# RET Ligands Mediate Endocrine Sensitivity via a Bi-stable Feedback Loop with ERα

Sachi Horibata,[1,2] Edward J. Rice,[1] Hui Zheng,[1] Lynne J. Anguish,[1] Chinatsu Mukai,[1] Brooke A. Marks,[1] Tinyi Chu,[1,3] Scott A. Coonrod,[1,4,*] and Charles G. Danko[1,4,5*]

[1]Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

[2]Department of Molecular Medicine, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

[3]Graduate Field of Computational Biology, Cornell University, Ithaca, NY 14853, USA

[4]Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

[5]Lead Contact

[*]Correspondence: sac269@cornell.edu (S.A.C.), dankoc@gmail.com (C.G.D.)

## Summary

The molecular mechanisms of endocrine resistance in breast cancer remain poorly understood. Here we used PRO-seq to map the location of hundreds of genes and thousands of distal enhancers whose transcriptional activities differ between endocrine sensitive and resistant MCF-7 cells. Our genome-wide screen discovered increased transcription of the glial-cell line derived neurotrophic factor (GDNF), a RET tyrosine kinase receptor ligand, which we validate as both necessary and sufficient for resistance in MCF-7 cells. GDNF caused endocrine resistance by switching the active state of a bi-stable feedback loop in the MCF-7 regulatory network from ERα signaling to GDNF-RET signaling. To cause this switch, GDNF downregulated ERα transcription and activated the transcription factor EGR1, which, in turn, induced GDNF. Remarkably, both MCF-7 cells and ER+ primary tumors appear poised for endocrine resistance via the RET signaling pathway, but lack robust RET ligand expression and only develop resistance upon expression of GDNF or other RET ligands.

## Highlights

- GDNF expression promotes endocrine resistance in MCF-7 cells.
- ER+ MCF-7 cells are poised for RET-mediated endocrine resistance, but lack expression of RET ligands.
- RET ligand expression predicts resistance to the aromatase inhibitor letrozole.
- GDNF regulatory network directly down-regulates ERα and indirectly up-regulates *GDNF*.

## Introduction

Estrogen receptor alpha (ERα) is the major driver of ~75% of all breast cancers. ERα is a transcription factor whose genomic actions are induced upon binding its cognate ligand, 17β-estradiol (E2). E2-liganded ERα activates and represses thousands of ERα target genes and non-coding RNAs (Carroll et al., 2006; Hah et al., 2011, 2013). Genes whose transcription is directly activated by ERα promote a mitogenic response in breast cancer cells, resulting in entry into the cell cycle, survival, and continued cell proliferation (Planas-Silva and Weinberg, 1997; Prall et al., 1998). Current therapies for patients with ER+ breast cancer are largely aimed at blocking the ERα signaling pathway. For example, tamoxifen blocks ERα function by competitively inhibiting E2/ERα interactions (Shiau et al., 1998) and fulvestraunt promotes ubiquitin-mediated degradation of ERα (Wakeling, 2000). Because ERα is such an important and pervasive breast cancer driver, endocrine therapies are estimated to have reduced breast cancer mortality by 25-30% (Early Breast Cancer Trialists' Collaborative, 2005; Emens and Davidson, 2009; Musgrove and Sutherland, 2009).

Despite the widespread success of endocrine therapies, approximately 40-50% of breast cancer patients will either present with endocrine resistant breast cancer at the time of diagnosis or progress into endocrine-resistant breast cancer during the course of treatment (Ma et al., 2009). Numerous studies have now identified growth factor signal transduction “escape pathways” that may provide mechanisms for cell growth and survival that are independent of E2. For example, increased signaling from the EGFR/HER2 (Benz et al., 1992), RET tyrosine kinase receptor signaling (Gattelli et al., 2013; Morandi et al., 2013; Plaza-Menacho et al., 2010), and IGFR (Chan et al., 2016) have each been associated with either experimental or clinical endocrine resistance. Importantly, novel therapies targeting these tyrosine kinase signaling pathways are now showing promise in phase 2 trials for certain cohorts of patients (Park et al., 2016), raising substantial interest in further deciphering the mechanisms by which anti-tyrosine kinase receptor therapies are acting to inhibit breast cancer cell growth. A better understanding

of the transcriptional targets of these signaling pathways as well as understanding how these pathways crosstalk with ERα signaling will likely aid in the development of new predictive biomarkers and new targets for therapeutic intervention.

Dissecting the transcriptional mechanisms underlying endocrine resistance has proven technically challenging owing to the large number of indirect target genes and complex nature of each growth factor escape pathway. An emerging strategy for dissecting transcriptional responses to stimuli involves measuring gene transcription at the level of nascent RNA production (Churchman and Weissman, 2011; Core et al., 2008; Mahat et al., 2016a; Nojima et al., 2015; Schwalb et al., 2016). These approaches are highly sensitive to immediate and transient transcriptional responses to stimuli, allowing the discovery of target genes within minutes of activation and hence distinguishing primary and secondary effects (Arner et al., 2015; Danko et al., 2013; Duarte et al., 2016; Hah et al., 2011; Mahat et al., 2016b). Moreover, these approaches can detect active transcriptional regulatory elements (TREs), including both promoters and distal enhancers, because these elements display distinctive patterns of transcription (Andersson et al., 2014a; Core et al., 2014; Danko et al., 2015; Hah et al., 2013; Kim et al., 2010) which are obscured in RNA-seq data owing to rapid degradation by the exosome complex (Andersson et al., 2014b; Core et al., 2014). Indeed, a recent method for detecting nascent transcription by mapping the location and orientation of actively transcribing RNA polymerase, called Precision Run-On and Sequencing (PRO-seq), serves as a powerful assay for both identifying TREs and measuring gene transcription levels (Danko et al., 2015).

Here, we used PRO-seq to comprehensively map RNA polymerase in an MCF-7 model of tamoxifen resistance (Gonzalez-Malerva et al., 2011). These maps provide the location of hundreds of genes and thousands of distal enhancers whose activities differ between tamoxifen sensitive (TamS) and resistant (TamR) MCF-7 cells. Results show that, whereas E2/ERα signaling dominates transcriptional activation in the sensitive lines, ERα signaling is suppressed

in the resistant lines due to the activation of GDNF-RET signaling which, in turn, modulates its output by activating the transcription factors SRF and AP-1. Remarkably, TamS MCF-7 cells express all of the necessary proteins to drive RET receptor signaling, with the exception of one of the RET ligands (GDNF, NRTN, ARTN, or PSPN). By manipulating GDNF expression in MCF-7 cells, we determined that high GDNF expression is both necessary and sufficient for tamoxifen resistance in our MCF-7 cell model. Mechanistically, we found that GDNF promotes tamoxifen resistance by switching the active state of a bi-stable feedback loop between ERα and a positive feedback loop between GDNF and the transcription factor EGR1. Our findings appear to be clinically relevant as we found that RET ligand expression is predictive of responsiveness to endocrine therapies in breast cancer patients. Taken together, our studies unravel the transcriptional regulatory circuitry that underlies RET-tyrosine kinase dependent resistance to endocrine therapies, and provides general insights into how escape pathways facilitate ERα-independent growth in ER+ breast cancers.

## Results

### Genome-wide maps of RNA polymerase in tamoxifen sensitive and resistant MCF-7 cells

Although MCF-7 cells are ER+ and largely require E2 for growth and proliferation, a subset of the heterogeneous MCF-7 cell population continues growing in the presence of anti-estrogens such as tamoxifen (Coser et al., 2009; Gonzalez-Malerva et al., 2011). We hypothesized that the *de novo* resistant cells display a unique transcriptional program which can be used to identify factors that play a causative role in tamoxifen resistance. We used PRO-seq to map the location and orientation of RNA polymerase in two sensitive and two *de novo* resistant MCF-7 cell lines that were clonally derived from parental MCF-7 cells (Gonzalez-Malerva et al., 2011). Consistent with the previous study, we found that the TamS lines (TamS; $B7^{TamS}$ and $C11^{TamS}$) were sensitive to as little as 1 nM of tamoxifen while the TamR lines (TamR; $G11^{TamR}$ and $H9^{TamR}$) were not affected at concentrations as high as 100 nM (**Figure 1A**). PRO-seq libraries were prepared from all four cell lines (**Figure 1B**) as previously described (Kwak et al., 2013; Mahat et al., 2016a) and sequenced to a combined depth of 87 million uniquely mapped reads (**Table S1**). We quantified the similarity of transcription in the MCF-7 cell subclones using the Pol II abundance in annotated gene bodies. Unbiased hierarchical clustering grouped $B7^{TamS}$ and $C11^{TamS}$ TamS lines into a cluster and left $G11^{TamR}$ and $H9^{TamR}$ TamR lines as more distantly related outgroups (**Figure 1C**). Although TamR cells clustered independently, all four MCF-7 clones are nevertheless remarkably highly correlated (Spearman's Rho > 0.95), suggesting that relatively few transcriptional changes are necessary to produce the tamoxifen resistance phenotype.

We identified 527 genes that are differentially transcribed between TamS and TamR MCF-7 cells (1% FDR, deSeq2 (Love et al., 2014)), 341 of which were transcribed more highly in TamS and 186 in TamR cell lines (**Figure 1D**). Several of the differentially transcribed genes, including, for example, *PGR*, *GREB1*, *IGFBP5*, *HOXD13*, and *GDNF,* were identified in other models of endocrine resistance (Esseghir et al., 2007; Ghoussaini et al., 2014; Mohammed et

al., 2013; Morandi et al., 2013; Plaza-Menacho et al., 2010; Zhong et al., 2015), supporting our hypothesis that transcriptional changes in the MCF-7 model are generally informative about endocrine resistance. In one example, the diagnostic marker *PGR* is transcribed uniquely in the $B7^{TamS}$ line and is largely absent from $G11^{TamR}$ (**Figure 1E**). To further confirm that transcriptional changes detected using PRO-seq lead to differences in mRNA abundance, we validated transcriptional changes in *PGR* and *GREB1* between the $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells using qPCR (**Figure 1F**).

Many of the differentially transcribed genes are targets of ERα signaling, including *PGR, GREB1*, *NOS1AP*, and *ELOVL2*, suggesting that changes between TamR and TamS MCF-7 cells can be explained in part by differences in the genomic actions of ERα. To test for an enrichment of ERα target genes, we asked whether immediate transcriptional changes following E2 treatment are correlated with genome-wide changes between TamS and TamR MCF-7 cells using an independent GRO-seq dataset (Hah et al., 2011). Indeed, genes up-regulated by 40 minutes of E2 treatment tend to be transcribed more highly in TamS MCF-7 cells, and genes down-regulated by E2 are higher in TamR cell lines (**Figure 1G**). Thus, our data implicates global changes in the genomic actions of ERα in tamoxifen resistance in this MCF-7 model system.

**Distal enhancer activities correlate with tamoxifen resistance**

To elucidate the mechanisms responsible for changes in gene transcription during tamoxifen resistance, we sought to discover the location of promoters and active distal enhancers, collectively called transcriptional regulatory elements (TREs). Nascent transcription is a highly sensitive way to identify groups of active enhancers (Andersson et al., 2014a; Core et al., 2014; Danko et al., 2015; Hah et al., 2013), and results in enhancer predictions that are highly similar to the canonical active enhancer mark, acetylation of histone 3 at lysine 27 (H3K27ac) (Azofeifa and Dowell, 2016; Core et al., 2014; Danko et al., 2015). We used the

dREG software package (Danko et al., 2015) followed by a novel peak refinement step that identifies the regions between divergent paused RNA polymerase (see STAR Methods; manuscript in preparation) to identify 39,753 TREs that were active in either the TamS or TamR MCF-7 lines. TREs discovered using dREG were highly enriched for other active enhancer and promoter marks in MCF-7 cells, especially H3K27ac (**Figure S1A**) as expected based on prior studies (Azofeifa and Dowell, 2016; Core et al., 2014; Danko et al., 2015; Hah et al., 2013). We selected a transcribed enhancer downstream of the *CCND1* gene for experimental validation using luciferase reporter gene assays, and confirmed luciferase activity in both $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells (**Figure S1B and S1C**).

We used the abundance of RNA polymerase recruited to each TRE as a proxy for its transcriptional activity in each MCF-7 subclone to identify differences in 1,452 TREs (812 increased and 640 decreased) (1% FDR, deSeq2) between TamS and TamR MCF-7 cells. Differentially transcribed TREs were frequently located near differentially expressed genes and undergo correlated transcriptional changes between the four MCF-7 subclones. *GREB1* and *PGR*, for example, are each located near several TREs, including both promoters (green) and enhancers (gray), which undergo changes between TamR and TamS MCF-7 cells that are similar in direction and magnitude to those of the primary transcription unit which encodes the mRNA (**Figure 1E**). These results are consistent with a broad correlation between changes at distal TREs and protein coding promoters (Hah et al., 2011, 2013).

We hypothesized that differential transcription at TREs reflect differences in the binding of specific transcription factors that coordinate changes between TamS and TamR lines. We identified 12 clusters of motifs enriched in TREs that are differentially active between TamS and TamR lines (Bonferroni corrected $p< 0.001$; RTFBSDB (Wang et al., 2016)). Remarkably, the top scoring motif in this analysis corresponds to an estrogen response element (ERE), the canonical DNA binding sequence which recruits ERα to estrogen responsive enhancers (**Figure 1H**). At least two of the top scoring motifs, putatively bound by NFIA and HOXC13, bind a

transcription factor that was itself differentially expressed between TamS and TamR MCF-7 cells (**Figure 1H**), suggesting a model in which transcriptional changes of a transcription factor elicit secondary effects on the activity of TREs, and downstream effects on gene transcription. Together, although largely correlative, these integrative analyses of gene and TRE transcriptional activity begin to reveal a transcriptional regulatory network that correlates with tamoxifen resistance in MCF-7 cells.

**ERα signaling remains functional in endocrine resistant lines**

*GREB1* and *PGR* play a critical role in ERα genomic activity in breast cancer cells (Mohammed et al., 2013, 2015). Our observation that transcription of these ERα co-factors was lost in the resistant lines (**Figures 1E and 1F**) suggests that ERα signaling may be defective in the TamR cell lines. Consistent with this expectation, several analyses (i.e., the enrichment of ERα target genes and EREs, **Figures 1G and 1H**) strongly implicate global changes in the genomic actions of ERα during the development of tamoxifen resistance. However, these analyses are correlative and do not directly test the immediate responses to E2 in TamR and TamS lines.

To directly test the hypothesis that the genomic actions of ERα are substantially altered in the TamR lines, we treated $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells for 40 minutes with either E2 or Tamoxifen, and monitored transcriptional changes using PRO-seq. As expected, RNA polymerase abundance increased sharply at ERα ChIP-seq peaks (Welboren et al., 2009) in $B7^{TamS}$ MCF-7 cells (**Figure 2A top**), consistent with E2 activating Pol II loading at estrogen-responsive TREs (Danko et al., 2013; Hah et al., 2013). Surprisingly, E2 also activated transcription in $G11^{TamR}$ lines (**Figure 2A bottom**), strongly suggesting that E2 signaling continued to function in TamR lines despite the almost complete lack of *GREB1* and *PGR*. Likewise, direct E2 target genes defined in a previous GRO-seq study (Hah et al., 2011) were largely up- or down- regulated as expected in both $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells (**Figure 2B**).

Notably, however, we observed a much more muted effect of E2 on both enhancer and gene transcription in G11$^{TamR}$ compared with B7$^{TamS}$ (**Figures 2A and 2B**), explaining the enrichment in E2 target genes and ERE motifs in differences between TamS and TamR lines, as described above. The reduced effect of E2 on transcription may reflect that the lack of *GREB1* or *PGR* in these lines reduces the effect that ERα has on transcriptional activation. Additionally, however, we also observed a 2.44-fold reduction in the abundance of ERα mRNA (**Figure 2C**). Thus, it appears that, while E2 signaling becomes less responsive in G11$^{TamR}$ MCF-7 cells, the E2 signaling pathway remains largely functional and able to affect gene transcription in a stimulus-dependent manner.

One current model of tamoxifen resistance posits that tamoxifen can function as an ERα agonist in resistant breast cancer cells (Osborne et al., 2003). If this hypothesis is correct, then tamoxifen should promote the activation of ERα target genes in the G11$^{TamR}$ cells. However, our results showed that tamoxifen had no effect on either enhancer or gene transcription in either B7$^{TamS}$ or G11$^{TamR}$ lines (**Figures 2A and 2B**). Looking genome-wide, the tamoxifen treated B7$^{TamS}$ and G11$^{TamR}$ MCF-7 cells are very highly correlated with untreated controls (Spearman's rank correlation $\rho > 0.99$; **Figure S2**). The lack of transcriptional differences in either line is consistent with ERα signaling having already been largely shut down under these conditions by three-days of growth in charcoal-stripped FBS, which depletes endogenous E2 from the media. Importantly, these results demonstrate that tamoxifen does not appear to function as an agonist in G11$^{TamR}$ cells contrary to one current model for endocrine resistance (Osborne et al., 2003).

