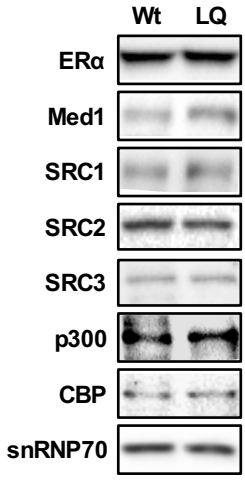
SUPPLEMENTAL MATERIAL

Dynamic Assembly and Activation of Estrogen Receptor α Enhancers Through Coregulator Switching

Murakami *et al.* (Kraus)

Contains:
- Supplemental Figs. S1 through S18
- Supplemental Materials and Methods
- Author Contributions
- Supplemental References

Supplemental Figures



Supplemental Figure S1. ERα coregulators of interest in this study are expressed at similar levels in 231/ERαWt and 231/ERαLQ cells.
Western blots showing the expression levels of ERα and its coregulators in 231/ERαWt and 231/ERαLQ cells.

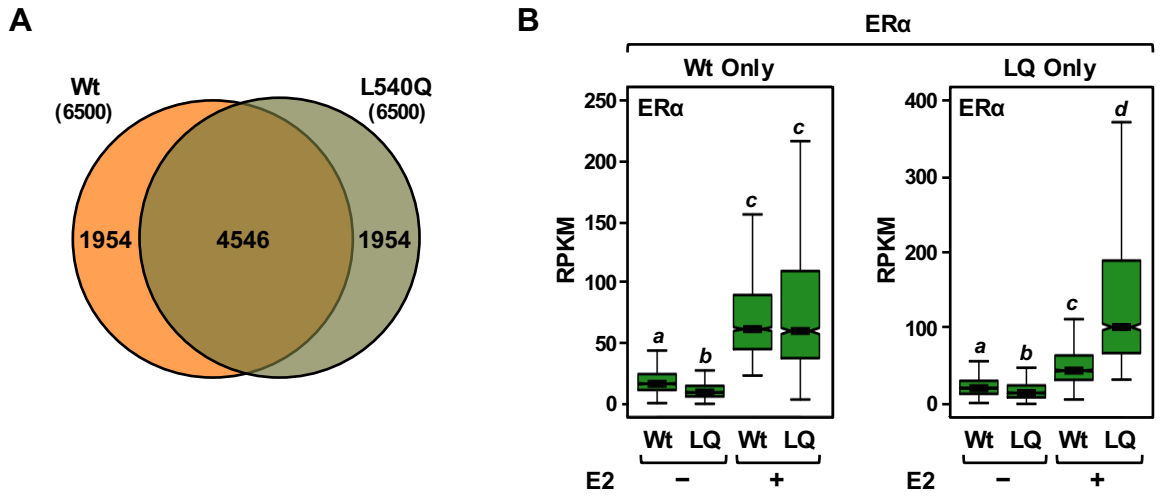


Figure S2. Considerable overlap between ERαWt and ERαLQ binding sites in 231/ERα cells.

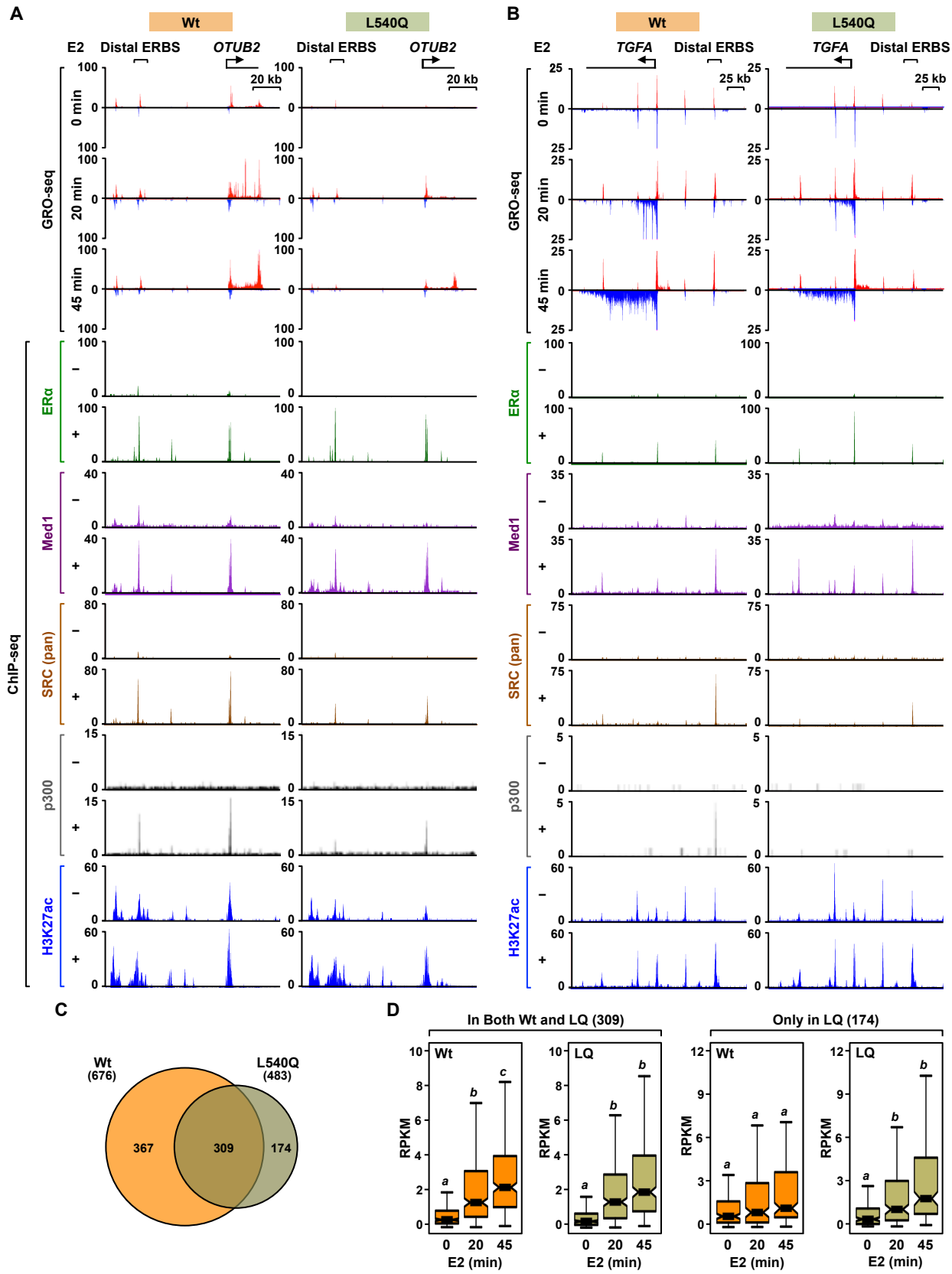
(A) A large subset of ERαWt and ERαLQ binding sites overlap in 231/ERα cells. Venn diagram showing the overall of significant ERαWt and ERαLQ peaks upon 45 min. of E2 treatment in 231/ERαWt and 231/ERαLQ cells, as determined by ChIP-seq, for the top 6,500 peaks.

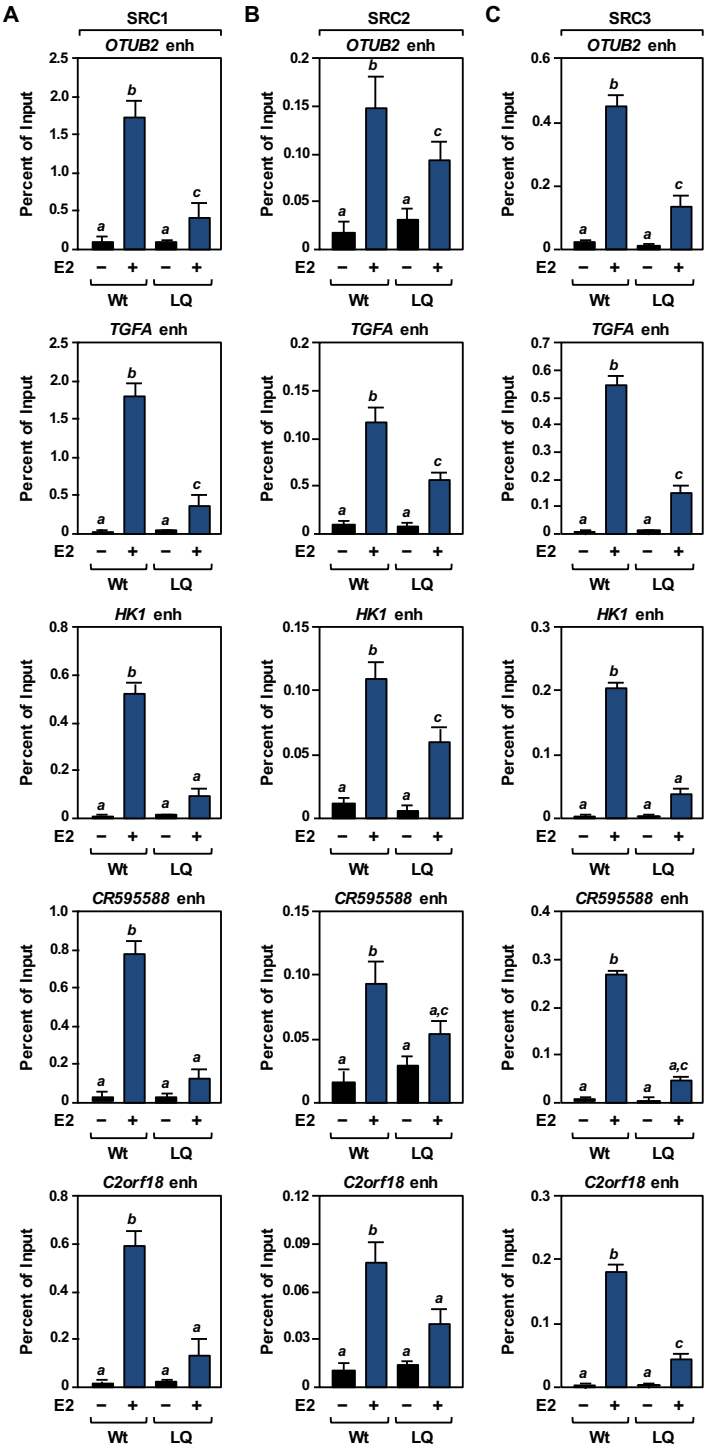
(B) Loci with significant peaks of ERαWt in 231/ERαWt cells are enriched in read counts for ERαLQ in 231/ERαLQ, and vice versa. Box plots of ERα ChIP-seq read counts ± 45 min. of E2 treatment at ERα binding sites unique to either ERαWt (Wt only) or ERαLQ (LQ only) after 45 min. of E2 treatment in 231/ERαWt and 231/ERαLQ cells. Box plots marked with different letters (*a*, *b*, *c, d*) are significantly different from each other ($p < 2.2 \times 10^{-16}$; Wilcoxon rank-sum test).

Supplemental Figure S3. Reduced SRC recruitment at ERα L540Q binding sites impairs enhancer formation and target gene transcription *(continued)*.
(A and B) Impaired recruitment of SRC and p300 at ERα L540Q binding sites results in reduced E2-stimulated transcription. GRO-seq browser tracks with a time course of E2 treatment (0, 20, and 45 min.) and ChIP-seq browser tracks for ERα, Med1, SRC (pan), p300, and H3K27ac ± 45 min. E2 at the *OTUB2* (A) and *TGFA* (B) loci in MBA-MB-231 cells expressing ERα wild-type or ERα L540Q (231/ERαWt and 231/ERαLQ cells, respectively).
(C) ERα L540Q exhibits impaired transcriptional activity. Venn diagram showing the number of genes significantly up-regulated by 45 min. of E2 treatment in 231/ERαWt and 231/ERαLQ cells as measured by GRO-seq. FDR = 0.05.
(D) Box plots showing the read counts for 309 genes up-regulated by both wild-type and L540Q ERα (*left*) or 174 genes up-regulated by the L540Q mutant, but not by the wild-type, (*right*) upon E2 treatment, as shown in panel (C). Box Plots marked with different letters (*a, b, c*) are significantly different from each other ($p < 1.54 \times 10^{-05}$; Wilcoxon rank-sum test).

[Supplemental Figure S3 is on the next page]

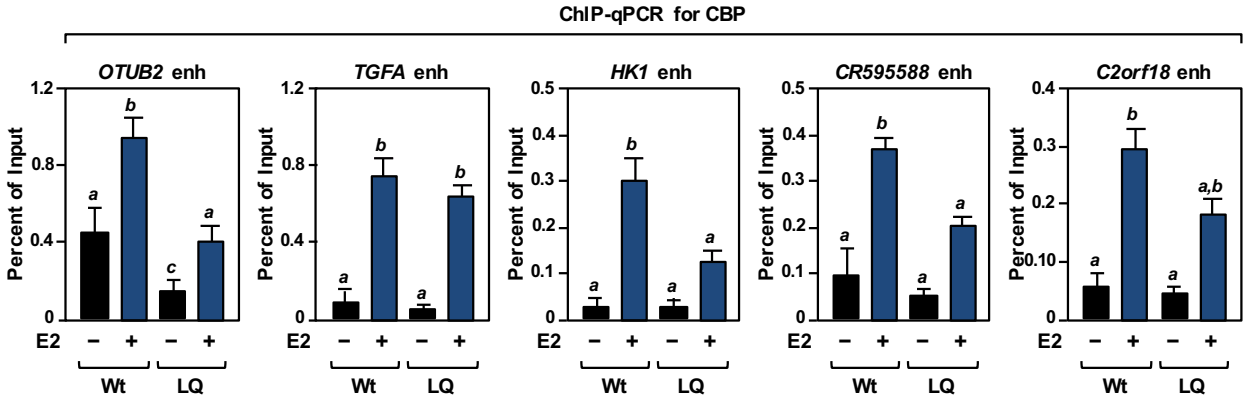
Supplemental Figure S3.



