# Dynamic ER Interactomes Control the Estrogen-Responsive Trefoil Factor (TFF) Locus Cell-Specific Activities.

Justine Quintin, Christine Le Péron, Gaëlle Palierne, Maud Bizot, Stéphanie Cunha, Aurélien A Sérandour, Stéphane Avner, Catherine Henry, Frédéric Percevault, Marc-Antoine Belaud-Rotureau, et al.

▶ **To cite this version:**

Justine Quintin, Christine Le Péron, Gaëlle Palierne, Maud Bizot, Stéphanie Cunha, et al.. Dynamic ER Interactomes Control the Estrogen-Responsive Trefoil Factor (TFF) Locus Cell-Specific Activities.. Molecular and Cellular Biology, 2014, 34 (13), pp.2418-2436. 10.1128/MCB.00918-13 . hal-00984591

**HAL Id: hal-00984591**

**https://univ-rennes.hal.science/hal-00984591**

Submitted on 28 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MCB Accepts, published online ahead of print on 21 April 2014
Mol. Cell. Biol. doi:10.1128/MCB.00918-13
Copyright © 2014, American Society for Microbiology. All Rights Reserved.

MCB Accepts published online ahead of print

# Dynamic ER Interactomes Control the Estrogen-Responsive Trefoil Factor (TFF) Locus Cell-Specific Activities

**Running Title:** ER-mediated regulations of the TFF cluster in 4D

**Keywords:** Estrogen receptor; Transcription; Interactome; Triplex forming oligonucleotides.

Justine Quintin[1], Christine Le Péron[1], Gaëlle Palierne[1], Maud Bizot[1], Stéphanie Cunha[1,*], Aurélien A. Sérandour[1,**], Stéphane Avner[1], Catherine Henry[2], Frédéric Percevault[1], Marc-Antoine Belaud-Rotureau[2,3], Sébastien Huet[1], Erwan Watrin[4], Jérôme Eeckhoute[1,5], Vincent Legagneux[6], Gilles Salbert[1], and Raphaël Métivier[1#]

[1] Equipe SP@RTE, UMR CNRS 6290. Equipe labellisée Ligue contre le Cancer Université de Rennes I. Campus de Beaulieu. 35042 RENNES CEDEX. FRANCE. [2] Cytogenetics and Cellular Biology Department, CHU RENNES. FRANCE. [3] BIOSIT, UMR CNRS 6290. Université de Rennes I. Faculté de Médecine. 2 avenue du Professeur Léon Bernard. 35043 RENNES CEDEX. [4] Equipe CC, UMR CNRS 6290. Université de Rennes I. Faculté de Médecine. 2 avenue du Professeur Léon Bernard. 35043 RENNES CEDEX. FRANCE. [5] INSERM U1011, Université Lille-Nord de France, Faculté de Médecine de Lille-Pôle Recherche. 59045 Lille. FRANCE. [6] Equipe EGD, UMR CNRS 6290. Université de Rennes I. Faculté de Médecine. 2 avenue du Professeur Léon Bernard. 35043 RENNES CEDEX. FRANCE.

* Present Address: Huntsman Cancer Institute. Salt Lake City, Utah. USA. ** Present Address: CRI, Cancer Research UK, Cambridge. UK.

**Correspondence footnote:**

Raphaël Métivier, Equipe SP@RTE, UMR CNRS 6290. Batiment 13. Université de Rennes 1. Campus de Beaulieu. 35042 Rennes Cedex. Tel: +33 (0)2 2323 5142 Fax: +33 (0)2 2323 6794; E-mail: raphael.metivier@univ-rennes1.fr

**Word Count :**

| | |
|---|---|
| **Material and Methods section:** | **3,995 words (21,833 characters)** |
| **Combined Introduction, Results and Discussion sections:** | **5,860 words (32,143 characters)** |
| **Figure Legends:** | **1,738 words (9,268 characters)** |

MCB Accepts published online ahead of print

**ABSTRACT**

Estradiol signaling is ideally suited for analyzing molecular and functional linkages between the different layers of information directing transcriptional regulations: DNA sequence, chromatin modifications and the spatial organization of the genome. Hence, estrogen receptor (ER) can bind at a distance from its target genes and engages timely and spatially coordinated processes to regulate their expression. In the context of the coordinated regulations of co-linear genes, identifying which ER binding sites (ERBSs) regulate a given gene still remains a challenging question. Here, we investigated the coordination of such regulatory events at a 2 Mb genomic locus containing the estrogen-sensitive TFF cluster of genes in breast cancer cells. We demonstrated that this locus exhibits a hormone and cohesin-dependent reduction in the plasticity of its three-dimensional organization that allows multiple ERBSs to be dynamically brought to the vicinity of estrogen-sensitive genes. Additionally, by using triplex forming oligonucleotides, we could precisely document the functional links between ER engagement at given ERBSs and the regulation of particular genes. Hence, our data evidence a formerly suggested cooperation of enhancers towards gene regulations, and also show that redundancy between ERBSs can occur.

**INTRODUCTION**

In Mammals, gene transcription relies on complex and highly organized regulatory processes, which include binding of transcription factors to cognate DNA sequences (*cis* elements), chromatin structure and epigenetic information, the action of additional factors in trans (cofactors and RNA Polymerase II (Pol II) machinery) and the spatial organization of the genome (1-5). Signaling pathways initiated by steroid hormones, such as 17β-estradiol (E2), provide model systems to study these different layers of transcription regulation in mammalian cells. Indeed, exposure to estrogens leads to transcriptional changes of cell-specific gene repertoires, which are mediated by E2-bound Estrogen Receptors (ESR1 -ER throughout the manuscript- and ESR2) (6). On model gene promoters, such as *TFF1*, ER together with a number of its cofactors associate with cognate binding sites (BS) in a cyclic manner to direct their transcription (7, 8). The spatial organization of the genome also determines the coordinated expression of genes (9, 10). This is notably the case for ER, where the existence of clusters of co-regulated genes can originate from genetic and epigenetic information or from chromatin dynamics itself. In some instances, such coordinated regulation of co-linear genes depends on a single regulating unit [*e.g.*, *HBB*, *Mrf4* and *Hox* clusters (11-13)].

Genome-wide analyses of ER binding sites (ERBSs) have demonstrated that ER binds only rarely to the proximal promoter of its target genes, but is mobilized onto intergenic and intronic sequences (14), which have been proposed to communicate with target genes via long-distance intrachromosomal interactions (15). Whether these distant elements are acting as global regulators for clustered E2-responsive genes is still an intriguing question. In addition, these genome-wide studies also showed that additional transcription factors are required for the accurate targeting of ER onto cognate sequences along the whole genome (16). These factors include FOXA1 (17), TFAP2C (19), and PBX1 (20). Among those, FOXA1 may act as an allosteric sensor for histone marks associated with active or poised chromatin (such as H3K4 mono/di-methylation), and it is therefore considered as a pioneer factor

MCB Accepts published online ahead of print

preparing chromatin for subsequent binding of ER (21-23).

We aimed here to obtain functional and mechanistic evidence that distant ERBSs elements actually constitute global regulators for clustered E2-responsive genes. To do so, we engaged an extensive analysis of mechanisms involved in the coordination of the estrogenic response of one cluster of E2-sensitive genes in breast carcinoma cells. These studies were performed in different breast cancer cell lines: MCF-7 cells that constitutively express both ER and FOXA1, and in MDA-MB231 cells that were engineered to constitutively express ER [cells named MDA::ER; (24)] but not FOXA1. Comparative observations made in these two cell lines allowed us to interrogate whether the introduction of ER in MDA-MB231 cells is sufficient to recapitulate regulatory processes observed in MCF-7 at the TFF locus. The combination of chromosome conformation capture methods (3C/4C) with ChIP-chip experiments and the use of triple helix forming oligonucleotides (TFOs), which allows testing the functional importance of individual enhancers, defined key molecular features specifying the transcriptional response induced by E2. We show that, in both cell types, ER engages similar mechanisms to regulate transcription of co-regulated gene clusters, in particular through long-range and dynamic interactions between multiple ERBSs and its target genes. By interfering specifically with the association of ER with given ERBSs, we were also able to determine the relative importance of these different BSs in the regulation of corresponding E2-dependent genes.

## MATERIALS AND METHODS

**Reagents.** All chemicals, restriction or modification enzymes were obtained from Sigma, Roche or New England Biolabs. All primers and siRNAs were purchased from Sigma. Antibodies were from from Abcam, Millipore or SantaCruz (Actin: sc-8432; CTCF: 07-729; ER: HC20 and ab10[TE111-SD1]; FOXA1: ab23738 and RAD21: ab992). The anti-Scc1/RAD21 was a gift from Dr. JM Peters and the anti-hCAPD2 Eg7.2 was previously published (25) BACs were purchased from Invitrogen (RP11-814F13, CTD-2337B13, RP11-35C4, CTD-260o11, RP11-113F1, CTD-1033M14).

**Triplex forming oligonucleotides (TFOs).** We developed a python algorithm (available upon request) following the rules defined in (26) to design putative TFOs targeting 15-30 bp long oligopyrimidine-oligopurine tracts included within ERBSs (Table 1), with one possible divergent base from a strict polyA/G sequence. Triplex formation was monitored *in vitro* by incubating increasing amounts of TFOs with DNA duplexes for 16h at 37°C in a buffer containing 10 mM MgCl2, 100 mM NaCl, 50 mM Tris-HCl (pH 7.4), 10% glycerol and 0.5 mg/ml tRNA. Complexes were separated by native electrophoresis on polyacrylamide gel containing 10 mM MgCl2 and 50 mM Tris-HCl (pH 7.4) and visualized by methylene blue staining.