Given that our findings also suggested that E2 signaling remains functional, but muted in the TamR line, we next tested whether ERα was required for the growth of our tamoxifen resistant cells. We found that the viability of both G11$^{TamR}$ and H9$^{TamR}$ MCF-7 cells was largely unaffected by treatment with the ER degrader, fulvestrant (**Figure 2D**). Therefore, endocrine resistance in G11$^{TamR}$ and H9$^{TamR}$ MCF-7 cells appears to occur independently of ERα signaling,

suggesting that these TamR lines are likely using an alternative pathway for cell survival and proliferation when grown in the presence of tamoxifen.

**GDNF is necessary and sufficient to confer endocrine resistance in MCF-7 cells**

Tyrosine kinase growth factor signaling pathways have been implicated in preclinical models of endocrine resistance (Benz et al., 1992; Gattelli et al., 2013; Plaza-Menacho et al., 2010). RET is a cell surface receptor that elicits cell survival signals when bound by one of four RET ligands, GDNF, NRTN, ARTN, and PSPN (Sariola and Saarma, 2003). Remarkably, one of these ligands, glial-cell derived neurotrophic factor (GDNF), was among the most highly up-regulated genes in both $G11^{TamR}$ and $H9^{TamR}$ MCF-7 lines (**Figure 3A**). We confirmed the transcriptional differences in *GDNF* between $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells using qPCR and found that GDNF mRNA levels were increased by ~25 fold in the resistant line (**Figure 3B**). Thus both *GDNF* transcription and mRNA abundance correlate with endocrine resistance in MCF-7 cells, suggesting that GDNF may contribute to the endocrine resistance phenotype.

We directly tested this hypothesis by manipulating GNDF levels in our MCF-7 model. We first examined the effects of 10 ng/mL of recombinant GDNF protein on the growth of $B7^{TamS}$ cells in the presence of antiestrogens. Remarkably, GDNF completely rescued $B7^{TamS}$ MCF-7 cells when challenged with both tamoxifen (**Figure 3C**) and fulvestrant (**Figure S3A**). Moreover, GDNF treatment without tamoxifen increased the proliferation rate of $B7^{TamS}$ MCF-7 cells by ~20% (**Figure 3C**), suggesting that the growth pathways activated by GDNF can work independently of ERα. Next we tested whether GDNF was necessary to confer endocrine resistance in our model system by using short hairpin RNAs (shRNA) to knockdown GDNF in $G11^{TamR}$ MCF-7 cells. Results show that GDNF depletion (GDNF-KD) reduced *GDNF* mRNA levels by 57.38% (**Figure 3D**) and that these cells were significantly more sensitive to tamoxifen treatment than G11 cells transfected with a scrambled control (**Figure 3E**). Moreover, endocrine resistance could be restored to GDNF-KD G11 cells by the addition of 5 ng/ mL recombinant

GDNF protein (**Figure 3E**), demonstrating that growth inhibition does not reflect an off-target effect of the GDNF shRNA. Taken together, these data demonstrate that *GDNF* plays a central and causal role in establishing endocrine resistance in $G11^{TamR}$ MCF-7 cells.

Having shown that *GDNF* expression promotes endocrine resistance in our MCF-7 cell model, we next asked whether *GDNF* mRNA abundance predicts poor relapse free survival (RFS) using publicly available microarray data (Györffy et al., 2010). Indeed, high *GDNF* expression significantly predicted poor RFS with a hazard ratio of 2.2 (p = 0.028) in one cohort of 88 breast cancer patients (**Figure 3F**). *GDNF* remained significantly correlated with RFS after controlling for expression of ESR1 (ERα), MKI67, and HER2 (ERBB2) using a multivariate analysis (HR = 2.27; $p$ = 0.027). Across 10 sufficiently powered cohorts of patients, GDNF had hazards ratios greater than 1 (i.e., high expression predicts poor RFS outcomes) in seven of these cohorts (mean = 1.758; $p$ = 0.03, two-sided Wilcoxon rank sum test). Moreover, the three studies with significant or borderline significant p-values all had hazards ratios greater than 1 (1.62, 1.75, and 2.2; **Supplementary Table 2**). Taken together, these results suggest a trend in which high transcription of *GDNF* predicts poor RFS in breast cancer patients, possibly suggesting that GDNF plays a role in endocrine resistance in the clinic.

**ER+ breast cancer cells are poised for RET mediated endocrine resistance, but lack RET ligand expression**

Increases in the expression RET tyrosine kinase or its co-receptor GFRα1 are implicated in endocrine resistance (Gattelli et al., 2013; Morandi et al., 2013; Plaza-Menacho et al., 2010). However, RET is itself transcriptionally activated by ERα and is highly abundant in endocrine sensitive ER+ breast cancer cell models (Hah et al., 2011). Analysis of mRNA-seq data from 1,177 primary breast cancers in the cancer genome atlas (TCGA) revealed that RET mRNA expression level was highest in ER+ breast cancer and correlates positively with expression level of *ESR1* (ERα) (Spearman's $\rho$ = 0.51, $p$ < 2.2e-16; **Figure 4A**), suggesting that it is a

direct transcriptional target of ERα *in vivo* as well. *GFRA1* mRNA encodes the GDNF co-receptor, GFRα1, and, together with RET, activates RET-ligand signaling. Further analysis of the mRNA-Seq data set found that *GFRA1* is also strongly correlated with *ESR1* mRNA in breast cancers (Spearman's $\rho$ = 0.67, $p$ < 2.2e-16; **Figure S4A**), suggesting that it is also a direct target of E2 signaling. In our MCF-7 endocrine resistance model, *GFRA1* transcription is 5-higher in TamS MCF-7 cells compared to TamR lines and *RET* transcription is not significantly different (**Figures 4B and 4C**), demonstrating that neither factor is overexpressed in TamR MCF-7 cells. These observations suggest that additional mechanisms beyond a high RET or *GFRA1* expression level cause endocrine resistance in cell models and *in vivo*.

Our finding that recombinant GDNF was sufficient for endocrine resistance in $B7^{TamS}$ MCF-7 cells demonstrates that GDNF is a key limiting factor, whose absence prevents TamS cells from taking on a resistant phenotype. To extend this hypothesis to primary breast cancers, we asked whether GDNF expression is low in general, such that it might limit RET pathway activation in most ER+ breast cancers. Indeed, GDNF expression was detectible in only 565 of 1,177 primary breast cancers (48%) analyzed by TCGA (**Figure S4B**). In principal, RET signaling may be activated by any of the four RET ligands (GDNF, NRTN, ARTN, and PSPN). However, only low levels of NRTN, ARTN, or their co-receptors were detected in primary breast tumors (**Figures 4D and 4E; Figure S4B**). Thus, we conclude that RET ligand expression is low compared with cell surface receptors, especially RET and GFRα1, which are activated in part by ERα. This contrast between RET receptors and ligands supports a model in which the RET signaling pathway is 'poised' for endocrine resistance by expression of the receptors and that limiting levels of *GDNF* expression, or possibly of other RET ligands, ensures endocrine sensitivity in most tumors.

Next we asked whether high RET ligand expression in a subset of ER+ tumors may explain some cases of endocrine resistance. A careful examination of the GDNF expression distribution in TCGA breast cancers revealed a long tail, indicating high GDNF expression in a

handful of cases in the TCGA dataset (**Figure 4E**). Our hypothesis that GDNF expression limits RET-dependent endocrine resistance implies that these GDNF-high samples should be prone to endocrine resistance. We devised a simple non-parametric computational approach, which we call the 'outlier score', to quantify the degree to which GDNF is highly expressed based on the symmetry of the empirical probability density function (see methods; **Figure 4E, blue line**). Based on this score, we conservatively estimate that, of 925 ER+ breast cancer patients in the TCGA dataset, 122 have high expression of at least one of the RET ligands (13%), 57 of which had high levels of GDNF (**Figure 4F**). If our proposed model that RET ligands are the limiting factor for endocrine resistance is accurate, cases with this long tail are those that are more likely to be resistant to endocrine therapies. To test this hypothesis, we analyzed expression microarray data collected prospectively by biopsies of patients that either respond, or do not respond, to the aromatase inhibitor letrozole (Miller et al., 2012). A score comprised of the sum of the outlier scores from all four RET ligands is significantly higher in patients that do not respond to letrozole treatment (p= 0.016, one-sided Wilcoxon rank sum test; **Figure 4G**). By contrast, RET shows no significant difference between patients that respond or do not respond to letrozole. These results suggest that RET ligand expression, but not RET itself, explain the differences in response to letrozole in this cohort of patients.

To further explore whether RET ligands contribute to endocrine resistance in primary breast cancers, we asked whether high expression of RET ligands predict RFS. We have already shown that patients with high expression of GDNF have poor clinical outcomes (**Figure 3F**), and we asked whether these results extend to the other three RET ligands. Expression in the upper quartile of ARTN and NRTN significantly predicts poor RFS with hazards ratios of 1.21 and 1.23, respectively ($p$ = 2.5e-3 and 8.5e-4), consistent with high expression predicting poor clinical outcomes (**Figure S4C**). Expression of PSPN was not significantly associated with RFS ($HR$ = 0.88; $p$ = 0.056). Our re-analyses of clinical samples support the hypothesis that elevated expression of RET ligands, especially *GDNF*, but possibly also *ARTN* or *NRTN*,

317 activate the RET signaling pathway and ultimately cause endocrine resistance in clinical
318 samples, as it does in our MCF-7 cell model. Taken together, our findings support the
319 hypothesis that RET ligands promote endocrine resistance in a clinical setting.
320
321 **GDNF-RET stimulation induces extensive transcriptional changes in MCF-7 cells**
322 We set out to identify the transcriptional targets activated by GDNF-induced RET
323 signaling. To identify both direct and indirect target genes that respond to GDNF-RET, and to
324 distinguish between them, we collected kinetic PRO-seq data following 0, 1, and 24 hours of
325 GDNF treatment in $B7^{TamS}$, $C11^{TamS}$, $G11^{TamR}$, and $H9^{TamR}$ MCF-7 cells. We sequenced PRO-
326 seq libraries to a high read depth (**Table S1**) and verified that biological replicates ($B7^{TamS}$ and
327 $C11^{TamS}$; $G11^{TamR}$ and $H9^{TamR}$) have highly correlated transcriptional patterns across the time
328 course (Spearman's rank correlation $\rho > 0.95$; **Figures S5A and S5B**).
329 We first compared transcriptional changes induced by GDNF between TamS and TamR
330 MCF-7 cells. Because GDNF is both sufficient for resistance in $B7^{TamS}$ and necessary for
331 resistance in $G11^{TamR}$ (**Figures 3C and 3E**), we hypothesized that its effects on gene
332 transcription are also likely to be highly similar in TamS and TamR MCF-7 cells. Consistent with
333 this expectation, transcriptional changes induced by GDNF were highly correlated between
334 TamS and TamR cell lines (Pearson's R > 0.73, $p < 2.2e\text{-}16$; **Figures S5C and S5D**). As
335 expected, transcriptional responses were lower in magnitude in TamR MCF-7 cells following
336 both 1 and 24 hours of GDNF treatment (**Figures S5C and S5D**), likely reflecting a dampened
337 GDNF response in TamR lines due to higher basal levels of GDNF. Given these observations,
338 we focused our downstream analyses on TamS MCF-7 cells.
339 We found that GDNF treatment changed the transcription of 4,921 genes, covering
340 ~15% of expressed transcripts (FDR < 0.01; **Figures 5A and 5B**) at either the 1 or 24 hrs time
341 points. Most targets were regulated immediately in a burst of transcription following 1 hr of
342 GDNF treatment (n = 3,849 at 1hr). Direct targets activated by 1 hr of GDNF treatment included

the immediate early transcription factors, *EGR1* and *ETS2,* which are highly responsive to growth factor signaling (Gregg and Fraizer, 2011; Roberson et al., 1995; Tarcic et al., 2012; Xie et al., 2005), and were up-regulated in this study 60-fold and 4-fold, respectively (**Figure 5A**). Likewise, transcription of *ESR1*, the gene which encodes ERα, was immediately down-regulated by GDNF signaling (**Figure 5A**), which might explain its lower expression in TamR lines. These immediate changes in transcription factor expression levels are likely to establish lasting secondary changes in MCF-7 cells in response to GDNF treatment.

**GDNF treatment stimulates SRF by activating of ERK phosphorylation**

We next asked which intracellular signaling pathways are responsible for immediate transcriptional changes following 1 hr of GDNF signaling. Discriminative motif discovery comparing TREs which change following 1 hr of GDNF treatment to those which are constant identified an 8.7-fold enrichment of motifs recognized by serum response factor (SRF) ($p$ < 2e-5, Fisher's Exact Test) (**Figure 5C**). Previous studies have shown that SRF contributes to the activation of immediate early genes following GDNF treatment in neurons and other cell models (Norman et al., 1988; Schratt et al., 2001), largely in response to ERK phosphorylation (Katz et al., 2007). Consistent with this model, Western blotting found that GDNF treatment robustly and rapidly stimulated ERK phosphorylation in both $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells (**Figure 5D**). In addition to SRF, a motif recognized by AP-1, a heterodimer comprised of FOS, JUN, and ATF family members, was also enriched 2.9-fold ($p$ < 1e-5, Fisher's Exact Test) in TREs which change following GDNF treatment (**Figure 5C**). The enrichment of AP-1 may reflect the direct effects of ERK signaling on AP-1 activation. Alternatively, this enrichment may be due to the upregulation of *FOSL1* (**Figure 5A**), an immediate early gene that was upregulated 16-fold following GDNF treatment. The gene body of *FOSL1* is short enough (8 kb) that the gene may be completely transcribed and translated within minutes of GDNF activation. Taken together, these findings support a model in which GDNF exerts its immediate transcriptional effects by the

369 activation of p-ERK and downstream effects on the SRF and AP-1 transcription factor
370 complexes.

371 Transcription factors may regulate transcription by changing the rates of several steps
372 early during gene transcription (reviewed by (Fuda et al., 2009)). Although Pol II densities
373 increase in the bodies of genes activated by GDNF, the pause peak slightly decreased in both
374 TamS cell lines (**Figure 5E**), suggesting that GDNF increases transcription, in part, by
375 stimulating the rate at which paused RNA Pol II transitions into productive elongation. To test
376 this hypothesis more rigorously, we computed changes in the pausing index between GDNF
377 treated (1 hr) and untreated TamS MCF-7 cells at genes up- or down-regulated by GDNF. To
378 avoid potentially confounding batch effects we enforced the assumption that global pausing
379 levels do not change between different samples, as we have described previously (Danko et al.,
380 2013) (see STAR Methods). Whereas genes that do not undergo changes in gene body
381 transcription had consistent pausing indices between conditions, up-regulated genes were
382 observed to have a lower pausing index after 1 hr of GDNF (**Figure 5F**; $p$ < 2.2e-16 Wilcoxon
383 rank sum test). Likewise, down-regulated genes were observed to have slightly but significantly
384 higher pausing indices ($p$ < 2.2e-16; Wilcoxon rank sum test). These results suggest that GDNF
385 treatment activates and represses genes in part by changing the rate at which Pol II is
386 transitions from a paused state into productive elongation.

387

388 **ESR1 and GDNF-EGR1 form a bi-stable feedback loop**

389 We set out to define the transcriptional regulatory network associated with GDNF-
390 dependent endocrine resistance. The dynamics of gene transcription can rigorously separate
391 direct and indirect target genes following a stimulus (Danko et al., 2013). Genes up-regulated
392 during the first 1 hour following *GDNF* treatment are largely assumed to be direct targets
393 because not enough time has elapsed for transcription, translation, and successive rounds of
394 transcriptional activation. We therefore defined all direct targets of E2 and GDNF signaling as

395 those genes responding by 40 min or 1 hr of treatment, respectively. Secondary targets, defined
396 as transcriptional changes following 24 hrs of GDNF treatment, were assigned to TFs whose
397 transcription changed following 1 or 24 hrs using ChIP-seq data in MCF-7 cells (Euskirchen et
398 al., 2007).

399 The resulting transcriptional regulatory network inferred from the data shows extensive
400 crosstalk between E2 and GDNF signaling programs (**Figure 6A**). We predict that GDNF/RET
401 and E2/ERα form a bi-stable feedback loop in which GDNF immediately (1 hr) inactivates the
402 transcription of ERα and activates transcription of *EGR1*, which, in turn, activates *GDNF*
403 transcription at 24 hrs (**Figures 6A and 6B**). Thus, *GDNF* is an indirect target of GDNF/ RET
404 signaling that reinforces its own activity through a positive feedback loop dependent on EGR1.
405 In turn, *EGR1* transcription is directly down-regulated following 40 min of E2. Thus, GDNF-RET
406 and ERα form a bi-stable feedback loop dependent on EGR1, in which either ERα or
407 GDNF/RET signaling can remain at a high level.

408

409 **GDNF-RET signaling down-regulates the E2/ ERα regulatory program**

410 To validate the transcriptional regulatory network inferred to underlie endocrine
411 resistance, we first focused on validating *ESR1*, which encodes the ERα protein, as a direct and
412 immediate GDNF target gene. PRO-seq data found that *ESR1* undergoes a two-fold decrease
413 in transcription following 1 hr of GDNF treatment and that this transcriptional change is stable
414 through 24 hrs (**Figures 6C and 6D**). These changes in *ESR1* transcription lead to a 2-fold
415 decrease in *ESR1* mRNA abundance following 4 and 24 hours of GDNF treatment (**Figures 6E**).
416 Although changes following 1 hour of GNDF treatment are unlikely to reflect indirect effects, it is
417 nevertheless plausible that a transcription factor encoded by a short gene, such as *FOSL1*, is
418 transcriptionally activated, translated, and responsible for inactivating *ESR1*.