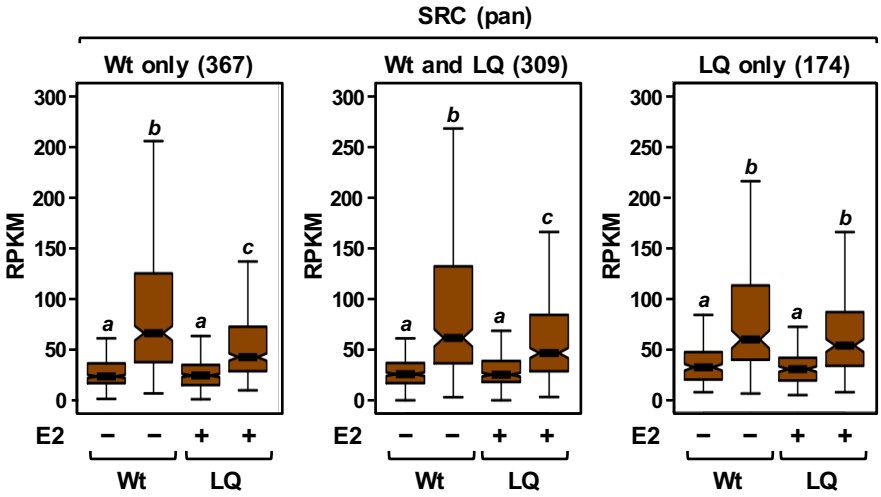


Supplemental Figure S4. The recruitment of all three SRC family members is impaired at ERα L540Q binding sites.

(A-C) ChIP-qPCR assays for (A) SRC1, (B) SRC2, and (C) SRC3 in 231/ERαWt and 231/ERαLQ cells treated ± E2 for 45 min. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b*, *c*) are significantly different from each other ($p < 0.05$; two-way ANOVA).

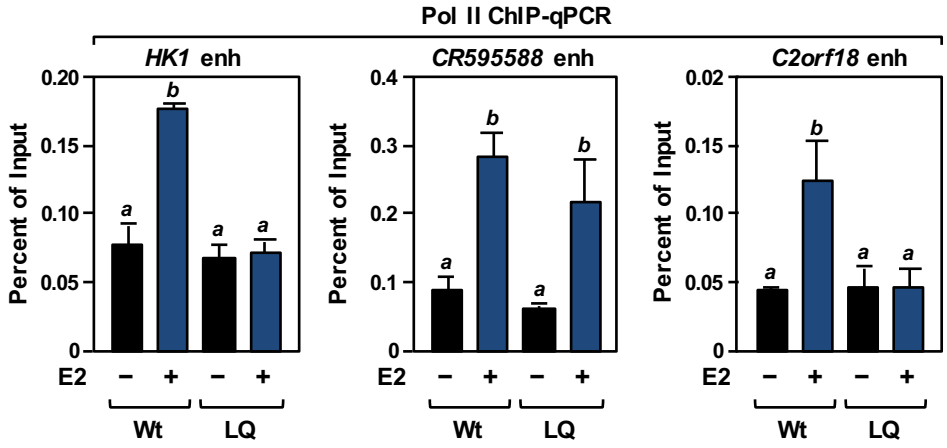


Supplemental Figure S5. Recruitment of CBP is impaired at ERα L540Q binding sites.
ChIP-qPCR assays for CBP in 231/ERαWt and 231/ERαLQ cells treated ± E2 for 45 min. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b*, c) are significantly different from each other ($p < 0.05$; two-way ANOVA).

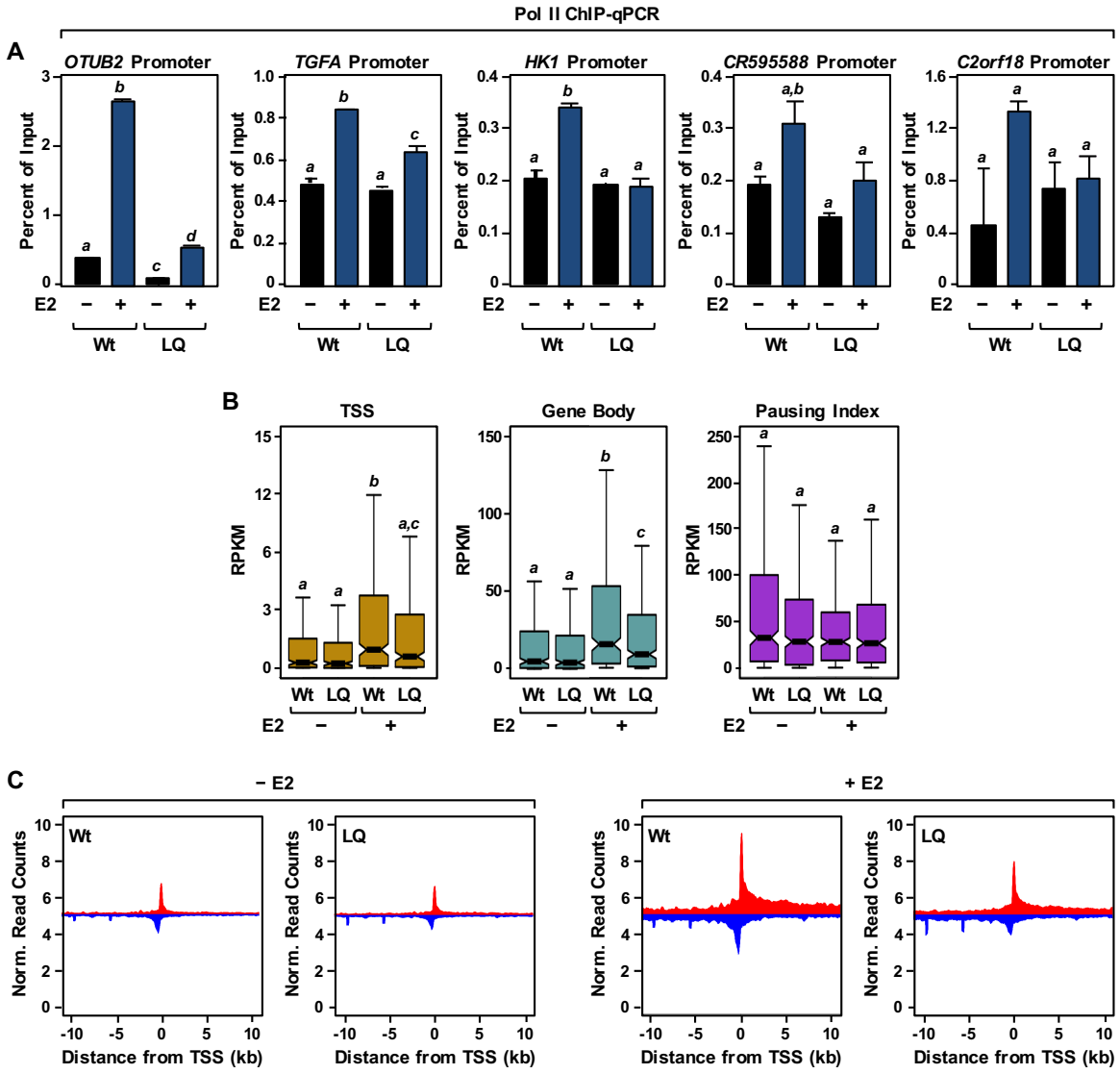


Supplemental Figure S6. Recruitment of SRC to various types of ERα L540Q genomic binding sites.

Box plots of SRC (pan) ChIP-seq read counts ± 45 min. of E2 treatment at ERα binding sites located nearby genes (1) uniquely up-regulated by ERαWt (*left*), (2) commonly up-regulated by ERαWt and ERαLQ (*middle*), or (3) uniquely up-regulated by ERαLQ (*right*). Box plots marked with different letters (*a*, *b*, *c*) are significantly different from each other ($p < 2.2 \times 10^{-5}$; Wilcoxon rank-sum test).



Supplemental Figure S7. Pol II recruitment is impaired at ERα L540Q binding sites.
ChIP-qPCR assays for Pol II in 231/ERαWt and 231/ERαLQ cells treated ± E2 for 45 min. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b*) are significantly different from each other ($p < 0.05$; two-way ANOVA).

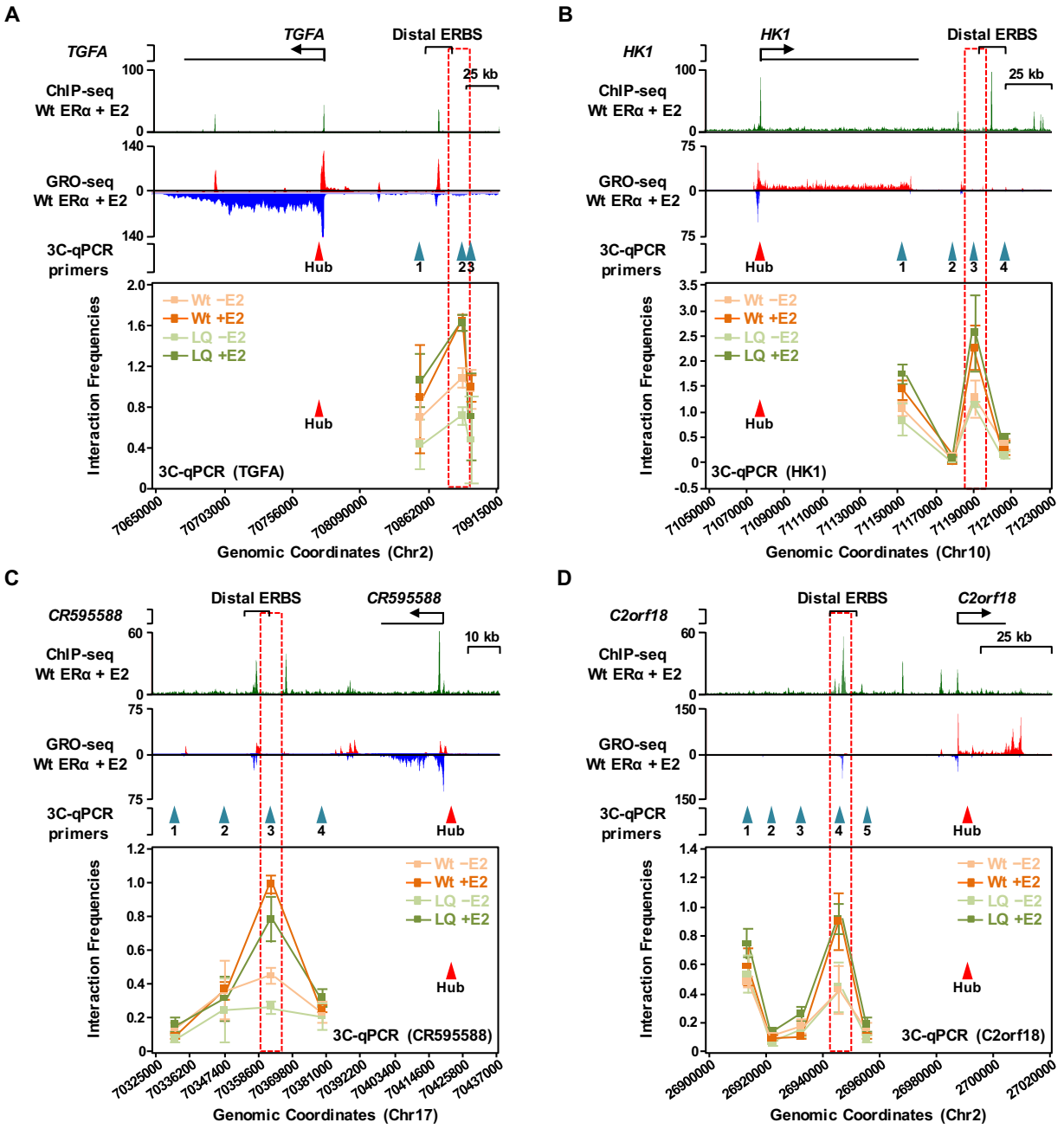


Supplemental Figure S8. Impaired E2-dependent Pol II loading and transcription initiation with ERα L540Q.

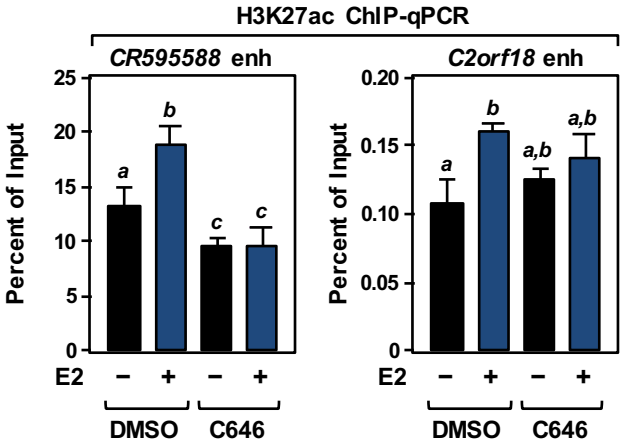
(A) Pol II recruitment is impaired at the promoters of ERα L540Q target genes. ChIP-qPCR assays for Pol II at the promoters of the indicated genes in 231/ERαWt and 231/ERαLQ cells treated ± E2 for 45 min. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b*, *c*, *d*) are significantly different from each other ($p < 0.05$; two-way ANOVA).