**Cell culture and reverse-transcription.** MCF-7, MDA-MB231 and MDA-MB231 cells stably expressing ERα [MDA::ER (24)] were maintained in DMEM (Gibco) containing 5% fetal calf serum (FCS, BioWest) and antibiotics (Roche) at 37°C under 5% CO2. MDA::ER media was supplemented with 0.8 mg/ml hygromycin (Calbiochem). For experiments requiring treatment with E2, cells were cultivated for 2 days in DMEM without phenol red containing 2% charcoal-

MCB Accepts published online ahead of print

stripped FCS (csFCS; BioWest) prior to the addition of E2 ($10^{-8}$M final concentration). Total RNAs from $10^7$ cells were purified using TrizolTM reagent (Life Technologies, Inc.) according to the manufacturer's instructions. Two µg of RNA served as template for M-MLV reverse transcriptase (Invitrogen) and Pd(N)6 random hexamers (Amersham Pharmacia Biosciences).

**Transfections.** $2.5x10^6$ cells were plated in 9cm dishes in DMEM/5%FCS for 16h and then grown for 24h in DMEM/2.5% csFCS. Media was then replaced with 4 ml of FCS and antibiotics free Opti-MEM (Sigma) and 1 µmol siRNAs (sense: Luciferase, AACACUUACGCUGAGUACUUCGA; CTCF, GGAGCCUGCCGUAGAAAUU; RAD21, GGUGAAAAUGGCAUUACGG) or 10 µmol TFOs were then transfected using oligofectamine as recommended by the manufacturer (Invitrogen). Following 6h of incubation, the media was completed with 125 µl of csFCS, and E2 stimulation ($10^{-8}$M) was done 36h later.

**Western blotting.** Half cells from confluent 9cm-diameter dishes were directly lysed in sample buffer, and subjected to classical SDS-PAGE. Proteins were transferred onto Hybond nitrocellulose membrane (Amersham) for 2h, which were subsequently blocked in PBS or TBS complemented with 0.1% Tween-20/4% dry milk for 1 h at 4°C. Membranes were then incubated overnight at 4°C with primary antibodies at appropriate concentrations (CTCF: 1/2000; Scc1/RAD21: 1/1000; ER: 1/2500; FOXA1: 1/2500; β-Actin: 1/5000 and 1/2000 for anti-hCAPD2 Eg7.2). Following three successive washes, blots were further incubated for 1h at room temperature using appropriate peroxidase-coupled secondary antibodies diluted at 1/10,000 in PBS or TBS plus 0.1% Tween-20/4% dry milk. Western blots were revealed by the ECL detection kit (Amersham).

**DNA-FISH.** Probes were produced by direct labelling of BACs clones through random priming (Bioprime array CGH genomic labeling system, Invitrogen) using fluorochrome conjugated nucleotides (dUTP-alexa fluor 488 from Invitrogen or dUTP-cyanine 3 from Perkin Elmer). Before use, probes were denatured 5min at 80°C and then 30min at 37°C. Cells were grown for two days on glass slides in DMEM without phenol red containing 2.5% csFCS. After addition of $10^{-8}$M E2 or ethanol (vehicle), slides were washed with PBS, and then fixed in 2% paraformaldehyde (PFA) for 10 min at 4°C. PFA fixed cells were permeabilized in 0.5% Triton X-100 and equilibrated in 1X SSC for 5 min. Slides were incubated one hour with 20ug/mL RNAseA in 1X SCC at 37°C and then sequentially washed 3 times with PBS, incubated in 2% PFA for 10 min at room temperature, in HCl 100mM for 10 min and then in 0.5% Triton for 10 min, with 3 washings with PBS between each step. Slides were then subjected to denaturation through sequential heating at 73°C in 70% formamide/ 30% 2X SSC for 7 min and then 3 min in 50% formamide / 50% 2X SSC. Hybridization with 600ng of labelled denatured DNA probes was performed overnight at 42°C in hybridization buffer (per 800µL: 200µL 25% Dextran sulfate; 100µL 20X SSC; 500µL deionised formamid) containing 150µg of Cot-I (Invitrogen) and 150µg of Salmon Sperm DNA. Slides were rinsed three times in 2X SSC, in 50% formamide/50% 2X

MCB Accepts published online ahead of print

SSC 20 min at 42°C and three times in 2X SCC again. Nuclei were stained with DAPI in 2X SSC for 5 min and then slides were mounted in ProLong Gold antifade reagent (Invitrogen) with a 22*40 coverslip.

**Cytogenetic analysis.** MCF-7 and MDA cells were plated in Lab-TekTM chamber slides (Nunc Thermo Scientific) and observed daily until they reached a stage of active division. Cells were then harvested using a MultiPrep Genie 205 apparatus (Genial Genetics) according to the recommendations of the manufacturer. After R-banding, twenty metaphases were captured and analysed. Complementary analysis using fluorescence in situ hybridization (FISH) was carried out according to standard procedures as described in (27). Slides were analysed with an epifluorescent microscope Olympus BX61 and images were captured using Isis® software (MetaSystems).

**Microscopy and image analysis.** Stacks were obtained with a 63x oil immersion objective of a DMRXA (LEICA) microscope or a Zeiss apotome (63x objective). Measurements of nuclear area and distance between the centroid of each probe were performed under Image J (http://rsbweb.nih.gov/ij/). Distances were determined in 2D, since pilot experiments did not evidence any qualitative difference between 2D and 3D-FISH experiments (data not shown). Entire stacks were taken for all selected nuclei (non-mitotic and containing the expected 3 pairs of hybridization signals), and the three channels (red, green and blue) were isolated using the «DeInterleave» plugin. Pictures in *z* (distance of 0.3 µm) containing maximum red or green signal intensities were selected for all channels, merged and then tresholded to eliminate background from specific signals for distance measures. We used the DAPI (blue channel) pictures to consistently determine the nucleus area, calculated following the determination of a threshold fluorescence value corresponding to an entry transition into the nucleus. This value was manually determined as the inflection point of a profile plotting DAPI signal measured in a 10 pixel large longitudinal window crossing the nucleus against pixel distance. Images from up to 100 nuclei were analyzed in each experiment. Significant variations between experimental conditions were tested by a Fisher *t-test* comparison for unpaired data, with a significance threshold set for p-values ≤0.05. To calculate 3D volumes, we first segmented automatically the 3-dimensional hybridization signals for each of the color channel using the triangle algorithm (28) implemented in ImageJ. After a cleaning step consisting of the successive application of an opening and closing filter, the 3D-volumes of the structures resulting from the union of the two segmentation masks were measured and expressed as voxels. To analyse the kinetic FISH experiment, we had to develop a custom Matlab (MathWorks, Natick, Massachusetts) image processing routine in order to quantitatively analyze the high number of images. For that, we used the maximum intensity projections of the 3-dimensional stacks acquired by fluorescence microscopy. The analysis steps were performed automatically to avoid potential bias associated to manual intervention. The nuclei were segmented on the DAPI channel using the Otsu approach (29) and a watershed algorithm (30) was applied to separate touching nuclei. For detecting the fluorescence spots on the images corresponding to each FISH probe, we used the algorithm developed by Sbalzarini and Koumoutsakos (31). A first filtering step was performed to remove the very dim spots and those located outside the nuclei. The intensities of the remaining spots were estimated and

MCB Accepts published online ahead of print

substracted to the local background. We only kept the spots whose intensity was exceeding the threshold TI, which was calculated as follows: TI = <I> + n*σI , with <I> and σI, the mean and the standard deviation of the spot intensities, respectively, and n a user-defined integer. On a subset of images, we compared the spots detected by the automatic analysis and those selected by manual inspection. By setting n to 5, we optimized the matching between the automatic and manual selections. The intensity spots corresponding to the two FISH probes were paired based on a nearest neighbor criterion. FISH pair distances exceeding 20 pixels were not considered. We noted that our approach was robust towards the choice of the integer n as similar distance distributions were obtained for $3 < n < 7$.

**Microarray and mRNA profiling data analyses.** RNAs for microarray analysis were isolated from $20^7$ MCF7 or MDA::ER cells treated with E2 or ethanol as vehicle control using the RNeasy Plus Mini Kit (Qiagen) with homogenization through QIAShredders (Qiagen). Integrity and purity of the RNAs were controlled on an Agilent Bioanalyzer using the RNA 6000 Nano Assay (Agilent). Ten µg of selected samples exhibiting a RIN >9.8 were then subjected to cDNA synthesis using the Superscript Double-Stranded cDNA Synthesis Kit (Invitrogen) and a mix of 50 pmol random hexamers and 50 pmol of Oligo-dT. cDNAs were then treated for 10 min at 37°C with 5 µg RNaseA (Invitrogen), purified through a phenol:chloroform:isoamyl alcohol extraction on Phase Lock light gels (Fisher Scientific) and then precipitated. Following a quality control on agarose/BET gels, all subsequent steps (labeling of cDNA, hybridization and scanning of the arrays) were performed at the NimbleGen service facilities (Rejkyavic, Island). For each experimental condition, three arrays (NimbleGen Homo sapiens 385K Array) were hybridized with independently prepared pools of cDNAs synthesized from experimental triplicates (independent experimental and biological triplicates). Quantile normalization of the data through the RMA algorithm and all primary analyses were performed using the ArraySTar software suite (DNAstar, Inc.). Data were filtered according to two criteria: i) expression values greater than the first quartile in all samples of at least one triplicate; and ii) triplicate standard deviations lower than the third quartile in all triplicates. Experimental groups were compared by analysis of variances (*t*-test) and *p*-values were adjusted by the Benjamini and Hochberg method. Genes were considered as differentially expressed between two experimental conditions when their adjusted *p*-value was lower than 0.05 and their fold change greater than 1.8. Estrogen-sensitive clusters were then defined by sliding a window of variable sizes and counting the number of E2-regulated genes within these windows. The best empirically defined parameters were to define clusters as regions comprising at least 3 regulated genes within a window of 7 genes.

