Published online 2 August 2019 *Nucleic Acids Research, 2019, Vol. 47, No. 18* **9557–9572**
doi: 10.1093/nar/gkz675

A CRISPR-Cas9 screen identifies essential CTCF anchor sites for estrogen receptor-driven breast cancer cell proliferation

Gozde Korkmaz[1,*], Zohar Manber[2], Rui Lopes[1], Stefan Prekovic[1], Karianne Schuurman[1], Yongsoo Kim[1], Hans Teunissen[3], Koen Flach[3], Elzo de Wit[3], Giorgio G. Galli[4], Wilbert Zwart[1,5], Ran Elkon[2,*] and Reuven Agami[1,6,*]

[1]Division of Oncogenomics, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands, [2]Department of Human Genetics, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel, [3]Division of Gene Regulation, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands, [4]Disease area Oncology, Novartis Institute for Biomedical Research, CH-4002 Basel, Switzerland, [5]Laboratory of Chemical Biology and Institute for Complex Molecular Systems, Department of Biomedical Engineering, Eindhoven University of Technology, PO Box 513, 5600MB, Eindhoven, The Netherlands and [6]Erasmus MC, Rotterdam University, Doctor Molewaterplein 40, 3015 GD Rotterdam, The Netherlands

Received January 29, 2019; Revised July 11, 2019; Editorial Decision July 22, 2019; Accepted July 24, 2019

ABSTRACT

Estrogen receptor α (ERα) is an enhancer activating transcription factor, a key driver of breast cancer and a main target for cancer therapy. ERα-mediated gene regulation requires proper chromatin-conformation to facilitate interactions between ERα-bound enhancers and their target promoters. A major determinant of chromatin structure is the CCCTC-binding factor (CTCF), that dimerizes and together with cohesin stabilizes chromatin loops and forms the boundaries of topologically associated domains. However, whether CTCF-binding elements (CBEs) are essential for ERα-driven cell proliferation is unknown. To address this question in a global manner, we implemented a CRISPR-based functional genetic screen targeting CBEs located in the vicinity of ERα-bound enhancers. We identified four functional CBEs and demonstrated the role of one of them in inducing chromatin conformation changes in favor of activation of PREX1, a key ERα target gene in breast cancer. Indeed, high PREX1 expression is a bona-fide marker of ERα-dependency in cell lines, and is associated with good outcome after anti-hormonal treatment. Altogether, our data show that distinct CTCF-mediated chromatin structures are required for ERα-driven breast cancer cell proliferation.

INTRODUCTION

Approximately 75% of all human breast tumors expresses ERα, a hormone receptor encoded by the *ESR1* gene that functions as a hormone-regulated transcription factor (1,2). ERα is generally activated upon binding of its endogenous ligand estradiol (E2), which results in the recruitment of a large spectrum of coregulators to form an active transcription complex and associate to the chromatin via Estrogen Responsive Elements (EREs) (3). Importantly, ERα rarely binds promoters but regulates gene expression by binding *cis*-regulatory enhancer elements, which interact in 3D genomic space with promoter regions of responsive genes through chromatin looping (4–6).

ERα is considered the main driver in these tumors, and current treatment of ERα-positive breast cancer resolve around endocrine therapies like tamoxifen, which prevents cofactor recruitment and transcription complex formation (7) or fulvestrant that induces receptor degradation (8). Other treatments are based on inhibiting the synthesis of estrogen with aromatase inhibitors, which prevents receptor activation. However, resistance to treatment is common and ∼80% of metastases with acquired resistance to endocrine therapeutics still express ERα (9–11). Multiple drivers of endocrine therapy resistance have been identified, including *ESR1* activating point mutants (12), ERα phosphorylation (13), overexpression of coregulators (14) and epigenetic reprogramming of the ERα cistrome (15,16). Nonetheless, other mechanisms of endocrine therapy resistance in breast cancer are likely to exist.

*To whom correspondence should be addressed. Tel: +31 20 512 2079; Email: r.agami@nki.nl
Correspondence may also be addressed to Gozde Korkmaz. Tel: +31 20 512 2033; Email: g.korkmaz@nki.nl
Correspondence may also be addressed to Ran Elkon. Tel: +31 20 512 2079; Email: ranel@tauex.tau.ac.il

Eukaryotic gene expression is tightly controlled by a set of steps that guarantee precise spatial and temporal gene expression patterns (17). The regulation of initiation of transcription by regulatory DNA elements is one of the most important events in gene expression (18). Regulatory DNA elements can be classified into two main branches according to their mode of activity as *trans*-acting factors and *cis*-regulatory elements (19). *Cis*-regulatory elements include promoters (20), enhancers (21), silencers (22) and insulators (23). They coordinate the timing and strength of gene expression during development and differentiation (24), growth and stress conditions. While promoters are located immediately upstream of their target genes, enhancers can be located far away from their target promoters on the linear chromosome template and rely on chromatin structure (DNA looping) to activate gene expression (25).

One of the most critical determinants of chromatin structure is the zinc-finger DNA binding protein CTCF (CCCTC-binding factor) (26,27). Together with cohesin (28), CTCF functions to insulate enhancers and facilitate correct enhancer/target-promoter interactions by creating chromatin territories dubbed as topologically active domains (TADs) (29,30). Two convergent linear CTCF sites (31) are held together by a cohesin ring and form 'insulated neighborhoods' (27) (chromatin loop structures) which come together to form megabase-scale TADs (32). Therefore, it is likely that genetic or epigenetic perturbations of CTCF binding elements (CBEs) may affect chromatin structure leading to local rewiring of promoter-enhancer contacts which potentially can change gene expression patterns (27,33,34) and as a result cause a phenotypic changes in cell behavior (35,36). Indeed, indications for a link between CTCF and ERα-mediated gene expression and phenotypic changes were recently made (15,37). By an integrative genomics approach, Chan and colleagues have partitioned the genome to CTCF blocks (genomic regions between two CBEs), and demonstrated that genes located within ERα-containing CTCF blocks are generally upregulated in response to estrogen while genes in blocks devoid of ERα not. In fact, the presence of an ERα binding site (EBS) within a CTCF block, and the distance to the nearest ERα, were shown to be the best predictive score for estrogen-responsive genes (37). Moreover, CTCF binding appeared not to only mark boundaries for accessibility and insulated transcriptional blocks, but also seemed to be required to shape ERα-mediated gene expression in breast cancer cells (15). However, the functional significance and the causal role of CTCF binding and associated gene expression changes to cancer progression are yet to be explored. Since CBEs and EBSs overlap in thousands of cases (15) and enhancer/target-gene interactions depend on proper chromatin interactions, we hypothesize here that CBEs can affect ERα function and induce tumor cell proliferation and resistance to anti-hormonal treatments.

The CRISPR-Cas9 system has been recently harnessed to delineate regulatory gene expression mechanisms mediated through promoter (38,39) and enhancer regions (40). Though applied to specific CTCF loci of interest before (27,34), a genome-wide CRISPR-Cas9 screen systematically targeting CTCF sites was not yet implemented. In this study, we designed a customized CRISPR-Cas9 library to target individual CBEs that reside in close proximity to EBSs and performed a drop-out genetic screen to identify CBEs essential for ERα-driven cell proliferation. Altogether, we validated four key CBEs, of which we further functionally characterized the one residing in the vicinity of *PREX1* (Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 1) promoter; a key driver of breast cancer cell proliferation.

MATERIALS AND METHODS

Cell lines and chemical reagents

HEK293-T, MDA-MB-231 and MCF-7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) medium (Gibco), supplemented with 10% Fetal Calf Serum (Hyclone) and 1% penicillin/streptomycin (Gibco). For the estrogen-stimulation experiments, MCF-7 cells were washed three times with phosphate-buffered saline and cultured in phenol red–free DMEM medium (Gibco) supplemented with 5% charcoal stripped serum (Gibco) for 72 h prior to E2 treatment (10−8 M). E2 (17β-estradiol) was purchased from Sigma. All cell lines were obtained from ATCC.

Custom single guide RNAs (sgRNAs) designed for this study:

Name	Oligo1	Oligo2
CTCF_ERa_MCF7_sgRNA680	CACCGTGGGCATAAGGAGTAACAGA	AAACTCTGTTACTCCTTATGCCCAC
CTCF_ERa_MCF7_sgRNA1117	CACCGCAATGCCAGGCTCTGCCATC	AAACGATGGCAGAGCCTGGCATTGC
CTCF_ERa_MCF7_sgRNA1118	CACCGTTCGCGCTGCTGCTGCCAGA	AAACTCTGGCAGCAGCGCAGCGCGAAC
CTCF_ERa_MCF7_sgRNA1659	CACCGAGACAGCGGTACCTGTCAGG	AAACCCTGACAGGTACCGCTGTCTC
CTCF_ERa_MCF7_sgRNA1661	CACCGAAAGACAGCGGTACCTGTCA	AAACTGACAGGTACCGCTGTCTTTC
CTCF_ERa_MCF7_sgRNA810	CACCGGACGCTTGGCTGCCACCTGG	AAACCCAGGTGGCAGCCAAGCGTCC
CTCF_ERa_MCF7_sgRNA811	CACCGGAGGACGCTTGGCTGCCACC	AAACGGTGGCAGCCAAGCGTCCTCC
CBE2 #1	CACCGGCTCGAGCATTGGTGCTCTC	AAACGAGAGCACCAATGCTCGAGCC
CBE2 #2	CACCGCGAGCATTGGTGCTCTCTGG	AAACCCAGAGAGCACCAATGCTCGC
EBE24365 #1	CACCGAGGAGCAGGCTGACCTGACC	AAACGGTCAGGTCAGCCTGCTCCTC
EBE24365 #2	CACCGGTGACTCTGTTCTCCTGGTC	AAACGACCAGGAGAACAGAGTCACC
intron #1	CACCGGATGGTGTCTGCGAACAGCG	AAACCGCTGTTCGCAGACACCATCC
intron #2	CACCGATTTTAGGGTTGTCCCCACG	AAACCGTGGGGACAACCCTAAAATC
PREX1 exon #1	CACCGGCGCGGCCGCGCTAGAATGG	AAACCCATTCTAGCGCGGCCGCGCC
PREX1 exon #2	CACCGTGGGCGCCTCCATTCTAGCG	AAACCGCTAGAATGGAGGCGCCCAC
PREX1 exon #3	CACCGGCAGAACGTGGCCGACTCAG	AAACCTGAGTCGGCCACGTTCTGCC
PREX1 exon #4	CACCGGCATCGCATCCGGCAGAACG	AAACCGTTCTGCCGGATGCGATGCC

Pooled library cloning

Standard de-salted DNA oligonucleotides were synthesized and purchased from IDT (Integrated DNA Technologies) to construct sgRNA library for CTCF sites in the vicinity of ERα (1709 sgRNAs+82 control sgRNAs). Complementary single-stranded oligos were phosphorylated and annealed by combining 100 μM oligos, 1× T4 PNK Buffer, 1 mM adenosine triphosphate (ATP), 5U T4 PNK and incubating the reaction at 37°C/30 min, 95°C/5 min followed by ramp down to 25°C at 5°C/min. Annealed oligos were diluted at 1:1000 in sterile water, and ligated to plasmid vector lentiCRISPR v2 (gift from Feng Zhang (Addgene plasmid #52961)) using the following parameters: 50 ng BsmBI (Fermentas) digested plasmid, 1 μl diluted oligo duplex, 1× Ligation Buffer (Roche), 5U T4 DNA Ligase (Roche) incubated at RT/30 min. We did five independent ligation reactions per pool and used them to transform highly competent *Escherichia coli* cells (EletroSHOX - Bioline, BIO-85038) according to manufacturer's protocol. In order to assess the complexity of our libraries, we plated 1 μl of cell transformation mixture on LB agar plates containing Ampicillin, incubated them overnight at 37°C, and counted individual bacterial colonies after 16 h. At this point, we estimated that each individual sgRNA is covered >100×, ensuring that our libraries have high-complexity and are suitable for pooled screening. Transformation mixtures were combined, grew in liquid LB until OD600 = 0.8 was reached, and plasmid DNA was harvested using Genopure Plasmid Maxi kit (Roche).