(B) ERαLQ shows impaired initiation for E2-dependent gene activation, which ultimately affects elongation. (*Left and middle*) Box plots showing the read counts for the transcription start site (TSS) (*left*) and the gene body (*middle*) for 367 genes up-regulated by ERαWt, but not by ERαLQ, upon E2 treatment as shown in Figure 3A. (*Right*) Box plots showing the pausing indexes for the same set of genes. Box plots marked with different letters (*a*, *b*, *c*) are significantly different from each other ($p < 2.2 \times 10^{-03}$; Wilcoxon rank-sum test).

(C) Metaplots of GRO-seq data ± 10 kb around the TSS of the same set of genes described in (B).



Supplemental Figure S9. E2-dependent enhancer-promoter chromatin loop formation is maintained with ERα L540Q, in spite of impaired SRC recruitment.
(Top) Browser tracks for ERα ChIP-seq and GRO-seq in 231/ERαWt cells after 45 min. of E2 treatment, shown with the location of the 3C primers. *(Bottom)* 3C-qPCR assays showing chromatin looping from a distal ERα binding site to the (A) *TGFA*, (B) *HK1*, (C) *CR595588*, or (D) *C2orf18* promoters in 231/ERαWt and 231/ERαLQ cells ± 45 min. of E2 treatment. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to Wt at eac h genomic location in each condition ($p < 0.05$; two-way ANOVA).



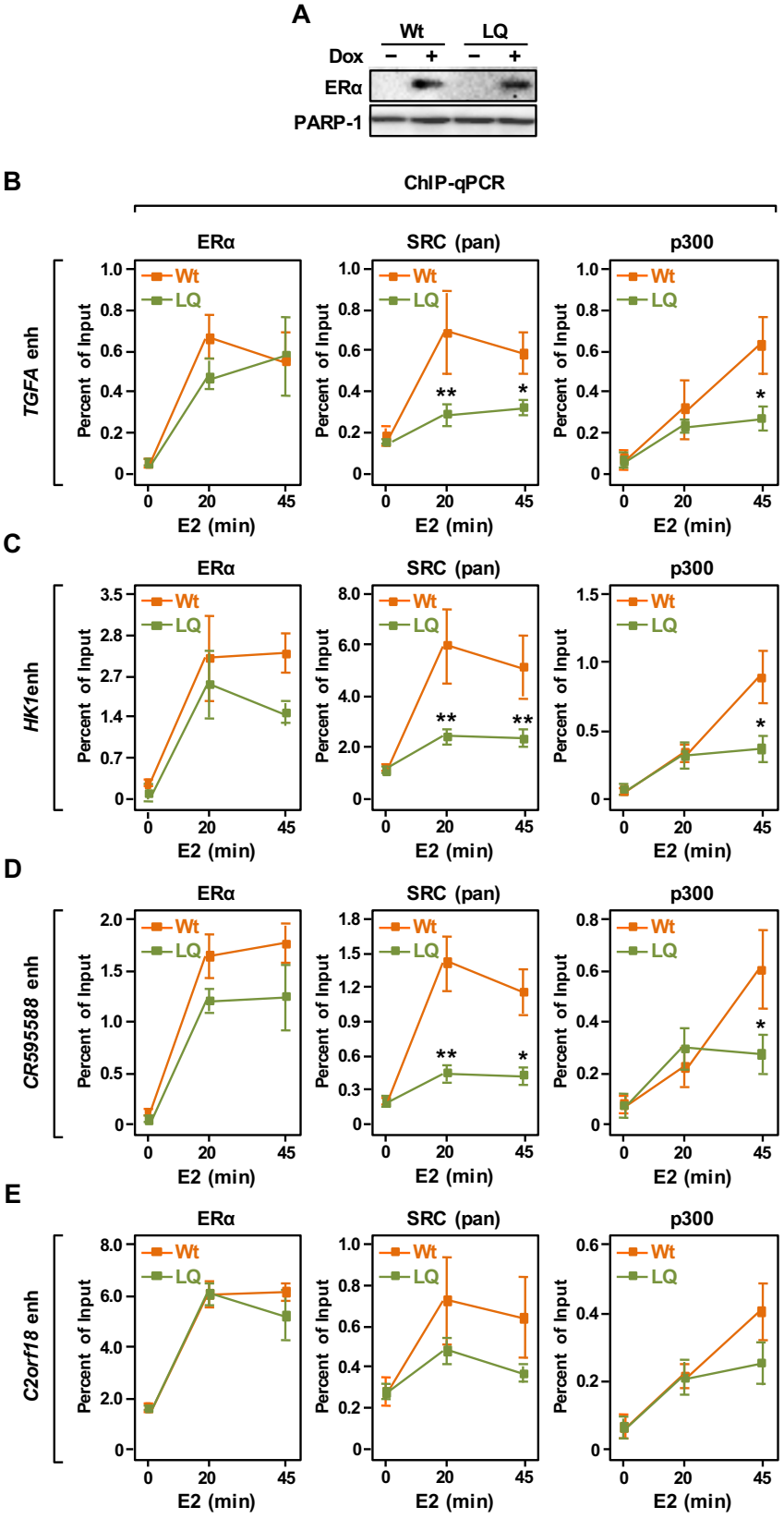
Supplemental Figure S10. p300 HAT activity is required for E2-stimulated increases in H3K27ac levels.

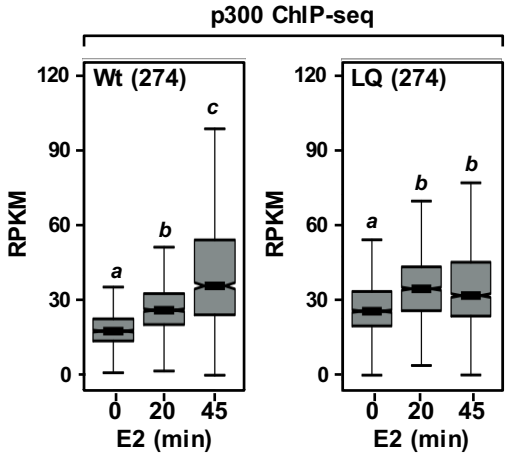
ChIP-qPCR assays for H3K27ac in 231/ERαWt cells treated ± E2 for 45 min. in the presence of the p300/CBP HAT inhibitor C646. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b*, *c*) are significantly different from each other ($p < 0.05$; two-way ANOVA).

Supplemental Figure S11. Impaired SRC recruitment with ERα L540Q causes abortive enhancer formation and target gene transcription *(continued)*.
(A) Western blot showing doxycycline (Dox)-dependent expression of ERα wild-type (Wt) and ERα L540Q (LQ) in MDA-MB-231 cells (231/ERαWt and 231/ERαLQ cells, respectively).
(B - E) p300 is recruited in an SRC-independent manner during the initial phase of enhancer formation ("enhancer priming"). ChIP-qPCR assays for ERα, SRC (pan), and p300 in 231/ERαWt and 231/ERαLQ cells treated with a time course of E2. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to Wt at each time point (*, $p < 0.05$; **, $p < 0.0005$; two-way ANOVA).

[Supplemental Figure S11 is on the next page]

Supplemental Figure S11.



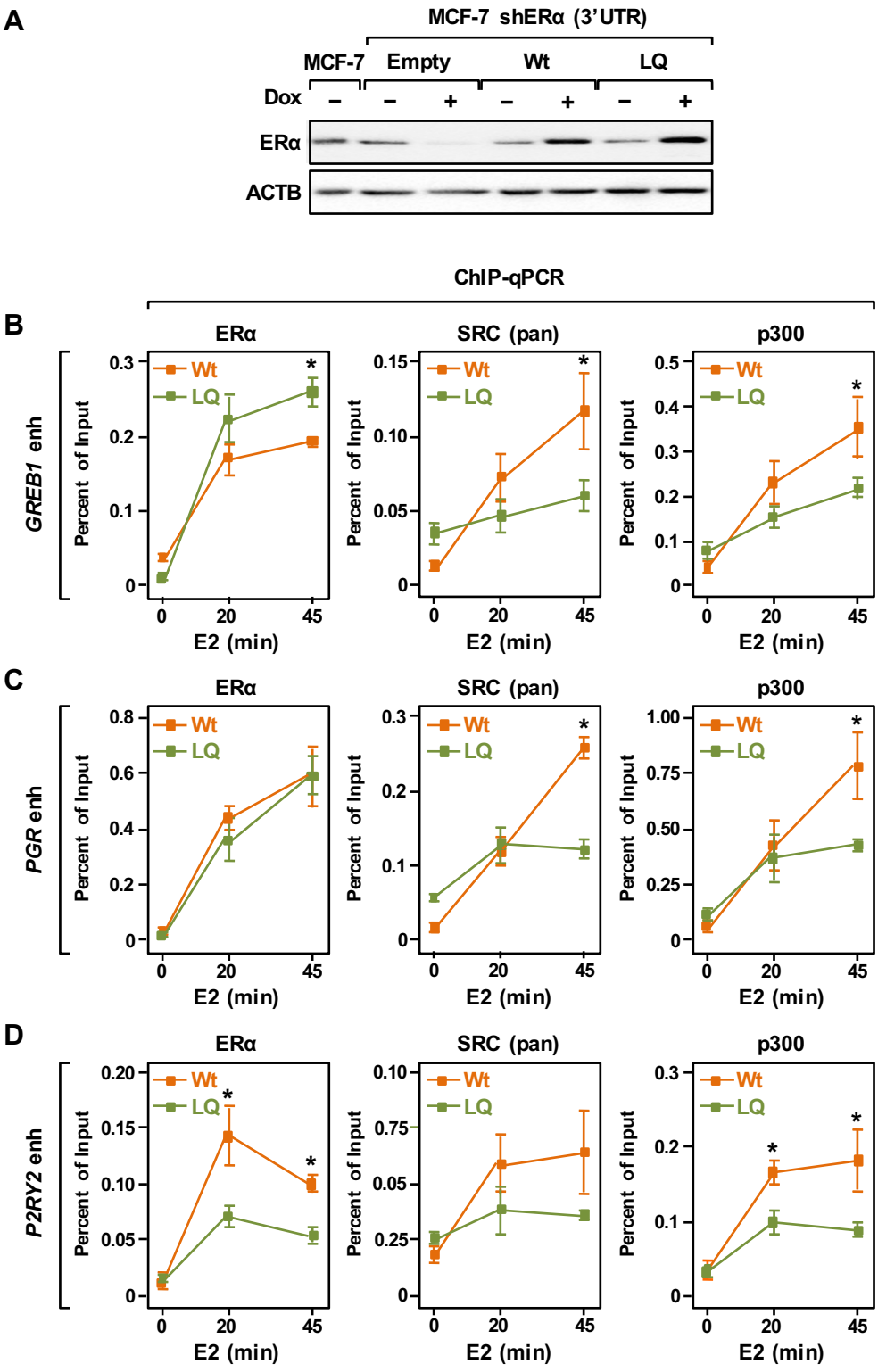


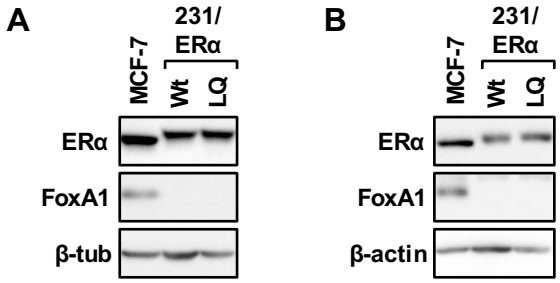
Supplemental Figure S12. p300 is recruited to ERαLQ binding sites similarly to ERαWt binding sites during enhancer priming, but not during enhancer maturation.
Box plots of p300 ChIP-seq read counts for a time course of E2 treatment at 274 ERα binding sites common to ERαWt and ERαLQ with >2-fold E2-dependent induction of p300 recruitment at 20 min. Box plots marked with different letters (*a*, *b*, *c*) are significantly different from each other ($p < 2.2 \times 10^{-16}$; Wilcoxon rank-sum test).

Supplemental Figure S13. Impaired SRC recruitment by ERα L540Q results in abortive p300 recruitment in MCF-7 cells.
(A) Western blotting for ERα in MCF-7 cells with Dox-dependent shRNA-mediated knockdown of endogenous ERα in combination with Dox-dependent re-expression of ERαWt or ERαLQ. The control sample with an empty re-expression vector shows the efficiency of knockdown.
(B-D) p300 recruitment at ERαLQ binding sites is induced during enhancer priming, but is attenuated during enhancer maturation in MCF-7 cells. ChIP-qPCR assays for ERα (*left*), SRC (pan) (*middle*), and p300 (*right*) at (B) *GREB1* enhancer, (C) *PGR* enhancer, and (D) *P2RY2* enhancer in the cell lines described in (A) treated with a time course of E2. The percent of input for SRC and p300 enrichment is normalized to the level of ERα binding in the corresponding conditions. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to Wt at each time point (*, $p < 0.05$; two-way ANOVA).

[Supplemental Figure S13 is on the next page]

Supplemental Figure S13.