**FAIRE assays.** Formaldehyde-assisted isolation of regulatory elements [FAIRE; (32)] methodology was conducted as previously (23).

**Chromatin immunoprecipitation.** Cells were washed twice with PBS, and cross-linked during 10 min at room temperature using 1.5% formaldehyde (Sigma) diluted in PBS. Following a subsequent incubation with 0.125 M

MCB Accepts published online ahead of print

glycin for 2 min, the cells were collected in 1 ml collection buffer [100 mM Tris-HCl (pH 9.4) and100 mM DTT]. Cells were then washed in 1 ml PBS, lysed for 15 min at room temperature in 300 µl of lysis buffer [10 mM EDTA, 50 mM Tris-HCl (pH 8.0), 1% SDS, 0.5% Empigen BB (Sigma)], and chromatin sonicated during 14 min using a BioRuptor apparatus (Diagenode), with 30 sec on/off duty cycles. Chromatin was then cleared through a 10 min centrifugation at 10,000 x g. ChIP experiments were then conducted with some modifications from previous protocol (2) using a tenth of the chromatin samples (30 µl) of the supernatants as inputs, and the remainder diluted 5-fold in IP buffer [2 mM EDTA, 100 mM NaCl, 20 mM Tris-HCl (pH 8.1), and 0.5%Triton X-100]. A $1/4^{th}$ of this fraction was subjected to immunoprecipitations overnight after a 3 hr preclearing at 4°C with 10 µg yeast tRNA and 80 µl of a 50% protein A-Sepharose bead (Amersham Pharmacia Biosciences) slurry. Complexes were recovered after 3 hr incubation at 4°C with 5 µg yeast tRNA and 40 µl of protein A-Sepharose. Precipitates were then serially washed, using 300 µl of Washing Buffers (WB) I [2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 0.1% SDS, 1% Triton X-100, 150 mMNaCl], WB II [2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 0.1% SDS, 1% Triton X-100, 500 mM NaCl], WB III [1 mM EDTA, 10 mM Tris-HCl (pH 8.1), 1% NP-40, 1% Deoxycholate, 0.25 M LiCl] and then twice with 1 mM EDTA, 10 mM Tris-HCl (pH 8.1). Precipitated complexes were removed from the beads through two sequential incubations with 50 µl of 1% SDS, 0.1 M $NaHCO_3$. Cross-linking was reversed by an overnight incubation at 65°C. DNA was purified on NucleoSpin™ columns (Macherey-Nagel) using 50 µl NTB buffer. Subsequent qPCR analysis used 1 µl of input material and 3 µl of ChIP samples.

**ChIP-on-chip assays and analysis of published datasets.** The ChIP procedure conducted on chromatin prepared from $15.10^6$ cells was similar to the one described above with the following modifications for final steps. Crosslinking was reversed by an overnight incubation at 65°C with 10 µg of Proteinase K (Sigma). Following a subsequent incubation of the samples with 2.5 µg RNAse (Sigma) for 1h at 37°C, the DNA was then purified on NucleoSpin™ columns (Macherey-Nagel) using NTB buffer and eluted in 40 µl of elution buffer. Efficiency of the ChIP assay was then evaluated using qPCR positive and negative controls. Experimental input and ChIP triplicates were then pooled by precipitation, resuspended in 25 µl H2O, and divided in two for amplification using the WGA whole genome amplification kit (Sigma). Following a quality control step, the amplified material was pooled and sent to NimbleGen services (Rejkyavic, Island) for hybridization on custom 385K arrays. These arrays were conceived by spotting genomic regions containing clusters of E2-regulated genes as defined from MDA::ER and MCF-7 estrogen-sensitive transcriptomes, as well as regions containing individual control cell-specific estrogen-sensitive genes (*cf*. Table 2). ChIP-chip signals normalization and peak calling steps were performed using the MA2C algorithm (33) on raw data obtained from two arrays hybridized with DNA prepared in two independent experiments. MA2C parameters were: robust normalization (C=1) and peak calling for a minimum of 4 probes (maximum gap set at 400 bp) in a sliding window of 300 bp half-width. All binding sites determined in our ChIP-chip experiments were confirmed by independent ChIP-qPCR assays, except the CTCFBS identified in the TFF1 promoter which was found to be antibody- and experiment-dependent. This site was therefore not included in the statistical

MCB Accepts published online ahead of print

examinations of our data. Analysis of published Affymetrix tiling arrays data (ER and Pol II ChIP-chips performed in MCF-7 cells) were analyzed under MAT (34). All genomic annotations were performed using algorithms present within the cistrome web-platform [http://cistrome.dfci.harvard.edu/ap; (35)]. MCF-7 ChIP-seq data for CTCF and RAD21 (concatenation of fastq obtained in duplicate experiments) and corresponding input were extracted from the GSE25710 series and aligned onto indexed chromosomes of genome hg18 using bowtie-0.12.7 (36) with parameters -p 7, --best, -m 1, -n 1, --sam and –l 28. The .sam files were then converted to .wig files, using samtools-0.1.12a (37) and MACS-1.3.7.1 (38). To compare RAD21 ChIP-seq datasets obtained in vehicle and estradiol-treated cells, we adjusted the bias of diverging sequencing depths through a linear normalization (factor of 2.1) of signal intensities of a given .wig so as to be comparable to the .wig file with the highest sequencing depth. Peak calling was then performed as previously described (39).

**Triplex capture experiments.** Triplex capture assays were performed on transfected cells which were subsequently cross-linked by 2% formaldehyde and lysed by sonication as described above. Sonicated chromatin was then incubated for 4h with streptavidin-coated magnetic beads (Dynal) that were blocked with 10 µg/ml BSA and 10 µg/ml yeast tRNA for 1h. Captured DNA was eluted by two rounds of elution in 0.1% SDS, purified following digestion with proteinase K and RNaseA and analyzed by qPCR.

**Chromosome conformation capture (3C) and circular 3C (4C).** Methods were adapted from (40), and used the *Dpn*II 4-base cutter as an enzyme of choice. Following a 5 min centrifugation at 2,000 x g, aliquots of $2.10^6$ cross-linked cells were washed with 200 µl of 1X *Dpn*II buffer, and then lyzed overnight at 37°C in 200 µl of 1X *Dpn*II buffer containing 0.3 % SDS with shaking. Following 2 passages through a syringe needle, 400 µl of 1X *Dpn*II buffer were sequentially added in 4 times, and SDS was sequestered with 67 µl of 20 % Triton X-100 at 37°C for 1h. 50 µl of the reaction mixture were then kept as input fraction for digestion efficiency controls. 550 µl of the chromatin preparation were then digested overnight with 400 U *Dpn*II at 37°c with shaking in a total volume of 500 µl of digestion buffer containing protease inhibitors (Roche). An additional step of 6h digestion with 150 U *Dpn*II was then performed. 50 µl of the digested chromatin was then kept for digestion efficiency controls, whilst the remaining was kept at 4°C during this step. To control the digestion efficiency, both input and digested aliquots were incubated with 9.5% SDS for 20 min at 65°C. Cell fragments were then eliminated by centrifugation at 12,000 x g for 5 min. 117 µl of TE buffer were then added together with 5 µl of 10 µg/µl RNAse A (Sigma) and the mixture was incubated at 37°C for 30 min before the addition of 8 µl of 5M NaCl and 10 µl of 10 µg/µl of Proteinase K (Sigma). Cross- linking was then reversed overnight at 65°C, and DNA purified on Macherey-Nagel columns. qPCR were then performed on input and digested fractions to calculate the digestion efficiency as follows: E%= $[1.9^{(Ct_{Input}-Ct_{sample})^{test\ region}}/1.9^{(Ct_{input}-Ct_{sample})^{control\ region}}]*df*100$, where the dilution factor (df) was =0.98360 [(50/610)/(50/600)], and the control region a region that contains no *Dpn*II fragment. If this % was > 85%, the remaining digested chromatin was subjected to a final lysis step by addition of 108 µl of 10% SDS and incubation at