419 To determine whether changes in *ESR1* transcription are a primary or secondary effect
420 of GDNF signaling, we set out to estimate the time at which *ESR1* is down-regulated following

421 the addition of GDNF to the cell culture media. To estimate the time at which the *ESR1*
422 promoter decreases transcriptional activity, we identified the end of the retreating wave of RNA
423 polymerase II 104,000 bp from the transcription start site at 60 min following GDNF treatment
424 (**Figures 6F**). We estimated the elongation rate of *ESR1* in MCF-7 cells to be 1.77 kb/min
425 between 10 and 40 min of E2 treatment using our previous time-course data (Hah et al., 2011).
426 At this elongation rate, we estimated that down-regulation of *ESR1* begins at approximately 1.13
427 min after adding GDNF to the MCF-7 culture media. Likewise, an alternative estimate using the
428 median elongation rate in MCF-7 cells of 2.1 kb/min (Danko et al., 2013) puts the start time at
429 ~10 min and 30 sec after the addition of *GDNF*. Thus, *ESR1* is a direct and immediate target of
430 *GDNF* signaling and is transcriptionally repressed in the minutes following the addition of GDNF
431 to the culture media.

432 To explore the dynamics with which changes in *ESR1* transcription lead to differences in
433 ERα protein level, we used Western blotting to track the abundance of ERα and
434 phosphorylated-ERα protein following the addition of GDNF in B7$^{TamS}$ MCF-7 cells. We found a
435 noticeable effect on ERα protein level as early as two hours after the addition of GDNF and that
436 changes reached their lowest level at 4 hrs (**Figures 6G**).

437 After 24 hrs of GDNF treatment, we found that the down-regulation of ERα protein likely
438 results in the transcriptional down-regulation of E2 target genes. For example, *PGR, GREB1,*
439 *ELOVL2,* and *NOS1AP* are unaffected at 1 hr, but transcriptionally down-regulated between two
440 and four fold following 24 hrs of GDNF treatment (**Figures 6H**). We conformed by qPCR that
441 the GDNF-induced decrease in *PGR* mRNA occurs at 24 hrs but not at 4 hrs (**Figures S6A**).
442 Several lines of genome-wide evidence support the indirect effects of GDNF on ERα target
443 genes as well. First, we find that E2 target genes are more than three-fold enriched in the set of
444 genes responding to GDNF at 24 hrs, but not at 1 hr (**Figure 6I**), and that transcriptional
445 changes at 24 hours of GDNF negatively correlate with 40 min of E2 (Pearson's R= -0.14; $p$ =
446 4.2e-3). Second, the binding motif that was most enriched in TREs that change Pol II

447 abundance following 24 hrs of GDNF treatment was the ERα binding site ($p$ < 1e-9, Fisher's
448 exact test; **Figure S6B**). Taken together, these results demonstrate that GDNF-RET signaling
449 down-regulates the E2 regulatory program within 6 hours of treatment by immediate effects on
450 the transcriptional activity of *ESR1* during the first 10 min of GDNF treatment.
451
452 **GDNF-EGR1 feedback loop results in GDNF activation**
453 We next focused on validating the activation loop between *GDNF* and the transcription
454 factor *EGR1*. Whereas GDNF transcription increased by 16-fold after 24 hrs of GDNF treatment
455 (**Figure 6B**), no changes were found in any of the earlier time points we examined, strongly
456 suggesting that *GDNF* transcription is indirectly activated by GDNF-induced RET signaling.
457 Regarding how GDNF induces its own expression, we predict that GDNF treatment promotes
458 the upregulation of the transcriptional activator, EGR1, which, in turn, binds to the GDNF
459 promoter thus activating GDNF expression. In support of this hypothesis, we found that *EGR1*
460 transcription was upregulated more that 30-fold following 1 hr of recombinant GDNF treatment
461 (**Figure 7A**). These changes in *EGR1* transcription led to an 83-fold increase in *EGR1* mRNA
462 abundance following 4 hrs of GDNF treatment (**Figure 7B**). In further support of our hypothesis,
463 we identified a TRE in the first intron of *GDNF* that is bound by *EGR1* in MCF-7 cells (**Figure**
464 **6B**). Our model also predicts that *EGR1* is directly activated by GDNF signaling, which is likely
465 mediated by an SRF binding site in the *EGR1* promoter (**Figure 7A**). This finding is consistent
466 with the enrichment of SRF binding motifs in TREs responding immediately to GDNF activation,
467 as well as previous reports of SRF activating *EGR1* through a binding site in its promoter in
468 other cell lines (Gregg and Fraizer, 2011). This data suggests that SRF activated by ERK
469 signaling directly up-regulates *EGR1* in MCF-7 cells, leading to a positive feedback loop with
470 GDNF.
471
472 **ERα downregulates *EGR1* transcription**

Our bi-stable feedback loop network model predicts that decreasing ERα activity by tamoxifen treatment should increase the transcription of *GDNF* by increasing *EGR1*. The network model prediction was based on the observation that *EGR1* decreased ~5-20-fold starting at just 10 min of E2 treatment (Hah et al., 2011). To test this prediction, we examined the abundance of *GDNF* mRNA following a time course of tamoxifen treatment. As predicted, tamoxifen significantly increased both *EGR1* and *GDNF* mRNA levels following 24 hours of tamoxifen treatment of $B7^{TamS}$ MCF-7 cells, but not at 40 min or 4 hrs (**Figures 7C and 7D**). Moreover, several lines of evidence suggest that this mutual suppressive relationship between ERα and EGR1 also holds in primary breast cancers. First, we note a highly significant negative correlation between *EGR1* and *ESR1* mRNA abundance among ER+ breast cancers analyzed using TCGA RNA-seq data (Pearson's R= -0.21; $p$ = 2.7e-10; **Figure 7E**). Second, we found that *EGR1* transcription increases substantially in primary tumor biopsies following treatment with the aromatase inhibitor letrozole ($p$ = 1.775e-06, Wilcoxon rank sum test; **Figure 7F**) (Miller et al., 2012). Taken together, several lines of evidence directly implicate ERα in suppressing *EGR1* expression in both MCF-7 cells as well as in primary tumors.

## Discussion

Here we have used genomic tools to reconstruct a gene regulatory network that we demonstrate is responsible for endocrine resistance in an MCF-7 breast cancer model. Our approach is uniquely able to distinguish primary from secondary target genes by using PRO-seq to measure nascent transcription over short (≤1 hr) and long (24 hrs) treatments with E2 and GDNF, two stimuli that are central to our proposed resistance network. Systematic experimental manipulation of GDNF expression in TamS and TamR cell lines all strongly support a causal role of this regulatory network in endocrine resistance. Systematic analysis of publicly available clinical data supports the involvement of this pathway in clinical cases of endocrine resistance and, most importantly, is, to our knowledge, the first study to suggest that expression of RET ligands (GDNF, ARTN, NRTN, and PSPN) are often responsible for RET mediated endocrine resistance in primary tumors. Overall, our study provides mechanistic insights into how growth factor 'escape pathways' become activated and facilitate ERα-independent growth in ER+ breast cancers.

The MCF-7 model of endocrine resistance that we studied here differs in important ways from prior work in other model systems. Most notably, resistance of $G11^{TamR}$ and $H9^{TamR}$ MCF-7 lines to ERα degradation by the small molecule fulvestrant demonstrates that endocrine resistance in our MCF-7 model works independently of ERα. This observation rules out several mechanisms that have been proposed to explain resistance to aromatase inhibitors, which have largely proposed E2-indepenent activation of ERα. For instance, somatic mutations or truncations in the *ESR1* protein coding sequence can result in the constitutive activation of ERα (Thomas and Ke Gustafsson, 2015). Alternatively, changes in the expression of the ERα transcriptional co-activator AIB1 has been associated with clinical and experimental tamoxifen resistance by switching tamoxifen, generally an ERα antagonist in mammary tissue, to an agonist (Osborne et al., 2003; Su et al., 2008). These mechanisms largely depend on the presence of ERα protein in breast cancer cells, and would not be resistant to ERα degradation

514 by fulvestrant. In addition, other lines of evidence rule out these models as well, including our
515 direct genome-wide experimental observations demonstrating that E2 remains an agonist and
516 tamoxifen an antagonist in our MCF-7 model (**Figures 2A and 2B**), as well as a complete lack
517 of genetic changes in ERα protein-coding sequence that are unique to either $G11^{TamR}$ or $H9^{TamR}$
518 cell lines (data not shown).

519 We are also the first to propose that RET-mediated endocrine resistance occurs when
520 ER+ breast cancer cells express the RET ligand GDNF. Work on the RET signaling pathway in
521 endocrine resistance has largely focused on amplifications or increases in the expression of
522 RET or its co-receptor GFRα1 in resistance to aromatase inhibitor inhibitors (Morandi et al.,
523 2013; Plaza-Menacho et al., 2010). However, RET is not significantly different in a cohort of
524 patients resistant to the aromatase inhibitor letrozole (**Figure 4G**), suggesting that other
525 mechanisms may occur more commonly in patients than differences in the expression of RET
526 itself. Indeed, we find that expression of RET and GFRα1 are both highest in ER+ breast
527 cancers, likely because of direct transcriptional activation of both genes by E2/ ERα (**Figure 4A;**
528 **Figure S4A**). Thus, we propose that ER+ breast cancer cells are intrinsically 'poised' for RET
529 mediated endocrine resistance by the activation of RET cell-surface receptors, but lack
530 expression of the ligand GDNF.

531 Based on these findings, we hypothesize that increased expression of any one of the
532 four RET ligands, GDNF, ARTN, NRTN, or PSPN confers endocrine resistance on cells
533 expressing the RET receptor. In support of this model, we report here that our scoring system
534 based on RET ligand overexpression in tumors significantly separates breast cancer patients
535 that respond to letrozole from those who do not (**Figure 4G**). Moreover, we found that RET
536 ligands are predictive of relapse free survival in other cohorts of patents, even after accounting
537 for the expression of other prognostic markers such as ER, PR, and HER2 (**Figure 3F**). Several
538 findings also strongly support the involvement of GDNF in endocrine resistance in our MCF-7
539 model, most notably the observations that GDNF rescues $B7^{TamS}$ lines and that GDNF

knockdown in G11 cells restores sensitivity to tamoxifen (**Figure 3E**). These observations are also supported by existing studies showing that another RET ligand, ARTN, contributes to tamoxifen resistance in MCF-7 cells (Kang et al., 2010), extending and supporting the findings reported here. However, there is one RET ligand that is notably an outlier. PSPN does not appear to have any predictive value in patients, and thus may not play the same role in resistance as the other three RET ligands. This may reflect the extremely low expression of its co-receptor, *GFRA4*, in primary breast cancers (**Figure S4B**), preventing PSPN from having much effect on breast cancer cells. Taken together, these findings suggest that RET ligand expression, especially GDNF, ARTN, and NRTN, explain endocrine resistance in many cases.

One finding that our current study cannot yet explain is that our proposed bi-stable regulatory network between ERα and GDNF/ EGR1 leads to the activation of GDNF in TamS as well as in TamR MCF-7 cells. Under our model, tamoxifen treatment in either TamS or TamR lines leads to the transcriptional activation of GDNF within 24 hours, a prediction of our model which we were able to validate by qPCR (**Figure 7D**). Thus, it remains unclear why endogenous transcription at the GDNF locus is not sufficient to confer endocrine resistance in $B7^{TamS}$ cells. One potential explanation is that higher basal GDNF expression in TamR MCF-7 cells grown in estrogen containing media (**Figures 3A and 3B**) gives TamR lines a “head start” when switching growth signaling programs from ERα to GDNF/RET. Testing this model will require up-regulation of endogenous GDNF in TamS cells, possibly through the use of emerging technologies like activating CRISPR (Perez-Pinera et al., 2013; Thakore et al., 2015).

It is also unclear how RET ligand expression is activated in primary tumors. The abundance of GDNF mRNA appears to be extremely low in primary breast tumors analyzed by TCGA (**Figures 4D, 4E and, S4B**), which were largely collected before therapeutic intervention (Ciriello et al., 2015; The Cancer Genome Atlas Network, 2012). Notably, GDNF is not natively expressed in ER+ TamS MCF-7 cells but rather becomes activated following extended GDNF treatments. This may suggest that GDNF expression is initiated in tumors by another stimulus-

dependent pathway or introduced by another cell type in the tumor microenvironment. This initial bolus of RET ligand might be required to 'prime' GDNF expression in tumor cells, activating the bi-stable feedback loop introduced here. Consistent with this, GDNF expression in tumors may require pro-inflammatory cytokines, such as tumor necrosis factor alpha (TNFα), to be transcribed in breast cancer cells (Esseghir et al., 2007). This finding may link poor survival outcomes in pro-inflammatory tumors (Franco et al., 2015; Zhou et al., 2005) with a GDNF-RET-mediated resistance to endocrine therapy.

Taken together, results reported in the present study reveal a regulatory network that is responsible for GDNF-RET mediated endocrine resistance in MCF-7 cells. Our work also supports this pathway in the development of resistance in primary breast cancers, and specifically supports a model in which RET ligands themselves, rather than expression of the RET receptors, are the primary determinants of resistance in breast cancer cells. Additional prospective clinical studies targeting larger cohorts of patients starting endocrine therapies will be required to fully validate our proposed mechanism of endocrine resistance.

## Author Contributions

The project was conceived by CGD, SAC, and SH. All cell culture and molecular experiments were done by SH, HZ, LJA, CM, and BAM. PRO-seq experiments were conducted by EJR and SH. Genomic data was analyzed by CGD, TC, and SH. Data collection, experiments, and analysis was supervised jointly by CGD and SAC. The paper was written by SH, CGD, and SAC with input from all authors.

## Acknowledgements

We thank X. Yao, L. Lan, as well as all members of the Danko and Coonrod labs for valuable insights and discussions. We also thank J. Lewis for input on the manuscript draft. Work in this publication was supported by NHGRI (National Human Genome Research Institute) grants from the US National Institutes of Health under award number R01 HG009309-01 to CGD. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

## References

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014a). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Andersson, R., Refsing Andersen, P., Valen, E., Core, L.J., Bornholdt, J., Boyd, M., Heick Jensen, T., and Sandelin, A. (2014b). Nuclear stability and transcriptional directionality separate functionally distinct RNA species. Nat. Commun. *5*, 5336.

Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drablos, F., Lennartsson, A., Ronnerblad, M., Hrydziuszko, O., Vitezic, M., et al. (2015). Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science (80-. ). *347*, 1010–1014.

Azofeifa, J.G., and Dowell, R.D. (2016). A generative model for the behavior of RNA polymerase. Bioinformatics btw599.

Benz, C.C., Scott, G.K., Sarup, J.C., Johnson, R.M., Tripathy, D., Coronado, E., Shepard, H.M., and Osborne, C.K. (1992). Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. Breast Cancer Res. Treat. *24*, 85–95.

Boulay, A., Breuleux, M., Stephan, C., Fux, C., Brisken, C., Fiche, M., Wartmann, M., Stumm, M., Lane, H.A., and Hynes, N.E. (2008). The Ret Receptor Tyrosine Kinase Pathway Functionally Interacts with the ER  Pathway in Breast Cancer. Cancer Res. *68*, 3743–3751.

Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., et al. (2006). Genome-wide analysis of estrogen receptor binding sites. Nat Genet *38*, 1289–1297.

Chan, J.Y., LaPara, K., and Yee, D. (2016). Disruption of insulin receptor function inhibits proliferation in endocrine-resistant breast cancer cells. Oncogene *35*, 4235–4243.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Ciriello, G., Gatza, M.L., Beck, A.H., Wilkerson, M.D., Rhie, S.K., Pastore, A., Zhang, H.,

McLellan, M., Yau, C., Kandoth, C., et al. (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell *163*, 506–519.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science (80-. ). *322*, 1845–1848.

Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of transcription start sites from nascent RNA supports a unified architecture of mammalian promoters and enhancers. Nat. Genet. *In Press*.

Coser, K.R., Wittner, B.S., Rosenthal, N.F., Collins, S.C., Melas, A., Smith, S.L., Mahoney, C.J., Shioda, K., Isselbacher, K.J., Ramaswamy, S., et al. (2009). Antiestrogen-resistant subclones of MCF-7 human breast cancer cells are derived from a common monoclonal drug-resistant progenitor. Proc. Natl. Acad. Sci. U. S. A. *106*, 14536–14541.

Danko, C.G., Hah, N., Luo, X., Martins, A.L. AL, Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. Mol. Cell *50*, 212–222.

Danko, C.G., Hyland, S.L., Core, L.J., Martins, A.L., Waters, C.T., Lee, H.W., Cheung, V.G., Kraus, W.L., Lis, J.T., and Siepel, A. (2015). Identification of active transcriptional regulatory elements from GRO-seq data. Nat. Methods *advance on*.

Duarte, F.M., Fuda, N.J., Mahat, D.B., Core, L.J., Guertin, M.J., and Lis, J.T. (2016). Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. Genes Dev. *30*, 1731–1746.