Supplemental Figure S14. Differential expression of the pioneer factor FoxA1 in MCF-7, 231/ERαWt, and 231/ERαLQ cells.
(A and B) Western blots showing the expression of ERα and FoxA1 in MCF-7, 231/ERαWt, and 231/ERαLQ cells. (A) 231/ERα cells with constitutive expression of ERαWt or ERαLQ. (B) 231/ERα cells with Dox-inducible expression of ERαWt or ERαLQ.

Supplemental Figure S15. SRC-independent ERα enhancer priming requires Mediator *(continued)*.

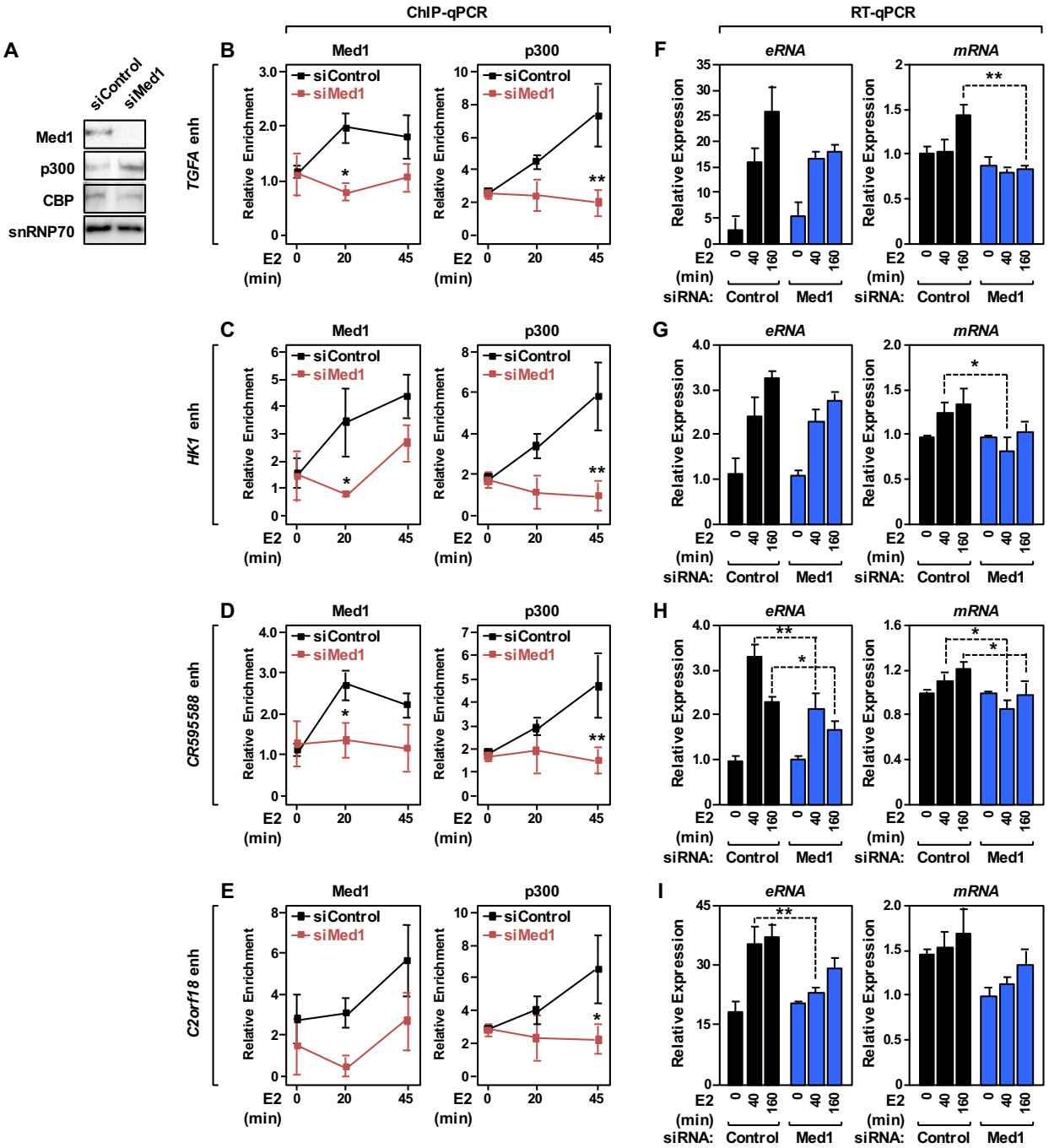
(A) Western blots showing the expression levels of Med1, p300, and CBP in 231/ERαWt cells with or without siRNA-mediated knockdown of Med1.

(B - E) SRC-independent recruitment of p300 requires Mediator. ChIP-qPCR assays for Med1 and p300 in Med1-depleted doxycycline (Dox)-inducible 231/ERαWt cells treated with a time course of E2. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the siRNA control at each time point (*, $p < 0.005$: **, $p < 0.0001$; two-way ANOVA).

(F - I) Mediator is required for E2-induced gene expression. RT-qPCR assays in Med1-depleted doxycycline (Dox)-inducible 231/ERαWt cells treated with a time course of E2. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the siRNA control at each time point (*, $p < 0.05$; **, $p < 0.005$; two-way ANOVA).

[Supplemental Figure S15 is on the next page]

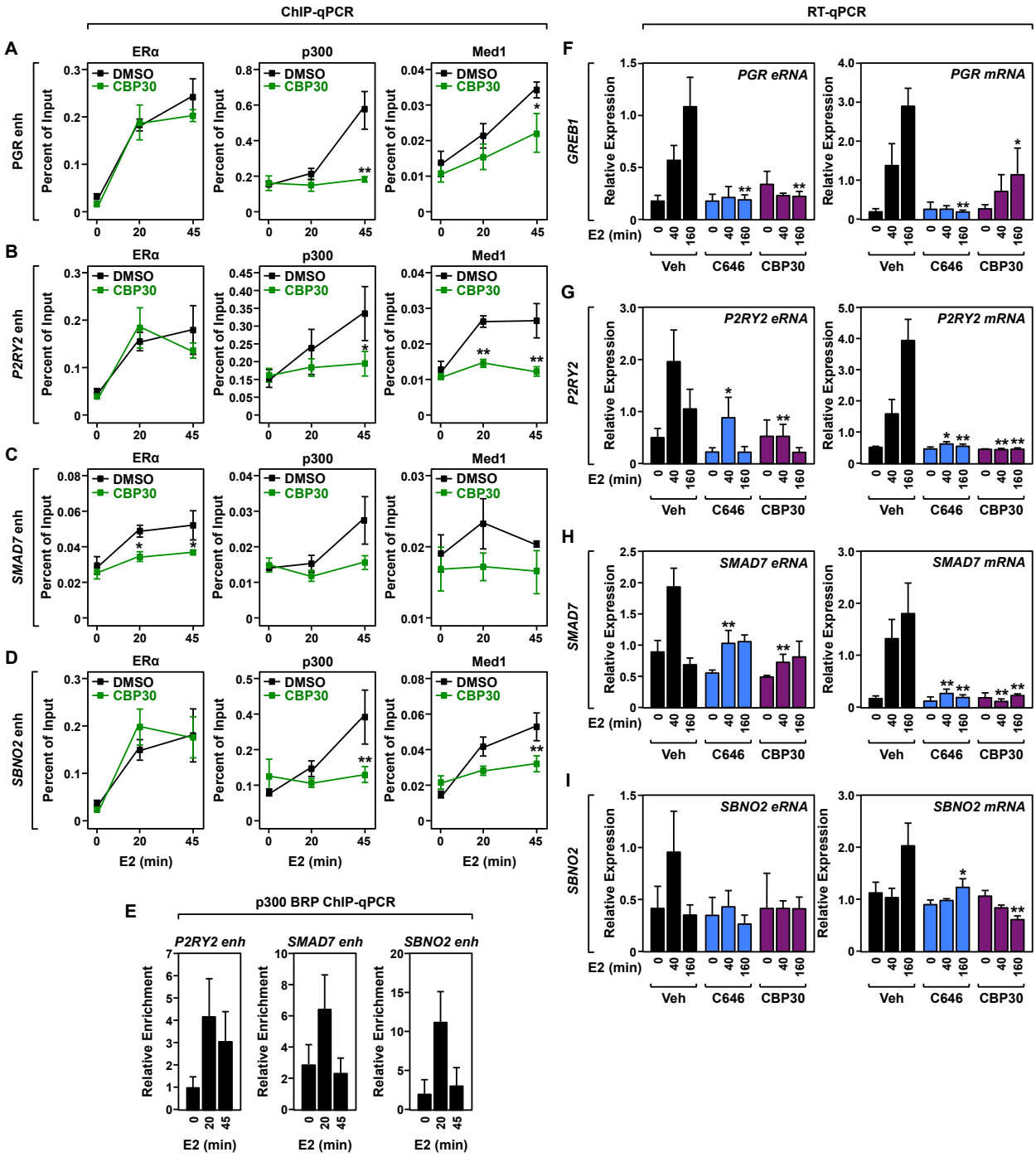
Supplemental Figure S15.



Supplemental Figure S16. The p300 bromodomain is required for ERα enhancer priming *(continued)*.
(A - D) The p300 bromodomain is required for p300 recruitment during enhancer priming. ChIP-qPCR assays for ERα, p300, and Med1 in MCF-7 cells treated with a time course of E2 in the presence of the p300/CBP bromodomain inhibitor SGC-CBP-30 (CBP30). Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the DMSO control at each time point (*, $p < 0.05$; **, $p < 0.005$; two-way ANOVA).
(E) The p300 bromodomain is recruited to ERα binding sites during enhancer priming. ChIP-qPCR assays for the IgG Fc-fused p300-BRP cassette in MCF-7 cells treated with a time course of E2. Each bar represents the mean + S.E.M. for at least three independent biological replicates.
(F - I) The p300 bromodomain and p300 acetyltransferase activity are required for E2-induced gene expression. RT-qPCR assays in MCF-7 cells treated with a time course of E2 in the presence of the p300 HAT inhibitor C646 or the bromodomain inhibitor SGC-CBP-30 (CBP30). Each bar represents the mean + S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the DMSO control at each time point (*, $p < 0.05$; **, $p < 0.005$; two-way ANOVA).

[Supplemental Figure S16 is on the next page]

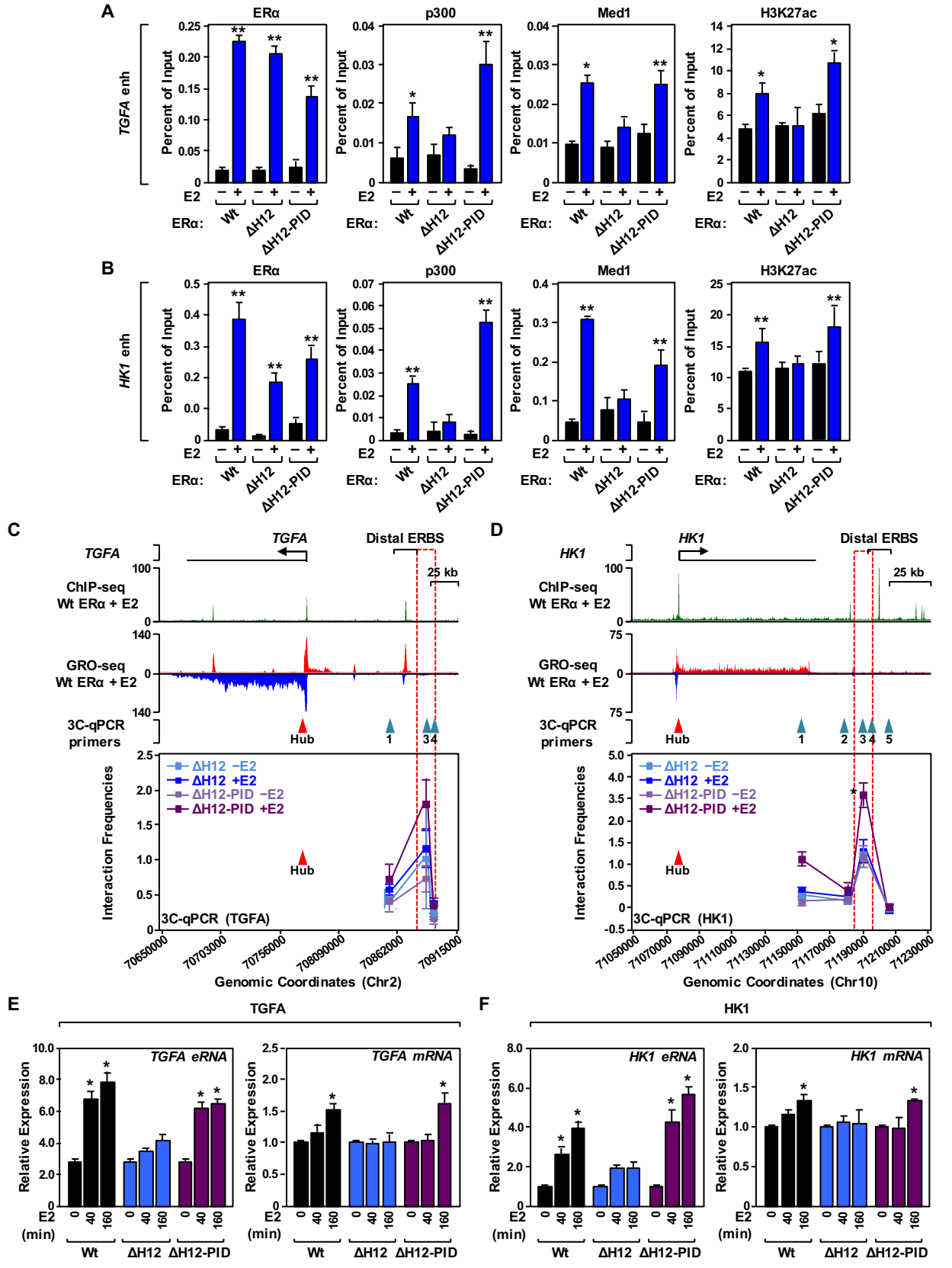
Supplemental Figure S16.