MCB Accepts published online ahead of print

65°C with shaking for 20 min. Two µg of digested chromatin, as evaluated from the amounts of DNA recovered in the digested fraction using a NanoDrop were then incubated for 1h at 37°C with 40 µl of 20 % Triton X-100 in a total volume of 694 µl of Ligation buffer (10 mM Tris-HCl pH=7.8; 0.1 µg/µl BSA and protease inhibitors). Ligations for 3C or 4C experiments were then performed at 16 °C with gentle agitation for 4h or 5 days respectively. For 3C, 10 µl of T4 DNA ligase were added in the reaction mixture together with 80 µl of its buffer, 8 µl of 10 µg/µl BSA and 8 µl of 100 mM ATP (total volume of 800 µl/DNA concentration=2.5 ng/µl). For the 4C samples, requiring a more diluted concentration of DNA, 906 µl of $H_20$ were first added and then 160 µl of 10X T4 DNA ligase buffer, 16 µl of 10 µg/µl BSA, 16 µl of 100 mM ATP and 15 µl of T4 DNA ligase (total volume of approx. 1800 ml). The ligation mix was replenished at days 2 and 4 with ATP (20 µl of 100 mM ATP). Following these 5 days of ligation, 4C samples were further incubated for 1.5h with 1 ul of T4 DNA ligase in order to ensure that the ligase fills any nicks in the circularized 3C products. The cross-linking of either 3C or 4C DNA products was then reversed overnight at 65°C following the addition of EDTA (1 mM final), NaCl (0.2 M final) and 10 or 20 µl of 10µg/µl proteinase K. Samples were then subjected to three successive phenol/chloroform/isoamylalcohol (25:24:1) extractions followed by a chloroform washing step, diluted 4 times in water and precipitated at -20°C for 2h by 2 volumes of isopropanol. Following centrifugation at 13, 000 x g, samples were then washed 3 times with 1 volume of 75% EtOH and resuspended in 50 µl of TE. 3C samples were then processed for analysis. In contrast, the circularized 3C products (4C) generated by the 5 day ligation were then purified from linear DNA by a combined digestion with 7 µl of exonuclease I and 2 µl of exonuclease III (New England BioLabs) in a total volume of 100 µl of 1X exonuclease I buffer. Circular DNA was then purified on Macherey-Nagel columns following a 25 min heating step at 85°C to inactivate the enzymes. Elution step was modified, by incubating the DNA bound on the columns with 50 µl of Tris-HCl (pH=7.8) containing 20 ng of ytRNA for 2h. 4C libraries were then amplified on a thermocycler using High-fidelity Taq polymerase (Invitrogen) using the following primers: TMPRSS2 5'-AACATAGTCCTCTTTGGCACA-3' and 5'-GTCAGTCTCGGGGAGGGACT-3'; RIPK4 5'-TTGGGAGCTTCCATCAAGAC-3' and 5'-GCTCCTTCATGGGTTCATTC-3'; TFF3 5'-GACCAGGGTGTGGTGTCC-3' and 5'-CAGCTCTGCTGAGGAGTACG-3'; TFF2 5'-CAGACCCTCATCCTCCAGAC-3' and 5'TATAAAGGCATCTGGCAATGTG-3'; TFF1 5'-GCTACATGGAAGGATTTGCTG-3' and 5'-CAGTGGAGATTATTGTCTCAGA-3'; TMPRSS3 5'-CATGGCTGCTCTGGGAAC-3' and 5'CCTCGGCTAAGGAGGTAGAG-3'; UBASH3A 5'-GTACGGCTTCCTGCCAAA CCGCTGCCATCTCTTCCT-3'. Amplification was made following a 30 sec denaturation step at 98°C as follows: (98°C 10 sec, 60°C 4 min, 68°C 5 min)x4 and (98°C 5 sec, 60°C 2 min, 68°C 5 min)x34 with a final incubation at 68°C for 10 min. In parallel of samples subjected to the whole procedure, additional aliquots of cells or of chromatin were processed to generate additional controls: minus cross-linking (entire procedure on cells not incubated with formaldehyde), minus ligase (entire procedure but with no ligase in the mix). These samples served to control the specificity of the ligation (opportunistic ligation by background proximity: minus cross-linking) and of the PCR (minus ligase). The relative frequencies of interactions detected by 3C were calculated as follows:

$$Freq = [1.9^{(Ct_{control\ region} - Ct_{test})} / \min(E\%_{test5'}; E\%_{test3'})] / (1.9^{(Ct_{control\ region} - Ct_{pos3Cctl})} / \min(E\%_{control}; E\%_{tpos3Cctl3'})],$$

MCB Accepts published online ahead of print

where the control region amplification (same as for the digestion controls) served to normalize over the inter-samples variations of DNA amounts, the pos3Ctl a region that was always ligating under any condition (ligation of two adjacent regions). All values were normalized for their availability for ligation, by taking into account the minimal (min) digestion efficiency measured for either extremities of the ligated product. As negative controls, we tested the minus ligase samples and assessed for two negative control interactions which were 1- the ligation of each fragments of interest with a fragment located within a *GAPDH* exon and 2- a negative region taken from our 4C screenings. We did not normalize the values obtained on test regions over those obtained in these negative controls as they were generally not producing any amplification signal. However, whenever this had to occur, the interaction detected with the test region was not taken into account. The 4C interactions were considered as being only qualitative, due to the amplification step present in the procedure. Calculations made to determine interacting regions were the same as for the 3C, except that in this case, the amounts of DNA were sufficient to allow further normalization over negative controls. In this case, the $Ct_{control\ region}$ used was the lower one (maximum interaction) obtained in the whole set of tested ligation-produced fragments. We considered values which were at least 2-fold higher than negative control regions as significant. Networks of long-distance chromatin interactions were generated under Cytoscape (41).

**Quantitative-PCRs (qPCR) and statistics.** All qPCRs used 1 µM of specific oligonucleotides (Sigma; sequences in Table S1 in the supplemental material) and were performed on BioRad MyiQ and BioRad CFX96 machines using BioRad iQ SYBR Green supermix with 50 rounds of amplification followed by determination of melting curves. Primers for RT-PCR were designed using the QuantPrime design tool (http://www.quantprime.de/) (42). Oligonucleotides for all other type of samples were designed under Primer3 (http://frodo.wi.mit.edu/primer3/). ChIP sample values were normalized in three steps: to inputs (ΔCt), then to control ChIP samples (beads alone or anti-H3 ChIP; ΔΔCt) and then to ΔΔCt values obtained on control DNA regions. FAIRE values were normalized to a positive control region (promoter of the *Rplp0* gene). Heatmaps of qPCR data were all generated under Mev (43), with values that were declared as significant from the control by Wilcoxon or Student t-tests (depending upon the number of values). To facilitate their reading, only values significantly differing from the control ones were included within the heatmaps.

**Microarray data**

Dataset were deposited at the NCBI's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/, (44)] under the GSE23850 and GSE32132 accession numbers (expression and ChIP-chip data, respectively).

**RESULTS**

**Cell-specific transcriptional regulations of the TFF cluster and ERBSs.** To consider processes governing the co-regulation of co-linear genes by E2, we first characterized the MDA::ER and MCF-7 repertoires of E2-sensitive genes

MCB Accepts published online ahead of print

by microarray analyzes. As shown within the Venn diagram in Fig. 1A, the estrogen-sensitive transcriptome of both cell-types was extremely divergent with only 19 genes in common after 4h of E2 treatment. This is consistent with previous studies (45-47). We next identified clusters of E2-regulated genes defined as regions comprising at least 3 regulated genes within a window of 7 genes. As a paradigm, we focused here on one cluster located within locus 21q22.3, previously identified in MCF-7 cells (48) and termed TFF (Fig. 1). This choice was guided by the facts that i) this region includes the prototypical estrogen-responsive gene *TFF1* and that ii) our transcriptome analysis indicated that this E2-sensitive region include different E2-sensitive genes in MCF-7 and MDA::ER cells, with only 4 genes regulated by E2 in MCF-7 and up to 7 in MDA::ER. This indicated that specific events may influence the transcriptional response of genes included in this genomic region, and thereby provided the opportunity to address these mechanisms. Note that we included the *TMPRSS2* gene into our definition of the estrogen-responsive TFF cluster, since it was found to be regulated by E2 in MDA::ER cells. We first performed RT-qPCR experiments to confirm these regulations. The results of these assays, illustrated within Fig. 1B showed *i*) a timely coordinated regulation of expected genes by E2; and *ii*) that ER is the main transcription factor responsible for their E2-responsiveness, since a 24h pre-treatment of the cells with the ER-targeting antiestrogen ICI abrogated E2 actions (Fig. 1B). Moreover, none of the tested genes were regulated in naïve MDA-MB231 cells (Fig. 1B), further confirming that the observed genes regulations by E2 strictly relied on the presence of ER.

These differing subsets of estrogen-sensitive genes between MCF-7 and MDA::ER cells might reflect the existence of different ER cistromes in the two cell lines. Hence, we characterized ER binding sites on this genomic region in MDA::ER cells. ER chromatin immunoprecipitation (ChIP) experiments were performed on chromatin prepared from MDA::ER cells treated for 50 min with $10^{-8}$M E2, with resulting samples hybridized on custom tiling DNA arrays designed to cover genomic regions containing E2-sensitive clusters (Table 2). These assays identified 17 ERBS within the TFF cluster in MDA::ER cells at a FDR<5% (Fig. 1C). When comparing these data to the ERBSs determined in MCF-7 cells by ChIP-chip technology [data from Carroll et al. (17)], only three binding sites were common to both cell lines (in green in Fig. 1C), including the comBS2 located within the *TFF1* promoter. This low overlap between MDA::ER and MCF-7 ERBSs is consistent with analyses made on the entirety of the genomic regions spotted on the arrays (Fig. 1D).

Independent anti-ER ChIP-qPCR assays confirmed the cell-specificity of these ERBSs. FAIRE (formaldehyde assisted isolation of regulatory elements) experiments further showed that MDA::ER specific ERBSs exhibited a condensed chromatin state in MCF-7 cells, and *vice-versa* (Fig. 1E). This observation was also made when evaluating the enrichment of these regions in canonical histone marks for enhancers (not shown). Thus, the chromatin condensation state of these genomic sequences confirmed their cell-specific function. Exception to this observation was made for the MCF-7 ERBS5 which also exhibited an “opened” conformation in MDA::ER and MDA-MB231 cells. This might illustrate the vicinity of this sequence with an annotated transcript. Recent data has shown that transcription can occur at enhancers (49-51). This prompted us to assess by ChIP-chip whether the RNA polymerase II (Pol II) was present on MDA::ER ERBSs. As shown within the Fig. 1F, we indeed found a general

MCB Accepts published online ahead of print

enrichment of MDA::ER ERBs in Pol II that further exhibited cell-specificity since it was not observed on sequences corresponding to MCF7 ERBSs. Concordantly, using MCF7 Pol II ChIP chip data from (17) in these analyzes also showed that the polymerase was enriched on MCF7 ERBSs but not MDA::ER ERBSs sequences (Fig. 1F).