Lentivirus production and purification

To produce lentivirus, 4×10^6 HEK293T cells per pool were seeded in five 100 mm dishes one day prior to transfection. For each dish, we diluted 10 μg of CRISPR-CTCF$^{ER\alpha}$ plasmid library, 3.5 μg of pVSV-G, 5 μg of pMDL RRE and 2.5 μg of pRSV-REV in 450 μl of 0.1× TE/H_2O, added 50 μl of $CaCl_2$ and incubated 5 min at RT. Plasmid DNA was precipitated by adding 500 μl 2× HEPES-buffered saline to the solution while vortexing at full-speed. The precipitate was added immediately to the plate and the cells were incubated for 14 h at 37°C, after which the medium was refreshed. Lentivirus-containing supernatants were collected 60 h post-transfection, filtered through a 0.45 μm membrane (Milipore Steriflip HV/PVDF) and stored at −80°C. All cell types and lentivirus batches tested were titrated in order to achieve a MOI of 0.4–0.5. Cell lines were infected with lentivirus supernatants supplemented with 8 μg/ml polybrene (Sigma). At 24 h post-infection, medium was replaced and cells were selected with 2 μg/ml puromycin (Gibco). Antibiotic selection was stopped as soon as no surviving cells remained in the no-transduction control plate. All experiments were performed within the first 15 days following the antibiotic selection.

CRISPR-Cas9 dropout screen in breast cancer cells

MCF-7 and MDA-MB-231 cells were infected with three independent pools of CRISPR-CTCF$^{ER\alpha}$. We infected ∼45 000 cells per vector, to ensure that every sgRNA was present in the cell population at the start of the experiment. Following antibiotic selection, cells were allowed to proliferate for 48 h to clear potentially toxic sgRNAs from the population. At this time point, we harvested half of the cells infected with CRISPR-CTCF$^{ER\alpha}$ pools ($T = 0$). The remaining cells were placed in culture, allowed to proliferate for 20 days, and then harvested ($T = 20$). Cell pellets were stored at −80°C and processed later on for further analysis. Genomic DNA was isolated from the harvested library-transduced cells, and amplified for the integrated vectors by polymerase chain reaction (PCR), and sequenced by Next-generation sequencing (NGS) to quantify the abundance of integrated sgRNAs present in each population. Read counts in each sample were normalized to 1M reads (RPM) and enrichment (depletion) scores were calculated for each sgRNA vector, in each cell line (MCF-7 and MDA-MB-231) by comparing its normalized frequency at $T = 20$ and $T = 0$ pools. (Only sgRNAs covered by at least 50 reads were considered. To avoid inflation of FC for sgRNAs covered by low number of reads, RPMs below 5 were set to 5). These enrichment/depletion scores were then standardized to *Z*-scores (raw counts, RPM, enrichment scores and *Z*-scores are provided in Supplementary Table S1). These standardized enrichment scores showed high correlation between replicates from the same cell line (Supplementary Figure S1A) Difference between the mean standardized enrichment scores in MCF-7 and MDA-MB-231 cells (which serve as control in the screen as they are not dependent on ERα) were calculated for each sgRNA. For validation, we selected sgRNA vectors whose repressive effect on proliferation was at least 30% stronger in MCF-7 cells compared to MDA-MB-231 cells.

Genomic DNA sequencing to identify sgRNAs

Frozen cell pellets were thawed and genomic DNA (gDNA) was isolated with DNeasy Blood and Tissue kit (Qiagen). Identification of sgRNAs was done by PCR in two steps. For the first PCR, the amount of input gDNA was calculated to achieve ∼500 coverage (assuming that 10^6 cells contain 6.6 μg gDNA), which resulted in 5 μg for CRISPR-CTCF$^{ER\alpha}$. For each sample, we performed two separate reactions (max. 1 μg gDNA per reaction) using Phusion DNA polymerase (Thermo Scientific) and combined the resulting amplicons. In the first PCR, we used the following primer sequences to amplify sgRNAs:

PCR1_F1

ACACTCTTTCCCTACACGACGCTCTTCCGATC TXXXXXXGGCTTTATATATCTTGTGGAAAGGA CG (XXXXXX represents a 6 bp barcode)

PCR1_R1

GTGACTGGAGTTCAGACGTGTGCTCTTCCGAT CTACTGACGGGCACCGGAGCCAATTCC

A second PCR was performed to attach Illumina adaptors and index samples. The second PCR was done in 50 μl reaction volume, including 7 μl of the product from the first PCR and using the following primers:

PCR2_P5

AATGATACGGCGACCACCGAGATCTACACTCT TTCCCTACACGACGCTCTTCCGATCT

PCR2_P7

CAAGCAGAAGACGGCATACGAGATXXXXXX GTGACTGGAGTTCAGACGTGTGCTCTTCCGAT CT (XXXXXX represents a 6 bp index)

Amplification was carried out with 15 cycles for both first and second PCR. After the second PCR, resulting amplicons were purified using Agencourt AMPure XP beads (Beckman Coulter), quantified in a Bioanalyzer 2100 (Agilent), mixed and sequenced in a HiSeq 2500 (Illumina).

Genomic DNA sequencing to identify CRISPR-induced mutations

Cell pellets were collected and gDNA was isolated with DNeasy Blood and Tissue kit (Qiagen). Amplification of target regions for sequencing was done by PCR in two steps. For each sample, we used 500 ng of gDNA as input for the first PCR (done in duplicate). Resulting amplicons were combined and we used 10 μl as input for the second PCR. Amplification was carried out with 20 cycles for both first and second PCR. After the second PCR, amplicons were purified using Agencourt AMPure XP beads (Beckman Coulter), quantified in a Bioanalyzer 2100 (Agilent), mixed and sequenced in a HiSeq 2500 (Illumina).

List of PCR primers:

Name	Primers
sgRNA680_F	CTTGACTGGCACCTGGAGAG
sgRNA680_R	ATTCCATGGTGTCTGCTGGG
sgRNA680_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT CGTGAT AGAGACCATGTTGTGCCAGG
sgRNA680_P7_SeqR	GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCTGTTTTGTGTGGG GGCCTACT
sgRNA1118_F	GTGAGAGGGTGAGAGTGTGC
sgRNA1118_R	CTGGGGAGTTCTGGCTCTTG
sgRNA1118_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT ACATCG CATGCACACACACACACACA
sgRNA1118_P7_SeqR	GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCTAGCTCAGTCATG TCAGCCAC
sgRNA1659_F	CACCAGGCTTGAGAGAAGCA
sgRNA1659_R	CATAGCCCTGTACCTCTGCG
sgRNA1659_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT GCCTAA TAGGTTATCTGCTGGGCCCA
sgRNA1659_P7_SeqR	GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCTGGGAATCGAGGC TCCCTAAG
sgRNA1661_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT TGGTCA TAGGTTATCTGCTGGGCCCA
sgRNA810_F	GTTGGAACTGGGAGGAGCTC
sgRNA810_R	GACTGGCTTCAGCATTGTGC
sgRNA810_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT CACTGT CAACCCGTTTTGGAAGCCAG
sgRNA810_P7_SeqR	GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCTGAGTGTGGCG AGCTGGATC
sgRNA811_P5_SeqF	ACACTCTTTCCCTACACGACGCTC TTCCGATCT ATTGGC CAACCCGTTTTGGAAGCCAG

Competitive proliferation assay

MCF-7 and MDA-MB-231 cells were infected with indicated sgRNAs to validate the results of the CRISPR-CTCF$^{ER\alpha}$ screen. Separately, we generated polyclonal MCF-7 and MDA-MB-231 cells stably expressing GFP using pLKO.1-GFP vector. GFP expressing cells were mixed in a 1:3 ratio with cells containing individual sgRNAs. The percentage of GFP expressing cells was assessed by flow cytometry at the beginning of the experiment ($T = 0$) and every 72 h onward ($T = 3$ days and $T = 6$ days). For every condition, 10 000 events were recorded and the data was analyzed using FlowJo software.

RNA isolation, reverse-transcription and quantitative real-time PCR (qPCR)

After the antibiotic selection was completed, cells were kept in cell culture at least 2–3 days in order to eliminate the adverse effect of antibiotic on cells. Total RNA was extracted using TRIsure (Bioline) reagent and following the manufacturers protocol. cDNA was produced with either SuperScript III (Invitrogen) or High-Capacity cDNA Reverse Transcription Kit (Applied Biosystem) using 2–4 μg of total RNA per reaction. qPCR reaction was performed with SYBR green I Master mix in a LightCycler 480 (Roche). TATA-binding protein (TBP) was used as an internal control.

List of qPCR primers:

	Forward	Reverse
PREX 1	TGGGACTCAT CCCTGAACTC	CTGTGCAACAAT GGCTTCAT
TBP	GGAGAGTTCT GGGATTGTAC	CTTATCCTCATGAT TACCGCAG

Western blot analysis

Whole-cell lysates were prepared as previously described (41). Membranes were immunoblotted with the following antibodies: PREX1 (ab183643, Abcam; 1:500), and beta-Actin (A2228, Sigma; 1:5000). Protein bands were visualized using corresponding secondary antibodies (Dako) and ECL reagent (GE Healthcare).

GRO-seq

GRO-seq was performed as described before with minor modifications. Briefly, 5 × 106 nuclei were isolated and incubated 5 min at 30°C with equal volume of reaction buffer (10 mM Tris–Cl pH 8.0, 5 mM MgCl2, 1 mM DTT, 300 mM KCL, 20 units of SUPERase In, 1% sarkosyl, 500 μM ATP, GTP and Br-UTP, 0.2 μM CTP+32P CTP) for the nuclear run-on. The reaction was stopped and total RNA was extracted with Trizol LS (Invitrogen) according to the manufacturer's instructions. RNA was fragmented using fragmentation reagents (Ambion) and the reaction was purified through p-30 RNAse-free spin column (BioRad). BrU-labeled RNA was immunoprecipitated with anti-BrU agarose beads (Santa Cruz), washed one time in binding

buffer, one time in low salt buffer (0.2× SSPE, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.05% Tween-20), one-time high salt buffer (0.25× SSPE, 1 mM EDTA, 0.05% Tween-20, 137.5 mM NaCl) and two times in TET buffer (TE with 0.05% Tween-20). RNA was eluted with elution buffer (20 mM DTT, 300 mM NaCl, 5 mM Tris–Cl pH 7.5, 1 mM EDTA and 0.1% sodium dodecyl sulphate) and isolated with Trizol LS. After the binding step, BrU-labbeled RNA was treated with tobacco acid pyrophosphatase (TAP, Epicenter) to remove 5′-methyl guanosine cap, followed by T4 polynucleotide kinase (PNK; NEB) to remove 3′-phosphate group. BrU-containing RNA was treated with T4 PNK again at high pH in the presence of ATP to add 5′-phosphate group. The reaction was stopped and RNA was extracted with Trizol LS. Sequencing libraries were prepared using TruSeq Small RNA kit (Illumina) following manufacturer's instructions. Briefly, end-repaired RNA was ligated to RNA 3′ and 5′ adapters, followed by RT-PCR amplification. cDNA was purified using Agencourt AMPure XP (Beckman Coulter) and amplified by PCR for 12 cycles. Finally, amplicons were cleaned and size-selected using Agencourt AMPure XP (Beckman Coulter), quantified in a Bioanalyzer 2100 (Agilent), and sequenced in a HiSeq 2500 (Illumina). Sequenced reads were aligned to the human genome (hg19) using bowtie2. Statistics on the number of sequenced and mapped reads is given in Supplementary Table S3.

RNA-seq

After the antibiotic selection was completed, cells were kept in cell culture until they reach 70–80% confluency. Total RNA was extracted using TRIsure (Bioline) reagent and following the manufacturers protocol. RNA-seq samples were processed with TruSeq RNA library prep kit v2 (Illumina) and sequenced in a HiSeq 2500 (Illumina). Sequenced reads were aligned to the human genome (hg19) using TopHat2 (42) and gene expression counts were calculated using HTseq (43) based on Ensembl's human gene annotations (v69) (44). Expression levels were normalized using quantile normalization. Statistics on the number of sequenced and mapped reads is given in Supplementary Table S3.

Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitations were performed as previously described (16,45). Cells were fixed with formaldehyde (1%) for 10 min and subsequently quenched with glycine. Following that, samples were lysed as described (46) and sonicated for at least 11 cycles of 30 s on, 30 s off using a Diagenode Bioruptor Pico. We used antibodies against ERα (MS-315-P, Thermo Scientific), FOXA1 (ab5089, Abcam) and CTCF (07–729, Merck Millipore). For ERα/FOXA1 ChIPs, 5 μg of antibody was conjugated with 50 μl Protein G magnetic beads. On the other hand, for CTCF ChIP 7.5 μl of the antibody was conjugated with 50 μl Protein A magnetic beads. Immunoprecipitated DNA was processed for library preparation by using TruSeq ChIP Library Preparation Kit. Samples were sequenced using an Illumina Hiseq2500 genome analyzer (65 bp reads, single end), and aligned to the Human Reference Genome (hg19, February 2009). Reads were filtered based on MAPQ quality (quality ≥ 20) and duplicate reads were marked. Peak calling over input control was performed using DFilter (47) and MACS 1.4 (48) peak callers. MACS was run with the following parameters: bw 300 -m 10,30 -p 1.00e-7 –nomodel. DFilter was run with bs = 50, ks = 30, refine, nonzero. The peaks shared by both peak callers were used for analysis. For pair-wise comparison of ChIP-seq, the binding signals in the union of the binding sites from the data pair were quantified using count per million reads using pybedtools (49,50). The binding profiles are compared and visualized using seaborn v0.7.0 (51) where the sites overlap with the target of sgRNAs are highlighted in the scatter plot.

4C

4C-seq was performed as described previously (52). Briefly, after obtaining a single cell suspension, crosslinking of the nuclei, chromatin was digested with MboI and subsequently re-ligated. Proteinase K was used to de-crosslink the DNA o/n at 65°C. DNA was isolated by phenol/chloroform extraction and subsequently digested with a Csp6I. To create circular DNA molecules, digested DNA was ligated under diluted conditions (7 ml). DNA was precipitated with 1/10 volume 3M sodium acetate and 1 volume isopropanol. DNA was quantified using Qbit (Thermo Fisher). 4C PCR was done using Expand long template polymerase (Sigma-Aldrich) in five separate PCR reactions using 100 ng per PCR.

After Illumina sequencing, reads were demultiplexed based on the primer sequence of the viewpoint (first 20 nt of the sequence read). Trimmed reads we mapped to the human genome (hg19) with BWA-MEM (53). Only reads whose start overlapped with an MboI restriction site were retained for further analysis. Read count files were normalized using peakC (54) before further analysis. CTCF motif positions are based on a previously done ChIPseq (15).

PREX1 expression data and patient survival

The performance of the PREX1 classifier for clinical outcome was assessed using the METABRIC patient series (55). Survival curves for PREX1 expression (lowest versus highest quartile) were constructed using the Kaplan–Meier method and compared using log-rank test. Adjusted Cox proportional hazard regression analyses were performed with the following covariates: PREX1 expression, age, tumor grade (grade 1 versus grade 2 versus grade 3), tumor size, tumor stage (1 versus 2 versus 3 versus 4 versus 5) and lymph node status (negative versus positive). When comparing the performance of the PREX1 classifier in the hormonal treatment setting, the group of patients receiving hormonal treatment was down sampled to match the patient numbers of the other group. Pearson's correlation coefficient was used to assess the correlation between PREX1 and ESR1 expression. A P-value < 0.05 was considered as a significant result.

Analysis of *PREX1* expression and ESR1 KD sensitivity in breast cancer cell

We compared the expression levels of *PREX1* (as measured by RNA-seq across the breast cancer cell lines of the Cancer Cell Line Encyclopedia) with the Ataris and Demeter scores for ESR1 knockdown previously reported by the projects DRIVE (56) and Achilles (57), respectively. The sensitivity scores of ESR1 knockdown correspond to an aggregate of 17 and 10 shRNAs for DRIVE and Achilles, respectively.

Analysis of *PREX1* and *ESR1* expression in tumor samples

The expression levels of PREX1 and ESR1 (log2 RNA-seq) were analyzed in 1095 breast cancer samples. The data was obtained from TCGA.

RESULTS

A functional genetic screen identifies CBEs essential for ERα-mediated mitogenic signaling

We hypothesized that correct chromatin structure is required to facilitate ERα-driven breast cell proliferation. As CTCF is one of the major determinants of chromatin structure, we setup a functional genetic screen in breast cancer cell lines to uncover CTCF-mediated DNA loops whose knockout attenuates ERα-mediated cell proliferation. We first constructed a CRISPR-Cas9 single guide RNA (sgRNA) library to target CBEs that are both engaged in DNA loop formation (anchor CTCF sites) and located in the vicinity of ERα-bound enhancers (Figure 1A). Initially, anchor CBEs were bioinformatically mapped using ENCODE's Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) data in MCF-7 cells (4). Then, we searched for evidence of direct CTCF-binding by crossing this result with ENCODE CTCF ChIP-seq data in MCF-7 and the presence of a CTCF-motif in each region (CTCF motif was derived by MEME and then used by FIMO (using cutoff of $P < 10–5$) to scan the peaks for its hits). This set of CTCF anchor sites was intersected with a map of ERα-bound active enhancers (identified by the overlap of ERα binding sites (5) and ENCODE p300 peaks in MCF-7) to select those that reside within a 100 Kb distance from any ERα-bound enhancer. This pipeline resulted in 668 CBEs to which we generated 1,709 targeting sgRNA vectors. We also added 82 non-targeting sgRNA vectors (negative controls) and an sgRNA vector targeting ERα-bound enhancer ($ERα^{1830}$) that we previously showed its requirement for mediating ERα-driven cell proliferation (positive control (40,58)). A lentiviral library was generated and named as $CRISPR\text{-}CTCF^{ERα}$.

Next, we transduced MCF-7 (ERα-positive) and MDA-MB-231 (ERα-negative) breast cancer cell lines with the $CRISPR\text{-}CTCF^{ERα}$ library (three biological replicates, 0.3 MOI, 500× coverage). Depletion scores were calculated for each sgRNA vector in each cell line (MCF-7 and MDA-MB-231) by comparing its normalized abundance in the $T = 20$ and $T = 0$ pools. These scores showed high correlation between the biological replicates of the same cell line (Figure 1B, Supplementary Figure S1A and Table S1 and 4). We aimed to find sgRNAs whose abundance was reduced specially in MCF-7 without observing any significant change in MDA-MB-231 cells. As expected, none of the negative controls showed any significant reduction in abundance (Supplementary Figure S1B), while the abundance of sgRNA-$ERα^{1830}$ was highly reduced only in MCF-7 cells (Figure 1C). Importantly, this experiment identified four CBEs (three of which were targeted by two independent sgRNAs) with a marked negative effect on the proliferation of MCF-7 but no significant effect on MDA-MB-231 cells (Figure 1C and Supplementary Figure S1B). Single transductions of these hits confirmed the functionality of the four CBEs by growth competition assays (Figure 1D; Supplementary Figure S2A and B). Following 9 days of culturing, each of the identified sgRNAs attenuated cell proliferation to a similar extent as observed with the positive control sgRNA-$ERα^{1830}$ (40,58). Thus, our functional genetic screen approach has yielded four CBEs that affect the proliferation of MCF-7 while having no effect on MDA-MB-231 cells.

Interestingly, examination of a publicly available dataset that profiled transcriptional responses to estradiol treatment in MCF-7 cells (using the GRO-seq technique) (59), showed that for all the four validated CBEs, the nearest gene was markedly responsive (induction above 2-fold) to the treatment ($P = 0.0003$; binomial test using the overall proportion of genes that were induced more than 2-fold in the dataset): *ESRP2, AKAP1, BICD2* and *PREX1* located, respectively, 7k, 35k, 48k and 36k bps from the CBEs targeted by sgRNA680, sgRNA810, sgRNA1659 and sgRNA1118 (Supplementary Table S2).

CBEs are required for local gene expression

We studied in a greater detail the impact of a subset of the validated sgRNAs (one from each of the four CBEs detected by the screen) on genome integrity and gene expression. First, we interrogated the level of DNA editing at the target sites by next-generation sequencing and found the expected >50% mutant allele frequencies, the majority of DNA alterations occurring within 15 nt from the planned cleavage sites (Figure 1E, left panel and Supplementary Figure S4A and B). Second, using mRNA sequencing, we assessed changes in gene expression between our hits and a non-targeting sgRNA. Notably, in all cases, this analysis detected a mild (∼10–50% reduction) but consistent negative impact on the expression of a handful of genes in the vicinity of the targeted regions (Figure 1E, right panel and Supplementary Figure S2C). Moreover, we used two publicly available GRO-seq datasets that profiled transcriptional responses to E2 in MCF-7 cells to examine the response of genes in the vicinity of the four validated CBEs (±400 kbp; Figure 1E) to estrogen (5,59). In both cases, nearby genes showed significant induction upon E2 treatment (Supplementary Figure S2D), suggesting that the selected CBEs are required for optimal local gene expression (27).

A key role for CBEs in controlling ERα-mediated PREX1 transcript activation

Next, we focused on the CBE region targeted by sgRNA1117 and sgRNA1118 (named as CBE1). This