Supplemental Figure S17. Forced recruitment of p300 to ERα binding sites promotes Mediator recruitment, enhancer formation, and target gene expression *(continued)*.
(A and B) Forced recruitment of p300 to an inactive ERα binding site through the SRC2 PID restores Mediator recruitment and H3K27ac enrichment. ChIP-qPCR assays for ERα, p300, Med1, and H3K27ac in the MDA-MB-231 cell lines expressing the ERαs described in Figure 6A ± E2 treatment for 45 min. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the vehicle control in each cell line (*, $p < 0.05$; **, $p < 0.005$; two-way ANOVA).
(C and D) Forced recruitment of p300 to inactive ERα binding sites restores enhancer-promoter chromatin looping. (*Top*) Browser tracks for ERα ChIP-seq and GRO-seq in 231/ERαWt and after 45 min. of E2 treatment, shown with the location of the 3C primers. (*Bottom*) 3C-qPCR assays showing the chromatin looping from a distal ERα binding site to the *TGFA* and *HK1* promoters in 231/ERαΔH12 or 231/ERαΔH12-PID cells. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to 231/ERαΔH12 cells for each treatment (*, $p < 0.0001$; two-way ANOVA).
(E) Forced recruitment of p300 to an inactive ERα binding site restores E2-responsive eRNA production and gene expression. RT-qPCR assays in 231/ERαWt, 231/ERαΔH12 or 231/ERαΔH12-PID cells treated in a time course of E2. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Asterisks indicate significant differences compared to the vehicle control in each cell line (*, $p < 0.05$; two-way ANOVA).

[Supplemental Figure S17 is on the next page]

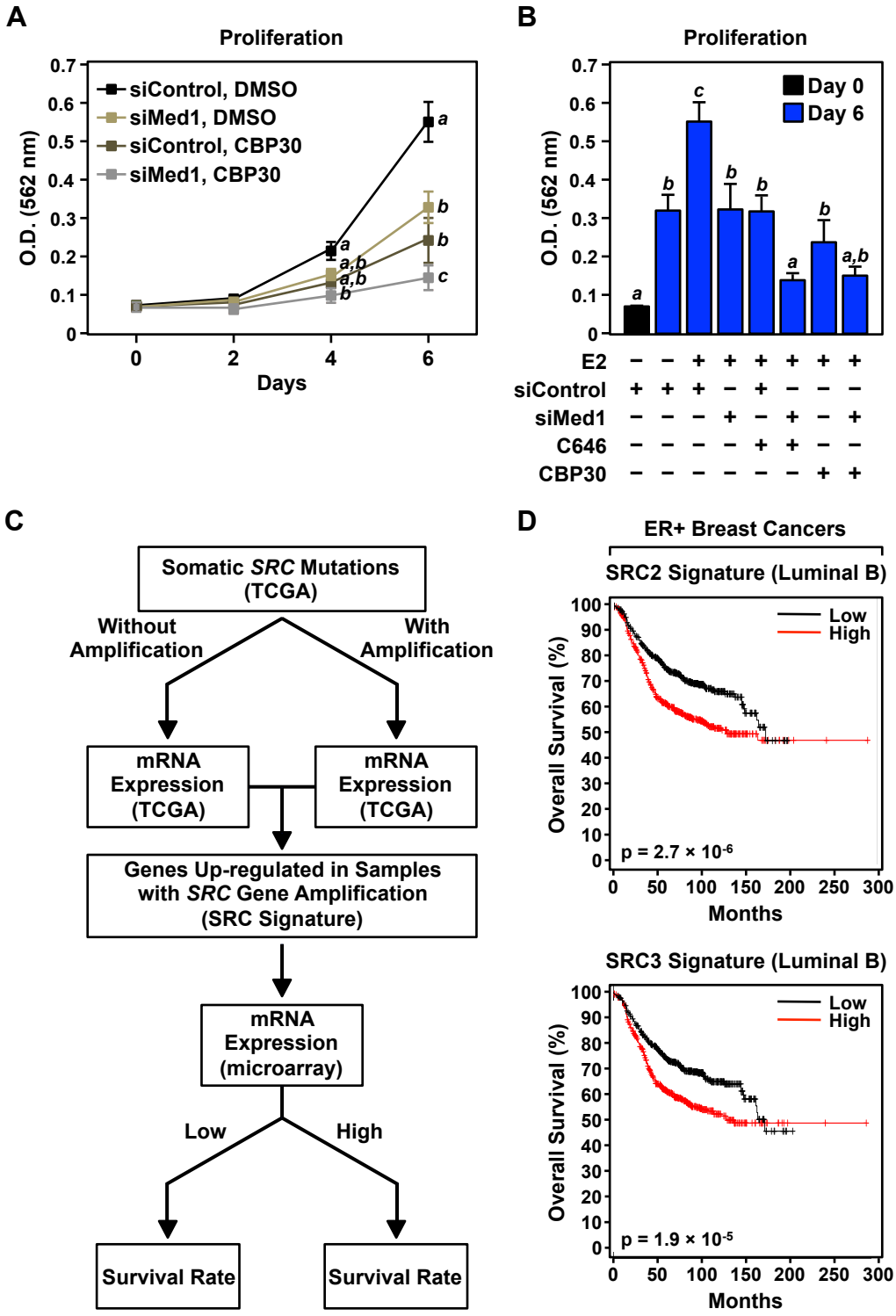
Supplemental Figure S17.



Supplemental Figure S18. p300, Mediator, and SRCs link ERα enhancer function to cell growth and clinical outcomes for ER-positive breast cancers *(continued)*.
(A) Growth curve showing the combinatorial effects of p300 inhibition and Med1 depletion on E2-dependent MCF-7 cell proliferation. Proliferation was measured after siRNA-mediated Med1-depletion ± the p300/CBP bromodomain inhibitor SGC-CBP-30 (CBP30) in the presence of E2. Each point represents the mean ± S.E.M. for at least three independent biological replicates. Points marked with different letters (*a*, *b, c*) are significantly different from each other ($p < 0.05$; two-way ANOVA).
(B) Proliferation assays showing the combinatorial effects of p300 inhibition and Med1 depletion on MCF-7 cell proliferation. The assays were performed as described in Fig. 7A and panel (A) above. The results are shown for 0 or 6 days of proliferation. Each bar represents the mean + S.E.M. for at least three independent biological replicates. Bars marked with different letters (*a*, *b, c*) are significantly different from each other ($p < 0.05$; one-way ANOVA).
(C) The pipeline of Kaplan-Meier analysis using a gene signature as collective genes up-regulated in tumor samples with *SRC2* or *SRC3* amplification compared to the tumor samples without amplification. Tumor samples on data from TCGA were divided in two groups depending on *SRC* gene amplification. Genes up-regulated in samples with *SRC* gene amplification relative to samples without amplification were denoted as signature genes. The expression of each signature gene (high or low) determined in curated microarray datasets were used to stratify ER-positive breast cancer patients into two groups. Clinical outcomes associated with the microarray datasets were plotted for patients in each group as in Fig. 7D.
(D) Kaplan-Meier plots for Luminal B ER-positive breast cancer patients using a set of genes (signature genes) whose expression is up-regulated in samples with *SRC2* or *SRC3* amplification compared to samples without amplification based on data from TCGA. The expression levels (high or low) of the signature genes determined in curated microarray datasets stratify patients into two groups. Their overall survival rates are shown in the plots.

[Supplemental Figure S18 is on the next page]

Supplemental Figure S18.



Supplemental Materials and Methods

Antibodies

The antibodies used were as follows: ERα (custom rabbit polyclonal antiserum generated in the Kraus Laboratory against the first 113 amino acids of human ERα (Kraus and Kadonaga 1998)); ERα (Enzo Biochem, ADI-SRA-1000-F); pan-SRC (rabbit polyclonal antiserum generated in the Kraus Laboratory against amino acids 624-1130 of mouse SRC2 (Acevedo et al. 2004)); p300 (Bethyl Laboratories, A300-358A); p300 (Santa Cruz Biotechnology, sc-584x); Med1 (Bethyl Laboratories, A300-793A); Med1 (Santa Cruz Biotechnology, sc-5334x); Pol II (Santa Cruz Biotechnology, sc-899x); H3K27ac (Abcam, ab4729); SRC1 (Santa Cruz Biotechnology, sc-8995); SRC2 (Santa Cruz Biotechnology, sc-8996); SRC3 (Santa Cruz Biotechnology, sc-9119); CBP (Santa Cruz Biotechnology, sc-369); β-actin (Cell Signaling, 3700S); snRNP70 (Abcam, ab83306); β-tubulin (Abcam, ab6046); and PARP-1 (custom rabbit polyclonal antiserum now available from Active Motif; cat. no. 39559).

Cell Culture and Treatments

MDA-MB-231 cells were obtained from the American Type Culture Collection (ATCC). Parental MDA-MB-231 cells, or MDA-MB-231 expressing ERα wild-type, L540Q, ΔH12, or ΔH12-PID, generated as described below, were maintained in phenol red-free Dulbecco's modified Eagle's medium/nutrient mixture F-12 Ham (DMEM/F-12; Sigma, D2906) supplemented with 10% charcoal-dextran treated calf serum (Sigma, C8056), 6 ng/mL human recombinant insulin (Sigma, I5500), 3.75 ng/mL hydrocortisone (Sigma, H0888), 16 μg/mL glutathione (Sigma, G6013), 100 units/mL penicillin/streptomycin (Gibco, 15140122), and 25 μg/mL gentamicin (Gibco, 15710064). The same conditions were used for experiments.

MCF-7 cells, kindly provided by Dr. Benita Katzenellenbogen (University of Illinois at Urbana-Champaign, IL), were maintained in minimal essential medium (MEM; Sigma, M1018) supplemented with 5% calf serum (Sigma, C8056), 100 units/mL penicillin-streptomycin (Gibco, 15140122), and 25 μg/mL gentamicin (Gibco, 15710064). Three days prior to experiments, the medium was switched to minimal essential medium Eagle (Sigma, M3024) supplemented with 5% charcoal-dextran treated calf serum (Sigma, C8056).

293T cells were obtained from the ATCC. There were maintained in Dulbecco's modified Eagle medium/high glucose (Sigma, D7777) supplemented with 10% FBS (Sigma, F2442), 100 units/mL penicillin-streptomycin (Gibco, 15140122), and 25 μg/mL gentamicin (Gibco, 15710064).

For experiments, the cells were treated with vehicle (DMSO) or 100 nM 17β-estradiol (E2) (Sigma, E8875) for the specified amount of time. Where indicated, the cells were treated with vehicle (DMSO), 25 μM of the p300/CBP acetyltransferase inhibitor C646 (Sigma, SML0002) (Bowers et al. 2010), or 50 μM the p300/CBP bromodomain inhibitor SGC-CBP30 (Sigma, SML1133) (Hay et al. 2014) for 30 min. prior to E2 treatment unless noted otherwise.

Preparation of Transgenic Cells with Ectopic Protein Expression

MDA-MB-231 cells with constitutive expression of ERα wild-type or L540Q were prepared as described previously (Acevedo et al. 2004). MDA-MB-231 cells with doxycycline-inducible expression of ERα wild-type, L540Q, ΔH12, or ΔH12-SRC2(PID) were generated by lentivirus-mediated transduction using the pINDUCER20 vector, kindly provided by Dr. Thomas Westbrook (Baylor College of Medicine, Houston, TX) (Meerbrey et al. 2011). Before use,

pINDUCER20 was modified to replace the Gateway® cloning sites with restriction enzyme sites suitable for conventional or Gibson cloning. cDNAs encoding the wild-type or variant ERαs noted above were constructed and cloned into the modified pINDUCER20 vector. A cDNA fragment encoding amino acids 1010-1130 of human SRC2 containing the p300-interacting domain (PID) (Kim et al. 2001; Acevedo and Kraus 2003) was cloned into the ERαΔH12 cDNA-containing vector to generate a C-terminal fusion. A cDNA fragment encoding the p300 Bromodomain-Ring-PHD domain (p300 BRP) (Delvecchio et al. 2013) was ligated to a cDNA fragment encoding the rabbit immunoglobulin G (IgG) Fc region and cloned into the modified pINDUCER20 vector. The protein expressed from this construct functions as an antibody-like p300 chromatin binding domain.

The pINDUCER20-based vectors described above were transfected into 293T cells using GeneJuice (Millipore, 70967) along with pCMV-VSV-G, pCMV-GAG-pol-Rev, and pAdVantage (Promega, E1711) for recombinant lentivirus production. The resulting viruses were filtered through a 0.44 μm filter and used to infect target cells. Infected cells were selected using 600 μg/mL G481 (Life Technologies, 11811031) to generate a pool of resistant cells stably harboring the transgene. After selection, ectopic expression of the transgenes was induced by treating the cells with 25 to 500 ng/mL doxycycline (Dox; titrated for each cell line to achieve similar expression levels) for 24 hours prior to experiments.