In conclusion, these results altogether show that the studied 2 Mb genomic region that covers the TFF cluster includes different sets of estrogen-responsive genes and ERBSs in MCF-7 and MDA::ER cells. Interestingly, the cell-specific ERBSs were not predominantly found at the proximity of cell-specific E2-regulated genes. In accordance with the fact that chromatin loops may place the promoter of these cell-specific genes in the vicinity of cell-specific ERBSs (14), this latter observation suggested that the three-dimensional organization of the TFF loci may differ between the two cell lines.

**Dynamic spatial reorganization of the TFF cluster upon E2 treatment.** We next envisioned that the coordinated regulations occurring at the level of the studied genomic region in both MCF-7 and MDA::ER cells (Fig. 1) may involve an E2-dependent spatial reorganization of this locus. To test this hypothesis, we sought to perform DNA-FISH experiments. Noteworthy, MDA::ER cells present three fluorescent TFF loci and there were a higher number of loci (at least 6) in our MCF-7 cell line. This originates from the complex hypertriploid karyotype of these MCF-7 cells harboring 3 chromosomes 21 and multiple non assignable chromosome parts that contain at least 4 TFF loci (see Fig. S1 in the supplemental material). As this high number of loci hindered the correct evaluation of the experimental results, we focused on MDA::ER cells for these specific analyzes. We first conducted experiments aiming to evaluate the spatial volume occupied by the 2 Mb genomic region encompassing the TFF cluster and thus performed DNA-FISH with a mix of fluorescent probes generated from multiple BACs covering a large part of the TFF region (Fig. 2A). The results of these experiments led us to evidence a compaction of the chromatin domain containing the TFF cluster after a 50 min treatment with E2 (Fig. 2B). This could be assigned to a smaller dispersion of the distances separating central (B5) and 5' (B1) probes following exposure to E2 (Fig. 2C). Interestingly, whilst there were no changes in the distribution of distances between probes generated using the central and the 3' BACs (B6), the distribution of the distances separating probes located at both extremities of the TFF region were again significantly different. Furthermore, quantile-quantile representations of data (Q-Q plots, Fig. 2C, bottom) indicated that observed changes mostly reflected a disappearance of large distances separating paired probes. Importantly, these variations were not an indirect consequence of a global reduction in nucleus volume upon E2 exposure (not shown). The 2 Mb region containing the TFF cluster thus undergoes spatial rearrangements under E2 treatment in MDA::ER cells, reflected by a more constrained three-dimensional conformation.

To gain further insights into how the cell-specific transcriptional activity of the TFF cluster was spatially organized, we characterized the spatial proximity of ERBS with the promoters of E2-regulated genes. We conducted 4C-qPCR on chromatin prepared from MCF-7, MDA::ER as well as MDA-MB231 cells all treated for 50 min with E2. Results of these experiments (Fig. S2 and Fig.S3) are schematized in Fig. 3A. All determined spatial proximities are given within the Table S2. These 4C assays recovered all but one of the interactions previously determined by ChiA-

MCB Accepts published online ahead of print

PET (15) as linking ERBSs to regulated promoters in MCF-7 cells, and further uncovered 38 new interactions (Fig. 3B), in agreement with the differing outcome of 4C and ChiA-PET techniques. These experiments indicated that major spatial constraints of this chromatin domain involve interactions between central and distant regions mainly located 5' of the TFF cluster (approx. 70% of the ERBS-promoters interactions) in both MCF-7 and MDA::ER cells. This is consistent with results obtained in DNA-FISH experiments. Interestingly, the *RIPK4* promoter was situated in the spatial vicinity of more ERBSs in MDA::ER than in MCF-7 cells (Fig. 3A and Fig. 3C). This might be correlated with the estrogenic regulation of this gene in the former cell line but not in the latter (Fig. 1). However, there was no strict correlation between the numbers of ERBS-promoter interactions and the amplitude of the estrogenic regulation of the gene, as evaluated by the RT-qPCR results from Fig. 1 (maximum $R^2$ of 0.22 and 0.51 observed following 2h of E2 treatment, in MCF-7 and MDA::ER cells, respectively). Importantly, 27% of the interactions detected in MDA::ER were also detected in ER negative MDA-MB231. This indicates that some of the MDA::ER ERBSs may be pre-existing enhancers that recruit additional transcription factors regulating the activity of all considered genes, with the exception of *TFF3*, as it did not establish any contact with the tested regions in MDA-MB231 cells (Fig. 3A and Fig. 3C). Conversely, since the remaining 73% of the interactions detected in MDA::ER were not detected in MDA-MB231, these results also demonstrate that the expression of ER in this system is sufficient, either directly or not, to remodel the spatial organization of this genomic region.

As illustrated within Fig. 3D and Fig. 3E, the MCF-7 and MDA::ER network of interactions between ERBSs and promoters of regulated genes is complex. One striking difference between these interactomes detected by 4C is that there are more singleton interactions in MCF-7 than in MDA::ER. Indeed, while 6 ERBSs were interacting with one single gene in MCF-7, there was only one of these exclusive contacts in MDA::ER cells (the BS15/*TFF3* association). Finally, this interactome study also showed that the ERBS located within the *TFF1* promoter (comBS2) directs more interactions than the others in both cell lines, suggesting that it plays a crucial role in the chromatin re-organization of the TFF cluster in response to E2.

**CTCF and cohesin are required for appropriate regulations and organization of the TFF cluster.** ER has been proven to modulate the frequencies of interactions between distant ERBSs and promoters of either up- or down-regulated genes (15, 52), in concert with CTCF and/or the cohesin complex (53). Hence, to better understand how E2 signaling impacts on chromatin organization, we mapped CTCF and RAD21 (a subunit of the cohesin complex) binding sites in MDA::ER through ChIP-chip analysis (Fig. 4A). The comparison of all the MDA::ER CTCFBSs and RAD21BSs identified on all the genomic regions spotted on our arrays with those previously determined in MCF-7 cells (53; restricted here to BSs contained within the spotted regions) indicated that most CTCFBSs were conserved between the two cell types on the contrary to RAD21BSs (Fig. 4B). This cell-specific RAD21 cistrome may represent a major source for the different E2-responses of the clustered TFF genes. On the other hand, the overlap between the BSs for ER, CTCF and RAD21 was much more important in MCF-7 cells than in MDA::ER within the TFF cluster (Fig. 4C). These differing ER/CTCF/RAD21 BS overlaps between MCF-7 and MDA::ER were reflected in the

MCB Accepts published online ahead of print

proportions of established contacts between ERBS and genes promoters (Fig. 4D). These data also indicated that the presence of CTCF and/or RAD21 at ERBSs does not correlate with the number of interactions made with gene promoters in the context of the TFF cluster (Chi2 and Fisher tests *p*-values>0.4). In agreement with this observation, we did not observe any significant increase in CTCF and RAD21 binding on respective ERBSs following the addition of E2 (Fig. 4E). This may signify that the dynamic modulations of the TFF three-dimensional organization that occurs upon E2 treatment involve other RAD21/CTCF sites. Alternatively, the physical contacts established between the ERBSs and the promoters may be directed by ER and occurring within a structurally fixed frame imposed by CTCF and/or RAD21. In favor of the latter hypothesis, the alignment of RAD21 ChIP-seq reads obtained in MCF-7 [data from (53)] on ER, CTCF and RAD21 shared binding sites within the TFF cluster did not evidence any significant change in RAD21 mobilization on ER/CTCF/RAD21, the sole ER/RAD21 BS and on CTCF/RAD21 BSs (*p*-value 0.0676 for this latter category of sites) (Fig. 4F).

The putative role of CTCF and RAD21 in establishing the spatial conformation of the TFF cluster genomic domain was next examined following the transfection of siRNAs targeting their expression. Control RT-qPCR and Western blots performed in MDA::ER cells are shown in Fig. 5A and 5B, with similar reductions observed in MCF-7 cells (not shown). We first performed 3C-qPCR experiments to evaluate the impact of these siRNA-mediated reductions in CTCF and RAD21 intracellular amounts on the frequencies of interaction between RAD21 and/or CTCF positive ERBSs and the promoters of E2-regulated genes. These experiments, summarized in Fig. 5C showed that the silencing of RAD21 diminished the frequency of interactions between ERBSs and their target gene promoters in both MDA::ER and MCF-7 cells. This was also observed following the transfection of siRNAs targeting CTCF, although to a lesser extent (Fig. 5C). DNA-FISH experiments further showed that RAD21 is essential for the global E2-induced constraints exerted on the TFF three-dimensional conformation (Fig. 5D). Unfortunately, the involvement of CTCF could not be addressed here due to its inefficient depletion (≈25%) in this particular experimental setup. Finally, disrupting CTCF and RAD21 expression by siRNAs drastically reduced both basal and induced transcriptional activity of E2-regulated genes in both cell lines (Fig. 5E). This ultimately led to a strong decrease of their fold inductions by E2 except for *TFF3* and *RIPK4* in MDA::ER cells and *UBASH3A* in MCF-7 cells.