Esseghir, S., Todd, S.K., Hunt, T., Poulsom, R., Plaza-Menacho, I., Reis-Filho, J.S., and Isacke, C.M. (2007). A Role for Glial Cell Derived Neurotrophic Factor Induced Expression by Inflammatory Cytokines and RET/GFR 1 Receptor Up-regulation in Breast Cancer. Cancer Res. *67*, 11732–11741.

Euskirchen, G.M., Rozowsky, J.S., Wei, C.-L., Lee, W.H., Zhang, Z.D., Hartman, S., Emanuelsson, O., Stolc, V., Weissman, S., Gerstein, M.B., et al. (2007). Mapping of

transcription factor binding regions in mammalian cells by ChIP: comparison of array- and sequencing-based technologies. Genome Res. *17*, 898–909.

Fang, F., Yu, M., Cavanagh, M.M., Hutter Saunders, J., Qi, Q., Ye, Z., Le Saux, S., Sultan, W., Turgano, E., Dekker, C.L., et al. (2016). Expression of CD39 on Activated T Cells Impairs their Survival in Older Individuals. Cell Rep. *14*, 1218–1231.

Franco, H.L., Nagari, A., and Kraus, W.L. (2015). TNFα Signaling Exposes Latent Estrogen Receptor Binding Sites to Alter the Breast Cancer Cell Transcriptome. Mol. Cell *58*, 21–34.

Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186–192.

Gattelli, A., Nalvarte, I., Boulay, A., Roloff, T.C., Schreiber, M., Carragher, N., Macleod, K.K., Schlederer, M., Lienhard, S., Kenner, L., et al. (2013). Ret inhibition decreases growth and metastatic potential of estrogen receptor positive breast cancer cells. EMBO Mol. Med. *5*, 1335–1350.

Ghoussaini, M., Edwards, S.L., Michailidou, K., Nord, S., Cowper-Sal Lari, R., Desai, K., Kar, S., Hillman, K.M., Kaufmann, S., Glubb, D.M., et al. (2014). Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation. Nat. Commun. *4*, 4999.

Gonzalez-Malerva, L., Park, J., Zou, L., Hu, Y., Moradpour, Z., Pearlberg, J., Sawyer, J., Stevens, H., Harlow, E., and LaBaer, J. (2011). High-throughput ectopic expression screen for tamoxifen resistance identifies an atypical kinase that blocks autophagy. Proc. Natl. Acad. Sci. U. S. A. *108*, 2058–2063.

Gregg, J., and Fraizer, G. (2011). Transcriptional Regulation of EGR1 by EGF and the ERK Signaling Pathway in Prostate Cancer Cells. Genes Cancer *2*, 900–909.

Győrffy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q., and Szallasi, Z. (2010). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res. Treat. *123*, 725–731.

Hah, N., Danko, C.G., Core, L., Waterfall, J.J., Siepel, A., Lis, J.T., and Kraus, W.L. (2011). A

672 rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer
673 cells. Cell *145*, 622–634.
674 Hah, N., Murakami, S., Nagari, A., Danko, C., and Kraus, W.L. (2013). Enhancer Transcripts
675 Mark Active Estrogen Receptor Binding Sites. Genome Res.
676 Kang, J., Qian, P.X., Pandey, V., Perry, J.K., Miller, L.D., Liu, E.T., Zhu, T., Liu, D.X., and Lobie,
677 P.E. (2010). Artemin is estrogen regulated and mediates antiestrogen resistance in mammary
678 carcinoma. Oncogene *29*, 3228–3240.
679 Katz, M., Amit, I., and Yarden, Y. (2007). Regulation of MAPKs by growth factors and receptor
680 tyrosine kinases. Biochim. Biophys. Acta *1773*, 1161–1176.
681 Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz,
682 M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-
683 regulated enhancers. Nature *465*, 182–187.
684 Kuhn, R.M., Haussler, D., and Kent, W.J. (2013). The UCSC genome browser and associated
685 tools. Brief. Bioinform. *14*, 144–161.
686 Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal
687 how promoters direct initiation and pausing. Science *339*, 950–953.
688 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
689 transform. Bioinformatics *25*, 1754–1760.
690 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
691 dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
692 Mahat, D.B., Kwak, H., Booth, G.T., Jonkers, I.H., Danko, C.G., Patel, R.K., Waters, C.T.,
693 Munson, K., Core, L.J., and Lis, J.T. (2016a). Base-pair-resolution genome-wide mapping of
694 active RNA polymerases using precision nuclear run-on (PRO-seq). Nat. Protoc. *11*, 1455–1476.
695 Mahat, D.B., Salamanca, H.H., Duarte, F.M., Danko, C.G., and Lis, J.T. (2016b). Mammalian
696 Heat Shock Response and Mechanisms Underlying Its Genome-wide Transcriptional
697 Regulation. Mol. Cell *62*, 63–78.

698 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
699 reads. EMBnet.journal *17*, 10–12.
700 Miller, W.R., Larionov, A., Anderson, T.J., Evans, D.B., and Dixon, J.M. (2012). Sequential
701 changes in gene expression profiles in breast cancers during treatment with the aromatase
702 inhibitor, letrozole. Pharmacogenomics J. *12*, 10–21.
703 Mohammed, H., D'Santos, C., Serandour, A.A., Ali, H.R., Brown, G.D., Atkins, A., Rueda, O.M.,
704 Holmes, K.A., Theodorou, V., Robinson, J.L.L., et al. (2013). Endogenous purification reveals
705 GREB1 as a key estrogen receptor regulatory factor. Cell Rep. *3*, 342–349.
706 Mohammed, H., Russell, I.A., Stark, R., Rueda, O.M., Hickey, T.E., Tarulli, G.A., Serandour,
707 A.A., Serandour, A.A.A., Birrell, S.N., Bruna, A., et al. (2015). Progesterone receptor modulates
708 ERα action in breast cancer. Nature *523*, 313–317.
709 Morandi, A., Martin, L.-A., Gao, Q., Pancholi, S., Mackay, A., Robertson, D., Zvelebil, M.,
710 Dowsett, M., Plaza-Menacho, I., and Isacke, C.M. (2013). GDNF-RET signaling in ER-positive
711 breast cancers is a key determinant of response and resistance to aromatase inhibitors. Cancer
712 Res. *73*, 3783–3795.
713 Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and
714 Proudfoot, N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription
715 Coupled to RNA Processing. Cell *161*, 526–540.
716 Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988). Isolation and properties of
717 cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response
718 element. Cell *55*, 989–1003.
719 Osborne, C.K., Bardou, V., Hopp, T.A., Chamness, G.C., Hilsenbeck, S.G., Fuqua, S.A.W.,
720 Wong, J., Allred, D.C., Clark, G.M., and Schiff, R. (2003). Role of the estrogen receptor
721 coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. J. Natl.
722 Cancer Inst. *95*, 353–361.
723 Park, J.W., Liu, M.C., Yee, D., Yau, C., van 't Veer, L.J., Symmans, W.F., Paoloni, M.,

724 Perlmutter, J., Hylton, N.M., Hogarth, M., et al. (2016). Adaptive Randomization of Neratinib in
725 Early Breast Cancer. N. Engl. J. Med. *375*, 11–22.
726 Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore,
727 P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by
728 CRISPR-Cas9–based transcription factors. Nat. Methods *10*, 973–976.
729 Planas-Silva, M.D., and Weinberg, R.A. (1997). Estrogen-dependent cyclin E-cdk2 activation
730 through p21 redistribution. Mol. Cell. Biol. *17*, 4059–4069.
731 Plaza-Menacho, I., Morandi, A., Robertson, D., Pancholi, S., Drury, S., Dowsett, M., Martin, L.-
732 A., and Isacke, C.M. (2010). Targeting the receptor tyrosine kinase RET sensitizes breast
733 cancer cells to tamoxifen treatment and reveals a role for RET in endocrine resistance.
734 Oncogene *29*, 4648–4657.
735 Prall, O.W.., Rogan, E.M., and Sutherland, R.L. (1998). Estrogen regulation of cell cycle
736 progression in breast cancer cells. J. Steroid Biochem. Mol. Biol. *65*, 169–174.
737 Prenzel, T., Begus-Nahrmann, Y., Kramer, F., Hennion, M., Hsu, C., Gorsler, T., Hintermair, C.,
738 Eick, D., Kremmer, E., Simons, M., et al. (2011). Estrogen-Dependent Gene Transcription in
739 Human Breast Cancer Cells Relies upon Proteasome-Dependent Monoubiquitination of Histone
740 H2B. Cancer Res. *71*, 5739–5753.
741 Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic
742 features. Bioinformatics *26*, 841–842.
743 Roberson, M.S., Misra-Press, A., Laurance, M.E., Stork, P.J., and Maurer, R.A. (1995). A role
744 for mitogen-activated protein kinase in mediating activation of the glycoprotein hormone alpha-
745 subunit promoter by gonadotropin-releasing hormone. Mol. Cell. Biol. *15*, 3531–3539.
746 Sariola, H., and Saarma, M. (2003). Novel functions and signalling pathways for GDNF. J. Cell
747 Sci. *116*.
748 Schratt, G., Weinhold, B., Lundberg, A.S., Schuck, S., Berger, J., Schwarz, H., Weinberg, R.A.,
749 Rüther, U., and Nordheim, A. (2001). Serum response factor is required for immediate-early

gene activation yet is dispensable for proliferation of embryonic stem cells. Mol. Cell. Biol. *21*, 2933–2943.

Schwalb, B., Michel, M., Zacher, B., Fruhauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. Science (80-. ). *352*, 1225–1228.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A., and Greene, G.L. (1998). The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell *95*, 927–937.

Su, Q., Hu, S., Gao, H., Ma, R., Yang, Q., Pan, Z., Wang, T., and Li, F. (2008). Role of AIB1 for Tamoxifen Resistance in Estrogen Receptor-Positive Breast Cancer Cells. Oncology *75*, 159–168.

Tarcic, G., Avraham, R., Pines, G., Amit, I., Shay, T., Lu, Y., Zwang, Y., Katz, M., Ben-Chetrit, N., Jacob-Hirsch, J., et al. (2012). EGR1 and the ERK-ERF axis drive mammary cell migration in response to EGF. FASEB J. *26*, 1582–1592.

Team, R.D.C. (2010). R: A language and environment for statistical computing. In R Foundation for Statistical Computing.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods *12*, 1143–1149.

The Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

Thomas, C., and Ke Gustafsson, J.-Å. (2015). Estrogen receptor mutations and functional consequences for breast cancer. Trends Endocrinol. Metab. *26*, 467–476.

776 Wakeling, A.E. (2000). Similarities and distinctions in the mode of action of different classes of
777 antioestrogens. Endocr. Relat. Cancer *7*, 17–28.
778 Wang, Z., Martins, A.L., and Danko, C.G. (2016). RTFBSDB: an integrated framework for
779 transcription factor binding site analysis. Bioinformatics btw338.
780 Welboren, W.-J., van Driel, M.A., Janssen-Megens, E.M., van Heeringen, S.J., Sweep, F.C.,
781 Span, P.N., and Stunnenberg, H.G. (2009). ChIP-Seq of ERalpha and RNA polymerase II
782 defines genes differentially responding to ligands. EMBO J. *28*, 1418–1428.
783 Xie, J., Bliss, S.P., Nett, T.M., Ebersole, B.J., Sealfon, S.C., and Roberson, M.S. (2005).
784 Transcript Profiling of Immediate Early Genes Reveals a Unique Role for Activating
785 Transcription Factor 3 in Mediating Activation of the Glycoprotein Hormone α-Subunit Promoter
786 by Gonadotropin-Releasing Hormone. Mol. Endocrinol. *19*, 2624–2638.
787 Zhong, Z., Shan, M., Wang, J., Liu, T., Xia, B., Niu, M., Ren, Y., and Pang, D. (2015). HOXD13
788 methylation status is a prognostic indicator in breast cancer. Int. J. Clin. Exp. Pathol. *8*, 10716–
789 10724.
790 Zhou, Y., Eppenberger-Castori, S., Marx, C., Yau, C., Scott, G.K., Eppenberger, U., and
791 Benz, C.C. (2005). Activation of nuclear factor-κB (NFκB) identifies a high-risk subset of
792 hormone-dependent breast cancers. Int. J. Biochem. Cell Biol. *37*, 1130–1144.

## Main figure legends

### Figure 1: ER target genes are uniquely expressed in TamS cells.

(A) Cell viability of tamoxifen sensitive (TamS; $B7^{TamS}$ and $C11^{TamS}$) and resistant (TamR; $G11^{TamR}$ and $H9^{TamR}$) MCF-7 cells upon treatment with 0 (vehicle; EtOH), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M of tamoxifen for 4 days. Data are represented as mean ± SEM (n=3).

(B) Experimental setup for PRO-seq. PRO-seq libraries were prepared from all four cell lines grown in the absence of tamoxifen for 3 days.

(C) Spearman's rank correlation of RNA polymerase abundance in the gene bodies (+1000 bp to the annotation end) of TamS and TamR cell lines.

(D) MA plot showing significantly changed genes (red) that are higher in TamS (top) or TamR (bottom) MCF-7 lines. Genes highlighted in the plots which are ERα targets are highlighted in blue.

(E) Transcription near the *PGR* and *GREB1* loci in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on the sense and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. Enhancers and promoters are shown in grey and light green shading, respectively. Arrows indicate the direction of gene annotations.

(F) *PGR* and *GREB1* mRNA expression levels in $B7^{TamS}$ and $G11^{TamR}$ cells. Data are represented as mean ± SEM (n=3 for *PGR*; n = 4 for *GREB1*). **** $p < 0.0001$. $G11^{TamR}$ is normalized to $B7^{TamS}$.

(G) Boxplots represent fold-change between TamS and TamR of genes that are either up- or down-regulated following 40 minutes of estrogen (E2) in Hah et. al. (2011). Spearman's Rho= 0.185, $p < 2.2e-16$.

(H) Motifs enriched in TREs that have different amounts of RNA polymerase between TamS and TamR cells compared with TREs that have consistent levels.

**Figure 2: Tamoxifen resistant lines have functional ERα signaling**

(A) Heatmap of changes in RNA polymerase abundance following 40 minutes of E2 or tamoxifen treatment near ERα bindings sites in $B7^{TamS}$ and $G11^{TamR}$ cells.

(B) Violin plots show fold changes in the indicated MCF-7 clone following 40 minutes of E2 or tamoxifen tamoxifen treatment at genes up- or down-regulated by E2 in Hah et. al. (2011). Up- and down-regulated genes are in green and blue, respectively.

(C) *ESR1* mRNA expression levels in $B7^{TamS}$ and $G11^{TamR}$ cells. Data are represented as mean ± SEM (n=3). **** $p < 0.0001$.

(D) Cell viability of TamS and TamR cells upon treatment with 0 (vehicle; DMSO), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M fulvestrant (ER degrader) for 4 days. Data are represented as mean ± SEM (n=3).

**Figure 3: *GDNF* is responsible for tamoxifen resistance in MCF-7 cells**

(A) Transcription near the *GDNF* locus in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. The region near the GDNF promoter is shown in light green shading. Arrow indicates the direction of gene annotations.

(B) *GDNF* mRNA expression levels in $B7^{TamS}$ and $G11^{TamR}$ cells. Data are represented as mean ± SEM (n=3). ** $p < 0.005$.

(C) Cell viability of $B7^{TamS}$ cells in the presence or absence of 10 ng/ml GDNF and/or 100 mM tamoxifen for 4 days. Data are represented as mean ± SEM (n=3). * $p < 0.05$, *** $p < 0.0005$.

(D) *GDNF* mRNA expression levels in G11$^{TamR}$ scrambled (SCR) and G11$^{TamR}$ GDNF knockdown (GDNF-KD) cells. Data are represented as mean ± SEM (n=3). **** $p < 0.0001$.

(E) Relative cell number of G11$^{TamR}$ scrambled (SCR) and G11$^{TamR}$ GDNF knockdown (GDNF-KD) cells after 4 days without or with 5 μM tamoxifen and/or 5 ng/ml GDNF treatment. Data are represented as mean ± SEM (n=9). * $p < 0.05$.

(F) Kaplan Meier (KM) plot of relapse free survival (RFS) in a cohort of 88 breast cancer patients with low (black) or high (red) GDNF expression.

**Figure 4: Expression of RET ligands contributes to endocrine resistance.**

(A) Density scatterplot showing *RET* and E*SR1* expression in mRNA-seq data from 1,177 primary breast cancer models in the cancer genome atlas (TCGA). Spearman's ρ = 0.51, *p* = 1.2e-60.

(B) Transcription near the *RET* locus in B7$^{TamS}$ and G11$^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. Enhancers and promoters are shown in grey and light green shading, respectively. Arrow indicates the directional movement of transcribed genes.

(C) Dot plot shows *RET* transcription levels in TamS and TamR MCF-7 cells.

(D) Density scatterplots show the expression of RET ligands (*GDNF*, *NRTN*, *ARTN*, and *PSPN*) versus *ESR1* based on mRNA-seq data from 1,177 primary breast cancers.

(E) RET ligand expression distribution in ER+ breast cancers. The dotted blue line represents 2.5 times the range between the 25$^{th}$ and 50$^{th}$ percentile.

(F) Fraction of ER+ breast cancers (n = 925) with at least one RET ligand exceeding the threshold shown in panel E (shown in dark blue, n = 122). Among the 4 RET ligands, GDNF was the most highly expressed (n = 60).