MCF-7 cells with simultaneous Dox-dependent shRNA-mediated knockdown of endogenous ERα and ectopic expression of ERα was prepared by sequential transduction of MCF-7 cells as follows. pTRIPZ-shRNA vectors were purchased from Dharmacon: shNegativeControl (RHS4743) and shERα targeting the *ESR1* 3'UTR (RHS4740-EG2099, V3THS_405935). The pTRIPZ-shRNA vectors were transfected into 293T cells using GeneJuice (Millipore, 70967) along with psPAX2 and pMD2.G for recombinant lentivirus production. The resulting viruses were filtered through a 0.44 μm filter, concentrated using Lenti-X™ concentrators (Clontech, 631232) according the manufacturer's protocol, and used to infect target cells. Infected cells were selected using 1 μg/mL Puromycin (Sigma, P8833) to generate a pool of resistant cells stably harboring the transgene. After selection, shRNA expression was induced by treating the cells with 200 to 500 ng/mL Dox for 72 hours prior to experiments. The knockdown efficiency was determined by Western blotting as described below. After confirming efficient knockdown, the MCF-7-pTRIPZ-shERα cells were transduced with the pINDUCER20-based vectors described above. After selection, the cells were treated with 10-500 ng/mL Dox for 72 hours to determine the Dox concentration that promotes efficient knockdown of endogenous ERα and re-expression of the ectopic ERα (Wt and L540Q mutant) from the transgenes at a level similar to endogenous ERα in MCF-7 cells.

siRNA-mediated Knockdown

siRNA-mediated knockdown was performed using Lipofectamine RNAiMax reagent (Invitrogen, 13778150) following the manufacturer's instructions using the following siRNAs: siControl (MISSION universal negative control #1; Sigma, SIC001); siMed1 (Sigma, SASI_Hs01_00089551); and siERα (Sigma, SASI_Hs01_00078598). The siRNAs were diluted to 100 nM in antibiotic-free, serum-free MEM and mixed by vortex in a tube. Lipofectamine RNAiMAX reagent was added to the diluted siRNAs and mixed by pipetting. The transfection complexes were transferred onto each well or plate by pipetting. The cells were seeded over the siRNA/lipofectamine transfection complex (typically 5 x 10^4 cells per well in a 6-well plate) to a

final siRNA concentration of 10 nM and incubated at 37°C, 5% CO_2 for 48 to 72 hours before experiments.

Western Blotting

The cells were collected by scraping in PBS containing complete protease inhibitor cocktail (Roche, 11697498001) and pelleted by gentle centrifugation. To prepare whole cell lysates, the cell pellets were resuspended in Lysis Buffer [20 mM HEPES pH 7.5, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 25% glycerol, 0.5% NP-40, 1 mM DTT, and 1x complete protease inhibitor cocktail (Roche, 11697498001)] and mixed by occasional gentle vortexing for 15 min at 4°C. The lysates were clarified by centrifugation at full speed in a microcentrifuge at 4°C. The supernatants were collected as whole cell lysates, snap frozen, and stored at -80°C until use.

Where indicated, the cells were separated into cytoplasmic and nuclear fractions. Briefly, the cells were lysed in Isotonic Buffer [10 mM Tris•HCl pH7.5, 300 mM sucrose, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 0.5% NP-40, 1 mM DTT, and 1x complete protease inhibitor cocktail (Roche, 11697498001)] by repeated pipetting. The nuclei were pelleted by gentle centrifugation and the supernatant was collected as cytoplasmic lysate, snap frozen, and stored at -80°C until used. The nuclear fraction was prepared by extracting the nuclei in Lysis Buffer with occasional gentle vortexing for 15 min at 4°C. The nuclear extracts were clarified by centrifugation at full speed in a microcentrifuge at 4°C, snap frozen, and stored at -80°C until used.

For assays, the cell lysates were thawed, and the protein concentration was measured by Bradford assay (BioRad, 5000006). The lysates were mixed with water and 5x SDS loading buffer, run on SDS-PAGE gels, and subjected to Western blotting.

Analysis of mRNA and eRNA Expression by RT-qPCR

RNA expression analysis by RT-qPCR was performed essentially as previously described (Hah et al. 2013). Total RNA was isolated using TRI Reagent (Sigma, T9424) and treated with RQ-1 DNase (Promega, M6101). 1 μg of total RNA was reverse transcribed using random hexamers (Roche. 11034731001) and M-MLV reverse transcriptase (Promega, M1705). mRNA and eRNA expression was analyzed by quantitative PCR (qPCR) with SYBR Green (Lonza, 50512) using a LightCycler 480 (Roche) and the following primers:

<u>MDA-MB-231 cells</u>

- *OTUB2* eRNA Forward 5’-GGAATTCCCAAAGAGCAAA-3’
- *OTUB2* eRNA Reverse 5’-TCTCGCCTGTGATGACTCAG-3’
- *OTUB2* mRNA Forward 5’-TCAGCAAAAGGTTCACCGC-3’
- *OTUB2* mRNA Reverse 5’-GTAGGAATAGCCCAAGGCCC-3’
- *TGFA* eRNA Forward 5’-TTTCTGTTCCTGGCTTGGCA-3’
- *TGFA* eRNA Reverse 5’-AGCCAGGTGACCTAGTGGTA-3’
- *TGFA* mRNA Forward 5’-GACTGGTCCCCCTTTCATGG-3’
- *TGFA* mRNA Reverse 5’-TCGTGAGCCCTCGGTAAGTA-3’
- *CR595588* eRNA Forward 5’-GCTCCAGGCAGTGTAGGAAG-3’
- *CR595588* eRNA Reverse 5’-AGACTCTGTTGGCCCTGTTG-3’
- *CR595588* mRNA Forward 5’-AGTCGGTGGGGTGTGAGTTA-3’
- *CR595588* mRNA Reverse 5’-TTGGGAAGCGTGGGTTATGT-3’
- *HK1* eRNA Forward 5’-AATTTCAGGGGAAGCCTGGG-3’

- *HK1* eRNA Reverse 5'-GACTCTCTGGCAGTCACACC-3'
- *HK1* mRNA Forward 5'-ACGTGTCCTTCCTCCTGTCT-3'
- *HK1* mRNA Reverse 5'-GATCCCGGACTCTTAGCTGC-3'
- *C2orf18* eRNA Forward 5'-TCCACATGGTTGTCTCTGCC-3'
- *C2orf18* eRNA Reverse 5'-TGCCTGAAGCTTGACCTCTG-3'
- *C2orf18* mRNA Forward 5'-ACCTGCCTGCCTAGAGAACT-3'
- *C2orf18* mRNA Reverse 5'-CAACGCCCAGGATACCAGAA-3'

MCF-7 cells

- *GREB1* eRNA Forward 5'-TGCTGGCTGCTTAAAAACCT-3'
- *GREB1* eRNA Reverse 5'-TGAAAACCCACACTTCCAAA-3'
- *GREB1* mRNA Forward 5'-CCTATTTTGGAATAAAAACTGACC-3'
- *GREB1* mRNA Reverse 5'-GGGGAGAATGACACAAAAGC-3'
- *P2RY2* eRNA Forward 5'-AAGGGCTAATGTTTGGCACA-3'
- *P2RY2* eRNA Reverse 5'-AGGGAGGTCCAGGAGGTCTA-3'
- *P2RY2* mRNA Forward 5'-CGGTGGACTTAGCTCTGAGG-3'
- *P2RY2* mRNA Reverse 5'-GCCTCCAGATGGGTCTATGA-3'
- *PGR* eRNA Forward 5'-CATTGAGTCATGGCCTTTGAT-3'
- *PGR* eRNA Reverse 5'-CCTTTCAGATGGGAGCTAGG-3'
- *PGR* mRNA Forward 5'-TTGCCAAGAAGGTGAAACTG-3'
- *PGR* mRNA Reverse 5'-CTTTGCATTGTCACCCCATC-3'
- *SBNO2* eRNA Forward 5'-CCTGTATTCTGGGGGCACTA-3'
- *SBNO2* eRNA Reverse 5'-CTCACCCCATCCAGTACACC-3'
- *SBNO2* mRNA Forward 5'-GACTGGGCACCCACAAGGGC-3'
- *SBNO2* mRNA Reverse 5'-GGAAGGGCTGGGGGAGGGAG-3'
- *SMAD7* eRNA Forward 5'-GGCATAGCTAGGACCTCACC-3'
- *SMAD7* eRNA Reverse 5'-GAGGGAGGAAAGTGGCTTCT-3'
- *SMAD7* mRNA Forward 5'-AAGAGAAGCATTCTCATTGGAAA-3'
- *SMAD7* mRNA Reverse 5'- TCAGGAGTCCTTTCTCTCTCAAA-3'

Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) was performed as previously described with some modifications (Hah et al. 2013). Cells were grown to ~80% confluence and were crosslinked with 1% formaldehyde in PBS at 37°C for 10 min. The crosslinking reaction was quenched by adding glycine to a final concentration of 125 mM. The cells were collected by scraping in 1x PBS containing 1x complete protease inhibitor cocktail (Roche, 11697498001). The cells were pelleted by brief centrifugation in a microcentrifuge and lysed by pipetting in Farnham Lysis Buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40, 1 mM DTT, 1x complete protease inhibitor cocktail). The nuclei were collected by brief centrifugation in a microcentrifuge and resuspended in SDS Lysis Buffer (Tris•HCl pH 7.9, 1% SDS, 10 mM EDTA, 50 mM, 1 mM DTT, 1x complete protease inhibitor cocktail) by pipetting and incubation on ice for 10 min. The chromatin was then sheared to ~200-500 bp DNA fragments by sonication using a Bioruptor Plus sonicator (Diagenode) for 25-30 cycles of 30 seconds on and 30 seconds off. Protein concentrations in the solubilized chromatin were determined using the BCA Protein Assay Kit (Pierce, 23225), and a normalized amount of soluble chromatin was precleared with Protein A or Protein G Dynabeads (Invitrogen, 10001D and 10003D,

respectively) or Protein A or Protein G agarose beads (Millipore, 16-125 and Invitrogen, 15920010, respectively), and incubated overnight with 8 μL of polyclonal antiserum or 2.5 to 5.0 μg of commercial antibody.

The immune complexes from the ChIP were precipitated by the addition of Protein A or Protein G Dynabeads or agarose beads (depending on the antibody used) and washed once with each of the following wash buffer in sequence: (1) Low Salt Wash Buffer (20 mM Tris•HCl pH 7.9, 2 mM EDTA, 125 mM NaCl, 0.05% SDS, 1% Triton X-100, 1x complete protease inhibitor cocktail); (2) High Salt Wash Buffer (20 mM Tris•HCl pH 7.9, 2 mM EDTA, 500 mM NaCl, 0.05% SDS, 1% Triton X-100, 1x complete protease inhibitor cocktail); (3) LiCl Wash Buffer (10 mM Tris•HCl pH 7.9, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1x complete protease inhibitor cocktail); and (4) 1x Tris-EDTA (TE) containing 1x complete protease inhibitor cocktail. The precipitated immune complexes were transferred to a new tube in 1x TE/complete protease inhibitor cocktail before elution of the genomic DNA fragments in Elution Buffer [40 mM Tris•HCl pH 7.9, 10 mM EDTA, 100 mM NaCl, 100 mM $NaHCO_3$, 1% SDS, and 50 μg Proteinase K (Life Technologies, 2542)] for 2 hours at 55°C. H3K27ac ChIP assays were performed in the presence of the deacetylase inhibitors nicotinamide (5 mM) and sodium butyrate (10 mM) until the elution of the ChIPed DNA. The crosslinks were reversed by incubating overnight at 65 °C, and the genomic DNA was purified using phenol:chloroform:isoamyl acid extraction (Sigma, P2069) followed by ethanol precipitation. The ChIPed DNA was (1) analyzed by qPCR with SYBR Green (Lonza, 50512) using a LightCycler 480 (Roche) and the primers listed below or (2) subjected to ChIP-seq library preparation as described below. Non-specific background signals in all ChIP assays were determined using IgG (for purified antibodies) or no antibody control (for antiserum). The data were expressed as the percent of input or relative enrichment (fold change).

ChIP using the p300 BRP-Fc fusion protein was performed using the standard ChIP-seq protocol described above except that the antibody addition was omitted. Instead, chromatin-associated p300 BRP-Fc was directly precipitated with Protein A Dynabeads using the rabbit Fc tag.