Altogether, these data indicate that cohesin and CTCF organize the E2-responsiveness of the genes included in the TFF cluster, in both cell lines, possibly by promoting a three-dimensional organization of the studied genomic locus which is propitious for the interaction between distant ERBS and promoters of activated genes.

**Dynamic 3D organization of the TFF cluster.** We questioned next whether one or a limited number of ERBSs within the TFF cluster could orchestrate the observed coordinated genes regulations through long-range interactions. This hypothesis would imply that one -or a few - ERBS is brought nearby promoters in a dynamic manner compatible with these transcriptional responses to E2. Hence, we evaluated by 3C-qPCR the dynamics of the interactions between ERBSs and genes promoters following treatment of the cells with E2. These data are summarized within Fig. 6 and Fig. 7; with circle areas being directly proportional to the fold changes in the relative frequency of

MCB Accepts published online ahead of print

interactions as compared to the initial situation (t0) (MCF-7 data are presented in Fig. S4). In x are the coordinates for the promoters of the genes and in y-axis are the coordinates for the ERBSs, whose positions are indicated. Importantly, all of the ERBS-promoter interactions were detected at t0 (MCF-7 controls in Fig. S4). In both cell types, a spatial reorganization was apparent as soon as 10 min following E2 addition with both increases and decreases in frequencies of interaction (Fig. 6 and Fig. 7). The existence of several TFF loci in both cell types may influence the interpretation of these kinetic 3C data, since we cannot ascertain that all of the interactions are occurring on the same genomic fragment. However, supporting the existence of dynamic variations in the three-dimensional organization of the TFF genomic region occur, we were able to further evidence such processes by DNA-FISH time-course experiments in MDA::ER cells (see Fig. S5). Some of the reorganizations evidenced by 3C-qPCR exhibited a relatively dynamic or even cyclical nature, such as those highlighted in orange within Fig. 6 and Fig. 7 subpanels. Interestingly, the dynamics of these spatial reorganizations differed between MDA::ER and MCF-7 cells. In MCF-7 cells, these variations were restrained to short-range interactions except for *TFF1*/BS1, *TFF1*/BS18 and *TMPRSS3*/BS8 contacts (Fig. 6). In contrast, the dynamic MDA::ER interactome highlighted important variations in long-range interactions between ERBSs and the promoters of regulated genes. This is notably illustrated by the interactions made by the ERBSs located in the 5' region of the cluster that climax at 50 min following the addition of E2 (Fig. 7). Furthermore, in MDA::ER cells, an apparent combination of local (BS2 and comBS1) and long-range interactions (comBS2 and even comBS3 located in the far 3' of the genomic region) could be correlated with the E2-mediated regulations of *TMPRSS2* and *RIPK4*. In contrast to what happens in MCF-7, the regulations of core TFF genes would thus be more influenced by distal than by local ERBSs in MDA::ER cells. Indeed, variations of the interactions made between ERBSs located within the *TFF3* to *TMPRSS3* (core TFF cluster) region appear less important than what was observed in the case of MCF-7 cells.

In conclusion, these data indicate that there is no single major ERBS that organizes E2 responsiveness within this genomic region. Keeping in mind that 3C-based assays have the intrinsic limitation to be unable to ascertain the co-occurrence of detected interactions in the same cell, we propose that the coordination of the transcriptional response of the TFF cluster to E2 mainly originates from a combinatorial engagement of ERBSs located within the *TFF1* promoter in MCF-7 cells with two nearby ones (BS14 or BS16), and with those located in the distant 5' region of the cluster in MDA::ER cells.

**E2 regulation of a given gene is driven by different ERBSs.** Although highly informative, the above 3C data did not allow us to establish the exact contribution of each ERBS toward the specific regulation of the genes that they contact. To investigate at the molecular level the contribution of given ERBS in regulating specific genes, we used small triplex forming oligonucleotides (TFOs) to interfere with ER binding at a given BS (Fig. 8A). Such oligonucleotides that form Hoogsteen or reverse Hoogsteen hydrogen bonds with the purine-rich strand of DNA have already been used to inhibit the transcription of genes such as *ets2* (54) or *c-myc* (55) [reviewed in (56, 57)]. We characterized 11 TFOs (Table 1) which were able *i*) to form DNA triplex *in vitro* (TFO anti-MCF-7 BS1 as example

MCB Accepts published online ahead of print

in Fig. 8B, otherwise see Fig. S6); *ii*) to specifically bind to target sequences but not *PKNOX1* promoter used here as a control (Fig. 8C and Fig. S6) and *iii*) to significantly disrupt ER binding on the corresponding ERBS (Fig. 8D and Fig. S6). Although all of the designed TFOs did not precisely target the center of each ERBS defined here as from the ChIP-chip profiles, we observed that their relative efficiency in disrupting ER binding was only mildly correlated to the distance separating the TFO target sequence to the center of the ERBS peak (Fig. 8E). The ability of TFOs to bind to their target sequence seemed also relatively independent of their chromatin status as evaluated from their relative enrichment by FAIRE (Fig. 8F).

Subsequent RT-qPCR experiments showed that in a majority of tested cases, decreasing ER binding on one ERBS did impact the transcriptional status of the genes to which it was spatially close (Fig. 8G), but not on control genes (Fig. S7). However, the converse was also observed, with no observable impact of ER binding disruption, as for instance the associations between the MCF-7 BS4 and *UBASH3A*, BS6 and *TMPRSS3* or MDA::ER BS10 and *UBASH3A.* This seems to imply that the binding of ER on some distant sites might not be essential for the regulation of the analyzed genes. Alternatively, this could also indicate a functional redundancy between the enhancers controlling the activity of the tested E2-sensitive genes. Abrogating ER binding on such site would be compensated for by the activity of the others. Additionally, the reduced ER mobilization provoked by TFOs diminished the frequencies of interactions linking the targeted ERBS with the promoters of their target genes (data not shown). Interestingly, these experiments also evidenced that decreasing ER binding on some sites affected the transcriptional status of genes that they did not contact. This was for instance observed in MCF-7 cells for the BS1 on *TFF3* levels, comBS1 on *TMPRSS2,* BS10 on *TFF2*, BS14 on *TMPRSS2* and *TFF3.* And this was also true in MDA::ER for the BS10 on *RIPK4* and BS12 on *TMPRSS2, RIPK4 and TFF3* amounts in MDA::ER. As shown by Fullwood et al. (15) in ChiA-PET assays, the three-dimensional organization of chromatin can place distant ERBSs in spatial vicinity. We therefore evaluated whether targeting the recruitment of ER on one ERBS by a specific TFO could reduce its mobilization on another ERBS (Fig. S8). Results from these experiments indicated that the observed transcriptional "collateral" effects were due -at least in part- to the establishment of additional interactions between ERBSs themselves. For instance, the reduction of *RIPK4* and *TFF3* expressions by the TFO targeting the ERBS12 in MDA::ER cells could reflect a contact made between this ERBS12 and the ERBS16 that controls these genes.

The use of TFOs therefore allowed us to demonstrate the functional relevance of the interactions linking ERBSs to E2-regulated promoters we have characterized. Although limited to the ERBSs on which TFOs were able to disrupt ER recruitment, these data illustrate that the comERBS2 located in the close vicinity of the *TFF1* promoter plays a central regulating role in both cell types. They also clearly indicate that the MDA::ER ERBS1 and MCF-7 ERBS6 play prominent roles in the transcriptional activity of the genes included in the TFF locus.

## DISCUSSION

We investigated here molecular processes allowing estradiol to co-regulate the transcriptional activity of genes clustered within a 2 Mb genomic region. Using a naturally E2-responsive breast cancer cell line (MCF-7) and a

MCB Accepts published online ahead of print

cellular system with a forced E2-sensitity (MDA::ER) we interrogated whether these mechanisms could be triggered by ER on its own. Despite different chromatin contexts, ER was found to drive tight regulations of the TFF cluster in both cell types, relayed by its mobilization on distinct genomic regions. Interestingly, the chromatin status of MDA::ER ERBSs in native MDA-MB231 cells indicated that a number of ERBSs were already exhibiting an opened conformation and characteristics of functional enhancers: enrichment in FAIRE experiments (Fig. 1), presence of marks for poised or active chromatin (unpublished observations) and spatial vicinity with the promoters of considered genes (Fig. 3). Preparation of chromatin to ER binding in MCF-7 cells involves the actions of the pioneer factor FOXA1 (17, 18) which is not expressed in MDA::ER cells. Accordingly, MCF-7 specific ERBSs are all but four FOXA1 positive. By contrast, ERBSs in MDA::ER were principally not located at sites bound by FOXA1 in MCF-7, except the common ERBS2 and ERBS3 and ERBS13. Hence, it may be that other factors act in a similar way than FOXA1 in the MDA system, or that ER acts on its own. Motifs analysis performed on the entire set of MDA::ER ERBSs identified in our ChIP-chip data indicated that in addition to ERE motifs, GATA sites were also significantly enriched (not shown). This is consistent with reports showing that factors of this family, and in particular GATA3 (47,58,59), are controlling ER activity.

