA for all other panels. Source data are provided as a Source Data file.

Fig. 2 ZEB1 enhances ER α recruitment at common binding sites. **a** ChIP-qPCR of ER α on the *GREB1* (+5 kb) and *TFF1* (+0.5 kb) binding sites in MCF7-V-ZEB1 cells. ER α ChIP values were normalized to a non-binding region and the input. Recruitment was compared to –DOX with the graph showing the means \pm SEM of $n \geq 5$ biologically independent experiments. **b, c** Volcano plots of ER α ChIP-seq with wild-type MCF7-V (data from our previously published dataset³⁵) and MCF7-V-ZEB1 cells showing the FDR values as a function of the fold-changes of the normalized ER α values of MCF7-V-ZEB1 cells compared to the ER α peaks of wild-type MCF7-V cells treated with E2 (**b**) or FI (**c**) from $n \geq 2$ biologically independent experiments. **d, e** Functional annotations for the biological functions of E2- or FI-only ER α binding sites, respectively, using GREAT. **f** Table summarizing the number of ER α (ESR1), ZEB1, and ESR1-ZEB1 shared motifs in the ER α ChIP-seq, as found with FIMO. **g, h** ChIP-qPCR of candidate ER α binding sites from top hits of the ChIP-seq data for E2 (**g**) and FI (**h**); $n \geq 3$ biologically independent experiments. **i** ZEB1 ChIP with MCF7-V-ZEB1 cells with (+DOX) or without (–DOX) ZEB1 ($n = 3$ biologically independent experiments). **j, k** Venn diagrams show the intersections between ZEB1 binding sites and E2- (**j**) or FI-induced ER α (**k**) binding sites of MCF7-V cells; the ZEB1 data are from published ChIP-seq datasets for MDA-MB-231 and Hs578T breast cancer cells ([GSE89206](#)). **l** Genome browser views of ZEB1 and ER α binding sites adjacent to the *GREB1* and *TFF1* genes. **m** Re-ChIP experiment showing that ZEB1 and ER α co-occupy the indicated shared binding sites ($n = 3$ biologically independent experiments). Veh, E2, and FI stand for vehicle, 17 β -estradiol, and forskolin + IBMX, respectively. All graphs with error bars represent the means \pm SEM. Asterisks indicate significant differences (ns = not significant, * $P < 0.05$, ** $P < 0.01$).

< 0.001, *** P = 0.0001, **** P < 0.0001). Statistical significance was determined with a two-way ANOVA. Source data are provided as a Source Data file and Supplementary Data 1 and 2.

Fig. 3. ER α interacts with ZEB1 and requires AP2 γ for an effective ZEB1-ER α interaction and transcriptional activity. **a** Immunoblots of an ER α and ZEB1 co-immunoprecipitation experiment. IPs with extracts from MCF7-V-ZEB1 without or with DOX treatment were performed with antibodies specific to the exogenously expressed ZEB1 (left) or to the endogenous ER α (right). A control immunoprecipitation was performed in parallel with IgG, blotted and exposed simultaneously. **b, c** Venn diagrams showing overlap of AP2 γ (**b**) or FOXA1(**c**) binding sites with ZEB1 and ERBSs from the ChIP-seq data. **d** Aggregation plot of the binding sites of ZEB1, ER α , AP2 γ , FOXA1, GATA3, and P300, and the open chromatin histone marks H3K27ac, H3K4me1, and H3K9me3. Except for ER α , ChIP-seq data were from published datasets: [GSE89206](#) (ZEB1), [GSE21234](#) (TFAP2C), [GSE25315](#) (FOXA1), [GSE60270](#) (GATA3, P300, H3K27ac, H3K4me1, and H3K9me3). **e** Genome browser snapshots of ZEB1, ER α , AP2 γ , and FOXA1 enhancers of indicated genes. **f** Luciferase reporter assays with ERE-Luc in HEK293T cells infected with lentiviral constructs for shRNAs targeting *FOXA1*, *TFAP2C*, or both mRNAs; scrambled shRNA (shScr) was used as negative control (mean \pm SEM, n = 3 biologically independent experiments). **g** Co-IPs with HEK293T cells co-transfected with ZEB1 and ER α expression vectors. A control IP was performed with an IgG antibody. **h, i** Cells infected with shScr or shFOXA1 (**h**) or shTFAP2C (**i**) viruses. ER α ChIP-qPCR values are represented as fold of the shScr in -DOX (mean \pm SEM, n = 3 biologically independent experiments). **j** Genome

browser snapshot of the ER α ChIP-seq peaks at the *TFAP2C* gene. **k** ChIP-qPCR of ERBS at the *TFAP2C* 5'UTR (mean \pm SEM, n = 4 biologically independent experiments). **l** Immunoblots of extracts from MCF7-V-ZEB1 cells. Veh, E2, and FI stand for vehicle, 17 β -estradiol, and forskolin + IBMX, respectively. Asterisks indicate significant differences compared to the –DOX data (ns = not significant, * P < 0.01, ** P = 0.001, **** P < 0.0001). Statistical significance was determined with a two-way ANOVA. Source data are provided as a Source Data file.