(G) Boxplots show RET ligands score and RET expression levels in patients that respond or do not respond to aromatase inhibitor letrozole. * $p$ = 0.016.

**Figure 5: GDNF activates thousands of target genes at the level of pause release.**

(A-B) MA plot shows significantly upregulated and downregulated genes (red) following 1 hour (A) or 24 hours (B) of GDNF treatment in TamS MCF-7 cells.

(C) Motifs enriched in TREs that have different amounts of RNA polymerase following 1 hour of GDNF treatment compared with TREs that have consistent levels.

(D) Immunoblot analysis of p-ERK in serum deprived B7$^{TamS}$ and G11$^{TamR}$ cells treatment with 10 ng/mL GDNF.

(E) Heatmap depicting changes in RNA polymerase density following 1 hour of GDNF treatment in B7$^{TamS}$ MCF-7 cells.

(F) Changes in pausing index between treated (1 hour) and untreated TamS MCF-7 cells at the indicated class of genes.

**Figure 6: Bi-stable feedback loop between *ESR1, EGR1,* and *GDNF*.**

(A) Transcriptional regulatory network of GDNF-dependent endocrine resistance highlighting the bi-stable feedback loop inferred between *ESR1*, *EGR1*, and *GDNF*. Each point represents a gene regulated following 1 or 24 hours of GDNF signaling. Only transcription factors or signaling molecules are shown. Blue and red edges represent activation or repression relationships, respectively.

(B) Transcription near the *GDNF* locus in B7$^{TamS}$ and G11$^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. The promoter is shown in light green shading. Arrows indicate the direction encoding annotated genes.

(C) Dot plots of transcription levels of *ESR1* B7$^{TamS}$ and G11$^{TamR}$ cells following GDNF

892 treatment.

893 (D) Transcription in the *ESR1* gene in B7$^{TamS}$ and G11$^{TamR}$ cells. PRO-seq densities on
894 sense strand and anti-sense strand are shown in red and blue, respectively. dREG
895 scores are shown in green. Enhancers and promoters are shown in grey and light green
896 shading, respectively. Arrow indicates the direction encoding annotated genes.

897 (E) *ESR1* mRNA expression levels in B7$^{TamS}$ cells following 10 ng/mL GDNF treatment. Data
898 are represented as mean ± SEM (n=3). **** p <0.0001.

899 (F) Difference in read counts in 3kb windows along *ESR1* between 1 hours of GDNF and
900 untreated TamS MCF-7 cells, The location of the wave of RNA polymerase along *ESR1*
901 was identified using a hidden Markov model and is represented by the yellow box.

902 (G) Immunoblot analysis of ERα and p-ERα in B7$^{TamS}$ and G11$^{TamR}$ cells treatment with 10
903 ng/mL for 0, 1, 2, and 4 hours.

904 (H) Dot plots representing transcription levels of *ERα* target genes (*PGR*, *GREB1*, *ELOVL2*,
905 and *NOS1AP*) following a timecourse of GDNF treatment.

906 (I) Bar plot showing the fraction of genes whose transcription is up-regulated by 40 min. of
907 E2 in all RefSeq annotated genes (left) or those which are downregulated by 1 (center)
908 or 24 hours (right) of GDNF treatment. E2 target genes were enriched in those down-
909 regulated following 24 hrs of GDNF treatment. The Y axis denotes the fraction of genes
910 that are direct up-regulated E2 targets (defined based on Hah et. al. (2011) and also up-
911 regulated in B7$^{TamS}$). # p = 1.098e-10, ## p= 6.556999e-19. Fisher's exact test was used
912 for statistical analysis.

913

914 **Figure 7: Validation of bi-stable feedback loop in MCF-7 cells and primary breast tumors**

915 (A) Transcription at the *EGR1* locus in B7$^{TamS}$ and G11$^{TamR}$ cells before and after treatment
916 with GDNF. PRO-seq densities on sense strand and anti-sense strand are shown in red
917 and blue, respectively. dREG scores are shown in green. The number of reads mapping

in EGR1 and SRF ChIP-seq data is shown in black. Arrow indicates the direction of annotated genes.

(B) *EGR1* mRNA expression level in $B7^{TamS}$ cell after treatment with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). ** $p < 0.01$, *** $p \leq 0.001$.

(C) *EGR1* mRNA expression level in $G11^{TamR}$ cells after treatment without (DMSO) or with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). * $p < 0.05$.

(D) *GDNF* mRNA expression levels in $G11^{TamR}$ cells after treatment without (DMSO) or with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). ** $p < 0.005$.

(E) Density scatterplots show the expression of *EGR1* versus *ESR1* based on mRNA-seq data from 1,177 primary breast cancers. ER+ breast cancers (n= 925), defined based on ESR1 expression (>1e-5), are highlighted in color. The trend line was calculated using Deming regression in the ER+ breast cancers (Pearson's R= -0.21; $p$ = 2.7e-10).

(F) Boxplots show *EGR1* expression level before or following 90 days of treatment with letrozole ($p$ = 1.8e-6, Wilcoxon Rank Sum Test).

## Supplemental figure legends

### Figure S1: dREG identifies highly enriched active enhancers and promoter marks in MCF-7 cells

(A) Heatmap depicting PRO-seq, Dnase-I-seq, H3K27ac, and H3K4me3 near 39,753 transcriptional regulatory elements (TREs) identified using dREG-HD from PRO-seq data (left) in TamS and TamR MCF-7 cells.

(B) Transcription and dREG scores in the locus near the *CCND1* gene in $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells.

(C) Luciferase activity in $B7^{TamS}$ and $G11^{TamR}$ MCF-7 cells in the presence of an enhancer located approximately 300kb downstream of *CCND1*. All data normalized to renilla control. Data are represented as mean ± SEM (n=3). ** $p < 0.01$, **** $p < 0.0001$.

### Figure S2: PRO-seq densities are unaffected after tamoxifen treatment

(A-B) Density scatterplot showing the correlation of PRO-seq densities between tamoxifen treated and untreated $B7^{TamS}$ (A) and $G11^{TamR}$ (B) MCF-7 cells.

### Figure S3: GDNF induces fulvestrant resistance in TamS cells

(A) Cell viability of $B7^{TamS}$ cells in the presence or absence of 10 ng/ml GDNF and/or 100 mM fulvestrant for 4 days. Data are represented as mean ± SEM (n=3). ** $p < 0.005$, **** $p < 0.0001$.

### Figure S4: RET ligand expression is low compared to RET and GFRα1 receptors

(A) Density scatterplot showing the relationship between *GFRA1* and E*SR1* expression levels in 1,177 primary breast cancer samples in the cancer genome atlas (TCGA). Pearson's R = 0.52; $p < 2.2e\text{-}16$.

(B) Violin plots depicting the absolute normalized expression level of receptor-tyrosine kinase receptors and ligands in 1,177 primary breast cancer samples (TCGA). For each color, the pair of genes represents receptor (left) and ligand (right). Gray represents the *RET* gene which encodes the RET tyrosine kinase receptor required for signal transduction of all four RET ligands.

(C) Kaplan Meier (KM) plots of survival probability in a cohort of breast cancer patients with low (black) or high (red) *NRTN* or *ARTN* expression. Patients are split based on the upper quartile of RET ligand expression.

**Figure S5: Highly correlated transcriptional patterns in biological replicates across the time course**

(A) Density scatterplot showing global transcription levels between TamS (B7 and C11; top) or TamR (G11 and H9; bottom) MCF-7 cell lines at 0, 1, or 24 hours GNDF treatment. The Spearman's rank correlation ($\rho$) values are shown for each plot.

(B) Heatmap shows Spearman's rank correlation of RNA polymerase abundance of TamS and TamR lines between the indicated samples. Sample order is determined by hierarchical clustering. Colorscales show 0, 1, or 24 hours of GDNF treatment (above) or TamS or TamR (right) as shown below the heatmap.

(C-D) Scatter plots depict transcriptional changes between TamS and TamR MCF-7 cells following 1 hour (C) or 24 hours (D) of GDNF treatment.

**Figure S6: GDNF causes decrease in *PGR* mRNA expression and ERα binding sites**

(A) *PGR* mRNA expression level in G11$^{TamR}$ cells after treatment without (water) or with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). **** p < 0.0001.

(B) Motifs enriched in TREs that have different amounts of RNA polymerase before and after 24 hours of GDNF treatment.

## STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHODS DETAILS
    - Cell Lines and Cell Culture
    - Cell Viability Assay
    - Cell Culture Set Up and Nuclei Isolation
    - Nuclear Run-on and PRO-seq Library Preparation
    - Mapping of PRO-seq Sequencing Reads
    - Identification of Active enhancers and Promoters Using dREG-HD
    - Differential Expression Analysis (DESeq2)
    - Motif Enrichment Analysis
    - TCGA Data Analysis
    - Letrozole Microarray Reanalysis
    - Pausing Analysis

1 **STAR Methods**

2 **KEY RESOURCES TABLE**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| anti-p-ERK | Cell Signaling | Cat# 4695 |
| anti-ERα | Santa Cruz | Cat# sc-543 |
| anti-p-ER | Cell Signaling | Cat# 2511 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| (Z)-4-Hydroxytamoxifen (4-OHT) | Sigma-Aldrich | Cat# H7904 |
| Fulvestrant | Sigma-Aldrich | Cat# I4409 |
| Recombinant human GDNF | PeproTech | Cat# 450-10 |
| SUPERase In RNase Inhibitor (20 U/L) | Life Technologies | Cat# AM2694 |
| Protease Inhibitor Cocktail | Roche | Cat# 11836153001 |
| Biotin-11-ATP | PerkinElmer | Cat# NEL544001EA |
| Biotin-11-GTP | PerkinElmer | Cat# NEL545001EA |
| Biotin-11-CTP | PerkinElmer | Cat# NEL542001EA |
| Biotin-11-UTP | PerkinElmer | Cat# NEL543001EA |
| Sarkosyl | Fisher Scientific | Cat# AC612075000 |
| Trizol | Life Technologies | Cat# 15596-026 |
| Trizol LS | Life Technologies | Cat# 10296-010 |
| GlycoBlue | Ambion | Cat# AM9515 |
| Hydrophilic streptavidin magnetic beads | NEB | Cat# S1421S |
| RppH | NEB | Cat# M0356S |
| T4 RNA Ligase 1 | NEB | Cat# M0204L |
| Critical Commercial Assays | | |
| RNeasy Kit | Qiagen | Cat# 74104 |
| High Capacity RNA-to-cDNA | Applied Biosystems | Cat# 4387406 |
| Power SYBR Green PCR Master Mix | Appplied Biosystems | Cat# 4367659 |
| Deposited Data | | |
| All genomic data was deposited in GEO and the sequence read archive | Herein | GSE93229 |
| | | |
| | | |

| | | |
|---|---|---|
| | | |
| Experimental Models: Cell Lines | | |
| MCF7-B7$^{TamS}$ | (Gonzalez-Malerva et al., 2011) | N/A |
| MCF7-C11$^{TamS}$ | (Gonzalez-Malerva et al., 2011) | N/A |
| MCF7-G11$^{TamR}$ | (Gonzalez-Malerva et al., 2011) | N/A |
| MCF7-H9$^{TamR}$ | (Gonzalez-Malerva et al., 2011) | N/A |
| Recombinant DNA | | |
| Plasmid for control shRNA | Sigma-Aldrich | Cat# SHC002 |
| Plasmid for *GDNF* shRNA | Sigma-Aldrich | Cat# SHCLND-NM_000514 |
| pLKO.1 shRNA plasmid | Addgene | Plasmid #1864 |
| psPAX2 packaging plasmid | Addgene | Plasmid #12260 |
| pMD2.G envelope plasmid | Addgene | Plasmid #12259 |
| Sequence-Based Reagents | | |
| Primers for *ACTB*, see STAR Methods | This paper | N/A |
| Primers for *PGR*, see STAR Methods | This paper | N/A |
| Primers for *GREB1*, see STAR Methods | This paper | N/A |
| Primers for *ESR1*, see STAR Methods | This paper | N/A |
| Primers for *GDNF*, see STAR Methods | Boulay et al., 2008 | N/A |
| Primers for *EGR1*, see STAR Methods | Fang et al., 2016 | N/A |
| Software and Algorithms | | |
| cutadapt | Martin, 2011 | |
| dREG | Danko et al., 2015 | https://github.com/Danko-Lab/dREG |
| dREG-HD | Manuscript in preparation; This paper | https://github.com/Danko-Lab/dREG.HD; |
| bigWig software package | | https://github.com/andrelmartins/bigWig |
| Visualization using R | Team, 2010 | |
| BedTools | Quinlan and Hall, 2010 | |
| bedGraphToBigWig program in the Kent Source software package | Kuhn et al., 2013 | |
| DEseq2 | Love et al., 2014 | |

| | | |
|---|---|---|
| RTFBSDB | Wang et al., 2016 | |
| Cytoscape software package | Shannon et al., 2003 | |
| GraphPad Prism | | |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Charles G. Danko, Ph.D.
Baker Institute for Animal Health
Cornell University
Hungerford Hill Rd.
Ithaca, NY 14853, USA
Phone: (607) 256-5620
E-mail: dankoc@gmail.com

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Cell Lines and Cell Culture

Tamoxifen sensitive (TamS; $B7^{TamS}$ and $C11^{TamS}$) and resistant (TamR; $G11^{TamR}$ and $H9^{TamR}$) MCF-7 cells (Gonzalez-Malerva et al., 2011) were a gift from Dr. Joshua LaBaer. TamS cells were grown in Dulbecco's Modified Eagle Medium supplemented with 5% fetal bovine serum and 1% Penicillin Streptomycin, and TamR cells were grown in the same media supplemented with 1 μM tamoxifen. Tamoxifen used throughout in this paper is (Z)-4-Hydroxytamoxifen (4-OHT; Sigma-Aldrich; Cat# H7904).

## METHODS DETAILS

### Cell Viability Assay

Briefly, 5 x $10^3$ TamS and TamR cells were grown in 24-well TC-treated plates in their specific culture media. After letting the cells adhere to the plate for 24 hours, cells were rinsed with PBS three times to remove any residual tamoxifen. The cells were treated with either increasing dosage of tamoxifen (0 (vehicle control; EtOH), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M) or

fulvestrant (Sigma-Aldrich; Cat# I4409) (0 (vehicle control; DMSO), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M).

For setting up the rescue experiment with GDNF (PeproTech; Cat# 450-10), 5 x $10^3$ $B7^{TamS}$ cells were grown in 24-well TC-treated plates in their specific culture media. After letting the cells adhere to the plate for 24 hours, cells were treated with either EtOH (vehicle), $10^{-7}$ M tamoxifen, $10^{-7}$ M tamoxifen and 10 ng/mL GDNF, or 10 ng/mL GDNF treatment. The same set up was performed for $10^{-7}$ M treatment of fulvestrant and using DMSO (vehicle) as a control.

After four days of endocrine treatment cells were fixed with 4% paraformaldehyde and stained with 0.5% crystal violet solution made in 25% methanol. After washing away non-specific crystal violet stain with PBS, we took pictures of each plate and the crystal violet stain from the fixed-cells was removed using 10% acetic acid. The absorbance was measured using the Tecan plate reader at $OD_{595nm}$. Samples were normalized to the untreated control. Three biological replicates were performed and data are represented as mean ± SEM.

**Cell Culture Set Up and Nuclei Isolation**

TamS and TamR lines were grown in 150mm TC-treated culture dish in their respective normal culture media. Cells were rinsed with PBS for at least three times 24 hours after plating. Both the TamS and TamR cells were grown in Dulbecco's Modified Eagle Medium supplemented with 5% fetal bovine serum and 1% Penicillin Streptomycin for an additional three days until ~80% confluency in the absence of tamoxifen, in order to measure the difference between TamS and TamR cells pre-treatment. For estrogen (E2) and tamoxifen treated TamS and TamR cells, cells were instead grown in phenol-red free Dulbecco's Modified Eagle Medium supplemented with 5% charcoal-stripped fetal bovine serum and 1% Penicillin Streptomycin for three days. The cells were then treated with either EtOH (vehicle control), 100 nM E2, or with 1 μM tamoxifen for 40 min. For GDNF treated TamS and TamR cells, the same experimental set

53 up as the pre-treatment was performed. In addition, the cells were treated with 10 ng/mL GDNF
54 for 0, 1, or 24 hours.

55 Nuclei were isolated as described in previously (Core et al., 2008). Briefly, cells were
56 rinsed three times with ice-cold PBS and lysed using lysis buffer (10 mM Tris-HCl pH 7.4, 2 mM
57 $MgCl_2$, 3 mM $CaCl_2$, 0.5% NP-40, 10% Glycerol, 1 mM DTT, 1X PIC (Roche; Cat#
58 11836153001), and 1 μl/10 mL SUPERase-In (ThermoFisher; Cat# AM2694) dissolved in DEPC
59 water). Cells were homogenized by gently pipetting at least 30 times and the nuclei were
60 harvested by centrifugation at 1000g for five minutes at 4°C. The isolated nuclei were washed
61 twice with lysis buffer and were resuspended in 100 μL freezing buffer (50 mM Tris HCl pH 8.3,
62 5 mM $MgCl_2$, 40% Glycerol, 0.1 mM EDTA pH 8.0, and 4 U/mL SUPERase-In). The isolated
63 nuclei were used for nuclear run-on and precision nuclear run-on sequencing (PRO-seq) library
64 preparation.