4C and 3C assays allowed the description of dynamic interactomes linking ERBSs to promoters of genes. Some of the detected interactions, however, engaged promoters of genes that were not regulated by E2 in our RT-qPCR or transcriptomic data, such as *TMPRSS2*, *RIPK4* or *TFF3* in MCF-7 cells. It is possible that these genes actually exhibit rapid transcriptional responses to estrogen as those evidenced by global run-on assays [GRO-seq (49-51)] that we would have missed in our analyses for sensitivity and timeliness reasons. In contrast with what would have been expected from an artificial cellular model as compared to a more “natural” one like MCF-7 cells, we found that the ER-mediated three-dimensional re-organization of the TFF cluster response to E2 is more important and more intricate in MDA::ER cells. Indeed, there were more singleton interactions in MCF-7 cells and the dynamics of the MCF-7 interactome following treatment with E2 was apparently lower than in MDA::ER cells. It could be hypothesized that the chromatin three-dimensional structure of this whole genomic region is already prepared for a response to E2 in MCF-7 cells, in contrast to the reconstituted E2-sensitive cellular model provided by MDA::ER cells. If true, this implies that, in MDA::ER cells, ER on its own is able to provoke important three-dimensional remodeling of the TFF locus to finely tune the transcription of target genes. Alternatively, the differing level of ploidy of our model cells (3 TFF loci in MDA::ER cells *vs.* 6 in MCF-7) may also impact the interpretation of the differences observed between both cell types. For instance, the presence of inactive or E2-insensitive loci may hamper and reduce the variations observed in either cell line. Unfortunately, we were unable to ascertain by RNA-FISH that all loci were transcribed and regulated in these cells, presumably due to the small size of the TFF genes. However, E2-induced variations of the three-dimensional organization of the studied genomic region were observed in each of the 3 loci in MDA::ER cells. This suggests that all 3 loci may transcriptionally respond to E2 in this cell type.

The regulatory unit that integrates the TFF cluster is of ~1 and ~2 Mb in size in MCF-7 and MDA::ER cells,

MCB Accepts published online ahead of print

respectively. These dimensions are coherent with those defined for Topologically Associating Domains (TADs) from Hi-C and 5C data (60-62). Hence, the different number of estrogen-sensitive genes between MDA::ER and MCF-7 cells could characterize the existence of cell-specific TAD geometries and differing boundaries. CTCF and RAD21/cohesin have been proposed to delineate regions of correlated transcriptional regulations (63-65), even if their presence might not systematically reflect a demarcation between insulated gene domains (66). Our data extend observations made in MCF-7 cells regarding the involvement of CTCF and RAD21 in the establishment of key connections between distant ERBSs and regulated promoters of the TFF cluster (53, 67). We further showed that RAD21 is required for the proper folding of this genomic region and its response to estrogen. Studies in MCF-7 or mice liver (53, 68) suggested that the main part of RAD21 actions on tissue-specific expression or estrogenic regulations would be CTCF independent. Accordingly, the limited overlap of RAD21 binding sites in MCF-7 and MDA::ER cells indicates that the MDA::ER specific cistrome of RAD21 engages cell-specific functions. Whether the cell-specific interactomes between ERBSs and gene promoters and the cell-specific size of the putative TFF TAD are directly linked to the differing RAD21 cistromes between MCF-7 and MDA::ER cells still remains an open question. One possible way to address this problematic would be to define the chromatin loops established between RAD21 BSs by ChiA-PET experiments in both cell lines.

Contrasting with the slight increase reported at the genome-scale by others (50, 51, 53), RAD21 recruitment to ERBSs in the TFF cluster was not significantly affected by E2. Whether this situation reflects an exception or a general behavior for clustered genes may constitute an interesting point to pursue the analysis of RAD21 role in organizing chromatin domains. On the other hand, our observations are coherent with a model in which active chromatin compartments are organized through constitutive loci (63-65). The characterization of physical contacts between RAD21 BSs through ChiA-PET may help in defining the cohesin complexes that organize such compartments where dynamic contacts between enhancers and genes would occur [as proposed in (69)]. The nature of the mechanisms that underlie the E2-mediated remodeling of the TFF domain may therefore be directly or indirectly under the sole control of ER or of the estrogenic response of the genes. For instance, Mediator, a protein complex loaded on active promoters can establish physical contacts between gene and promoters (70). In accordance with models of proximity ligation proposed by Gavrilov and coworkers (69), the recruitment of multiple proteins provoked by ER on its sites and its affinity with components of the transcriptional machinery would stabilize interactions that occur otherwise. It can also be inferred from this hypothesis that mobilization dynamics of the proteins recruited by ER on chromatin (7, 8) may at least partly be responsible for the dynamic property of ERBS-promoters physical contacts. Such a process was evidenced in the case of the *CDKN1A* gene promoter placed under the transcriptional control of another nuclear receptor, VDR (vitamin D3 receptor) (71). Accordingly, our kinetic 3C dissection of the three-dimensional reorganization of the TFF cluster indicates that all interactions between ERBSs and gene promoters do exist already in the absence of E2. Hormone and ER binding would then have to be considered as signals that remodel pre-existing conformations; a conclusion that seems to emerge from recent Hi-C data, which compared the global organization of MCF-7 chromatin in the absence of hormone and

MCB Accepts published online ahead of print

following 50 min of treatment (72).

A general problematic that emerged from genome-wide studies is the evidence that the number of interactions made between the BSs of a transcription factor and genes promoters are not systematically reflected at the transcriptional level (73). The use of TFOs allowed us to give further indications on how binding sites for a TF are mobilized in space and time in order to regulate the transcription of its target genes. Although already hypothesized (14, 17, 74), we demonstrate here the validity of the concept of functional redundancy between ER-bound enhancers. Indeed, a single promoter can establish contacts with several ERBSs, and we demonstrated that the resulting buildout might be in some cases insensitive to a particular ERBS inactivation, providing robustness to the regulatory system. In other cases, the mobilization of ER on distinct master regulatory regions appears sufficient to provoke the transcriptional response of the gene. Comparing the enrichments of main *vs.* secondary ERBSs in particular histone modifications, ER, Pol II, CTCF or RAD21 proteins, or even their relative frequencies of interaction with their target promoters, did not evidence a particular segregating characteristic. The differential recruitment of FOXA1, PBX1, GATA proteins or yet uncharacterized cofactors may account for the specific use of a given ERBS or the collaborative use of several ERBSs. Alternatively, the existence of a particular master regulatory region may originate from a coincident propitious folding of local chromatin due to high-order organization and the presence of cognate DNA sequences mobilizing ER. Finally, our dynamic 3C experimental data taken together with results obtained following transfection of TFOs also indicate that the role of one redundant enhancer towards that of another one may depend or shift over time or upon experimental condition. The latter was for instance suggested from data comparing the recruitment of ER when activated by EGF and E2 (75).

Although limited by the target sequence requirements, the nucleic acid composition of 2 to 7% of the ERBSs identified by ChIP-seq appears compatible for the binding of TFOs. Extending the experimental workflow used in this study to this whole sub-population of ERBSs would give rise to a functionalized partial interactome that could greatly enhance our knowledge of the links existing between the functions of enhancers and the organization of the genome. Furthermore, as for ER, BSs for many transcription factors have been shown to be grouped around responsive genes in clusters of enhancers (76). Hence TFOs or modified TFOs with increased efficiencies such as PNAs (peptide nucleic acids) (77), LNAs (locked nucleic acids) (78) or bisLNAs (79) may constitute powerful molecular tools to assess for the generality of enhancer redundancy.

**ACKNOWLEDGMENTS**

This work was supported by the CNRS, the University of Rennes I, and benefited from grants of the ARC, the Ligue Contre le Cancer (Equipe Labellisée Ligue 2009), Région Bretagne (CREATE 4793) and the the Agence Nationale pour la Recherche (ANR-09-BLAN-0268-01). J.Q. was a recipient of a joint fellowship from the CNRS and the Region Bretagne. A.S was supported by a fellowship from the French Ministère de l'enseignement supérieur et de la recherche, and S.C. was funded by a CNRS post-doctoral fellowship. We greatly acknowledge Dr. Denis Habauzit for all his advices and help on the design and analysis of TFOs.

MCB Accepts published online ahead of print

**REFERENCES**


1. **Harmston N1, Lenhard B.** 2013. Chromatin and epigenetic features of long-range gene regulation. Nucleic Acids Res. 41:7185-7199.
2. **Razin SV, Gavrilov AA, Ioudinkova ES, Iarovaia OV.** 2013. Communication of genome regulatory elements in a folded chromosome. FEBS Lett. **587:**1840-1847.
3. **Wei Z, Huang D, Gao F, Chang WH, An W, Coetzee GA, Wang K, Lu W.** 2013. Biological implications and regulatory mechanisms of long-range chromosomal interactions. J Biol Chem. 288:22369-22377.
4. **Cavalli G, Misteli T.** 2013. Functional implications of genome topology. Nat Struct Mol Biol. 20:290-299.
5. **Van Bortle K, Corces VG.** 2013. The role of chromatin insulators in nuclear architecture and genome function. Curr Opin Genet Dev. 23:212-218.
6. **McEwan IJ.** 2009. Nuclear receptors: one big family. Methods Mol. Biol. **505:**3-18.
7. **Métivier R, Penot G, Hubner MR, Reid G, Brand H, Kos M, Gannon F.** 2003. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. Cell **115:**751-763.
8. **Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M.** 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell **103:**843-852.
9. **van Driel R, Fransz PF, Verschure PJ**. 2003. The eukaryotic genome: a system regulated at different hierarchical levels. J. Cell Sci. **116:**4067-4075.
10. **Hurst LD, Pal C, Lercher MJ.** 2004. The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. **5:**299-310.
11. **Li Q, Barkess G, Qian H**. 2006. Chromatin looping and the probability of transcription. Trends Genet. **22:**197-202.
12. **Carvajal JJ, Keith A, Rigby PW.** 2008. Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. Genes Dev. **22:**265-276.
13. **Tang Y, Huang Y, Shen W, Liu G, Wang Z, Tang XB, Feng DX, Liu DP, Liang CC.** 2008. Cluster specific regulation pattern of upstream regulatory elements in human alpha- and beta-globin gene clusters. Exp. Cell Res. **314:**115-122.
14. **Carroll JS, Brown M**. 2006. Estrogen receptor target gene: an evolving concept. Mol. Endocrinol. **20:**1707-1714.
15. **Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y.** 2009. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature **462:**58-64.
16. **Zaret KS, Carroll JS.** 2011. Pioneer transcription factors: establishing competence for gene expression. Genes

MCB Accepts published online ahead of print

Dev. **25:**2227-2241.