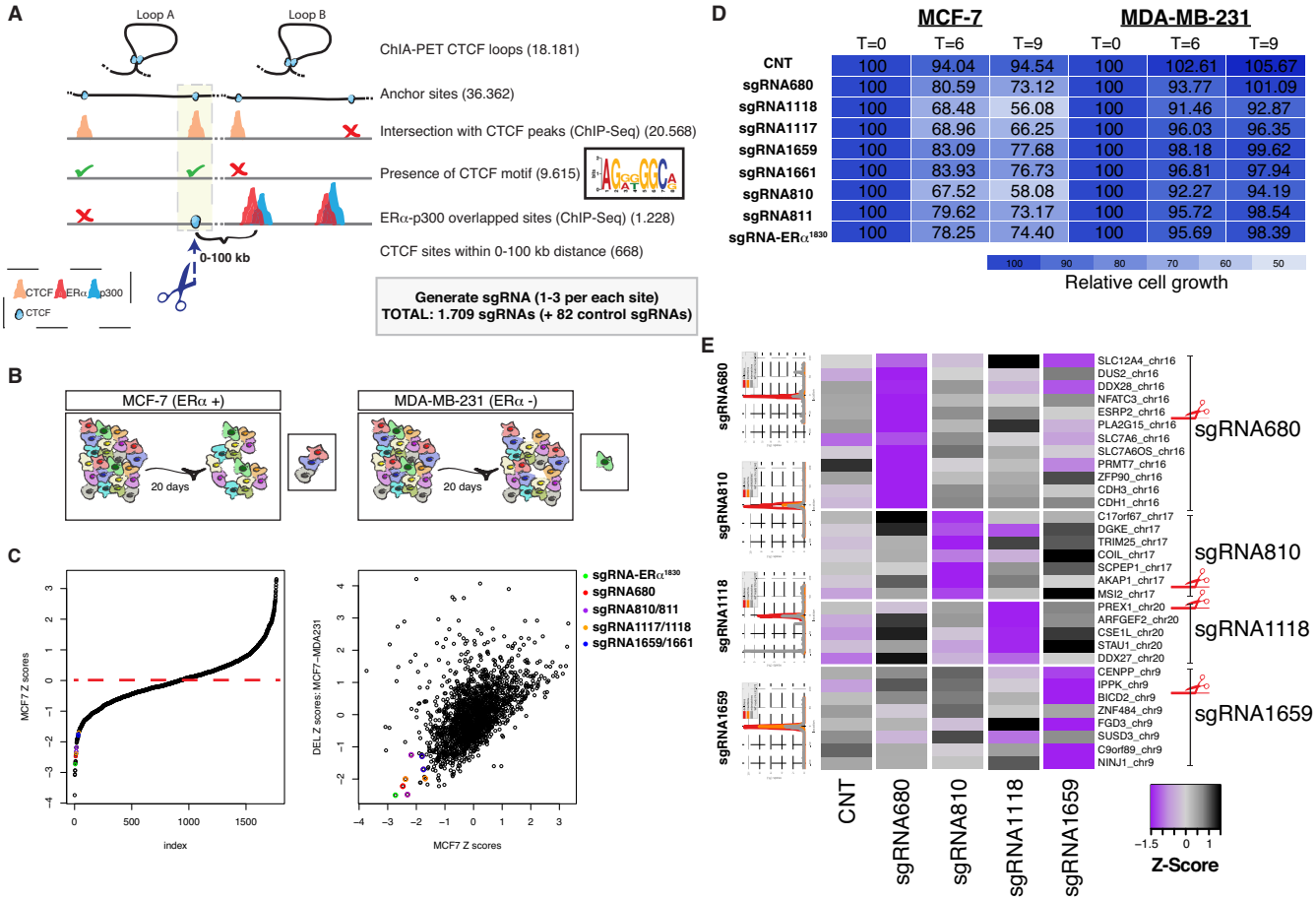

Figure 1. A focused CRISPR-Cas9 genetic screen identifies CBEs required for mitogenic function of ERα. (**A**) The selection criteria and a summary of the pipeline used to generate a customized CRISPR-Cas9-based library (named as CRISPR-CTCF$^{ER\alpha}$). The number in brackets indicates remaining CTCF binding sites. (**B**) Screening strategy based on negative selection of sgRNAs. (**C**) Distribution of sgRNAs' standardized enrichment scores (*Z*-scores) in only MCF-7 (left) and in MCF-7 relative to MDA-MB-231 cells (right). A green dot represents the positive control sgRNA targeting enhancer1830 (labeled as enh1830) (40). (**D**) Validation of candidate hits by a competitive proliferation assay in MCF-7 and MDA-MB-231 cells. CNT indicates a non-targeting sgRNA. sgRNA targeting enhancer1830 was used as a positive control (40). GFP positive cells were determined by flow cytometry. For complete plots see Figure S1A. (**E**) RNA-seq showed a local downregulation of the genes residing within 400 kbps away from each target site. CNT indicates a nontargeting sgRNA. Heatmap represents *Z*-score of expression fold-change (between the test and control sgRNAs). Mutation profile of each CBE target site was determined by DNA-seq experiment, and is presented at the left-hand side. In all samples the frequency of CRIPSR-Cas9 meditated mutations at the target CBE was more than 50%.

CBE has a strong effect on *PREX1* expression, a target gene that shows a robust transcriptional activation following estradiol treatment in MCF-7 cells (∼2-fold, Figure 2E). Expression analysis of all genes residing within a 2 MB window around the target site (CBE1) indicated a significant effect on the mRNA level of *PREX1* (Figures 1E and 2A). PREX1 is a Rac guanine nucleotide exchange factor, whose activity promotes breast cancer cell proliferation through the activation of the extracellular signal-regulated kinases (ERK1/2) (60). The PREX1 locus is frequently amplified in various cancer types (61,62), especially in breast cancer tumors (up to 30%) with a preference for ERα-positive cases (63), suggesting an oncogenic potential within the ERα pathway. Using qRT-PCR, we validated the changes in *PREX1* gene expression induced by both CBE1 sgRNAs (Figure 2B). Notably, CRISPR-mediated targeting of *PREX1* exons (Figure 2D and Supplementary Figure S2E), but not introns (Figure 2C and Supplementary Figure S4D), resembled the effect of targeting CBE1 in MCF-7 cells (Figure 2D). Intriguingly, in the same sgRNA1117 or sgRNA1118-transduced MCF-7 cell populations, we observed marked reduction in the protein levels of PREX1 isoforms (Supplementary Figure S3). It should be noted that due to the amplification of *PREX1* locus in MCF-7 cells, a subtle effect of CRISPR-Cas9 targeting was observed when introns were targeted (64). In contrast to the effect in MCF-7 cells, targeting *PREX1* exons did not affect the proliferation capacity of MDA-MB-231 cells (Figure 2D)—though *PREX1* locus is amplified in MDA-MB-231 too.

To examine the estradiol (E2) responsiveness, we cultured MCF-7 cells (wildtype and CBE1 knockout) in hormone-deprived medium and subsequently treated with E2 or DMSO control for 30 min, 60 min and 6 h and measured *PREX1* mRNA expression. Indeed, *PREX1* expression was significantly increased after E2 treatment (Figure 2E and Supplementary Figure S2F), indicating that *PREX1* is a primary target gene of ERα. Furthermore, knocking out

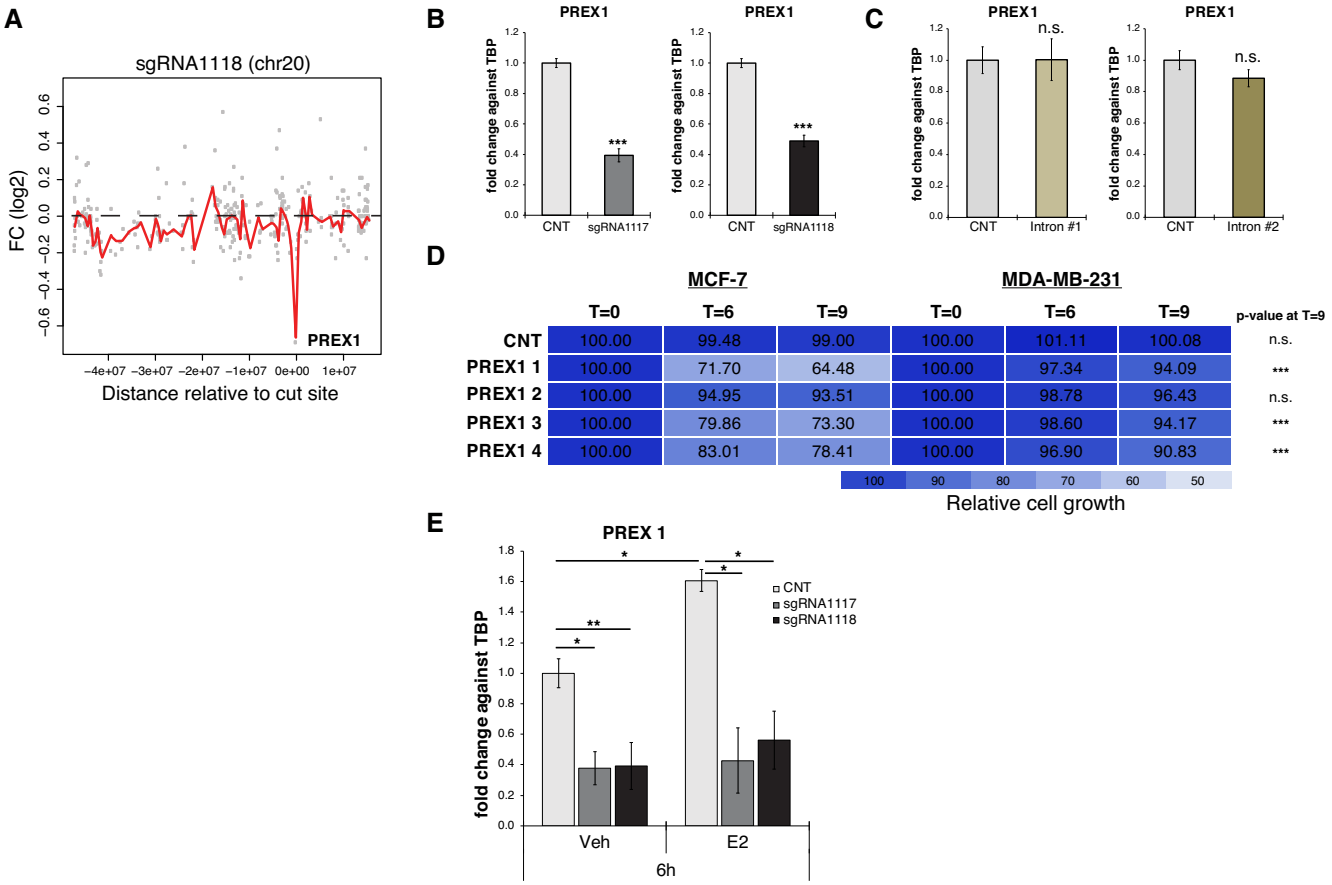

Figure 2. A key role for CBE1 in ERα-mediated stimulation of *PREX1* activity. (**A**) Gene expression fold-change (log2) between sgRNA1118 (targeting CBE1) and non-targeting control depicted for genes in chr20 (the location of sg1118 CRISPR-Cas9 cut site is set to 0). The red line is a lowess (locally weighted scatterplot smoothing) regression curve fitted to the data. (**B** and **C**) Relative expression levels of *PREX1* was measured by qRT-PCR in sgRNA1117 and sgRNA1118 (targeting CBE1) and intron targeted MCF-7 cells, compared with control non-targeting sgRNA (CNT). Gene expression levels were normalized to TBP. $n = 3$ $***P < 0.005$, $*P < 0.05$, two-tailed Student's *t*-test. (**D**) Competitive proliferation assays in MCF-7 and MDA-MB-231 after transduction with individual sgRNAs targeting *PREX1* gene. Values on day 6 ($T = 6$) and day 9 ($T = 9$) were normalized to day 0 ($T = 0$). $***P < 0.005$, $**P < 0.01$, $*P < 0.05$, two-tailed Student's *t*-test. n.s., not significant.

CBE1 significantly compromised the induction of *PREX1* expression by E2 (Figure 2E and Supplementary Figure S2F). Moreover, examining genes in the vicinity of *PREX1* gene indicated that regardless of their induction by E2, CBE1 compromised their expression, substantiating its impact on local gene expression (Supplementary Figure S2G). Altogether, our experiments revealed a key role for CBE1 in ERα-mediated stimulation of *PREX1* levels.

An interplay between several CBEs regulate *PREX1* expression

To further assess the binding of CTCF to CBE1 upon sgRNA introduction, we performed CTCF chromatin-immunoprecipitation and sequencing (ChIP-seq) experiments. This analysis revealed a moderate binding of CTCF to CBE1, which was significantly reduced in cells transduced with either sgRNA1117 or 1118, compared to control (Figure 3A, left and middle panels). This result validates CBE1 as a binding site of CTCF, and more importantly demonstrates a strong and specific effect of CBE1-targeting vectors. Then, we examined the chromatin architecture surrounding CBE1, and designed circularized chromosome conformation capture (4C-seq) experiments with CBE1 as a viewpoint (52). This analysis uncovered an interaction of CBE1 with an upstream CBE (termed here as CBE2) and the promoter of *PREX1* (Figure 3B and Supplementary Figure S5A). Interaction between CBE1 and CBE2 is in the expected convergent orientation (31). In CBE1-targeted cells, we detected a mild reduction in the interaction frequency with CBE2. To confirm the observed interaction patterns, we included two additional 4C viewpoints; CBE2 (Figure 3B and Supplementary Figure S5B) and the *PREX1* promoter (Figure 3B and Supplementary Figure S5D). In the control condition, CBE2 interacts with another downstream CTCF site (termed here CBE3) (Figure 3B and Supplementary Figure S5C). However, CBE1 knockout induced a strong interaction frequency between CBE2 and the *PREX1* promoter, suggesting that CBE1 is important to keep the promoter region in a more accessible chromatin configuration for transcriptional activation by impeding the interaction between CBE2 and *PREX1* promoter and by favoring interaction between CBE2 and CBE3. Therefore, we targeted CBE2 using two different sgRNAs and exam-