65

66 **Nuclear Run-on and PRO-seq Library Preparation**

67 Nuclear run-on experiments were performed according to the methods described
68 previously by (Kwak et al., 2013; Mahat et al., 2016a). $1x10^7$ nuclei in 100 μL freezing buffer
69 were mixed with 100 μL of 2x nuclear run-on buffer (10 mM Tris-HCl pH 8.0, 5 mM MgCl2, 1
70 mM DTT, 300 mM KCl, 50 μM biotin-11-ATP (Perkin Elmer; Cat# NEL544001EA), 50 μM biotin-
71 11-GTP (Perkin Elmer; Cat# NEL545001EA), 50 μM biotin-11-CTP (Perkin Elmer Cat#
72 NEL542001EA), 50 μM biotin-11-UTP (Perkin Elmer; Cat# NEL543001EA), 0.4 units/μL
73 SUPERase In RNase Inhibitor (Life Technologies; Cat# AM2694), 1% Sarkosyl (Fisher
74 Scientific; Cat# AC612075000). The mixture was incubated at 37 °C for five minutes. The biotin
75 run-on reaction is stopped using Trizol (Life Technolgies; Cat# 15596-026), Trizol LS (Life
76 Technologies; Cat# 10296-010) and pelleted. The use of GlycoBlue (Ambion; Cat# AM9515) is
77 recommended for higher pellet yield. RNA pellets were re-dissolved in DEPC water and

denatured in 65 °C for 40 seconds and hydrolyzed in 0.2 N NaOH on ice for 10 minutes to have a hydrolyzed RNA length with that range ideally of 40 to 100 nts. Bead binding (NEB; Cat# S1421S) is performed to pull down nascent RNAs followed by 3' RNA adaptor ligation (NEB; Cat# M0204L). Another bead binding is performed followed by 5' de-capping using RppH (NEB; Cat# M0356S). 5' phosphorylation is performed followed by 5' adaptor ligation. The last bead binding is performed before generation of cDNA by reverse transcription. PRO-seq libraries were prepared according to manufacturers' protocol (Illumina) and were sequenced using the Illumina NextSeq500 sequencing.

**Mapping of PRO-seq Sequencing Reads**

PRO-seq reads failing Illumina quality filters were removed. Adapters were trimmed from the 3' end of remaining reads using cutadapt with a 10% error rate (Martin, 2011). Reads were mapped with BWA (Li and Durbin, 2009) to the human reference genome (hg19) and a single copy of the Pol I ribosomal RNA transcription unit (GenBank ID# U13369.1). The location of the RNA polymerase active site was represented by a single base which denotes the 3' end of the nascent RNA, which corresponds to the position on the 5' end of each sequenced read. Mapped reads were normalized to reads per kilobase per million mapped (RPKM) and converted to bigWig format using BedTools (Quinlan and Hall, 2010) and the bedGraphToBigWig program in the Kent Source software package (Kuhn et al., 2013). Downstream data analysis was preformed using the bigWig software package, available from: https://github.com/andrelmartins/bigWig. All data processing and visualization was done in the R statistical environment (Team, 2010).

**Identification of Active Enhancers and Promoters using dREG-HD**

We identified TREs using dREG (Danko et al., 2015). Data collected from all four cell lines (TamR and TamS MCF-7 cells) or between different time points (GDNF treatment) was combined to increase statistical power for the discovery of a superset of TREs active during any of the conditions examined.

The precise coordinates of TREs were refined using a strategy that we call dREG-HD (available at https://github.com/Danko-Lab/dREG.HD; manuscript in preparation). Briefly, dREG-HD uses an epsilon-support vector regression (SVR) with a Gaussian kernel to map the distribution of PRO-seq reads to DNase-I signal intensities. Training was conducted on randomly chosen positions within dREG peaks in K562 cells (GEO ID# GSM1480327) extended by 200bp on either side. We selected the optimal set of features based on maximizing the Pearson correlation coefficient between the imputed and experimental DNase-I signal intensity over an independent validation set. Before DNase-I imputation, PRO-seq data was preprocessed by normalizing read counts to the sequencing depth and scaled such that the maximum value is within the 90 percentile of the training examples. To identify peaks, we smoothed the imputed DNase-I signal using a cubic spline and identified local maxima. We tuned the performance of the peak calling by empirically optimizing two free parameters that control the (1) smoothness of spline curve fitting, and (2) a threshold level on the intensity of the imputed DNase-I signal. Parameters were optimized to achieve <10% false discovery rates on a K562 training dataset by a grid optimization over free parameters. We tested the optimized dREG-HD model (including both DNase-I imputation and peak calling) a GRO-seq dataset completely held out from model training and parameter optimization in on GM12878 lymphoblastoid cell lines (GSM1480326). Testing verified that dREG-HD identified transcribed DNase-I hypersensitive sites with 82% sensitivity at a 10% false discovery rate.

Additional genomic data in MCF-7 cells generated by the ENCODE project was downloaded from Gene Expression Omnibus. TREs discovered using dREG-HD were

compared with ChIP-seq for H3K27ac and H3K4me3 (accession numbers: GSM945854 and GSM945269) and DNase-1 hypersensitivity (GSM945854).

**Differential Expression Analysis (DESeq2)**

We compared between treatment conditions or cell lines using gene annotations (GENCODE v19). We counted reads in the interval between 1,000 bp downstream of the annotated transcription start site to the end of the gene for comparisons between TamS and TamR cell clones. When comparing gene expression between GDNF treated and untreated MCF-7 cells we counted reads in the window between 1,000 bp downstream of the transcription start site and the end of the annotation or 60,000 bp into the gene body (whichever was shorter). This window was selected to avoid (1) counting reads in the pause peak near the transcription start site, and (2) to focus on the 5' end of the gene body affected by changes in transcription during 60 minutes of GDNF treatment assuming a median elongation rate of 2 kb/ minute. We limited analyses to gene annotations longer than 2,000 bp in length. To quantify transcription at enhancers, we counted reads on both strands in the window covered by each dREG-HD site. Differential expression analysis was conducted using deSeq2 (Love et al., 2014) and differentially expressed genes were defined as those with a false discovery rate (FDR) less than 0.01.

**Motif Enrichment Analysis**

Motif enrichment analyses were completed using the default set of 1,964 human motifs in RTFBSDB (Wang et al., 2016) clustered into 622 maximally distinct DNA binding specificities (see ref Wang et. al. (2016)). We selected the motif to represent each cluster whose canonical transcription factor is most highly transcribed in MCF-7 cells. We fixed the motif cutoff log odds ratio of 7.5 (log *e*) in a sequence compared with a third-order Markov model as background. We

identified motifs enriched in dREG-HD TREs that change transcription abundance between two conditions using Fisher's exact test with a Bonferroni correction for multiple hypothesis testing. TREs were compared to a background set of >1,500 GC-content matched TREs that do not change transcription levels (<0.25 absolute difference in magnitude (log-2 scale) and $p > 0.2$) using the enrichmentTest function in RTFBSDB (Wang et al., 2016).

**TCGA Data Analysis**

Processed and normalized breast cancer RNA-seq data was downloaded from the International Cancer Genome Consortium (ICGC) data portal website (https://dcc.icgc.org). Data profiling each gene was extracted using shell scripts. Processing and visualization was done in R.

**Letrozole Microarray Reanalysis**

We reanalyzed Affymetrix U133A microarray data profiling mammary tumor biopsies before and after treatment with letrozole (Miller et al., 2012). Miller et. al. (2012) collected data from mammary tumor biopsies prior to letorozle treatment, 10-14 days following the start of treatment, and 90 days following the start of treatment. Samples were annotated as a "responder" (i.e., responds to letrozole treatment), a "non-responder" (i.e., no benefit from letrozole treatment), or "not assessable" (i.e., unknown). The Series Matrix Files were downloaded from Gene Expression Omnibus (GSE20181) and each gene of interest was extracted and processed into a text file. We used the following Affymetrix ID numbers 221359_at, 210683_at, 210237_at, 221373_x_at, 211421_s_at, and 201694_s_at to represent *GDNF*, *NRTN*, *ARTN*, *PSPN*, *RET*, and *EGR1*, respectively. We found no evidence of differences in RET or RET ligand expression across the three time points, and we therefore used the average expression of each RET ligand in each sample when comparing between responsive and non-responsive patients in order to decrease assay noise.

Outlier scores were designed to score the degree to which each sample falls within the tail of the distribution representing high expression levels of each RET ligand (as shown in **Fig. 4E**). Because endocrine resistance could, in principal, be caused either by high expression of any individual RET ligand on its own, or by moderately high expression of multiple RET ligands in combination, we devised a data transformation and sum approach to score the degree to which all four of the RET ligands were highly expressed in each sample. In our data transformation, expression levels were centered by the median value and scaled based on the lower tail of the expression distribution (between quartile 0 and 50). This approach is similar in concept to a Z-score transform, but uses the lower tail to estimate the variance in order to avoid having high expression levels, which we hypothesize here may contribute to endocrine resistance, from contributing to the denominator used to standardize the distribution of each RET ligand. After transforming scores from all four RET ligands separately, we took the sum of the scores to represent our final 'outlier score'. Because our hypothesis specifically predicted an increase in the RET ligand score to correlate with letrozole resistance, and because the number of patients was small, we designed the analysis to use a one-tailed Wilcoxon rank sum test. However, in practice using a two-tailed Wilcoxon rank sum test did not change the results of our analysis. Data was processed and visualization was completed using R.

**Pausing Analysis**

Pause and gene body densities were quantile normalized across all GDNF time course PRO-seq data before pausing analysis in order to avoid potential unknown confounding effects, as described by Danko et. al. (2013). Pausing indices were defined as the ratio of quantile normalized RNA polymerase densities in 500 bp centered on the annotated GENCODE (v19) transcription start sites and the gene body (+1kb to +60kb, as defined above). In the pausing analysis we compared the log *e* transformed ratio of pausing indices between 1 hour of GNDF

and untreated TamS MCF-7 cells. All computations were preformed using the R statistical package.

**Reconstructing Tamoxifen Resistance Regulatory Network**

We defined direct targets of E2 and GDNF signaling as all of those genes undergoing transcriptional changes following short durations of ligand treatment (<40-60 minutes). We used existing GRO-seq data following 40 minutes of E2 treatment (GSE27463). Data following GDNF treatment were collected during the course of this study. Secondary targets were defined as transcriptional changes following 24 hours of GDNF treatment. Secondary targets were assigned to transcription factors (TFs) with binding sites located nearby (<50 kb from the transcription start site) genes that changed following 1 or 24 hours of GDNF. Binding sites were derived from ENCODE ChIP-seq data in MCF-7 cells using BroadPeak peak calls for CEBP, EGR1, ELF1, FOSL2, FOXM1, GABPA, GATA3, JUND, MAX, NR2F2, NRSF, PML, SRF, TAF1, TCF12, and TEAD4. Data for each TF was downloaded from the ENCODE DCC (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeHaibTfbs/). Networks were visualized using the Cytoscape software package (Shannon et al., 2003).

**RNA Isolation and Quantitative Real-Time PCR**

RNA was purified using RNeasy Kit (Qiagen; Cat# 74104) and 1µg of purified RNA was reverse-transcribed using High Capacity RNA-to-cDNA kit (Applied Biosystems; Cat# 4387406) according to the manufacturers' protocols. Real-time quantitative PCR analysis was performed using the following primers: *ACTB* Forward (5'-CCAACCGCGAGAAGATGA-3') and Reverse (5'- CCAGAGGCGTACAGGGATAG-3'); *PGR* Forward (5'-GTCAGGCTGGCATGGTCCTT-3') and Reverse (5'-GCTGTGGGAGAGCAACAGCA-3'); *GREB1* Forward (5'-GTGGTAGCCGAGTGGACAAT-3') and Reverse (5'-ATTTGTTTCCAGCCCTCCTT-3') (Prenzel et al., 2011); *ESR1* Forward (5'- TTACTGACCAACCTGGCAGA-3') and Reverse (5'-

ATCATGGAGGGTCAAATCCA-3'); *GDNF* Forward (5'- TCTGGGCTATGAAACCAAGGA-3') and Reverse (5'- GTCTCAGCTGCATCGCAAGA-3') (Boulay et al., 2008); *EGR1* (5'-AGCCCTACGAGCACCTGAC-3') and Reverse (5'- GTTTGGCTGGGGTAACTGGT-3') (Fang et al., 2016); and Power SYBR Green PCR Master Mix (Applied Bioystems; Cat#4367659). The samples were normalized to β-actin. At least three biological replicates were performed and data are represented as mean ± SEM. All statistical analysis for qPCR were performed using GraphPad Prism. Groups were compared using two-tailed unpaired Student's t-test.

**Generation of GDNF Knockdown G11 Cells**

*GDNF* expression was stably knocked down in $G11^{TamR}$ cells by transduction with lentivirus expressing either shRNA scrambled control or *GDNF* shRNA. Mission shRNA lentivirus plasmids for control shRNA (Cat# SHC002) and *GDNF* shRNA (Cat# SHCLND-NM_000514) from Sigma-Aldrich were used. Specifically, 1.5 µg pLKO.1 shRNA plasmid (Addgene; Plasmid #1864), 0.5 µg psPAX2 packaging plasmid (Addgene; Plasmid #12260), and 0.25 µg pMD2.G envelope plasmid were used for packaging (Addgene; Plasmid #12259). The lentiviruses were generated and transduced according to the manufacturer's instructions (Sigma-Aldrich). Clones were selected in 2 µg/ml of puromycin.

**Cell Proliferation Assay**

Approximately $1x10^6$ G11-scrambled (G11-SCR) and G11-GDNF-knockdown (G11-GDNF-KD) cells were plated in T25 TC-flask. The cells were grown in either 0, 1 or 10 µM tamoxifen in the presence or absence of 5 ng/mL GDNF for 7 days. The cell number was counted for quantification and was normalized to the untreated group. Three biological replicates were performed.

### Immunoblot Analysis

Whole cell lysates were resolved by SDS-PAGE followed by transfer to PVDF membrane. The membranes were stained with Ponceau to visualize the total bound-protein. The membranes were incubated overnight with primary antibodies diluted in TBST in 4 °C using the following antibody concentrations: anti-p-ERK (1:1000; Cell Signaling; Cat# 4695), anti-ERα (1:1000; Santa Cruz; Cat# sc-543) and anti-p-ER (1:1000; Cell Signaling; Cat# 2511). The primary antibodies were detected with HRP-conjugated secondary antibodies and were exposed to ECL reagents.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical parameters include the exact number of biological replicates (n), standard error of the mean (mean ± SEM), and statistical significance are reported in the Figure legends. Data are reported statistically significant when $p < 0.05$ by two-tailed Student's t-test. In figures, asterisks and pound signs denote statistical significance as calculated by Student's t-test. Specific p-values are indicated in the Figure legends. Statistical analysis was performed using GraphPad PRISM 6.

## DATA AND SOFTWARE AVAILABILITY

### Data Resources

Raw data files for the PRO-seq analysis have been deposited in Gene Expression Omnibus under Accession Number GSE93229.

### Software Availability

Software and scripts used in all analyses are publicly available without restriction on GitHub at

278 https://github.com/Danko-Lab/mcf7tamres. At the time of submission, the most recent commit

279 was version number: 855156ad07c042c88089cb4f31bf9d544487a1b2.