Primers for ChIPing non-histone proteins

• *OTUB2* enhancer Forward	5'-GGAATTCCCAACAGAGCAAA-3'
• *OTUB2* enhancer Reverse	5'-TCTCGCCTGTGATGACTCAG-3'
• *TGFA* enhancer Forward	5'-GGAGAAAGGAGGTGGAACGG-3'
• *TGFA* enhancer Reverse	5'-GACTCAAAGTGACAGGGGCA-3'
• *CR595588* enhancer Forward	5'-GTCACTTGTTCTCCTGCGTG-3'
• *CR595588* enhancer Reverse	5'-GGGAAGCAGTGCTCATCCAG-3'
• *HK1* enhancer Forward	5'-CCCTCCTGAATGACAGATGG-3'
• *HK1* enhancer Reverse	5'-CTGCCTGACTCACACTGGAA-3'
• *C2orf18* enhancer Forward	5'-CATGTGACCCCAAAGAGGAG-3'
• *C2orf18* enhancer Reverse	5'-CATCCAGGCTTAACCAGAGG-3'
• *GREB1* enhancer Forward	5'-TGCTGGCTGCTTAAAAACCT-3'
• *GREB1* enhancer Reverse	5'-TGAAAACCCACACTTCCAAA-3'
• *P2RY2* enhancer Forward	5'-AAGGGCTAATGTTTGGCACA-3'
• *P2RY2* enhancer Reverse	5'-AGGGAGGTCCAGGAGGTCTA-3'
• *PGR* enhancer Forward	5'-ATGCAGAGCCATTGCAAAAT-3'

• *PGR* enhancer Reverse	5'-ATGCAGAGCCATTGCAAAAT-3'
• *SBNO2* enhancer Forward	5'-CCTGTATTCTGGGGGCACTA-3'
• *SBNO2* enhancer Reverse	5'-CTCACCCCATCCAGTACACC-3'
• *SMAD7* enhancer Forward	5'-GGCATAGCTAGGACCTCACC-3'
• *SMAD7* enhancer Reverse	5'-GAGGGAGGAAAGTGGCTTCT-3'

Primers for ChIPing H3K27ac

• *OTUB2* enhancer Forward	5'-CCGAGCCTCTCCTCATTTCC-3'
• *OTUB2* enhancer Reverse	5'-CCATCAATGGTGGCAGGAGA-3'
• *TGFA* enhancer Forward	5'-TTTCTGTTCCTGGCTTGGCA-3'
• *TGFA* enhancer Reverse	5'-AGCCAGGTGACCTAGTGGTA-3'
• *CR595588* enhancer Forward	5'-ACAGGGCCAACAGAGTCTTG-3'
• *CR595588* enhancer Reverse	5'-CATGCTGCACACAGATCACG-3'
• *HK1* enhancer Forward	5'-TGCTGACAATCCAGCAAGGAA-3'
• *HK1* enhancer Reverse	5'-GATTTACTCGGAGAGTGCCCC-3'
• *C2orf18* enhancer Forward	5'-AACACAGGACAAGGGAGCAG-3'
• *C2orf18* enhancer Reverse	5'-GGGGTCAGGCAGACACATAC-3'

ChIP-seq Library Preparation

ChIP-seq libraries were prepared as previously described with some slight modifications (Luo et al. 2014). Briefly, input DNA or ChIPed DNA was subjected to additional purification using Agencourt AMPure XP beads (Beckman Coulter, A63881). Two and a half to 10 ng of purified genomic DNA was subjected to end repair using an end-repair mix (Enzymatics, Y9140-LC-L) and 0.1 mM dNTPs. A single dA base was added to the end repaired DNA using Klenow 3′→5′ Exo-minus (Enzymatics, P7010-HC-L), 1x Blue Buffer (Enzymatics, B0110), and 0.2 mM dATP to facilitate adaptor ligation. TruSeq DNA Sample Prep Kit adaptors (custom synthesized by IDT and HPLC purified) were partially annealed to form double-stranded adaptors on one terminus, and annealed to add DNA through double-stranded DNA-DNA ligation using T4 DNA ligase (Enzymatics, L6030-HC-L) in DNA Rapid Ligase Buffer (Enzymatics, B1010). The adaptor-ligated DNA was amplified by PCR for 6 to 11 cycles and purified by electrophoresis on a 1% agarose gel containing SYBR Gold (Invitrogen, S11494). The DNA was excised from the agarose gel and eluted using a QIAquick Gel Extraction Kit (Qiagen, 28706). The quality of the library was assessed using a D1000 ScreenTape (Agilent, 5067-5582) on a 2200 TapeStation (Agilent) and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Q32851). The libraries with unique adaptor barcodes were multiplexed and sequenced on an Illumina HiSeq 2000 (single-end, 50 base reads).

Analysis of ChIP-seq Data

Quality Control and Alignment. Quality of the ChIP-seq datasets was assessed using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The ChIP-seq reads were aligned to the human reference genome (hg19) using BOWTIE (version 0.12.7) with the default parameters (Langmead et al. 2009). Uniquely mapped reads were visualized on the UCSC genome browser as bigWig files prepared using BEDTools (Quinlan and Hall 2010) and custom R scripts (available on request).

Peak Calling. ERα peak calling from a total of four replicates per condition in two sequencing runs was performed using MACS software (Zhang et al. 2008) using the default p-

value and input condition as a control. Peak calling for wild-type (Wt) and L540Q (LQ) ERα was performed separately. The peak calling was performed separately for each sequencing run, and the peak calls from the two runs were compared to yield a set of common peaks. The final universe of ERα peaks from Wt and LQ were then compared using the mergePeaks function in HOMER (Heinz et al. 2010) to identify ERα peaks shared between the Wt and LQ conditions. The read counts ± 500 bp around the summit of the common ERα peaks were used to calculate the fold change (FC) where,

$$FC = Wt/LQ$$

In the subsequent analyses, a set of common ERα peaks with approximately equivalent binding strength in Wt and LQ was selected using the cut off $0.75 < FC < 1.25$. This set of ERα peaks was used to determine differences between Wt and LQ in the recruitment of coregulators or the enrichment of H3K27ac. For boxplot representations of ChIP-seq data, read counts ± 500 bp surrounding the summit of the common ERα peaks were counted and visualized.

Comparing Called Peaks. Peaks for wild-type (Wt) and L540Q (LQ) ERα were rank ordered separately by the number of read counts in a fixed 1 kb window (± 500 bp) surrounding each ERα peaks. The genomic coordinates of the top 6500 peaks for each ERα (i.e., ERαWt or ERαLQ) were compared. The overlap of the ERαWt and ERαLQ peaks were then visualized in a Venn diagram.

Chromosome Conformation Capture (3C) Assays

Chromosome conformation capture (3C) assays were performed as previously described with some modifications (Dekker et al. 2002; Wang et al. 2009). The cells in one 10 cm diameter plate at ~80% confluence (~4 x 10^6 cells) were crosslinked in 1% formaldehyde in PBS at 37°C for 10 min. Crosslinking was quenched by the addition of glycine to a final concentration of 125 mM. The cells were scraped and collected in PBS containing 1x complete protease inhibitor cocktail (Roche, 11697498001), and then pelleted by centrifugation in a microfuge at 4°C. The cells were resuspended in 3C Lysis Buffer (10 mM Tris•HCl pH7.5, 10 mM NaCl, 0.2% Triton X-100, 1 mM DTT, 1x complete protease inhibitor cocktail) and lysed by pipetting and gentle mixing at 4°C for 30 min. to release the nuclei, which were collected by centrifugation in a microcentrifuge at 500 x g for 5 min. The pelleted nuclei were resuspended in 3C Digestion Buffer [1.2x NEB Restriction Enzyme Buffer 3 (NEB, B7003S), 0.3% SDS] and incubated at 37°C for 1 hour with gentle mixing. Triton X-100 was then added to the nuclear suspension to a final concentration of 2%. After continued incubation at 37°C for 1 hour with gentle mixing, 400 units *BglII* restriction enzyme (NEB, R0144S) were added and the chromatin was digested overnight at 37°C with gentle mixing. SDS was added to the *BglII*-digested chromatin to a final concentration of 1.6% and the mixture was incubated in 65°C for 20 min. to quench the digestion. The digested chromatin was then transferred to a 15 mL conical tube, diluted with 6.5 mL of 3C Ligation Buffer (50 mM Tris•HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 1% Triton X-100, 10 mM DTT), and incubated at 37°C for 1 hour with gentle mixing.

The digested chromatin was then ligated with 2000 units of T4 DNA Ligase (NEB, M0202L) at 16°C for 4 hours, followed by 30 min. at room temperature. The ligated chromatin was deproteinized by treatment with 300 μg Proteinase K (Sigma, P2308) at 37°C for 1 hour, reverse crosslinked at 65°C overnight, purified by phenol:chloroform:isoamyl acid (Sigma, P2069) extraction, and precipitated with ethanol. The 3C DNA was collected by centrifugation,

resuspended in 1x TE, and further purified using a NucleoSpin Gel and PCR Clean-Up Kit (Takara, 740609.250) with buffer NTB (Takara, 740595.150) following the manufacturer's instructions. Bacterial artificial chromosomes (BAC) spanning the loci to be tested were subjected to digestion with *BglII*, ligation, phenol:chloroform:isoamyl acid purification, and ethanol precipitation to generate a standard curve to normalize the amplification efficiency for each primer set.

The purified 3C DNA was quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Q32851) and then normalized to (1) 2.5 ng/μL for qPCR assays with SYBR Green (Lonza, 50512) or (2) 25 ng/μL for qPCR assays with TaqMan probes (Thermo Fisher Scientific) using a QuantiTech Probe PCR Kit (Qiagen, 204343) with the primers and probes listed below. qPCR assays with SYBR Green were performed on the 3C DNA to quantify: (1) the input of the locus to be tested using the loading primer mix, (2) the input of the *GAPDH* locus using the *GAPDH* loading primer mix, and (3) the interaction frequencies of the *GAPDH* locus for normalization using the *GAPDH* interaction primer mix. A serial dilution of the 3C DNA was used to generate a standard curve for each primer set. qPCR assays with TaqMan probes were performed to quantify the interaction frequencies in the locus to be tested using the hub and each test primer. A standard curve for the TaqMan qPCR assay was generated using a serial dilution of the digested/ligated BAC control DNA. The interaction frequencies determined by the TaqMan qPCR assay were normalized to the input of the test locus, the input of the *GAPDH* locus, and the interaction frequencies of the *GAPDH* locus.

BAC templates:

Locus	*BAC Clone ID*
• *OTUB2*	RP11-666E24
• *TGFA*	RP11-771E16, RP11-36L20
• *CR595588*	CTD-2363K16
• *HK1*	RP11-652D17
• *C2orf18*	CTD-2246L24

TaqMan probes:

Locus	*Sequence (5' to 3')*
• *OTUB2*	TTCCTCTCCGGGCCTGACCT
• *TGFA*	ATCTAGGAAACCTCCGTGGGGCTAGTCT
• *CR595588*	CTCGTGAGGCTTATTCACTACCATGAGAACAGG
• *HK1*	TCTAGGATCACAGCTTGGATCTGTGAGTC
• *C2orf18*	ATCTTCAGTGTCCAGGAAGAAGGTACGG

Primers:

Locus	*Name*	*Sequence (5' to 3')*
• *GAPDH*	Interaction Forward (Wang et al. 2009)	CCTTCTCCCCATTCCGTCTT
	Interaction Reverse (Wang et al. 2009)	TGTGCGGTGTGGGATTGTC
	Loading Forward (Wang et al. 2009)	ACAGTCCATGCCATCACTGCC
	Loading Reverse (Wang et al. 2009)	GCCTGCTTCACCACCTTCTTG
• *OTUB2*	Loading Forward	CAGTAATGTTCTCAGACTTC
	Loading Reverse	CTAGAGCTCTGACTCCAC
	Hub	AGTCAGAGCTCTAGGGA

	Test 1	AGTCAGGCAGGGAGAT
	Test 2	GGGATTCACACCCAGAT
	Test 3	AAAGGAGGTGTCGTCTAG
	Test 4	ATGGGTTTGGAGCAGAT
	Test 5	CGAAGAAAGAAGCCCTTTATAA
	Test 6	CTGGGGGCTTCTTAGAT
• *TGFA*	Loading Forward	CAATGCTCAGGTTCCAAGTA
	Loading Reverse	CTGTTAGGAGTCTCGGTTAATG
	Hub	AGACTCCTAACAGCCAGTT
	Test 1	CAGAGAGAAGGTGCTGTG
	Test 2	TGGATTCAAATCCAGGATCC
	Test 3	TGAATCAGTAGTCGGAATATAGA
• *CR595588*	Loading Forward	ACAAAAGAGGAATGATGGCT
	Loading Reverse	CCAGGAAAGATGAGAAGCAT
	Hub	TGAATCATGGGGACAGTTTC
	Test 1	GCGACCTCTTTCTACCA
	Test 2	CCGCCTGGGATAAAAGTT
	Test 3	CATCTTTGGCCTTTCCAGA
	Test 4	GACCTTCTGCTCTTAAGAAAAC
• *HK1*	Loading Forward	GCATCTAAGCTCCTCCTTTT
	Loading Reverse	CTCTACCCTAGCTCTTGACT
	Hub	TCTATACAACTGGGACCAC
	Test 1	ATAGCTTCTCTTGAAAGATTTAGA
	Test 2	GGTAGTAGACACTTCTAAACAAC
	Test 3	CCTTGCATGAGCCACAC
	Test 4	CTCACTGTCAAGTTATCAAGAA
• *C2orf18*	Loading Forward	TCTTCTGTGTCCTTTCTGTG
	Loading Reverse	ACTCTACTACACTGTCCTCC
	Hub	TGTAAAATGGACTTGGTGAT
	Test 1	ATTCTGTGCCTGCAAAGAA
	Test 2	GAAACTACGGAGTGTGTTTG
	Test 3	ATAGGATGGAAACCACCAGT
	Test 4	TCTTTGATCAGGGTCAGGT
	Test 5	CCCTGCTCACATCTCCTAA

Preparation of Global Run-on Sequencing (GRO-seq) Libraries

GRO-seq libraries were prepared as previously described with some modifications (Luo et al. 2014). Approximately 2 x 10^6 cells were seeded on a 15 cm^2 plate and grown to ~80% confluence. The cells were washed twice with ice-cold PBS and collected by scraping in Swelling Buffer [10 mM Tris•HCl pH 8.0, 2 mM MgOAc, 3 mM $CaCl_2$, 0.5 mM DTT, 1x complete protease inhibitor cocktail, and SUPERase•In (Ambion)] on ice. The cells were pelleted by centrifugation and resuspended in Lysis Buffer [10 mM Tris•HCl pH 8.0, 2 mM MgOAc, 3 mM $CaCl_2$, 10 mM NaCl, 0.5% NP-40, 300 mM sucrose, 0.5 mM DTT, 1x complete protease inhibitor cocktail, and SUPERase•In (Ambion)]. To isolate the nuclei, the swollen cells

were lysed by pipetting up and down 60 times through a narrow bore tip. The nuclei were pelleted by brief centrifugation, equilibrated, and dispersed in Freezing Buffer to 5 x 10^6 nuclei/100 µL [50 mM Tris•HCl pH 8.3, 5 mM $MgCl_2$, 40% glycerol, 0.1 mM EDTA, and SUPERase•In (Ambion)], aliquoted, and stored at -80°C.