Fig. 4 ZEB1 reprograms the ER α transcriptome to promote an EMT phenotype of breast cancer cells. **a** Venn diagram showing the overlap of the differentially expressed genes for MCF7-V-ZEB1 cells with (+DOX) or without (–DOX) ZEB1 expression in presence of indicated treatments. Fold changes were calculated relative to values from –DOX (veh) (n = 2 biologically independent experiments). **b, c** Venn diagrams demonstrate the comparison of uniquely altered genes induced by E2 (**b**) or FI (**c**) in each group of cells with (+DOX) or without (–DOX) ZEB1 expression. **d-g** Bar plots of a GSEA showing the $-\log_{10}$ (P value) of the top 20 unique GO terms associated with significantly altered mRNA expression levels upon E2- or FI-induced ER α activation in cells with (+DOX) or without (–DOX) ZEB1. E2 and FI stand for 17 β -estradiol and forskolin + IBMX, respectively. In all panels genes with fold changes (FC) of expression ≥ 1.5 were considered upregulated and those with the FC ≤ 1.5 were considered downregulated. Only the genes that had a FDR < 0.05 were included in the analyses. Source data are provided in Supplementary Data 3-5.

Fig. 5 ZEB1-induced novel ER α target genes are implicated in invasion. a

Heatmap showing the gene expression levels of the top 20 genes from the RNA-seq data analysis, grouped to highlight treatment-specific signatures (as indicated on the side). Red rectangles around gene names indicate the candidate genes selected for validation and functional assays. **b, c** mRNA levels of candidate genes determined by RT-qPCR (mean \pm SEM, $n = 3$ biologically independent experiments). **d** Bar graphs representing the changes in the 3D invasion ability of spheroids formed from cells with (+DOX) or without (–DOX) ZEB1 expression upon knocking down the expression of *MUC16*, *ESR2*, *DSCAM*, and *P2RX7* with the respective shRNAs. The means of area values from $n \geq 3$ biologically independent experiments were normalized to the mean of the values from the corresponding –DOX shScr condition treated with vehicle. Veh, E2, FI, and ICI stand for vehicle, 17 β -estradiol, forskolin + IBMX, and fulvestrant, respectively. The error bars are not shown to simplify the graph. Source data are provided as a Source Data file.

Fig. 6 ZEB1 induces different EMT states in breast cancer cells with CD151 as a marker of early EMT and as a potentially interesting target in combination with EGFR inhibitors. Panels **a-j** show different aspects of a scRNA-seq analysis of MCF7-V cells expressing ZEB1. **a** Visualization of the distribution of the EpCAM^{high} and EpCAM^{low} cell populations based on the comparison of the transcriptomes of individual cells; the graph was generated by a Uniform Manifold Approximation and Projection (UMAP); the clusters are color-coded and each dot represents a single cell. **b** Single cells expressing *EPCAM* in different clusters. **c** Violin plot of the *EPCAM* transcript levels across different clusters; note that the steps at low levels of expression are due to rounding off after normalization. **d** Single cells expressing

ESR1 in different clusters. **e** Violin plot for *ESR1* transcripts. **f-h** Expression of top markers of epithelial (**f**), hybrid EMT (**g**) and mesenchymal (**h**) states in breast cancer cells upon induction of an EMT by ZEB1 expression. Legend shows a color gradient of normalized expression. The accompanying scheme of the EMT on the right was created with BioRender.com. **i** Expression of the mRNA for the *CD151* cell surface marker across various subpopulations. **j** Violin plot of the *CD151* transcript levels across different clusters. **k, l** Relative viability of –DOX and +DOX cells with or without *CD151* knockdown exposed to the EGFR inhibitors erlotinib at 10 μ M (**k**) or gefitinib at 20 μ M (**l**) for 72 h (mean \pm SEM, $n \geq 4$ biologically independent experiments). Statistical significance was determined with unpaired and two-tailed Student's *t*-test. Asterisks indicate significant differences compared to the –DOX (veh) (* $P < 0.01$, ** $P < 0.001$, *** $P = 0.0006$, **** $P < 0.0001$) **m, n** Kaplan-Meier plots for overall survival (OS) and distant metastasis-free survival (DMFS) of ER α ⁺ (**m**) and ER α [–] (**n**) breast cancer patients classified as tumors expressing CD151^{high}/ZEB1^{high} (red line) and CD151^{low}/ZEB1^{low} (gray line). Source data are provided as a Source Data file.

Fig. 7 ZEB1 promotes metastasis *in vivo*. **a** Bioluminescent images (BLI) of primary tumors of mice injected with one million ZsGreen- or ZEB1-expressing wild-type MCF7 cells into the mammary fat pad. The panel on the right is the corresponding quantitative analysis of the average BLI counts of multiple primary tumors. The statistical significance was determined with an unpaired and two-tailed Student's *t*-test ($n \geq 3$ biologically independent experiments, mean \pm SEM); the asterisk indicates the statistically significant difference compared to the ZsGreen group with $P < 0.05$. **b and c** Left panels are BLI images showing metastatic lesions

in bones and lungs, respectively. The corresponding panels on the right are bar graphs of the metastatic indices (ratio of BLI counts of metastases over BLI counts of the corresponding primary tumor) for bones and lungs ($n \geq 3$ animals, error bars = mean \pm SEM). The corresponding graphs with the BLI counts of bone and lung metastases are shown in Supplementary Fig. 8i. Note the log₁₀ scales of the Y-axes. **d** The table summarizes the incidence of metastases ($n \geq 3$ animals), considering only those with areas with BLI counts $> 10^2$. **e** Multiphoton confocal microscopy images of an *ex vivo* bone invasion assay showing the bone explant stained in red with alizarin red and the invading ZsGreen-labeled (green) control (Ctl) or ZEB1-expressing MCF7-V cells. Without ZEB1, the cells remain on top of the bone tissue whereas ZEB1-expressing cells invade it. The upper left of each panel shows a projection of a Z-stack with at least 20 images, and the upper right of each panel a slice of the area delimited by a yellow rectangle. Images were processed and generated with the software Imaris 9.7 to illustrate representative slide and 3D views of bone tissue. Scale bar = 50 μ m. **f** Transwell assay to evaluate the migration of ZsGreen-labeled control (Ctl) or ZEB1-expressing MCF7-V cells stimulated by bone and muscle tissues. Crystal violet dye was used to stain the migrated cells. Decellularized bone served as negative control. Images are representative of three independent experiments. Source data are provided as a Source Data file.