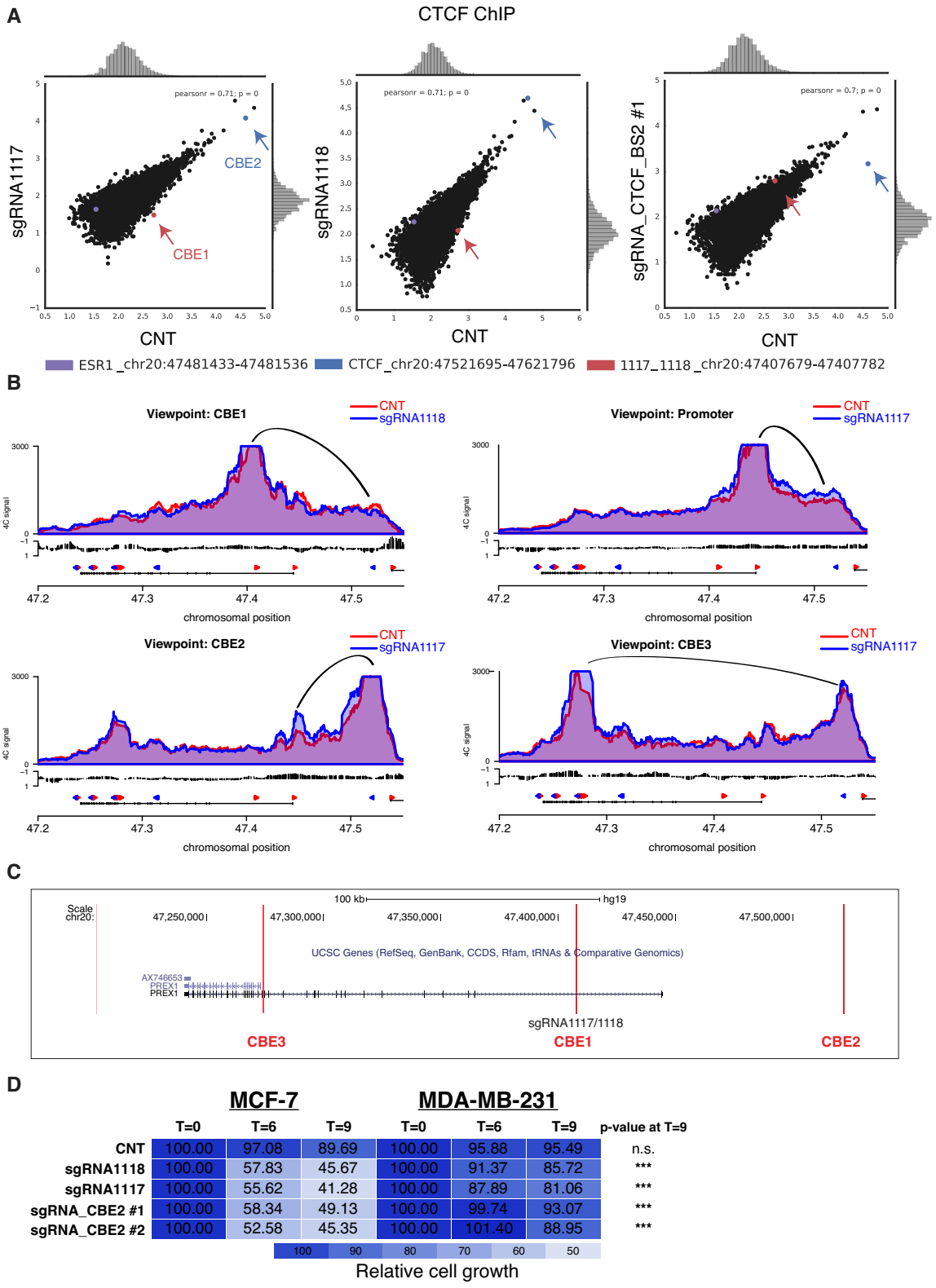

Figure 3. Loss of CTCF binding to CBE1 modulates local chromatin interactions with *PREX1* promoter. (**A**) MCF-7 cells stably expressing nontargeting sgRNA (CNT), sgRNA1117 and sgRNA1118 (targeting CBE1), and sgRNA-CBE2 were examined by CTCF-ChIP-seq. Red arrow for CBE1, blue arrow for CBE2 and purple arrow for ERα binding site 24365 (EBE[24365]). (**B**) Relative chromatin interaction intensity of PREX1 locus is indicated with viewpoints from CBE1, CBE2, CBE3 and promoter by 4C-seq data aligned with CTCF ChIP-seq. RefSeq gene panel was shown at the bottom of the panel. CNT indicates a nontargeting sgRNA. For complete data see Supplementary Figure S3. (**C**) Genomic coordinates of CBE1 and CBE2. (**D**) Competitive proliferation assays of MCF-7 cells transduced with individual sgRNAs targeting CBE2. Values on day 6 ($T = 6$) and day 9 ($T = 9$) normalized to day 0 ($T = 0$). ***$P < 0.005$, **$P < 0.01$, *$P < 0.05$, two-tailed Student's *t*-test. n.s., not significant.

ined cell proliferation and interaction profile of each CBE. Similar to targeting CBE1 and *PREX1*, targeting CBE2 markedly reduced cell proliferation in MCF-7, but not in MDA-MB-231 cells (Figure 3C and D). Interestingly, targeting CBE2 triggered an interaction between CBE3 and CBE1, which are in a non-convergent orientation (65), while decreasing the interaction frequency between CBE2 and CBE3 (Supplementary Figure S5). The interaction frequency between CBE2 and *PREX1* promoter was preserved after targeting CBE2, suggesting that this is CTCF independent (Supplementary Figure S5). Thus, our analysis indicates an intricate interplay between several CBEs for maintaining a chromatin structure in the PREX1 locus that supports its induction and enhance stimulation of cell proliferation by ERα.

A chromatin conformation structure around *PREX1* promoter allows efficient transcriptional activation by ERα

As our results demonstrate that CBE1 is required for *PREX1* expression, our next goal was to identify enhancers whose activity is attenuated upon targeting CBE1. Building on eRNAs as a quantitative measure of enhancer activity (66), we used Global Run-On sequencing (GRO-seq) (67) to measure changes in nascent transcription in CBE1-knockout cells compared to control. The *dREG* (discriminative Regulatory Element detection from GRO-seq (dREG)) (68) tool detected 25 putative transcriptional regulatory elements (TREs) in the *PREX1* locus. Notably, the majority of these elements showed reduced transcriptional activity throughout the whole *PREX1* locus in CBE-targeted cells (Figure 4A; Supplementary Figure S6A and B). Intersection with publicly available ERα ChIP-seq data showed that one of the elements that showed the most pronounced attenuation of transcription (interval number 24365 (Figure 4A), which resides between *PREX1* promoter and CBE2) has a strong ERα binding site (named here EBE^{24365}) (Supplementary Figure S4C), and contains key marks of a transcriptionally active ERα-bound enhancer (p300 binding and H3K27Ac; Supplementary Figure S6C) (40). We validated reduced transcriptional activity by measuring the expression level of enhancer-associated RNAs (eRNAs) in CBE1-targeted and control cells using qRT-PCR (Figure 4B). This effect was specific, since eRNA levels of ERα-bound enhancer 1830 (40) showed no change after targeting CBE1 (Figure 4B). Moreover, CRISPR-targeting of EBE^{24365} resulted in ∼2-fold reduction in PREX1 expression (Supplementary Figure S6D and E), and attenuated cell cycle progression of MCF-7 cells, while having a much smaller impact on the proliferation of MDA-MB-231 cells (Figure 4C). This indicates that this regulatory element (transcription interval 24365) is a functional unit linking ERα activity to *PREX1* expression. Therefore, this result also suggests that CBE1 is required for optimal *PREX1* expression at least in part through EBE^{24365}.

Next, we investigated whether chromatin conformation changes induced by CBE1 are required for optimal activation of PREX1 transcription by EBE^{24365}. Therefore, we first checked the recruitment of ERα to EBE^{24365} in the presence or absence of CBE1 by ChIP-seq analysis. Surprisingly, we found no clear difference in ERα binding between CBE1-knockout cells compared to control (sgRNA1117 and 1118 induced cells, Figure 4D left and middle panel), suggesting that CBE1 is required for transcriptional activity rather than ERα binding. As a positive control, we directly targeted EBE^{24365} and obtained the expected significant reduction in ERα binding and specific inhibition of MCF-7 cell proliferation (Figure 4D, right panel, and C). Additionally, targeting CBE2 did not affect ERα binding to EBE^{24365} (Supplementary Figure S7B). Altogether, our results indicate that CTCF binding to its cognate elements establishes a chromatin conformation structure around the *PREX1* promoter that allows efficient transcriptional activation by ERα (Figure 4E).

PREX1 expression is a bona-fide marker of ER-dependency and sensitivity to anti-hormonal treatment of breast cancers

At last, we examined whether *PREX1* expression can serve as a biomarker and indicate sensitivity to anti-hormonal treatment of breast cancer. To explore this, we initially examined the expression of *PREX1* in the METABRIC cohort (containing 2000 breast cancer cases (55)), and in the TCGA dataset (1095 breast cancer samples). This analysis demonstrated a positive correlation between *PREX1* and *ESR1* in both datasets (Figure 5A and B). Second, we interrogated DRIVE (56) and Achilles (57) datasets and found out that *PREX1* expression levels are associated with ERα-dependency in breast cancer cell lines (Figure 5C). At last, we examined the correlation between PREX1 and disease-free survival of breast cancer patients in the METABRIC dataset. This analysis revealed that *PREX1* expression correlates significantly with disease-free survival of ERα-positive breast cancer patients, with low *PREX1* levels indicating poor prognosis. (Figure 5D). Moreover, higher levels of *PREX1* were associated with better outcome only in endocrine therapy treated cases, and not for patients who did not receive any adjuvant endocrine therapeutics (Figure 5E and F), indicating that PREX1 could be a predictive biomarker for the efficacy of adjuvant endocrine therapy response.

DISCUSSION

In summary, we present here an approach to uncover and characterize the functional role of genetic elements important for chromatin structure in human breast cancer cells. We focused on ERα signaling, being the key-driver in the majority of breast cancers (69,70) and because ERα mainly binds to distal transcriptional enhancers whose activity is likely to be dependent on local chromatin conformation for proper activation of its target genes (15,71). Although genome-wide chromatin conformation capture (Hi-C) (72) and ChIA-PET (4) data support the chromatin interactions mediated by ERα, causal evidence for it is lacking. Therefore, we aimed to target CTCF binding sites in the vicinity of ERα-bound enhancers by using a CRISPR-Cas9-based screening approach. This allowed us to identify four functional CBEs, each of which is required for local transcriptional activity, suggesting that one or more genes in proximity of the target sites can be vital for ERα function. Though our study was limited to two cell lines, it demonstrates the