Run-on was performed as previously described with some modifications (Luo et al. 2014). The nuclei were incubated in the presence of 5'-bromo-UTP and α-^{32}P radiolabeled-CTP for 5 min., with subsequent quenching by the addition of RQ-1 DNase (Promega, M6101). RNA was isolated from the run-on reaction mixture and hydrolyzed by incubating with 0.2 N NaOH on ice for 15 min. The fragmented RNA was treated with T4 PNK (NEB, M0201S) in the absence of ATP to dephosphorylate the 3'-terminus. Nascent transcripts were isolated from the total RNA with two rounds of affinity purification using anti-BrdU antibody-conjugated agarose beads (Santa Cruz Biotechnology, sc-32323). PolyA tails were added to the nascent transcripts using 1 mM ATP and 0.5 U/µL *E. coli* Poly(A) Polymerase (NEB, M0276S) at 37°C for 8 min. cDNA was generated using the oNTI223HIseq primer (custom synthesized by IDT and HPLC purified) (Ingolia et al. 2009) and purified on an 8% polyacrylamide TBE-urea gel. The purified cDNA was then circularized using CircLigase (Epicentre, CL4111K), relinerized using APE1 (NEB, M0282S), and amplified using TruSeq small RNA-seq PCR primers (custom synthesized by IDT and HPLC purified) to generate the GRO-seq library. The library quality was assessed using a D1000 ScreenTape (Agilent, 5067-5582) on a 2200 TapeStation (Agilent) and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Q32851). The libraries with unique adaptor barcodes were multiplexed and sequenced on an Illumina HiSeq 2000 (single-end, 50 bp reads) for a total of ~47 million raw reads per biological replicate. The experiment has two biological and two technical replicates. The technical replicates were merged into a single data file before further analysis after confirming strong positive correlations between them.

Analysis of GRO-seq Data and Integration with ChIP-seq Data

GRO-seq data analyses were performed as previously described (Hah et al. 2013; Nagari et al. 2017) using the groHMM software package available from Bioconductor (http://bioconductor.org/packages/release/bioc/html/groHMM.html) (Danko et al. 2014; Chae et al. 2015).

Quality Control. Quality of the GRO-seq datasets was assessed using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The GRO-seq reads were subjected to trimming using Cutadapt (Martin 2011) to remove the polyA tails and adapter sequences to maximize the mappability.

Alignment and Gene Annotation. The trimmed GRO-seq reads were aligned to the human reference genome (hg19) and one complete copy of an rDNA repeat (GenBank ID: U13369.1) using the BWA aligner (Li and Durbin 2010). The mapped reads were visualized as UCSC genome browser tracks after conversion to bigWig files using the groHMM package (Danko et al. 2014; Chae et al. 2015). A complete set of annotated coding genes was assembled from the RefSeq, UCSC, and Gencode databases. Overlapping annotations were removed from the assembled set of genes to avoid multiple counting. The set of annotated long non-coding RNA (lncRNA) genes was downloaded from the LNCipedia 2.0 database. The lists of coding and lncRNA gene annotations were combined and used in the subsequent analysis.

Calling Differential Gene Expression. Differential gene expression between experimental conditions was determined using edgeR (Robinson et al. 2010). Expression levels

in a given condition were determined using the number of read counts between +1 to +13 kb from the 5' end of the gene annotations (Hah et al. 2011). Significantly regulated genes were determined using a false discovery rate (FDR) cutoff of 5%. The read counts +1 to +13 kb from the 5' end of each gene were visualized in boxplots as Reads Per Kilobase of gene per Million mapped reads (RPKM).

Calculation of Pausing Indexes. The read counts in windows from the TSS to +1 kb and from +1 kb to the 3' end to the end of the gene were collected as "TSS" and "gene body," respectively, and represented in boxplots as Reads Per Kilobase of gene per Million mapped reads (RPKM). Pausing indexes were calculated using the pausing index function in the groHMM software package (Chae et al. 2015).

Box Plots. For ChIP-seq, box plot representations were used to quantitatively represent the read distribution in a fixed 1 kb window (± 500 bp) surrounding each ERα binding site using the box plot function in R. Wilcoxon rank sum tests were performed to determine the statistical significance of all comparisons.

Metaplots. Metaplots of GRO-seq data were generated using the metagene function in the groHMM software package (Chae et al. 2015)

Analysis of ERα Binding Sites Nearest to Regulated Genes. The ERα peaks from the ChIP-seq analyses that were (1) present in both the Wt and LQ conditions and (2) located nearest to the genes up-regulated in the Wt 45 min. E2 condition were obtained using custom Perl scripts (available on request). The GRO-seq RPKM values of the gene (i.e., gene transcription) and the ERα peaks (i.e., enhancer transcription) were represented as box plots and line plots (average RPKM).

Cell Proliferation Assays

On day -2 (minus 2), MCF-7 cells were seeded at 4,000 cells per well in 24-well plates in minimal essential medium Eagle (Sigma, M3024) supplemented with 5% charcoal-dextran treated calf serum (Sigma, C8056) and allowed to rest at 37°C, 5% CO_2 for 16 hours. On day -1 (minus 1), the specified siRNAs were transfected using Lipofectamine RNAiMax (Thermo Fisher Scientific, 13778150) according to the manufacturer's protocol. On day 0, the cells were treated with control vehicle (DMSO), 4 μM C646, or 10 μM SGC-CBP-30 ± 100 nM E2. The cells were collected every two days, fixed in 10% formaldehyde for 10 min., and stored at 4°C. After the final time point was collected, all samples were stained with 0.1% crystal violet in 75 mM phosphoric acid for 30 min. After washing with a large volume of water, the crystal violet was extracted from the cells using 10% acetic acid and measured as absorbance at 562 nm.

Analysis of Somatic Mutations, Copy Number Alterations, and Gene Expression Using TCGA Data Sets

Relevant gene sets and expression data were accessed from The Cancer Genome Atlas (TCGA) (http://www.cbioportal.org) (Cerami et al. 2012; Network 2012; Gao et al. 2013). Bar graphs showing the somatic mutations and copy number alterations in the *NCOA1* (SRC1), *NCOA2* (SRC2), and *NCOA3* (SRC3) genes were generated using the OncoPrint tool on cBioPortal (http://www.cbioportal.org/index.do) and the METABRIC breast cancer cohort of TCGA datasets (Network 2012; Pereira et al. 2016). The ER-positive dataset was selected for ER status in the clinical attributes on the study summary page, the genes of interest were specified in the query window, and OncoPrint was generated in response to the query (Cerami et al. 2012; Network 2012; Gao et al. 2013).

For mRNA expression analysis in relation to copy number alterations, the clinical attributes, copy number alterations, and mRNA expression data from the METABRIC breast cancer cohort of TCGA (Network 2012; Pereira et al. 2016) were downloaded from cBioPortal (Cerami et al. 2012; Gao et al. 2013) as separate files. The following analysis was performed using custom R scripts (available on request). The datasets on separate files were merged based on patient IDs, and ER status in the clinical attributes was used to sort and extract the data for ER-positive samples. The datasets of the ER-positive breast cancer samples were binned based on the copy number of the *NCOA2* (SRC2) or *NCOA3* (SRC3) genes, and the z-score for the expression of the corresponding mRNA were visualized in boxplots. Due to the low number of samples with *NCOA1* (SRC1) copy number alterations, the mRNA expression analysis for *NCOA1* based on the copy number alteration was omitted.

Kaplan-Meier Analyses

Kaplan-Meier analyses were performed using the Genotype 2 Outcome (G-2-O) online tool (http://www.g-2-o.com/) (Pongor et al. 2015). First, a list of genes more highly expressed in Luminal A or B breast cancer patient samples with amplification of *SRC2* or *SRC3*, compared to samples without amplification of *SRC2* or *SRC3*, was prepared using TCGA datasets, namely *SRC2* or *SRC3* signature gene sets. Next, the expression of the genes in the *SRC2* or *SRC3* signature gene sets were determined using curated publicly available microarray gene expression datasets linked to clinical outcomes (Pongor et al. 2015). The patients were stratified based on the top 50% (high) or the bottom 50% (low) of expression for each of the *SRC2* or *SRC3* signature genes, and the survival rates were calculated for each group of patients for each signature gene. The overall survival rates were plotted as a metagene (average) of all the signature genes to give higher statistical power in the analysis (Pongor et al. 2015).

Genomic Data Sets

The ChIP-seq and GRO-seq datasets generated from MDA-MB-231-ERα Wt and MDA-MB-231-ERα L540Q cells for the current study were deposited in the NCBI's Gene Expression Omnibus (GEO) database under the super series accession number GSE95123. The ChIP-seq and GRO-seq datasets can be found under subseries GSE95121 and GSE95122, respectively.

Author Contributions

S.M. and W.L.K. developed the questions, hypotheses, and model systems for this project based on prior observations from the Kraus lab and preliminary observations from S.M. S.M. and W.L.K. designed the experiments and S.M. executed them with input from W.L.K. S.M. performed all of the wet lab experiments (cell-based and genomic assays), as well as the Kaplan-Meier analyses. A.N. performed the computational analyses. W.L.K. secured funding and provided intellectual support. S.M. and A.N prepared the figures and wrote the methods, and S.M. wrote the remaining text. The figures and text were edited and finalized by W.L.K. with input from S.M. and A.N.

Supplemental References

Acevedo ML, Kraus WL. 2003. Mediator and p300/CBP-steroid receptor coactivator complexes have distinct roles, but function synergistically, during estrogen receptor alpha-dependent transcription with chromatin templates. *Mol Cell Biol* **23**: 335-348.

Acevedo ML, Lee KC, Stender JD, Katzenellenbogen BS, Kraus WL. 2004. Selective recognition of distinct classes of coactivators by a ligand-inducible activation domain. *Mol Cell* **13**: 725-738.

Bowers EM, Yan G, Mukherjee C, Orry A, Wang L, Holbert MA, Crump NT, Hazzalin CA, Liszczak G, Yuan H et al. 2010. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol* **17**: 471-482.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**: 401-404.

Chae M, Danko CG, Kraus WL. 2015. groHMM: a computational tool for identifying unannotated and cell type-specific transcription units from global run-on sequencing data. *BMC Bioinformatics* **16**: 222.

Danko CG, Chae M, Martins A, Kraus WL. 2014. groHMM: GRO-seq Analysis Pipeline. in *Bioconductor*. Bioconductor http://bioconductor.org/packages/release/bioc/html/groHMM.html.

Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306-1311.

Delvecchio M, Gaucher J, Aguilar-Gurrieri C, Ortega E, Panne D. 2013. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat Struct Mol Biol* **20**: 1040-1046.

Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E et al. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**: pl1.

Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL. 2011. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* **145**: 622-634.

Hah N, Murakami S, Nagari A, Danko CG, Kraus WL. 2013. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res* **23**: 1210-1223.

Hay DA, Fedorov O, Martin S, Singleton DC, Tallant C, Wells C, Picaud S, Philpott M, Monteiro OP, Rogers CM et al. 2014. Discovery and optimization of small-molecule ligands for the CBP/p300 bromodomains. *J Am Chem Soc* **136**: 9308-9319.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589.

Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218-223.

Kim MY, Hsiao SJ, Kraus WL. 2001. A role for coactivators and histone acetylation in estrogen receptor alpha-mediated transcription initiation. *EMBO J* **20**: 6084-6094.

Kraus WL, Kadonaga JT. 1998. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev* **12**: 331-342.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589-595.

Luo X, Chae M, Krishnakumar R, Danko CG, Kraus WL. 2014. Dynamic reorganization of the AC16 cardiomyocyte transcriptome in response to TNFalpha signaling revealed by integrated genomic analyses. *BMC Genomics* **15**: 155.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**.

Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, Herschkowitz JI, Burrows AE, Ciccia A, Sun T et al. 2011. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc Natl Acad Sci U S A* **108**: 3665-3670.

Nagari A, Murakami S, Malladi VS, Kraus WL. 2017. Computational Approaches for Mining GRO-Seq Data to Identify and Characterize Active Enhancers. *Methods Mol Biol* **1468**: 121-138.

Network CGA. 2012. Comprehensive molecular portraits of human breast tumours. *Nature* **490**: 61-70.

Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ et al. 2016. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7**: 11479.

Pongor L, Kormos M, Hatzis C, Pusztai L, Szabo A, Gyorffy B. 2015. A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients. *Genome Med* **7**: 104.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140.

Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J, Chen Z, Beroukhim R, Wang H, Lupien M et al. 2009. Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell* **138**: 245-256.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.