Fig. 8 Proposed mechanism of transcriptional regulation by the ZEB1-ER α complex during early/hybrid EMT stages of breast cancer. Non-invasive primary breast tumors are formed by the abnormal proliferation of luminal mammary epithelial cells with increased ER α activity. FOXA1 acts as a pioneer factor for the recruitment of ER α for transcription of genes involved in cell proliferation. In

early/hybrid states of EMT, AP2 γ becomes a determining pioneer factor recruiting a ZEB1-ER α complex to ER α binding sites. This complex reprograms the ER α cistrome and transcriptome towards activation of genes involved in early EMT and metastatic dissemination. The illustration was created with BioRender.com.

Fig. 1

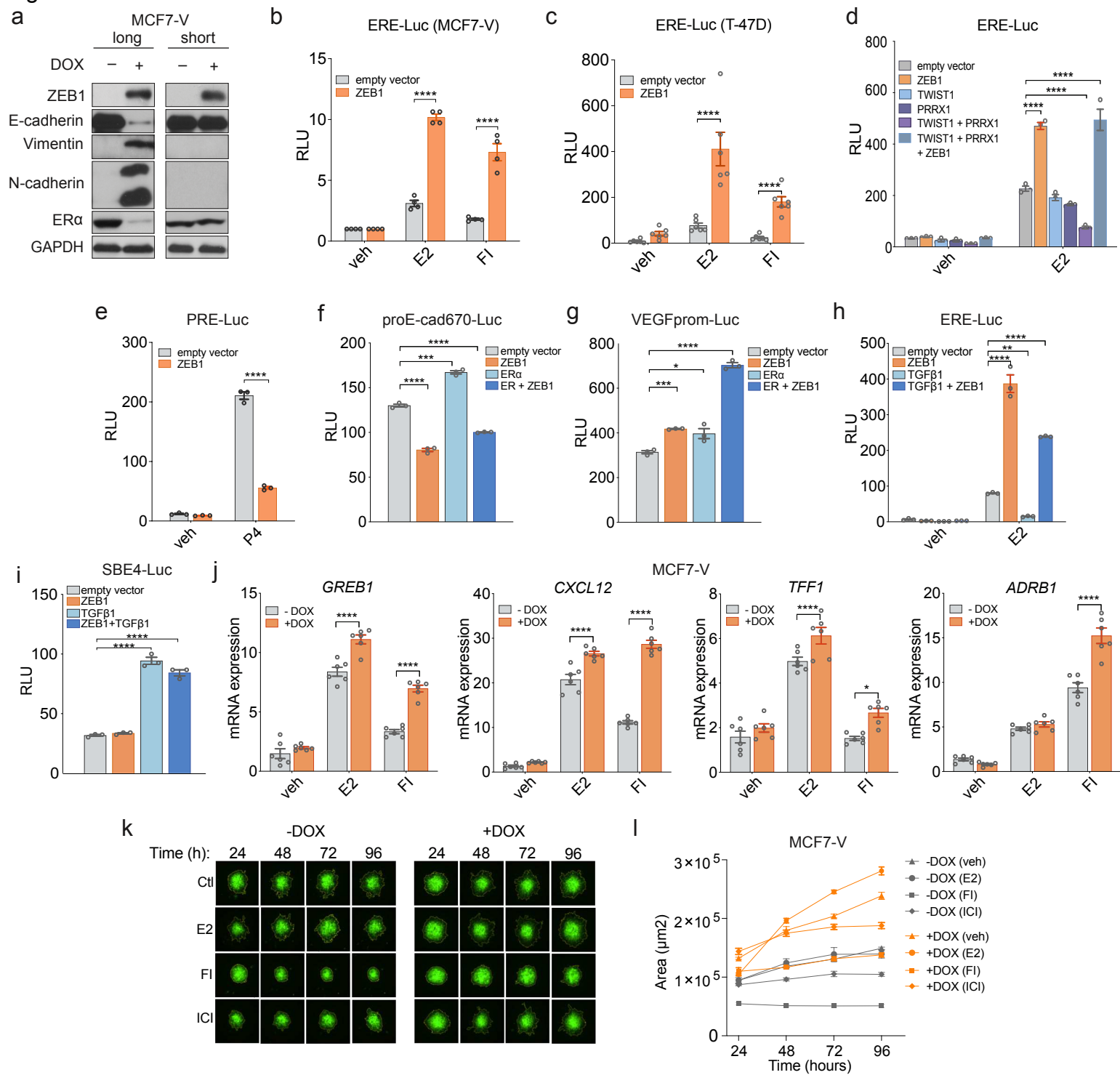


Fig. 2

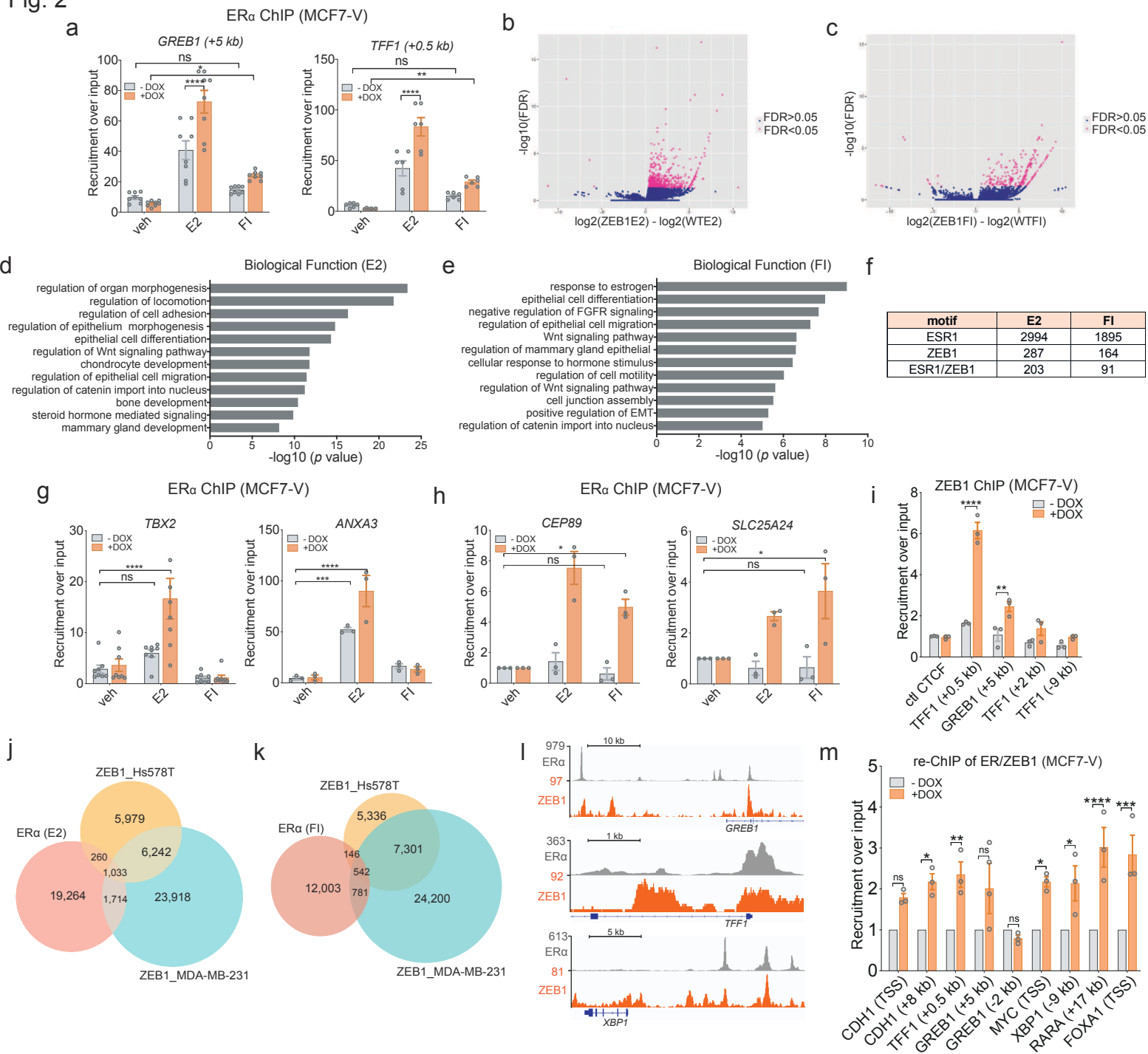
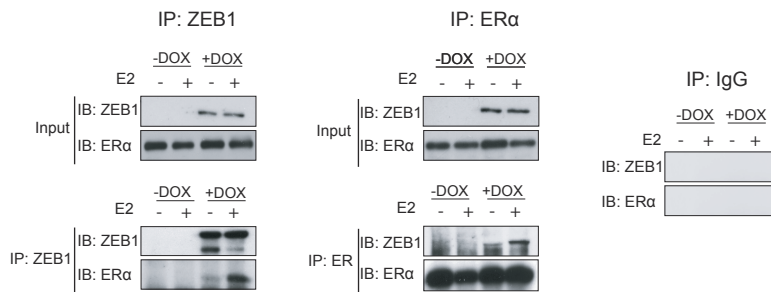