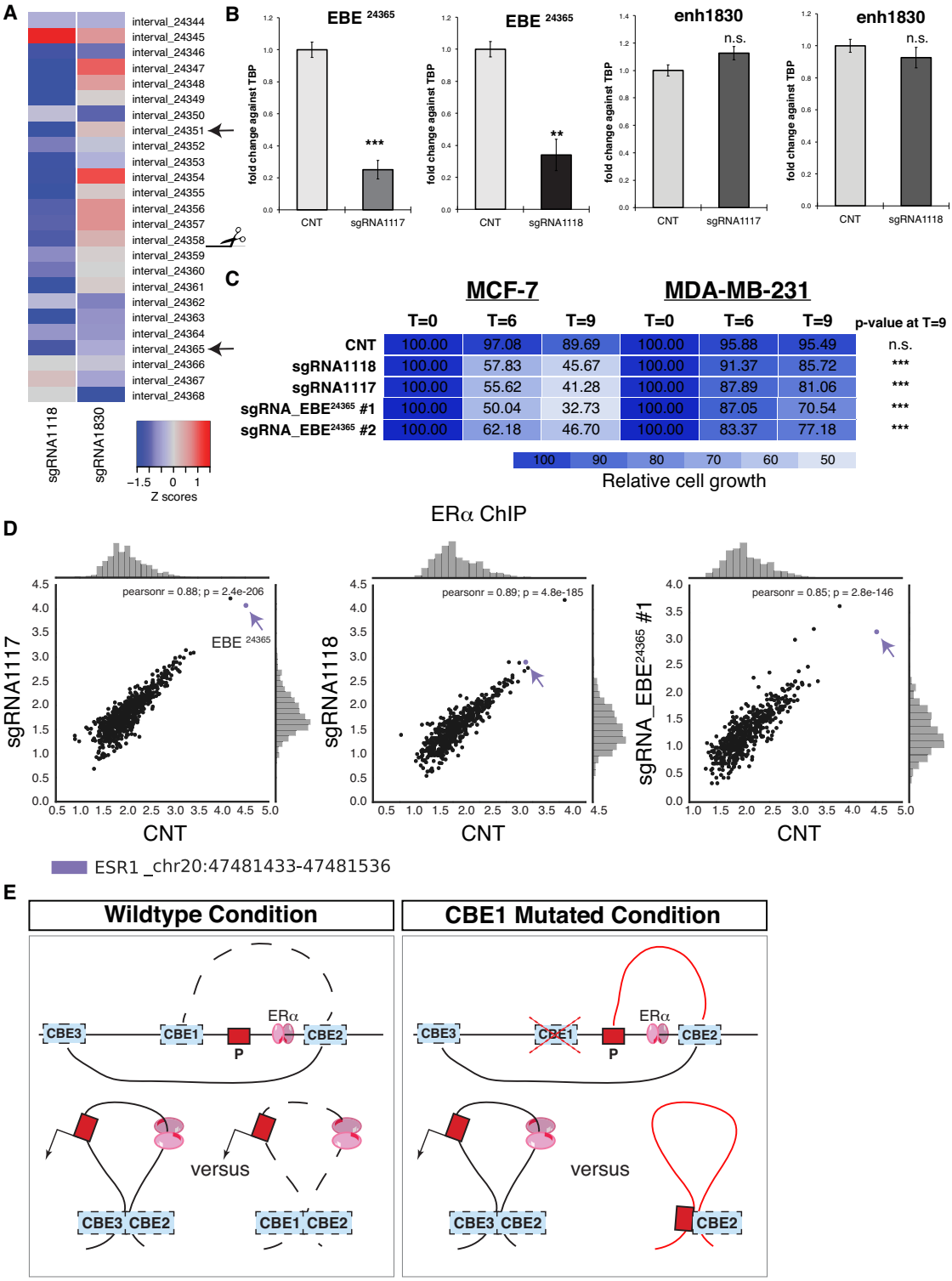

Figure 4. Enhancer activity but not ERα binding is affected by the local chromatin structure at PREX1 locus. (**A**) Transcriptional activity measured by GRO-seq in MCF-7 cells transduced with sgRNA1118 (targeting CBE1), sgRNA1830 here serving as a negative control or nontargeting sgRNA for normalization. Heat-blot was generated according to the *Z*-score. The heatmap shows the change (*Z*-score of fold-change relative to nontargeting control) in the activity of the putative TREs (intervals showing bi-directional transcription) that were identified by dREG in the PREX1 locus (see also Figure S5B). (**B**) Relative expression level of eRNA at EBE^{24365} site was measured by qRT-PCR in MCF-7 cells. Gene expression levels were normalized to TBP. $n = 3$ $^{***}P < 0.005$, $^{*}P < 0.05$, two-tailed Student's *t*-test. n.s., not significant. CNT indicates a non-targeting sgRNA. (**C**) MCF-7 cells stably expressing nontargeting sgRNA (CNT), sgRNA1117 and sgRNA1118 (targeting CBE1), and sgRNA-EBE^{24365} were examined by ChIP-seq. (**D**) MCF-7 cells were assayed by competitive proliferation assay after transduced with individual sgRNAs targeting EBE^{24365}. Values on day 6 ($T = 6$) and day 9 ($T = 9$) normalized to day 0 ($T = 0$). $^{***}P < 0.005$, $^{**}P < 0.01$, $^{*}P < 0.05$, two-tailed Student's *t*-test. n.s., not significant. (**E**) Model summarizes the role of chromatin conformation on *PREX1* gene regulation.

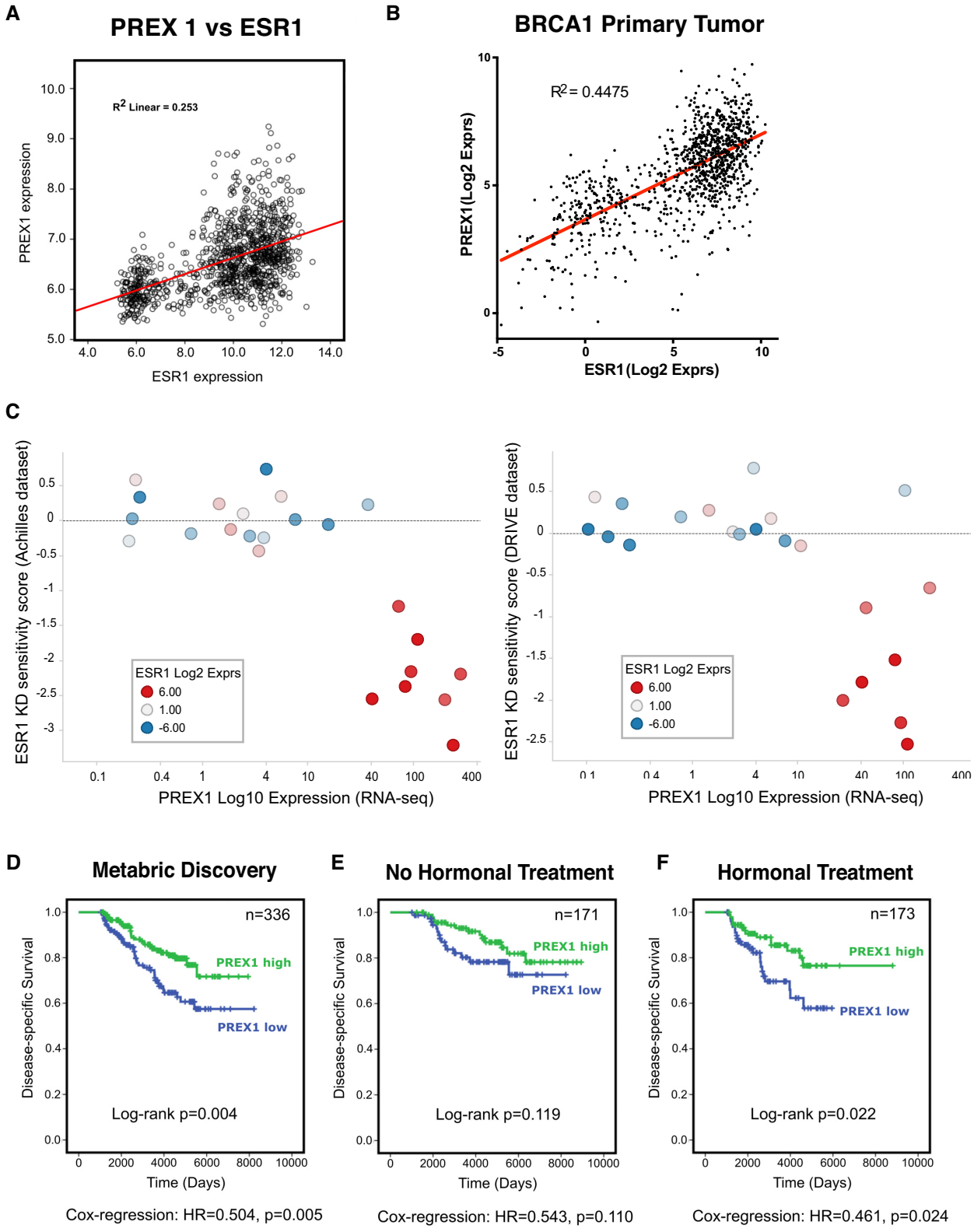

Figure 5. Clinical parameters associated with *PREX1* gene in the METABRIC cohort of breast cancer patients. (**A**) Correlation plot of *PREX1* versus *ESR1* expression data. (**B**) Scatter plot representing the expression of *ESR1* and *PREX1* (log2 RNA-seq) in 1095 primary breast tumor samples from TCGA. R2, correlation coefficient. (**C**) Dot plot representing the expression levels of *PREX1* (X-axys, log2 RNA-seq) compared to the sensitivity to ESR1 knockdown according to the DRIVE (left panel) and Achilles (right panel) datasets. Each dot represents a breast cancer cell line and they are color-coded based on the expression level of *ESR1* (red = maximum, white = median and blue = minimum). (**D**) Disease-specific survival of ERα-positive breast cancer patients from the METABRIC cohort categorized according to PREX1 expression levels. Log-rank and adjusted Cox regression is shown. (**E** and **F**) As in (D), but now patients were additionally stratified for their hormonal treatment status.

robustness of functional genetic screens in uncovering the functional chromatin structure elements in ERα pathway.

Our results pinpoint a key role of CBEs in mediating chromatin interactions that established robust expression of *PREX1* by ERα. *PREX1* is located at chr20q.13 which is amplified in 8–29% of breast tumors (63). While elevated PREX1 expression levels are observed in luminal type breast cancer cell lines and tumors, basal type cancer cell and tumors, as well as normal breast tissue, have either lower or no PREX1 expression at all (63). PREX1 functions as a Rac1 activator in cancer cells and downregulation of PREX1 levels repressed breast cancer cell proliferation and tumorigenesis (63). It has also been documented that PREX1 levels are inversely correlated with PI3K activation, and positively correlated with ERα expression in human breast tumors (73). Here, we validated the expression correlation of *PREX1* and ERα, show that *PREX1* is highly expressed in ERα−driven breast cancer cell lines, and correlates with patient survival and sensitivity to anti-hormonal treatment of breast cancer.

Although a link between ERα and PREX1 was described before, the exact mechanism how ERα regulates *PREX1* remained elusive. For example, Barrio-Real *et al.* have suggested that the methylation status of PREX1 promoter can determine responsiveness to ERα activity (74). Nevertheless, our study provides a mechanistic explanation to the PREX1 and ERα link in breast cancer, and proposes a key role for CTCF and chromatin conformation in the regulation of *PREX1* expression in breast cancer.

Since the *PREX1* locus is amplified in MCF-7 cells (seven copies; (75)), loss of genomic DNA by generating two simultaneous cuts emerges as a possibility. However, CRISPR-mediated targeting of CBE1 did not affect neither CTCF binding to CBE2 nor ERα-binding to EBE^{24365} (Figure 3A and 4D). Moreover, targeting CBE1 did not change the interaction frequency between CBE2 and CBE3 (Supplementary Figure S4). In addition, *PREX1* locus is also amplified in MDA-MB-231 cells (four copies; (75)) where we do not observe any effect on proliferation upon targeting CBE1 (Figure 1D and Supplementary Figure S2B). Therefore, we conclude that the proliferation effect seen in MCF-7 cells upon CBE1 targeting is due to the interference with CTCF binding.

CTCF has a ubiquitous origin in its function. Its contribution to gene expression, heterochromatin spread and higher-order genomic compartmentalization is still under investigation. Nora *et al.* used an auxin-inducible degron system to elaborate the role of CTCF in gene regulation and organization of chromatin structure by triggering acute CTCF degradation (76). Their study revealed the involvement of CTCF in insulation of TADs and in the regulation of gene expression, while higher-order genomic compartmentalization seemed like a CTCF-independent event. Focusing on CTCF-mediated long-range enhancer-promoter interactions, Ren *et al.*, used RNAi-mediated CTCF downregulation to demonstrate the effect of CTCF on cell type specific genes (77). Moreover, a CRISPR-mediated deletion of CTCF at the *TFF1* locus clearly supported the notion that CTCF has a critical role in enhancer–promoter interaction of cell type specific genes (77). Our results here support a critical role of CTCF in ERα-mediated regulation of gene expression by shaping enhancer–promoter interactions in breast cancer cells.