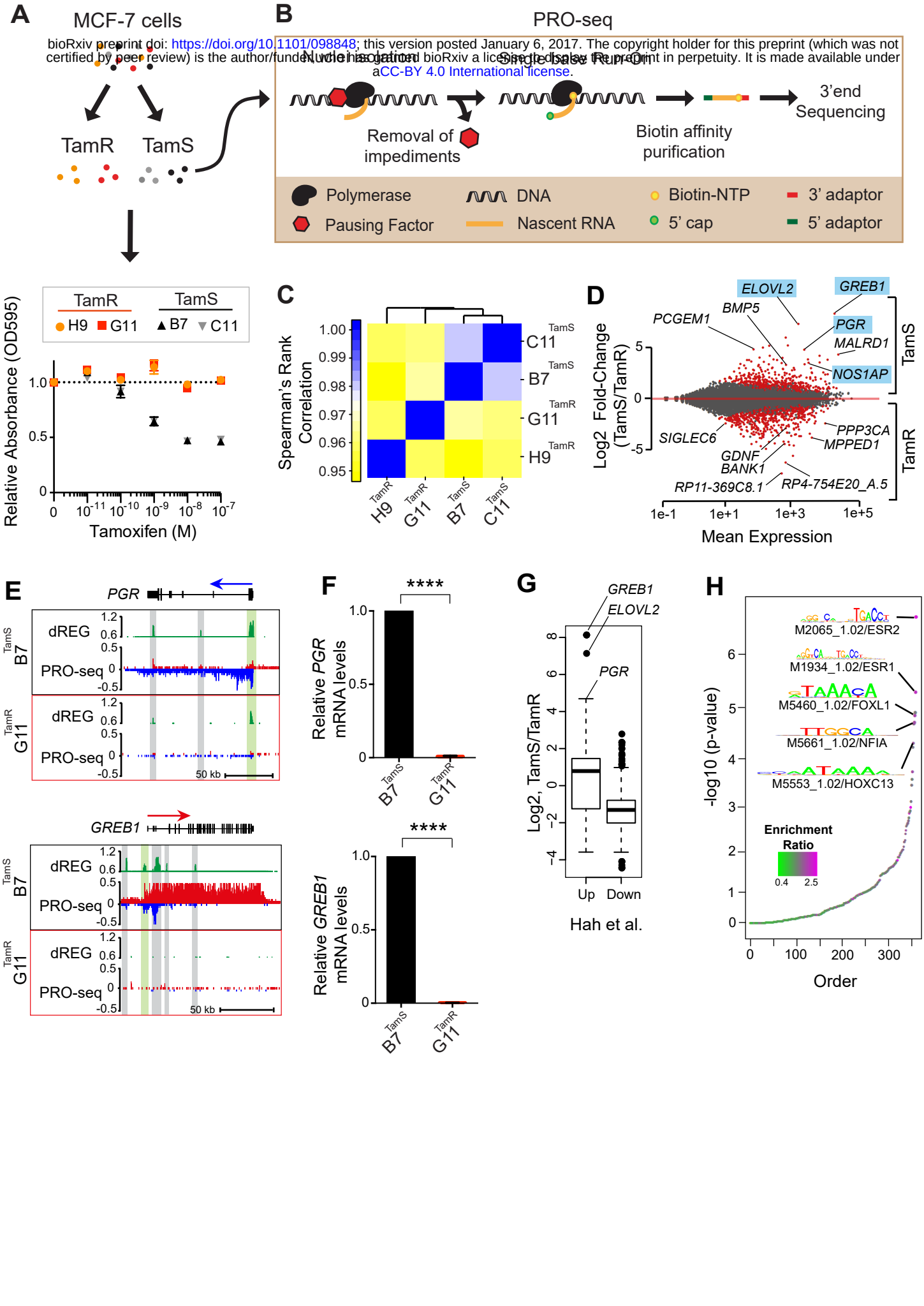

**Figure 1: ER target genes are uniquely expressed in TamS cells.**

(A) Cell viability of tamoxifen sensitive (TamS; $B7^{TamS}$ and $C11^{TamS}$) and resistant (TamR; $G11^{TamR}$ and $H9^{TamR}$) MCF-7 cells upon treatment with 0 (vehicle; EtOH), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M of tamoxifen for 4 days. Data are represented as mean ± SEM (n=3).

(B) Experimental setup for PRO-seq. PRO-seq libraries were prepared from all four cell lines grown in the absence of tamoxifen for 3 days.

(C) Spearman's rank correlation of RNA polymerase abundance in the gene bodies (+1000 bp to the annotation end) of TamS and TamR cell lines.

(D) MA plot showing significantly changed genes (red) that are higher in TamS (top) or TamR (bottom) MCF-7 lines. Genes highlighted in the plots which are ERα targets are highlighted in blue.

(E) Transcription near the *PGR* and *GREB1* loci in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on the sense and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. Enhancers and promoters are shown in grey and light green shading, respectively. Arrows indicate the direction of gene annotations.

(F) *PGR* and *GREB1* mRNA expression levels in $B7^{TamS}$ and $G11^{TamR}$ cells. Data are represented as mean ± SEM (n=3 for *PGR*; n = 4 for *GREB1*). **** $p < 0.0001$. $G11^{TamR}$ is normalized to $B7^{TamS}$.

(G) Boxplots represent fold-change between TamS and TamR of genes that are either up- or down-regulated following 40 minutes of estrogen (E2) in Hah et. al. (2011). Spearman's Rho= 0.185, $p < 2.2e-16$.

(H) Motifs enriched in TREs that have different amounts of RNA polymerase between TamS and TamR cells compared with TREs that have consistent levels.

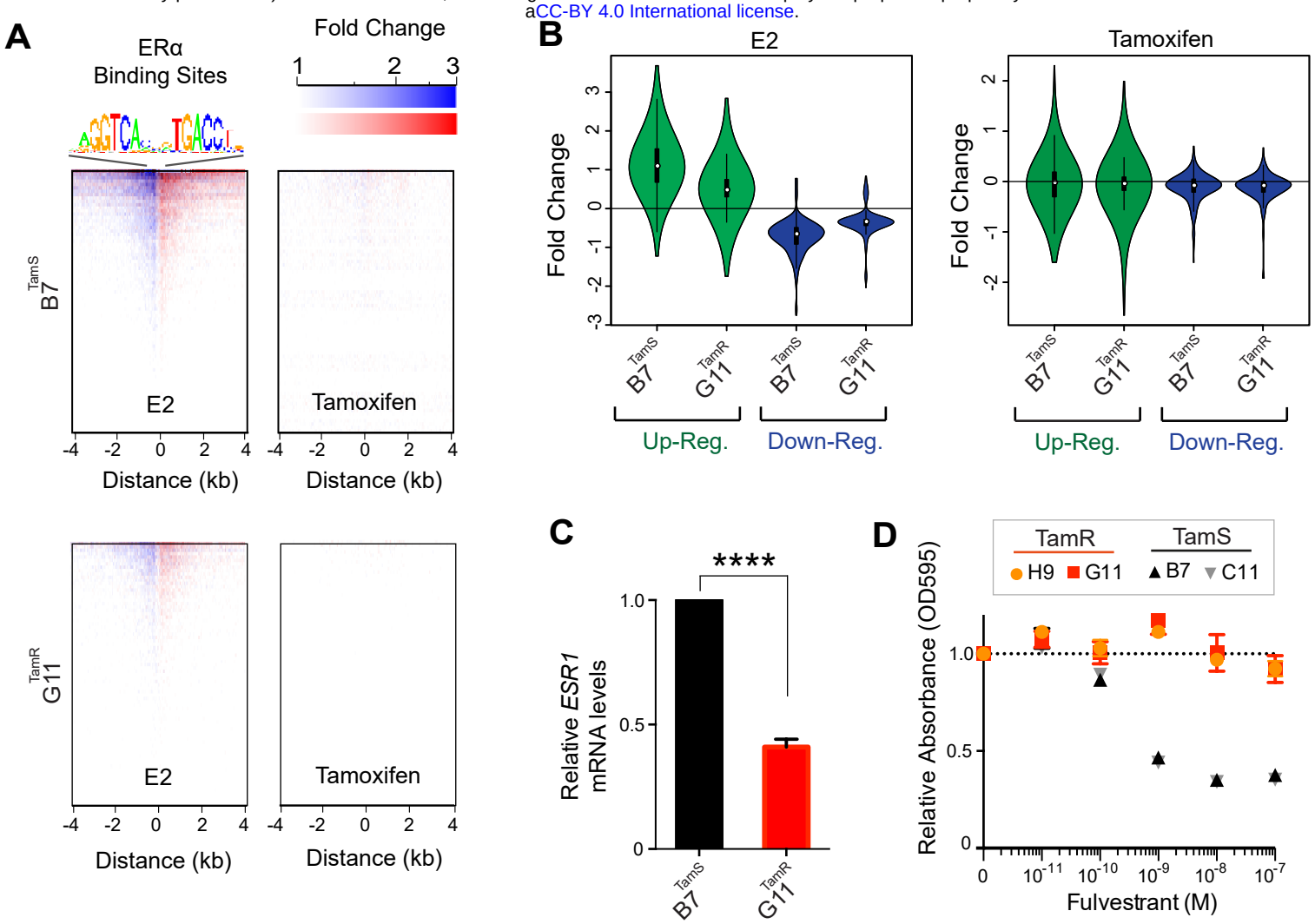


**Figure 2: Tamoxifen resistant lines have functional ERα signaling**

(A) Heatmap of changes in RNA polymerase abundance following 40 minutes of E2 or tamoxifen treatment near ERα bindings sites in $B7^{TamS}$ and $G11^{TamR}$ cells.

(B) Violin plots show fold changes in the indicated MCF-7 clone following 40 minutes of E2 or tamoxifen tamoxifen treatment at genes up- or down-regulated by E2 in Hah et. al. (2011). Up- and down-regulated genes are in green and blue, respectively.

(C) *ESR1* mRNA expression levels in $B7^{TamS}$ and $G11^{TamR}$ cells. Data are represented as mean ± SEM (n=3). **** $p < 0.0001$.

(D) Cell viability of TamS and TamR cells upon treatment with 0 (vehicle; DMSO), $10^{-11}$, $10^{-10}$, $10^{-9}$, $10^{-8}$, or $10^{-7}$ M fulvestrant (ER degrader) for 4 days. Data are represented as mean ± SEM (n=3).

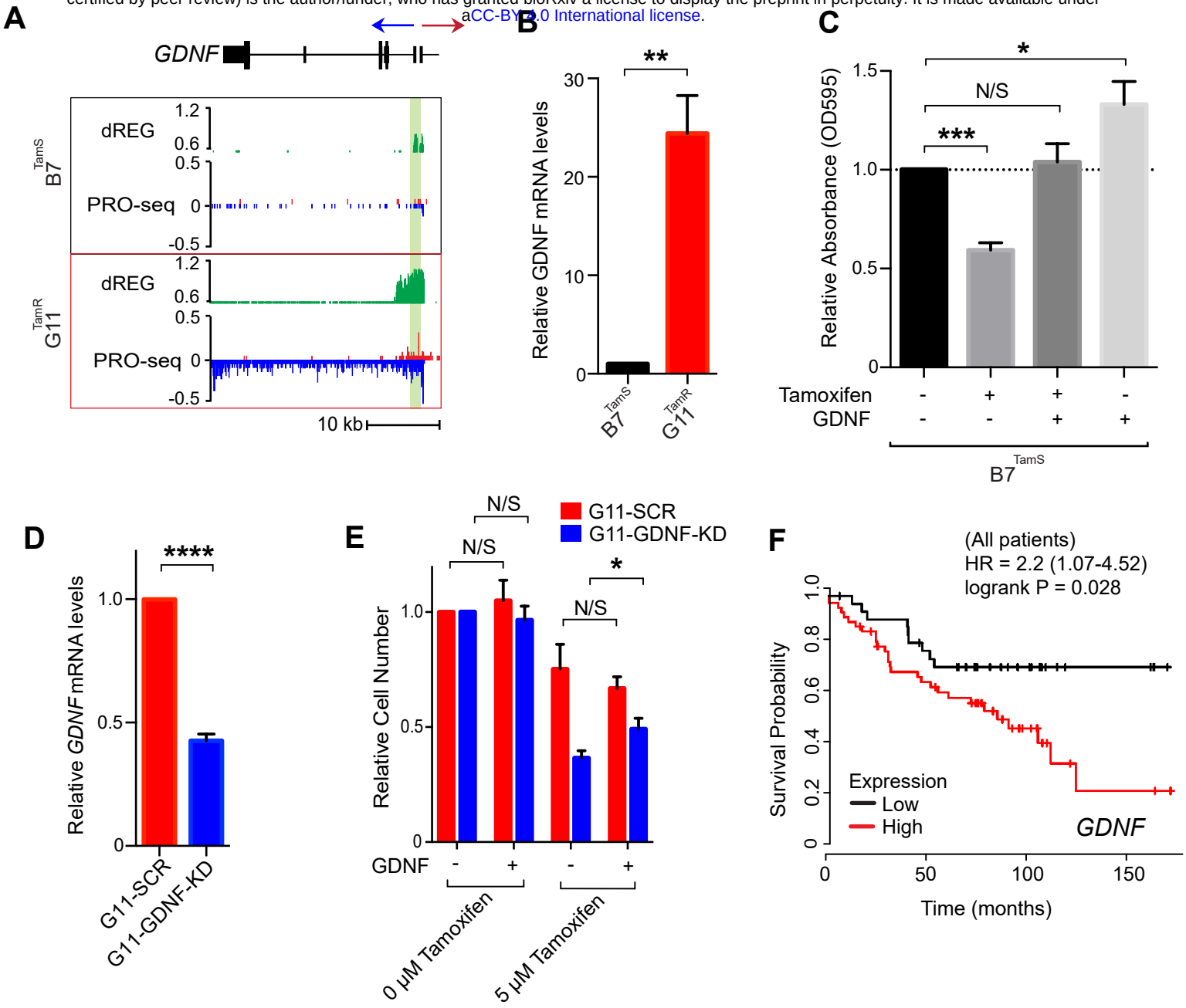


**Figure 3: *GDNF* is responsible for tamoxifen resistance in MCF-7 cells**

(A) Transcription near the *GDNF* locus in B7$^{TamS}$ and G11$^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. The region near the GDNF promoter is shown in light green shading. Arrow indicates the direction of gene annotations.

(B) *GDNF* mRNA expression levels in B7$^{TamS}$ and G11$^{TamR}$ cells. Data are represented as mean ± SEM (n=3). ** $p < 0.005$.

(C) Cell viability of B7$^{TamS}$ cells in the presence or absence of 10 ng/ml GDNF and/or 100 mM tamoxifen for 4 days. Data are represented as mean ± SEM (n=3). * $p < 0.05$, *** $p < 0.0005$.

(D) *GDNF* mRNA expression levels in G11$^{TamR}$ scrambled (SCR) and G11$^{TamR}$ GDNF knockdown (GDNF-KD) cells. Data are represented as mean ± SEM (n=3). **** $p < 0.0001$.

(E) Relative cell number of G11$^{TamR}$ scrambled (SCR) and G11$^{TamR}$ GDNF knockdown (GDNF-KD) cells after 4 days without or with 5 μM tamoxifen and/or 5 ng/ml GDNF treatment. Data are represented as mean ± SEM (n=9). * $p < 0.05$.

(F) Kaplan Meier (KM) plot of relapse free survival (RFS) in a cohort of 88 breast cancer patients with low (black) or high (red) GDNF expression.

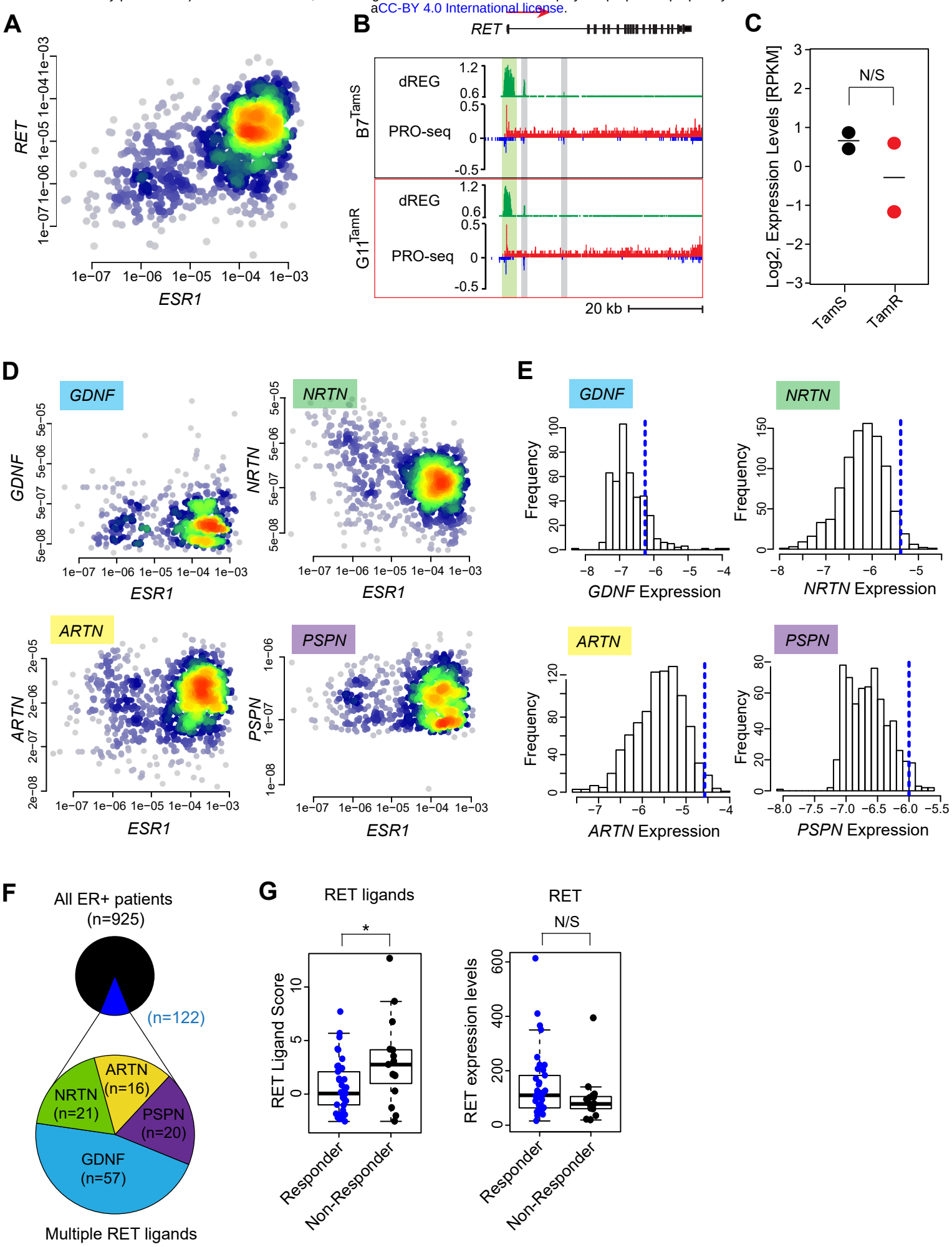

**Figure 4: Expression of RET ligands contributes to endocrine resistance.**

(A) Density scatterplot showing *RET* and E*SR1* expression in mRNA-seq data from 1,177 primary breast cancer models in the cancer genome atlas (TCGA). Spearman's ρ = 0.51, *p* = 1.2e-60.