Fig. 3

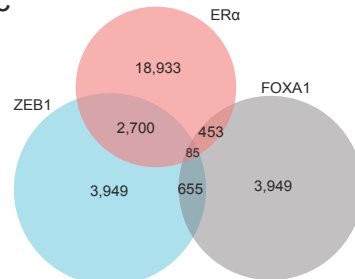
a



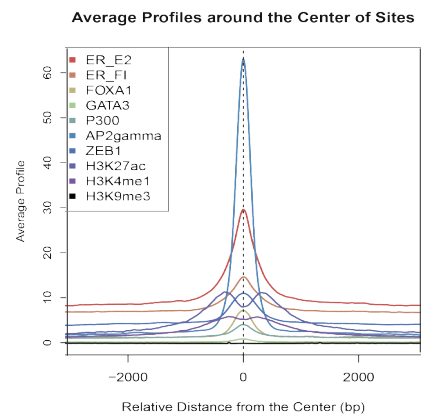
b



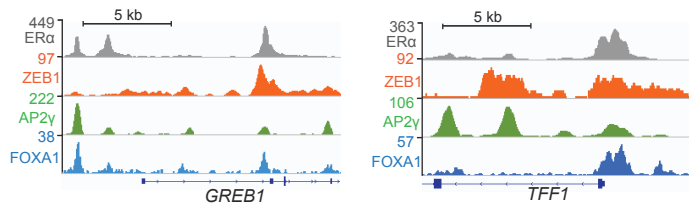
c



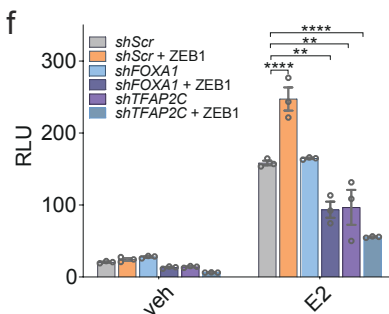
d



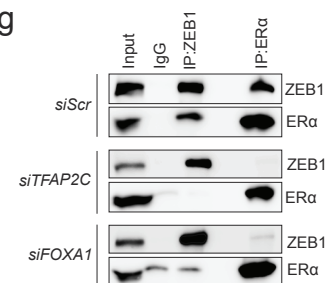
e