Our results indicate the importance of both CTCF and ERα binding for achieving an optimal transcription activation of *PREX1*. We observed that *PREX1* expression level was attenuated in the absence of estrogen signaling (Figure 2E), but more severely affected estrogen deprivation was combined with loss of CTCF binding to CBE1. Moreover, even 6 h of estrogen signaling was not sufficient to stimulate the *PREX1* transcription level in the absence CTCF-binding to CBE1 (Figure 2E). Therefore, CTCF-mediated chromatin conformation seems crucial to maintain ERα-mediated gene expression in the *PREX1* locus. This conclusion is supported by the data obtained by Zhang *et al.* who also observed that CTCF binding is required for ERα-mediated transcription of the estrogen responsive gene *TFF1* (78). In their work, ERα chromatin binding was blocked upon CTCF knockdown. Differently, we observed that while eRNA expression of the most prominent ERα-bound enhancer (EBE^{24365}) in the *PREX1* locus was compromised, ERα-binding to this site was not altered (Figure 4A, B and D). It is known that tightly enclosed chromatin interactions are required to maintain optimal transcriptional levels by inducing efficient cycling of the transcriptional machinery (4,79,80). Additionally, CTCF interacts directly with the large subunit of RNAPII (81) to modulate transcription, allow accurate transcription initiation, and induce transcription initiation pausing. Thus, ERα binding at the *PREX1* locus requires CTCF binding to effectively activate transcription. Whether transcriptional initiation, or elongation, or both is affected by CTCF remains to be elucidated.

The most prominent structural interaction in the *PREX1* locus is observed between CBE2 and CBE3 (see Figure 4E). An additional interaction with lower frequency is observed between CBE1 and CBE2. CRISPR-Cas9 mediated mutations at CBE1 stimulated the interaction frequency between CBE2 and PREX1 promoter, and led to downregulation of *PREX1* transcription (Figure 2A). We therefore speculate that the increased CBE2- PREX1 promoter interaction frequency following CBE1 targeting is inhibitory to *PREX1* transcriptional activity either because of reduced accessibility of transcription factors to the promoter region, reduced binding of RNAPII or direct inhibition.

Interestingly, also preventing CTCF binding to CBE2 (Figure 3A) caused a strong downregulation of *PREX1* expression (Supplementary Figure S7A). This effect could not be explained by changes in ERα binding to sites in the region (Supplementary Figure S7B). In contrast, we did observe changes in chromatin conformation which can explain the requirement of CTCF binding to CBE2 for *PREX1* expression (Figure 4D). Two interactions are favored as a result of targeting CBE2: between CBE2 and PREX1 promoter and between CBE3 and CBE1 (Supplementary Figure S7C). We speculate that both of these interactions are inhibitory. While CBE3-CBE1 interaction may affect progression of transcription CBE2-PREX1 promoter interaction may affect initiation leading to a greater effect on *PREX1* gene expression than targeting CBE1 or EBE^{24365}.

When we examined T47D, another ERα positive breast cancer cell line; we found that CTCF did not bound CBE1

in this cell line (Supplementary Figure S8A). In line with this, when CBE1 was targeted either by sgRNA1117 or sgRNA118, no change in *PREX1* expression was detected (Supplementary Figure S8B). Also, ERα binding pattern and eRNA activity show differences between MCF-7 and T47D (Supplementary Figure S8C). Therefore, as noted before (15), cell type specific patterns of CTCF binding can shape gene expression changes in response to E2.

Our detailed biochemical analysis of PREX1 locus indicated that a correct chromatin structure is required for transcriptional activity of a nearby key ERα-bound enhancer (as determined by eRNA expression), but it is dispensable for ERα binding. This can be explained by reduced interference with promoter activity through a competitive mechanism between an upstream and a downstream CTCF sites. Recently, a similar mechanism was observed for c-MYC, where the promoter of *PVT1* competes with the promoter of c-MYC for engagement with nearby enhancer elements, pinpointing *PVT1* promoter as a DNA element with potential tumor suppressing activity (82). This suggests that competitive chromatin regulatory mechanisms are common and are employed by cancer cells to stimulate tumor proliferation.

Altogether, our study shows that the CRISPR-Cas9 technology is a robust tool to identify and characterize functional regulatory nodes of chromatin conformation in an unbiased fashion. Furthermore, our study adds more evidence for the importance of CTCF-mediated chromatin conformation for gene expression and oncogenic signaling. As our approach here was based on a selected group of CTCF sites residing in the vicinity of ERα-enhancers, it is possible that some ERα-binding sites function to regulate target gene expression independently of CTCF-mediated chromatin conformation. To address this issue, a genome wide CTCF-binding site screen is required. Indeed, it is expected that the approach we presented here will be scaled up and be used to uncover novel functions of chromatin regulatory elements of the human genome under normal and pathological conditions. Nevertheless, our study highlights the importance of chromatin conformation in gene regulation and supports the notion that mutations in these regions can influence local gene expression and cellular phenotypes such as cancer progression.

DATA AVAILABILITY

The sequencing results of GRO-seq (GSE124572), RNA-seq (GSE124573) and ChIP-seq (GSE124667) on MCF-7 cells induced with sgRNA vectors targeting CTCF binding elements are available in the NCBI GEO database.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank the current and previous members of the Agami lab for valuable and insightful discussions.
Authors' contributions: G.K. performed most of the experiments. Z.M. conducted all the bioinformatics analyses. R.L. helped the drop-out screening and GRO-seq experiment. G.G. and K.F. performed correlation analysis. R.E created the bioinformatic pipeline. H.T. and E.d.W performed and analyzed the 4C data. S.P., K.S., Y.K. and W.Z. carried out and analyzed the ChIP-seq data, respectively. G.K. and R.A. wrote the manuscript.

FUNDING

Israeli Cancer Association (ICA) (to R.E.); Marguerite Stolz Research Fellowship Fund; Gad, Nava and Shye Shtacher Fellowship (to Z.M.) (in part); Edmond J. Safra Center for Bioinformatics at Tel Aviv University Fellowship (to R.E.); enhReg ERC-AdV and NWO [NGI 93512001/2012 to R.A.]. Funding for open access charge: Institutional Funds.
Conflict of interest statement. None declared.

REFERENCES

1. Blows,F.M., Driver,K.E., Schmidt,M.K., Broeks,A., van Leeuwen,F.E., Wesseling,J., Cheang,M.C., Gelmon,K., Nielsen,T.O., Blomqvist,C. *et al.* (2010) Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med.*, **7**, e1000279.
2. Lapidus,R.G., Nass,S.J. and Davidson,N.E. (1998) The loss of estrogen and progesterone receptor gene expression in human breast cancer. *J. Mammary Gland Biol. Neoplasia*, **3**, 85–94.
3. 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.
4. Fullwood,M.J., Liu,M.H., Pan,Y.F., Liu,J., Xu,H., Mohamed,Y.B., Orlov,Y.L., Velkov,S., Ho,A., Mei,P.H. *et al.* (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*, **462**, 58–64.
5. Li,W., Notani,D., Ma,Q., Tanasa,B., Nunez,E., Chen,A.Y., Merkurjev,D., Zhang,J., Ohgi,K., Song,X. *et al.* (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature*, **498**, 516–520.
6. Hah,N., Murakami,S., Nagari,A., Danko,C.G. and Kraus,W.L. (2013) Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.*, **23**, 1210–1223.
7. 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.
8. Zwart,W., Rondaij,M., Jalink,K., Sharp,Z.D., Mancini,M.A., Neefjes,J. and Michalides,R. (2009) Resistance to antiestrogen arzoxifene is mediated by overexpression of cyclin D1. *Mol. Endocrinol.*, **23**, 1335–1345.
9. Schrijver,W.A., van Diest,P.J., Moelans,C.B. and Dutch Distant Breast Cancer Metastases Consortium (2017) Unravelling site-specific breast cancer metastasis: a microRNA expression profiling study. *Oncotarget*, **8**, 3111–3123.
10. Schrijver,W., Schuurman,K., van Rossum,A., Droog,M., Jeronimo,C., Salta,S., Henrique,R., Wesseling,J., Moelans,C., Linn,S.C. *et al.* (2018) FOXA1 levels are decreased in pleural breast cancer metastases after adjuvant endocrine therapy, and this is associated with poor outcome. *Mol. Oncol.*, **12**, 1884–1894.
11. Early Breast Cancer Trialists' Collaborative, G. (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet.*, **365**, 1687–1717.
12. Toy,W., Shen,Y., Won,H., Green,B., Sakr,R.A., Will,M., Li,Z., Gala,K., Fanning,S., King,T.A. *et al.* (2013) ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.*, **45**, 1439–1445.