(B) Transcription near the *RET* locus in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. Enhancers and promoters are shown in grey and light green shading, respectively. Arrow indicates the directional movement of transcribed genes.

(C) Dot plot shows *RET* transcription levels in TamS and TamR MCF-7 cells.

(D) Density scatterplots show the expression of RET ligands (*GDNF*, *NRTN*, *ARTN*, and *PSPN*) versus *ESR1* based on mRNA-seq data from 1,177 primary breast cancers.

(E) RET ligand expression distribution in ER+ breast cancers. The dotted blue line represents 2.5 times the range between the $25^{th}$ and $50^{th}$ percentile.

(F) Fraction of ER+ breast cancers (n = 925) with at least one RET ligand exceeding the threshold shown in panel E (shown in dark blue, n = 122). Among the 4 RET ligands, GDNF was the most highly expressed (n = 60).

(G) Boxplots show RET ligands score and RET expression levels in patients that respond or do not respond to aromatase inhibitor letrozole. * *p* = 0.016.

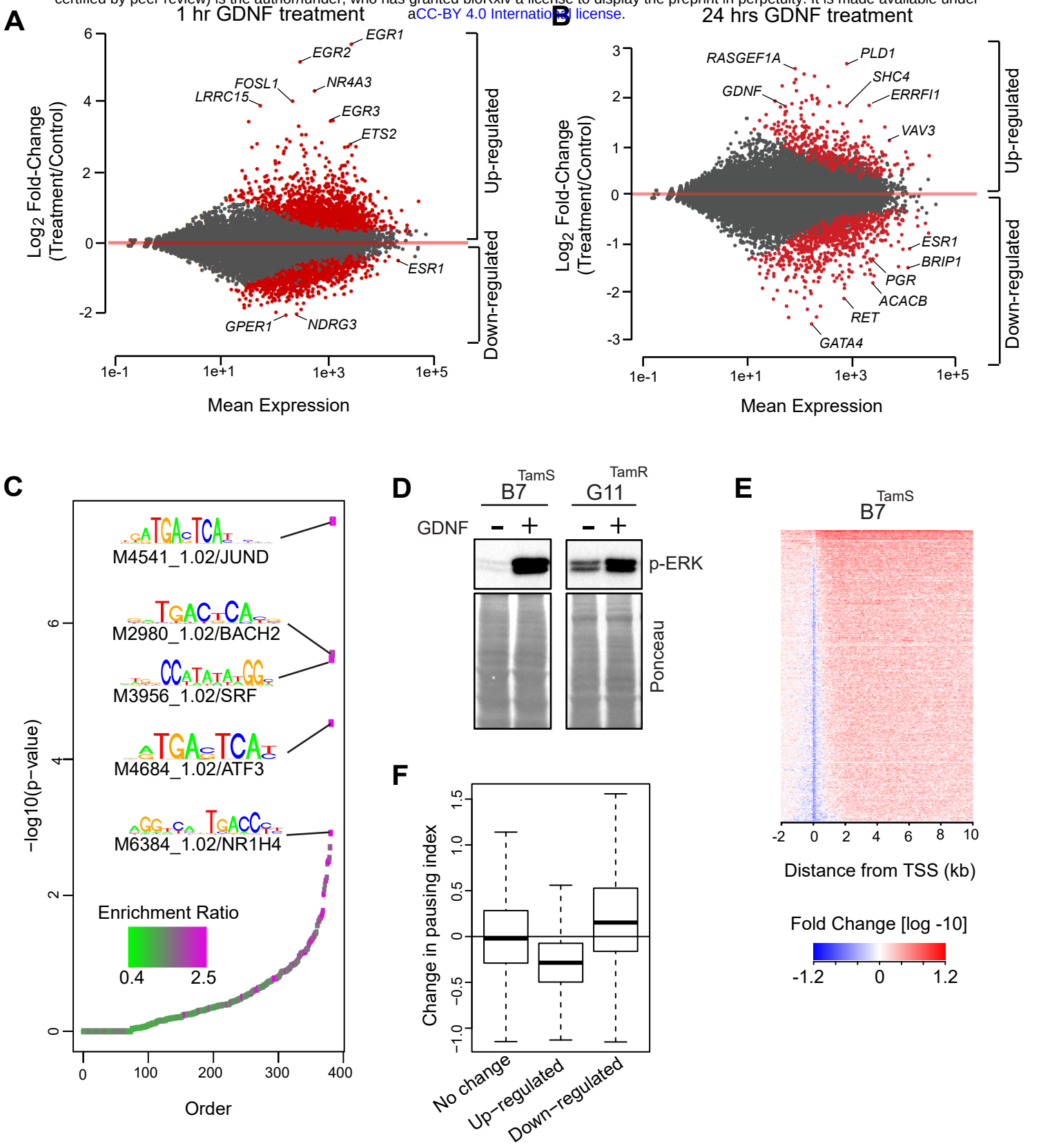


**Figure 5: GDNF activates thousands of target genes at the level of pause release.**

(A-B) MA plot shows significantly upregulated and downregulated genes (red) following 1 hour (A) or 24 hours (B) of GDNF treatment in TamS MCF-7 cells.

(C) Motifs enriched in TREs that have different amounts of RNA polymerase following 1 hour of GDNF treatment compared with TREs that have consistent levels.

(D) Immunoblot analysis of p-ERK in serum deprived $B7^{TamS}$ and $G11^{TamR}$ cells treatment with 10 ng/mL GDNF.

(E) Heatmap depicting changes in RNA polymerase density following 1 hour of GDNF treatment in $B7^{TamS}$ MCF-7 cells.

(F) Changes in pausing index between treated (1 hour) and untreated TamS MCF-7 cells at the indicated class of genes.

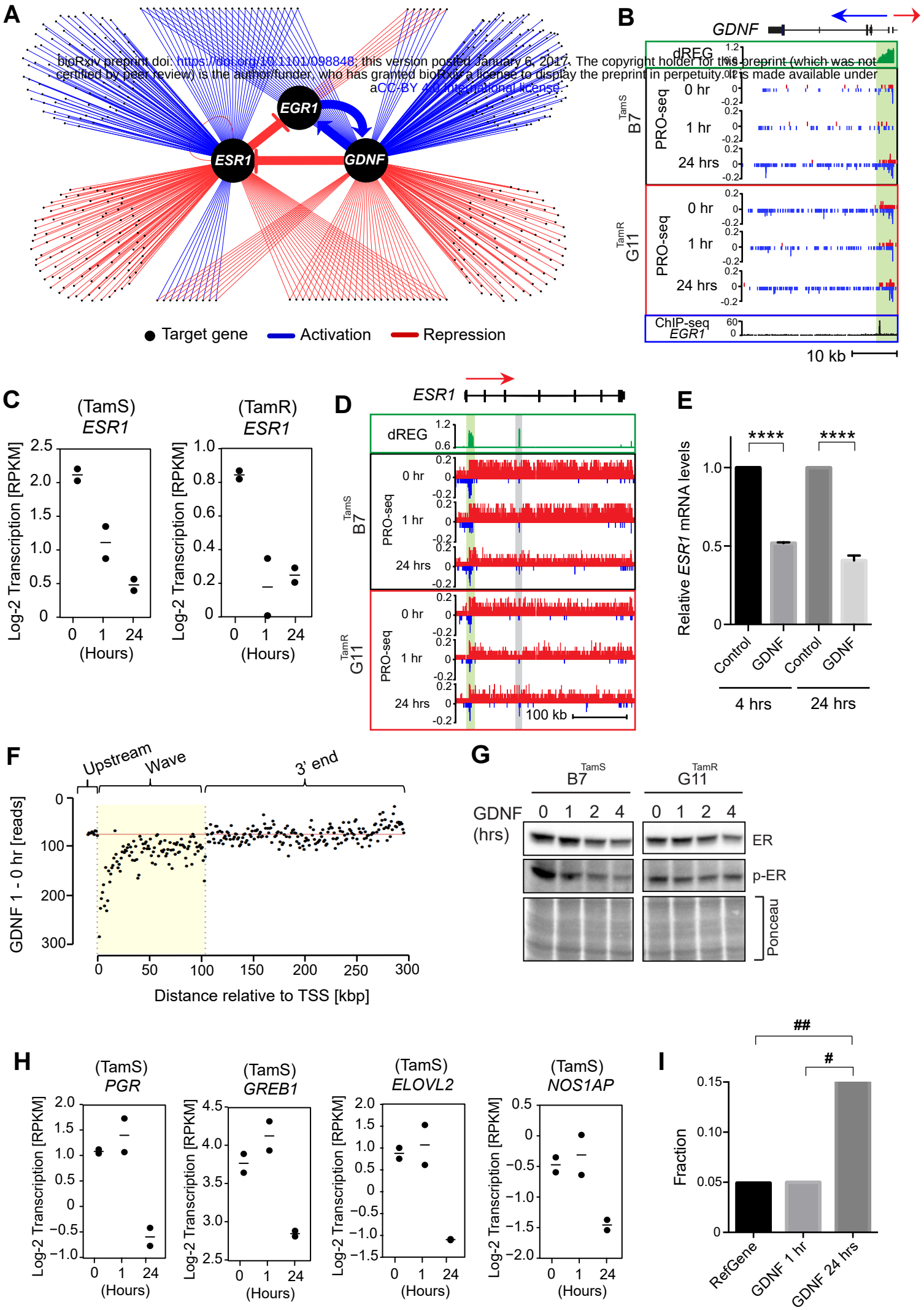

**Figure 6: Bi-stable feedback loop between *ESR1, EGR1,* and *GDNF*.**

(A) Transcriptional regulatory network of GDNF-dependent endocrine resistance highlighting the bi-stable feedback loop inferred between *ESR1*, *EGR1*, and *GDNF*. Each point represents a gene regulated following 1 or 24 hours of GDNF signaling. Only transcription factors or signaling molecules are shown. Blue and red edges represent activation or repression relationships, respectively.

(B) Transcription near the *GDNF* locus in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. The promoter is shown in light green shading. Arrows indicate the direction encoding annotated genes.

(C) Dot plots of transcription levels of *ESR1* $B7^{TamS}$ and $G11^{TamR}$ cells following GDNF treatment.

(D) Transcription in the *ESR1* gene in $B7^{TamS}$ and $G11^{TamR}$ cells. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. Enhancers and promoters are shown in grey and light green shading, respectively. Arrow indicates the direction encoding annotated genes.

(E) *ESR1* mRNA expression levels in $B7^{TamS}$ cells following 10 ng/mL GDNF treatment. Data are represented as mean ± SEM (n=3). **** $p < 0.0001$.

(F) Difference in read counts in 3kb windows along *ESR1* between 1 hours of GDNF and untreated TamS MCF-7 cells, The location of the wave of RNA polymerase along *ESR1* was identified using a hidden Markov model and is represented by the yellow box.

(G) Immunoblot analysis of ERα and p-ERα in $B7^{TamS}$ and $G11^{TamR}$ cells treatment with 10 ng/mL for 0, 1, 2, and 4 hours.

(H) Dot plots representing transcription levels of *ERα* target genes (*PGR*, *GREB1*, *ELOVL2*, and *NOS1AP*) following a timecourse of GDNF treatment.

(I) Bar plot showing the fraction of genes whose transcription is up-regulated by 40 min. of E2 in all RefSeq annotated genes (left) or those which are downregulated by 1 (center) or 24 hours (right) of GDNF treatment. E2 target genes were enriched in those down-regulated following 24 hrs of GDNF treatment. The Y axis denotes the fraction of genes that are direct up-regulated E2 targets (defined based on Hah et. al. (2011) and also up-regulated in $B7^{TamS}$). # p = 1.098e-10, ## p= 6.556999e-19. Fisher's exact test was used for statistical analysis.

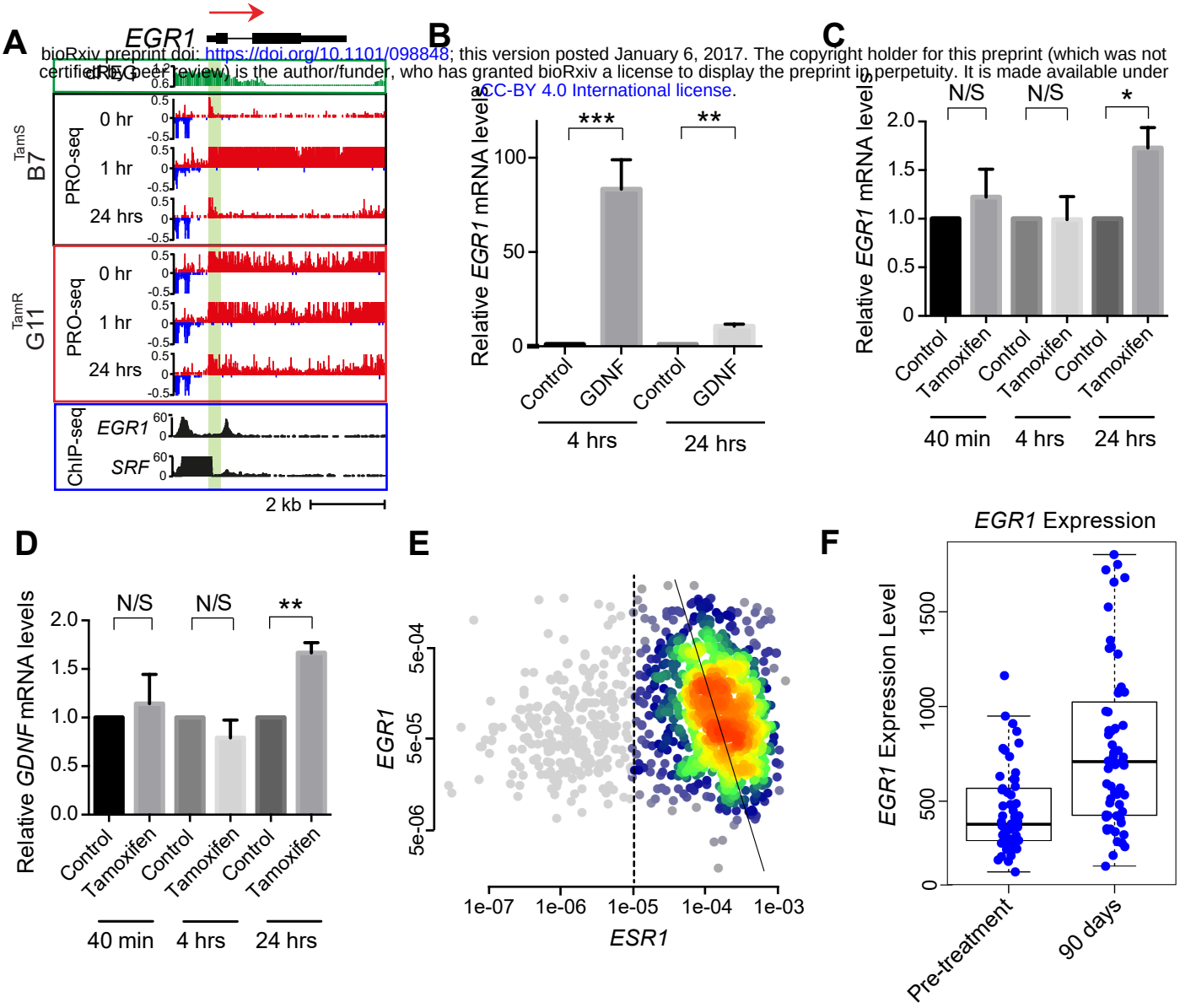


**Figure 7: Validation of bi-stable feedback loop in MCF-7 cells and primary breast tumors**

(A) Transcription at the *EGR1* locus in $B7^{TamS}$ and $G11^{TamR}$ cells before and after treatment with GDNF. PRO-seq densities on sense strand and anti-sense strand are shown in red and blue, respectively. dREG scores are shown in green. The number of reads mapping in EGR1 and SRF ChIP-seq data is shown in black. Arrow indicates the direction of annotated genes.

(B) *EGR1* mRNA expression level in $B7^{TamS}$ cell after treatment with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). ** $p < 0.01$, *** $p \leq 0.001$.

(C) *EGR1* mRNA expression level in $G11^{TamR}$ cells after treatment without (DMSO) or with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). * $p < 0.05$.

(D) *GDNF* mRNA expression levels in $G11^{TamR}$ cells after treatment without (DMSO) or with 10 ng/mL GDNF for 4 or 24 hrs. Data are represented as mean ± SEM (n=3). ** $p <$ 0.005.

(E) Density scatterplots show the expression of *EGR1* versus *ESR1* based on mRNA-seq data from 1,177 primary breast cancers. ER+ breast cancers (n= 925), defined based on ESR1 expression (>1e-5), are highlighted in color. The trend line was calculated using Deming regression in the ER+ breast cancers (Pearson's R= -0.21; $p$ = 2.7e-10).

(F) Boxplots show *EGR1* expression level before or following 90 days of treatment with letrozole ($p$ = 1.8e-6, Wilcoxon Rank Sum Test).