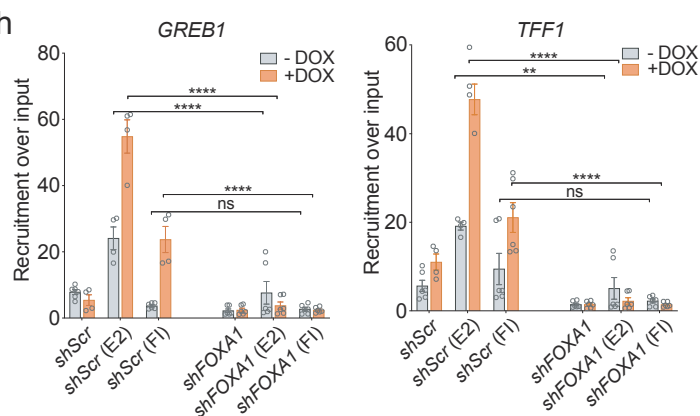
f



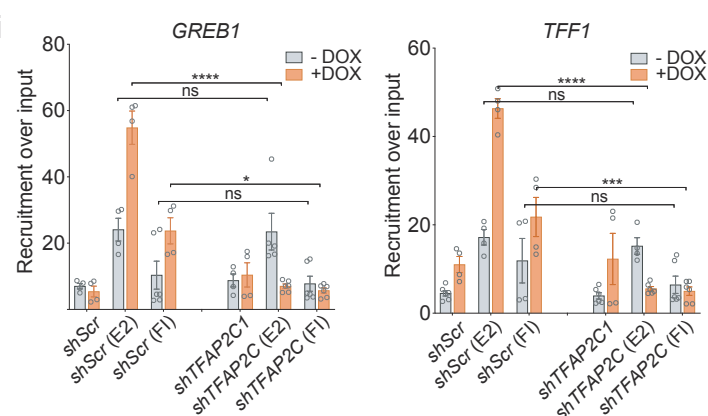
g



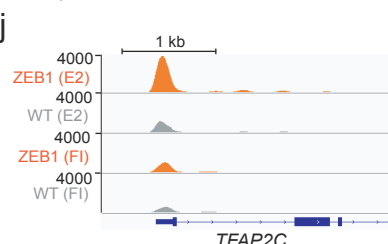
h



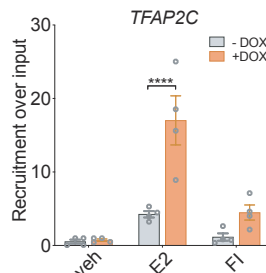
i



j



k



l



Fig. 4

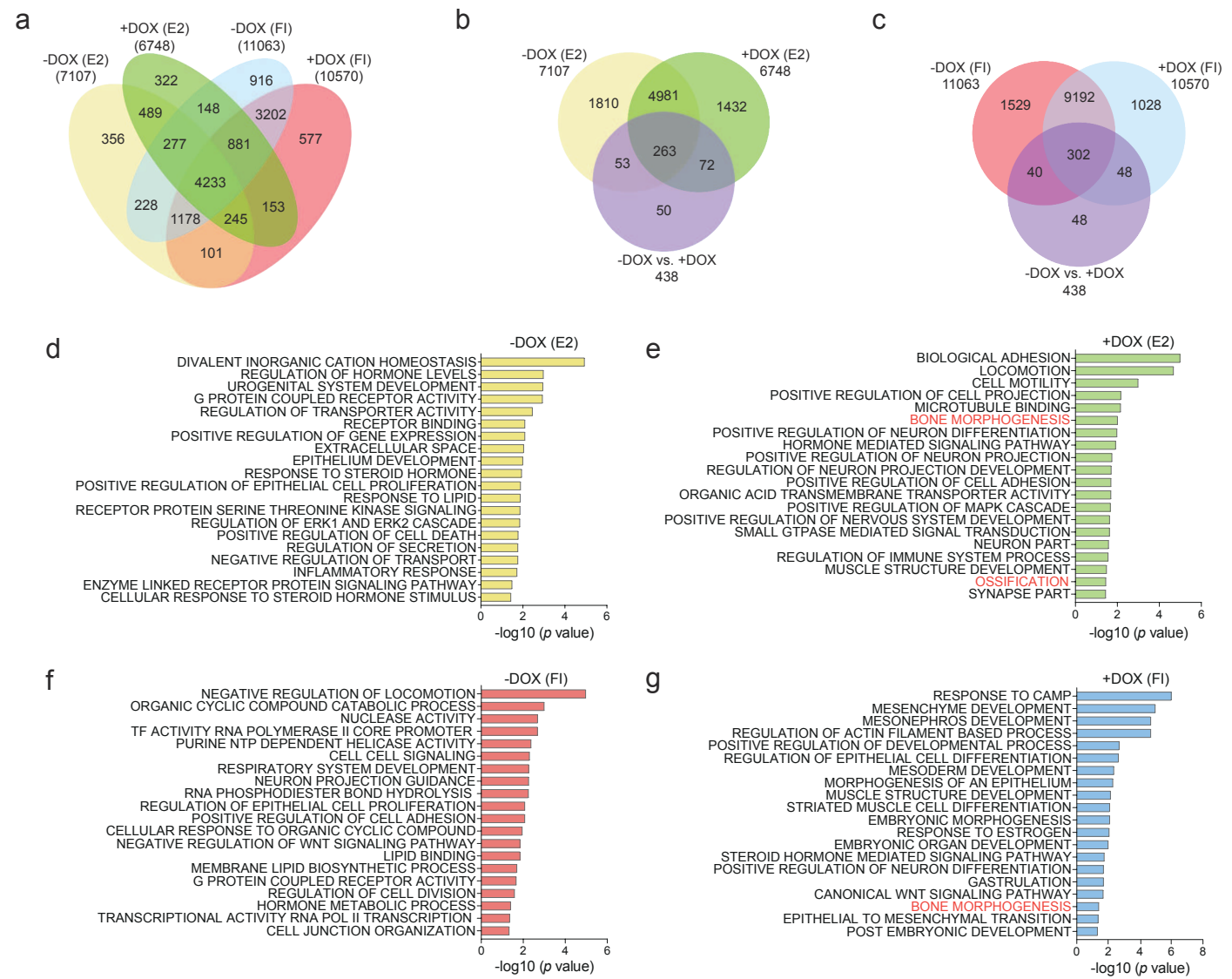