13. Zwart,W., Griekspoor,A., Berno,V., Lakeman,K., Jalink,K., Mancini,M., Neefjes,J. and Michalides,R. (2007) PKA-induced resistance to tamoxifen is associated with an altered orientation of ERalpha towards co-activator SRC-1. *EMBO J.*, **26**, 3534–3544.
14. Zwart,W., Griekspoor,A., Rondaij,M., Verwoerd,D., Neefjes,J. and Michalides,R. (2007) Classification of anti-estrogens according to intramolecular FRET effects on phospho-mutants of estrogen receptor alpha. *Mol. Cancer Ther.*, **6**, 1526–1533.
15. Ross-Innes,C.S., Brown,G.D. and Carroll,J.S. (2011) A co-ordinated interaction between CTCF and ER in breast cancer cells. *BMC Genomics*, **12**, 593.
16. Jansen,M.P., Knijnenburg,T., Reijm,E.A., Simon,I., Kerkhoven,R., Droog,M., Velds,A., van Laere,S., Dirix,L., Alexi,X. *et al.* (2013) Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. *Cancer Res.*, **73**, 6632–6641.
17. Maston,G.A., Evans,S.K. and Green,M.R. (2006) Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.*, **7**, 29–59.
18. Consortium,E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
19. Arnone,M.I. and Davidson,E.H. (1997) The hardwiring of development: organization and function of genomic regulatory systems. *Development*, **124**, 1851–1864.
20. Raab,J.R. and Kamakaka,R.T. (2010) Insulators and promoters: closer than we think. *Nat. Rev. Genet.*, **11**, 439–446.
21. Maston,G.A., Landt,S.G., Snyder,M. and Green,M.R. (2012) Characterization of enhancer function from genome-wide analyses. *Annu. Rev. Genomics Hum. Genet.*, **13**, 29–57.
22. Li,L.M. and Arnosti,D.N. (2011) Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. *Curr. Biol.*, **21**, 406–412.
23. Bulger,M. and Groudine,M. (2010) Enhancers: the abundance and function of regulatory sequences beyond promoters. *Dev. Biol.*, **339**, 250–257.
24. Frasch,M. and Levine,M. (1987) Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in Drosophila. *Genes Dev.*, **1**, 981–995.
25. Levine,M. (2010) Transcriptional enhancers in animal development and evolution. *Curr. Biol.*, **20**, R754–R763.
26. Lobanenkov,V.V., Nicolas,R.H., Adler,V.V., Paterson,H., Klenova,E.M., Polotskaja,A.V. and Goodwin,G.H. (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5′-flanking sequence of the chicken c-myc gene. *Oncogene*, **5**, 1743–1753.
27. Dowen,J.M., Fan,Z.P., Hnisz,D., Ren,G., Abraham,B.J., Zhang,L.N., Weintraub,A.S., Schujiers,J., Lee,T.I., Zhao,K. *et al.* (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, **159**, 374–387.
28. Merkenschlager,M. and Odom,D.T. (2013) CTCF and cohesin: linking gene regulatory elements with their targets. *Cell*, **152**, 1285–1297.
29. Dixon,J.R., Selvaraj,S., Yue,F., Kim,A., Li,Y., Shen,Y., Hu,M., Liu,J.S. and Ren,B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
30. Nora,E.P., Lajoie,B.R., Schulz,E.G., Giorgetti,L., Okamoto,I., Servant,N., Piolot,T., van Berkum,N.L., Meisig,J., Sedat,J. *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, **485**, 381–385.
31. de Wit,E., Vos,E.S., Holwerda,S.J., Valdes-Quezada,C., Verstegen,M.J., Teunissen,H., Splinter,E., Wijchers,P.J., Krijger,P.H. and de Laat,W. (2015) CTCF binding polarity determines chromatin looping. *Mol. Cell*, **60**, 676–684.
32. Gibcus,J.H. and Dekker,J. (2013) The hierarchy of the 3D genome. *Mol. Cell*, **49**, 773–782.
33. Phillips-Cremins,J.E., Sauria,M.E., Sanyal,A., Gerasimova,T.I., Lajoie,B.R., Bell,J.S., Ong,C.T., Hookway,T.A., Guo,C., Sun,Y. *et al.* (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*, **153**, 1281–1295.
34. Hnisz,D., Weintraub,A.S., Day,D.S., Valton,A.L., Bak,R.O., Li,C.H., Goldmann,J., Lajoie,B.R., Fan,Z.P., Sigova,A.A. *et al.* (2016) Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, **351**, 1454–1458.
35. Lupianez,D.G., Kraft,K., Heinrich,V., Krawitz,P., Brancati,F., Klopocki,E., Horn,D., Kayserili,H., Opitz,J.M., Laxova,R. *et al.* (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**, 1012–1025.
36. Kemp,C.J., Moore,J.M., Moser,R., Bernard,B., Teater,M., Smith,L.E., Rabaia,N.A., Gurley,K.E., Guinney,J., Busch,S.E. *et al.* (2014) CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.*, **7**, 1020–1029.
37. Chan,C.S. and Song,J.S. (2008) CCCTC-binding factor confines the distal action of estrogen receptor. *Cancer Res.*, **68**, 9041–9049.
38. Shalem,O., Sanjana,N.E., Hartenian,E., Shi,X., Scott,D.A., Mikkelson,T., Heckl,D., Ebert,B.L., Root,D.E., Doench,J.G. *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343**, 84–87.
39. Wang,T., Wei,J.J., Sabatini,D.M. and Lander,E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.
40. Korkmaz,G., Lopes,R., Ugalde,A.P., Nevedomskaya,E., Han,R., Myacheva,K., Zwart,W., Elkon,R. and Agami,R. (2016) Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.*, **34**, 192–198.
41. Korkmaz,G., le Sage,C., Tekirdag,K.A., Agami,R. and Gozuacik,D. (2012) miR-376b controls starvation and mTOR inhibition-related autophagy by targeting ATG4C and BECN1. *Autophagy*, **8**, 165–176.
42. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
43. Anders,S., Pyl,P. T. and Huber,W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
44. Cunningham,F., Amode,M.R., Barrell,D., Beal,K., Billis,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fitzgerald,S. *et al.* (2015) Ensembl 2015. *Nucleic Acids Res.*, **43**, D662–D669.
45. Zwart,W., Koornstra,R., Wesseling,J., Rutgers,E., Linn,S. and Carroll,J.S. (2013) A carrier-assisted ChIP-seq method for estrogen receptor-chromatin interactions from breast cancer core needle biopsy samples. *BMC Genomics*, **14**, 232.
46. Schmidt,D., Wilson,M.D., Spyrou,C., Brown,G.D., Hadfield,J. and Odom,D.T. (2009) ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods*, **48**, 240–248.
47. Kumar,V., Muratani,M., Rayan,N.A., Kraus,P., Lufkin,T., Ng,H.H. and Prabhakar,S. (2013) Uniform, optimal signal processing of mapped deep-sequencing data. *Nat. Biotechnol.*, **31**, 615–622.
48. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
49. Dale,R.K., Pedersen,B.S. and Quinlan,A.R. (2011) Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics*, **27**, 3423–3424.
50. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
51. Waskom,M., Botvinnik,O., drewokane, Hobson,P., Halchenko,Y., Lukauskas,S., Warmenhoven,J., Cole,J.B., Hoyer,S. and Vanderplas,J. (2016) seaborn: v0.7.0 (January 2016). https://zenodo.org/record/45133#.XT7Lti2B1Gw.
52. van de Werken,H.J., Landan,G., Holwerda,S.J., Hoichman,M., Klous,P., Chachik,R., Splinter,E., Valdes-Quezada,C., Oz,Y., Bouwman,B.A. *et al.* (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods*, **9**, 969–972.
53. Li,H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. https://arxiv.org/abs/1303.3997.
54. Geeven,G., Teunissen,H., de Laat,W. and de Wit,E. (2018) peakC: a flexible, non-parametric peak calling package for 4C and Capture-C data. *Nucleic Acids Res.*, **46**, e91.
55. Curtis,C., Shah,S.P., Chin,S.F., Turashvili,G., Rueda,O.M., Dunning,M.J., Speed,D., Lynch,A.G., Samarajiwa,S., Yuan,Y. *et al.* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, **486**, 346–352.
56. McDonald,E.R. 3rd, de Weck,A., Schlabach,M.R., Billy,E., Mavrakis,K.J., Hoffman,G.R., Belur,D., Castelletti,D., Frias,E., Gampa,K. *et al.* (2017) Project DRIVE: a compendium of cancer

dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell*, **170**, 577–592.
57. Tsherniak,A., Vazquez,F., Montgomery,P.G., Weir,B.A., Kryukov,G., Cowley,G.S., Gill,S., Harrington,W.F., Pantel,S., Krill-Burger,J.M. *et al.* (2017) Defining a cancer dependency map. *Cell*, **170**, 564–576.
58. Lopes,R., Korkmaz,G., Sonia Aristin Revilla,S.A., Vliet,R., Nagel,R., Lars Custers,L., Kim,Y., Breugel,P.C., Zwart,W., Moumbeini,B. *et al.* (2018) CUEDC1 is a primary target of ERalpha essential for the growth of breast cancer cells. *Cancer Lett.*, **436**, 87–95.
59. Hah,N., Danko,C.G., Core,L., Waterfall,J.J., Siepel,A., Lis,J.T. and Kraus,W.L. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell*, **145**, 622–634.
60. Liu,H.J., Ooms,L.M., Srijakotre,N., Man,J., Vieusseux,J., Waters,J.E., Feng,Y., Bailey,C.G., Rasko,J.E., Price,J.T. *et al.* (2016) PtdIns(3,4,5)P3-dependent Rac Exchanger 1 (PREX1) Rac-Guanine Nucleotide Exchange Factor (GEF) activity promotes breast cancer cell proliferation and tumor growth via activation of extracellular Signal-regulated kinase 1/2 (ERK1/2) signaling. *J. Biol. Chem.*, **291**, 17258–17270.
61. Gont,A., Daneshmand,M., Woulfe,J., Lavictoire,S.J. and Lorimer,I.A. (2017) PREX1 integrates G protein-coupled receptor and phosphoinositide 3-kinase signaling to promote glioblastoma invasion. *Oncotarget*, **8**, 8559–8573.
62. Montero,J.C., Seoane,S., Ocana,A. and Pandiella,A. (2011) P-Rex1 participates in Neuregulin-ErbB signal transduction and its expression correlates with patient outcome in breast cancer. *Oncogene*, **30**, 1059–1071.
63. Sosa,M.S., Lopez-Haber,C., Yang,C., Wang,H., Lemmon,M.A., Busillo,J.M., Luo,J., Benovic,J.L., Klein-Szanto,A., Yagi,H. *et al.* (2010) Identification of the Rac-GEF P-Rex1 as an essential mediator of ErbB signaling in breast cancer. *Mol. Cell*, **40**, 877–892.
64. Aguirre,A.J., Meyers,R.M., Weir,B.A., Vazquez,F., Zhang,C.Z., Ben-David,U., Cook,A., Ha,G., Harrington,W.F., Doshi,M.B. *et al.* (2016) Genomic copy number dictates a Gene-Independent cell response to CRISPR/Cas9 targeting. *Cancer Discov.*, **6**, 914–929.
65. Haarhuis,J.H.I., van der Weide,R.H., Blomen,V.A., Yanez-Cuna,J.O., Amendola,M., van Ruiten,M.S., Krijger,P.H.L., Teunissen,H., Medema,R.H., van Steensel,B. *et al.* (2017) The cohesin release factor WAPL restricts chromatin loop extension. *Cell*, **169**, 693–707.
66. Andersson,R., Gebhard,C., Miguel-Escalada,I., Hoof,I., Bornholdt,J., Boyd,M., Chen,Y., Zhao,X., Schmidl,C., Suzuki,T. *et al.* (2014) An atlas of active enhancers across human cell types and tissues. *Nature*, **507**, 455–461.
67. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.
68. 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*, **12**, 433–438.
69. Hayashi,S.I., Eguchi,H., Tanimoto,K., Yoshida,T., Omoto,Y., Inoue,A., Yoshida,N. and Yamaguchi,Y. (2003) The expression and function of estrogen receptor alpha and beta in human breast cancer and its clinical application. *Endocr. Relat. Cancer*, **10**, 193–202.
70. Dahlman-Wright,K., Cavailles,V., Fuqua,S.A., Jordan,V.C., Katzenellenbogen,J.A., Korach,K.S., Maggi,A., Muramatsu,M., Parker,M.G. and Gustafsson,J.A. (2006) International Union of Pharmacology. LXIV. Estrogen receptors. *Pharmacol. Rev.*, **58**, 773–781.
71. Fiorito,E., Sharma,Y., Gilfillan,S., Wang,S., Singh,S.K., Satheesh,S.V., Katika,M.R., Urbanucci,A., Thiede,B., Mills,I.G. *et al.* (2016) CTCF modulates Estrogen Receptor function through specific chromatin and nuclear matrix interactions. *Nucleic Acids Res.*, **44**, 10588–10602.
72. Mourad,R., Hsu,P.Y., Juan,L., Shen,C., Koneru,P., Lin,H., Liu,Y., Nephew,K., Huang,T.H. and Li,L. (2014) Estrogen induces global reorganization of chromatin structure in human breast cancer cells. *PLoS One*, **9**, e113354.
73. Dillon,L.M., Bean,J.R., Yang,W., Shee,K., Symonds,L.K., Balko,J.M., McDonald,W.H., Liu,S., Gonzalez-Angulo,A.M., Mills,G.B. *et al.* (2015) P-REX1 creates a positive feedback loop to activate growth factor receptor, PI3K/AKT and MEK/ERK signaling in breast cancer. *Oncogene*, **34**, 3968–3976.
74. Barrio-Real,L., Benedetti,L.G., Engel,N., Tu,Y., Cho,S., Sukumar,S. and Kazanietz,M.G. (2014) Subtype-specific overexpression of the Rac-GEF P-REX1 in breast cancer is associated with promoter hypomethylation. *Breast Cancer Res.*, **16**, 441.
75. Tym,J.E., Mitsopoulos,C., Coker,E.A., Razaz,P., Schierz,A.C., Antolin,A.A. and Al-Lazikani,B. (2016) canSAR: an updated cancer research and drug discovery knowledgebase. *Nucleic Acids Res.*, **44**, D938–D943.
76. Nora,E.P., Goloborodko,A., Valton,A.L., Gibcus,J.H., Uebersohn,A., Abdennur,N., Dekker,J., Mirny,L.A. and Bruneau,B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.
77. Ren,G., Jin,W., Cui,K., Rodrigez,J., Hu,G., Zhang,Z., Larson,D.R. and Zhao,K. (2017) CTCF-Mediated Enhancer-Promoter interaction is a critical regulator of Cell-to-Cell variation of gene expression. *Mol. Cell*, **67**, 1049–1058.
78. Zhang,Y., Liang,J., Li,Y., Xuan,C., Wang,F., Wang,D., Shi,L., Zhang,D. and Shang,Y. (2010) CCCTC-binding factor acts upstream of FOXA1 and demarcates the genomic response to estrogen. *J. Biol. Chem.*, **285**, 28604–28613.
79. Metivier,R., Penot,G., Hubner,M.R., Reid,G., Brand,H., Kos,M. and Gannon,F. (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell*, **115**, 751–763.
80. Reid,G., Hubner,M.R., Metivier,R., Brand,H., Denger,S., Manu,D., Beaudouin,J., Ellenberg,J. and Gannon,F. (2003) Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling. *Mol. Cell*, **11**, 695–707.
81. Chernukhin,I., Shamsuddin,S., Kang,S.Y., Bergstrom,R., Kwon,Y.W., Yu,W., Whitehead,J., Mukhopadhyay,R., Docquier,F., Farrar,D. *et al.* (2007) CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol. Cell Biol.*, **27**, 1631–1648.
82. Cho,S.W., Xu,J., Sun,R., Mumbach,M.R., Carter,A.C., Chen,Y.G., Yost,K.E., Kim,J., He,J., Nevins,S.A. *et al.* (2018) Promoter of lncRNA gene PVT1 is a Tumor-Suppressor DNA boundary element. *Cell*, **173**, 1398–1412.