Fig. 5

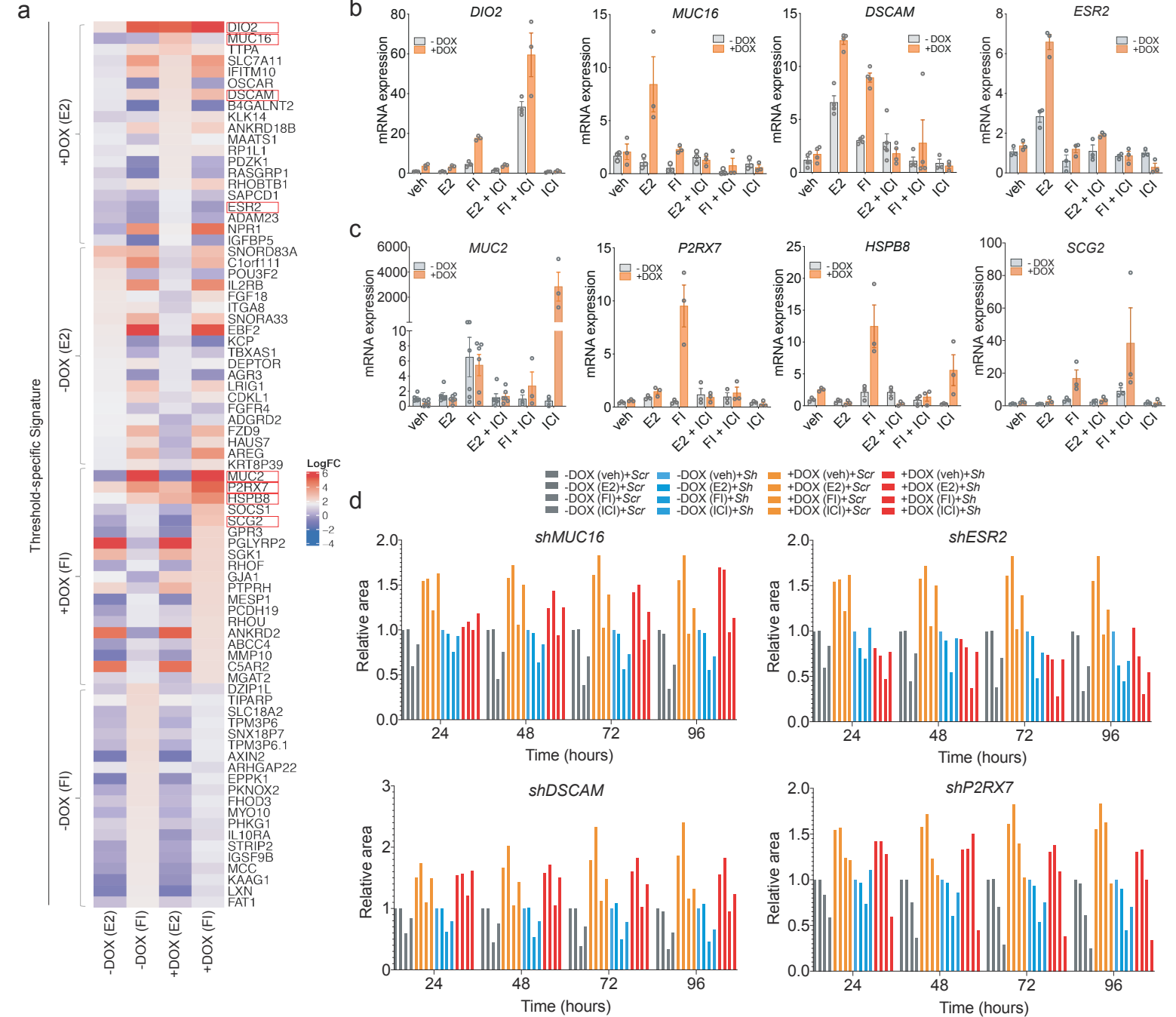


Fig. 6

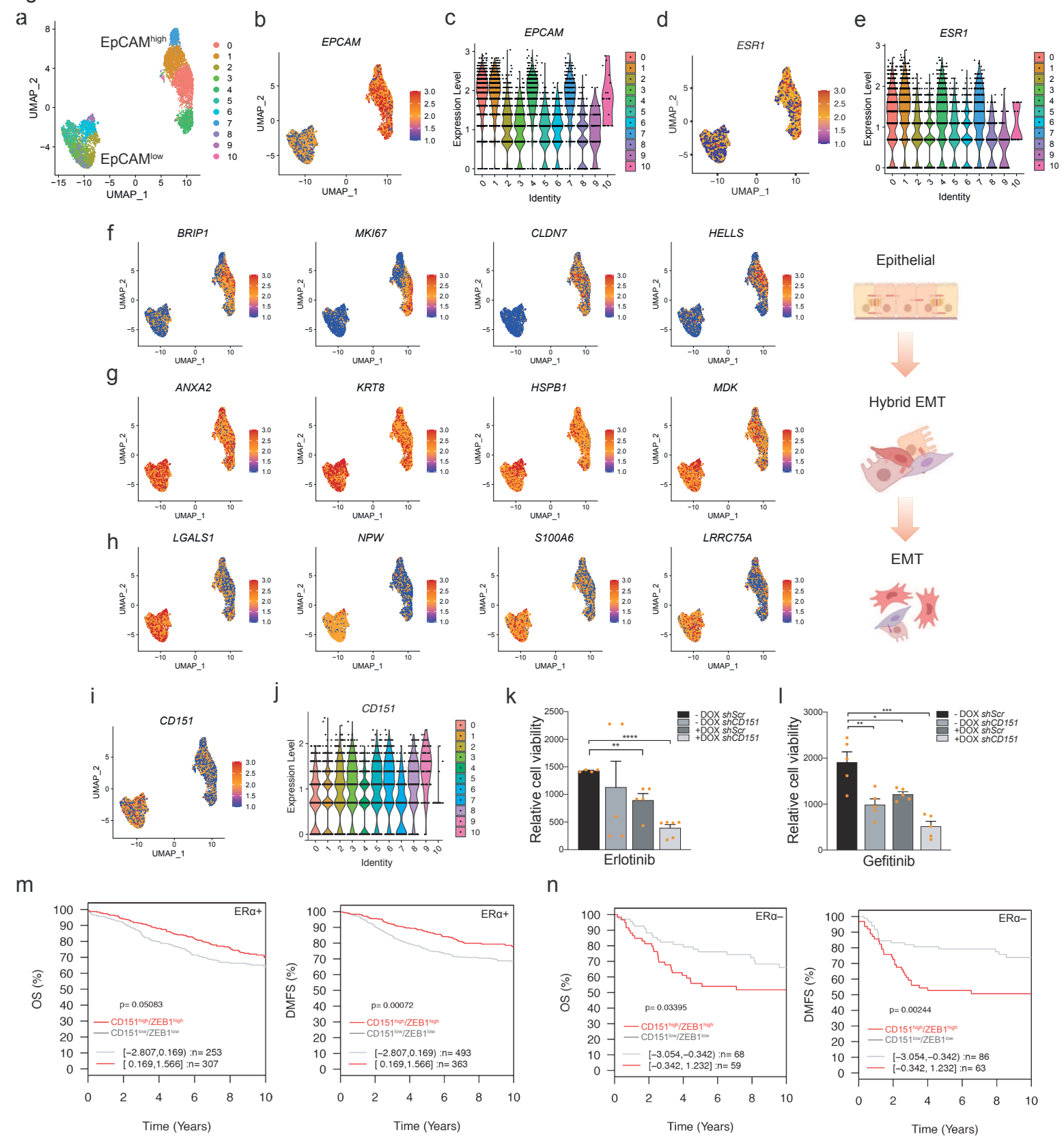


Fig. 7

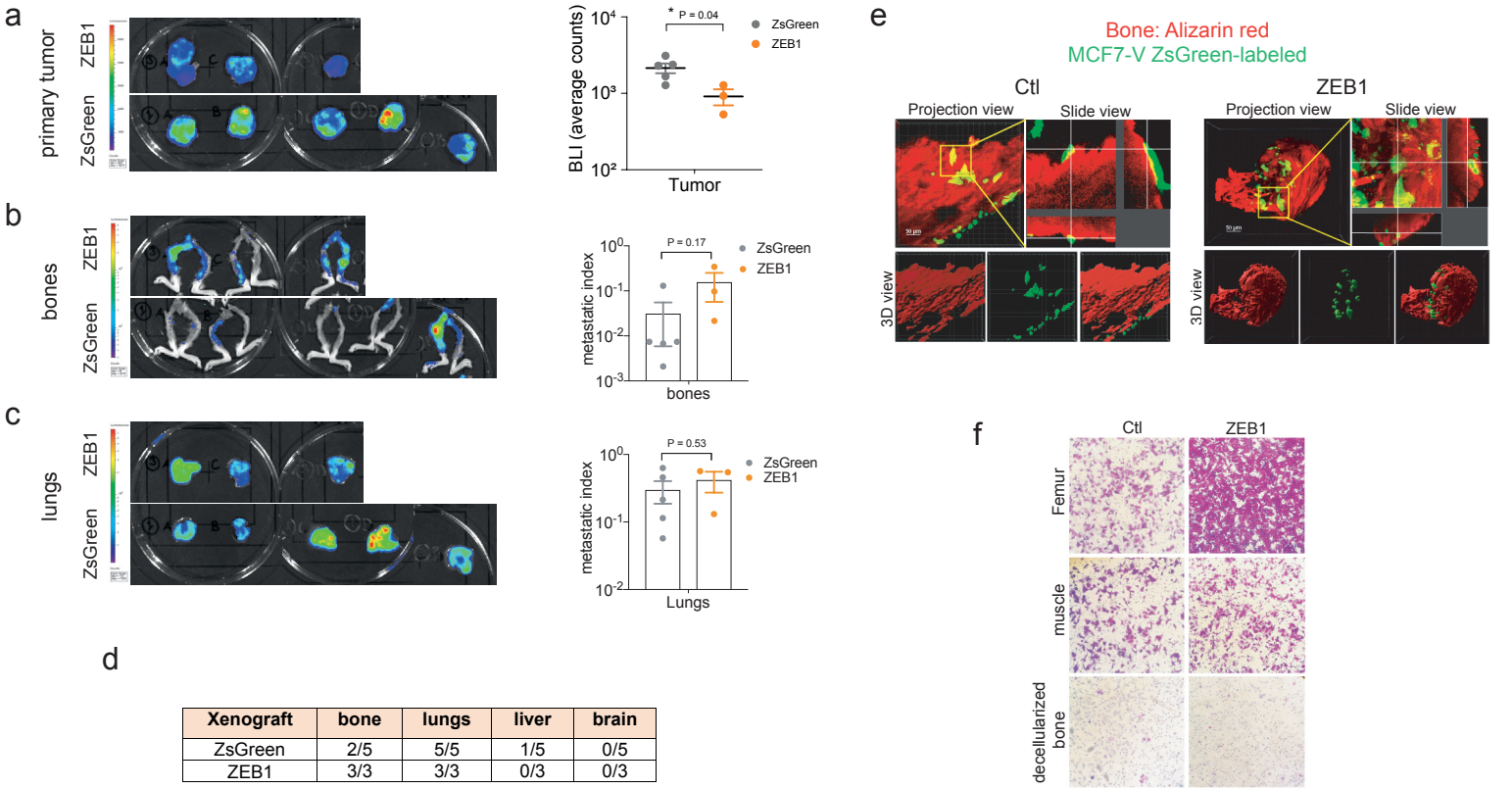
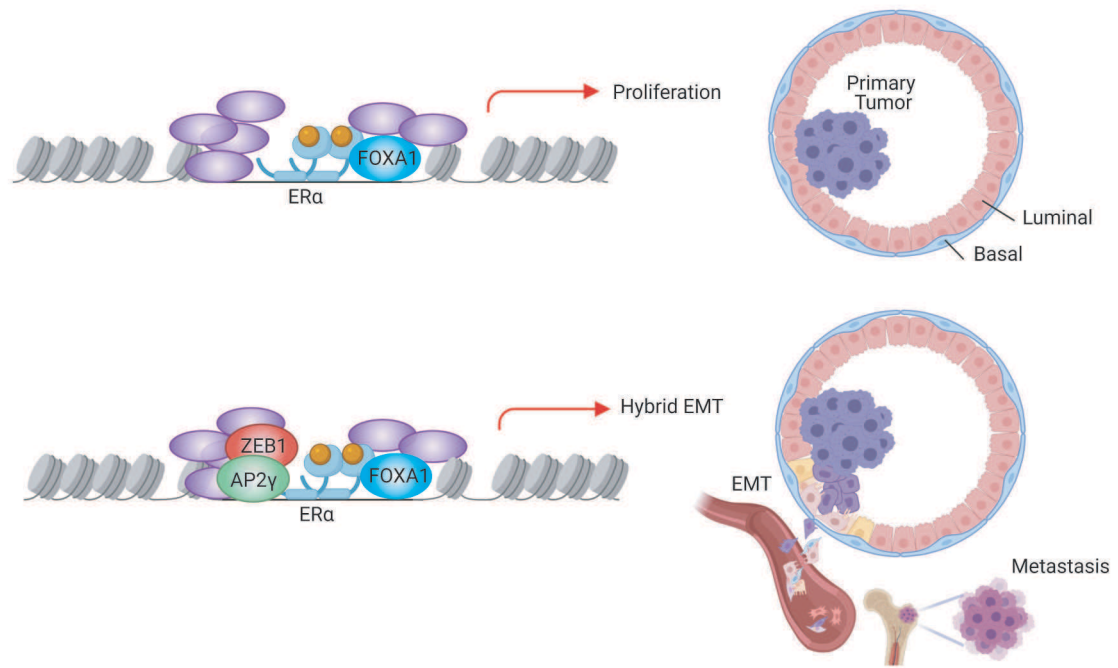


Fig. 8



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [DescriptionofSupplementaryDataFiles.docx](#)
- [SI.pdf](#)
- [SupplementaryData1.xlsx](#)
- [SupplementaryData2.xlsx](#)
- [SupplementaryData3.xlsx](#)
- [SupplementaryData4.xlsx](#)
- [SupplementaryData5.xlsx](#)