17. **Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M.** 2006. Genome-wide analysis of estrogen receptor binding sites. Nat. Genet. **38:**1289-1297.
18. **Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS.** 2011 FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat. Genet. **43:**27-33.
19. **Tan SK, Lin ZH, Chang CW, Varang V, Chng KR, Pan YF, Yong EL, Sung WK, Cheung E.** 2011. AP-2γ regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. EMBO J. **30:**2569-2581.
20. **Magnani L, Ballantyne EB, Zhang X, Lupien M.** 2011. PBX1 genomic pioneer function drives ERα signaling underlying progression in breast cancer. PLoS Genet. **7:**e1002368.
21. **Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M.** 2008. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell **132:**958-970.
22. **Magnani L, Eeckhoute J, Lupien M.** 2011. Pioneer factors: directing transcriptional regulators within the chromatin environment. Trends Genet. 27:465-474.
23. **Serandour AA, Avner S, Percevault F, Demay F, Bizot M, Lucchetti-Miganeh C, Barloy-Hubler F, Brown M, Lupien M, Métivier R, Salbert G, Eeckhoute J.** 2011. Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. Genome Res. **21:**555-565.
24. **Reid G, Hubner MR, Metivier R, Brand H, Denger S, Manu D, Beaudouin J, Ellenberg J, 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.
25. **Watrin E, Legagneux V.** 2005. Contribution of hCAP-D2, a non-SMC subunit of condensin I, to chromosome and chromosomal protein dynamics during mitosis. Mol. Cell Biol. **25:**740-750.
26. **Vekhoff P, Ceccaldi A, Polverari D, Pylouster J, Pisano C, Arimondo PB.** 2008. Triplex formation on DNA targets: how to choose the oligonucleotide. Biochemistry **47:**12277-12289.
27. **Eimer S, Dugay F, Airiau K, Avril T, Quillien V, Belaud-Rotureau MA, Belloc F.** 2012. Cyclopamine cooperates with EGFR inhibition to deplete stem-like cancer cells in glioblastoma-derived spheroid cultures. Neuro. Oncol. **14:**1441-1451.
28. **Zack GW, Rogers WE, Latt SA.** 1977. Automatic measurement of sister chromatid exchange frequency. J. Histochem. Cytochem. **25:**741–753.
29. **Otsu N**. 1979. A threshold selection method from gray-level histograms. IEEE Trans Sys. Man. Cyber. **9:**62-66.
30. **Meyer, F.** 1994. Topographic distance and watershed lines. Signal Processing. **38:**113-125.
31. **Sbalzarini IF, Koumoutsakos P.** 2005. Feature point tracking and trajectory analysis for video imaging in cell biology. J. Struct. Biol. **151:**182-195.
32. **Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD**. 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory

MCB Accepts published online ahead of print

730 Elements) isolates active regulatory elements from human chromatin. Genome Res. **17:**877-885.
731 33. **Song JS, Johnson WE, Zhu X, Zhang X, Li W, Manrai AK, Liu JS, Chen R, Liu XS.** 2007. Model-based analysis of
732 two-color arrays (MA2C). Genome Biol. **8:**R178.
733 34. **Johnson WE, Li W, Meyer CA, Gottardo R, Carroll JS, Brown M, Liu XS.** 2006. Model-based analysis of tiling-
734 arrays for ChIP-chip. Proc. Natl. Acad. Sci. U.S.A. **103:**12457-12462.
735 35. **Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, Veeraraghavan N, Albert I, Miller W, Makova**
736 **KD, Hardison RC, Nekrutenko A.** 2007. A framework for collaborative analysis of ENCODE data: making large-
737 scale analyses biologist-friendly. Genome Res. **17:**960-964.
738 36. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA
739 sequences to the human genome. Genome Biol. **10:**R25.
740 37. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome**
741 **Project Data Processing Subgroup**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics
742 **25:**2078-2079.
743 38. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu**
744 **XS.** 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. **9:**R137.
745 39. **Sérandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, Palierne G, Gheeraert C, Barloy-**
746 **Hubler F, Péron CL, Madigou T, Durand E, Froguel P, Staels B, Lefebvre P, Métivier R, Eeckhoute J, Salbert G.**
747 2012. Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers.
748 Nucleic Acids Res. **40:**8255-8265.
749 40. **Gondor A, Rougier C, Ohlsson R.** 2008. High-resolution circular chromosome conformation capture assay. Nat.
750 Protoc. **3:**303-313.
751 41. **Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T.** 2003.
752 Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res.
753 **13:**2498-2504.
754 42. **Arvidsson S, Kwasniewski M, Riano-Pachon DM, Mueller-Roeber B.** 2008. QuantPrime--a flexible tool for
755 reliable high-throughput primer design for quantitative PCR. BMC Bioinformatics **9:**465.
756 43. **Saeed AI, Sharov V, White J, Li J, Liang W, hagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A,**
757 **Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V,**
758 **Quackenbush J.** 2003. TM4: a free, open-source system for microarray data management and analysis.
759 Biotechniques **34:**374-378.
760 44. **Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M,**
761 **Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Edgar R.** 2009. NCBI GEO: archive for high-throughput
762 functional genomic data. Nucleic Acids Res. **37:**D885-890.
763 45. **Moggs JG, Murphy TC, Lim FL, Moore DJ, Stuckey R, Antrobus K, Kimber I, Orphanides G.** 2005. Anti-
764 proliferative effect of estrogen in breast cancer cells that re-express ERalpha is mediated by aberrant

MCB Accepts published online ahead of print

765 regulation of cell cycle genes. J. Mol. Endocrinol. **34**:535-551.
766 46. **Nott SL, Huang Y, Li X, Fluharty BR, Qiu X, Welshons WV, Yeh S, Muyan M.** 2009. Genomic responses from
767 the estrogen-responsive element-dependent signaling pathway mediated by estrogen receptor alpha are
768 required to elicit cellular alterations. J. Biol. Chem. **284**:15277-15288.
769 47. **Kong SL, Li G, Loh SL, Sung W-K, Liu ET.** 2011. Cellular reprogramming by the conjoint action of ERα, FOXA1,
770 and GATA3 to a ligand-inducible growth state. Mol Syst Biol **7**: 526. *doi: 10.1038/msb.2011.59.*
771 48. **Chinery R, Williamson J, Poulsom R**. 1996. The gene encoding human intestinal trefoil factor (TFF3) is located
772 on chromosome 21q22.3 clustered with other members of the trefoil peptide family. Genomics **32:**281-284.
773 49. **Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL.** 2011. A rapid, extensive, and transient
774 transcriptional response to estrogen signaling in breast cancer cells. Cell **145:**622-634.
775 50. **Hah N, Murakami S, Nagari A, Danko C, Kraus WL.** 2013. Enhancer Transcripts Mark Active Estrogen Receptor
776 Binding Sites. Genome Res. doi: 10.1101/gr.152306.112.
777 51. **Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass**
778 **CK, Rosenfeld MG.** 2013. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional
779 activation. Nature. **498:**516-520.
780 52. **Hsu PY, Hsu HK, Singer GA, Yan PS, Rodriguez BA, Liu JC, Weng YI, Deatherage DE, Chen Z, Pereira JS, Lopez**
781 **R, Russo J, Wang Q, Lamartiniere CA, Nephew KP, Huang TH.** 2010. Estrogen-mediated epigenetic repression
782 of large chromosomal regions through DNA looping. Genome Res **20:**733-744.
783 53. **Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT.** 2010. A CTCF-
784 independent role for cohesin in tissue-specific transcription. Genome Res. **20:**578-588.
785 54. **McGuffie EM, Catapano CV.** 2002. Design of a novel triple helix-forming oligodeoxyribonucleotide directed to
786 the major promoter of the c-myc gene. Nucleic Acids Res. **30:**2701-2709.
787 55. **Jain A, Magistri M, Napoli S, Carbone GM, Catapano CV.** 2010. Mechanisms of triplex DNA-mediated
788 inhibition of transcription initiation in cells. Biochimie **92:**317-320.
789 56. **Knauert MP, Glazer PM.** 2001. Triplex forming oligonucleotides: sequence-specific tools for gene targeting.
790 Hum. Mol. Genet. **10:**2243-2251.
791 57. **Kalish JM, Glazer PM.** 2005. Targeted Genome Modification via Triple Helix Formation. Ann. N.Y. Acad. Sci.
792 **1058:**151-161.
793 58. **Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M.** 2007. Positive cross-regulatory loop ties
794 GATA-3 to estrogen receptor alpha expression in breast cancer. Cancer Res. **67:**6477-6483.
795 59. **Theodorou V, Stark R, Menon S, Carroll JS.** 2013. GATA3 acts upstream of FOXA1 in mediating ESR1 binding
796 by shaping enhancer accessibility. Genome Res. 23, 12-22.
797 60. **Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J,**
798 **Gribnau J, Barillot E, Blüthgen N, Dekker J, Heard E.** 2012. Spatial partitioning of the regulatory landscape of
799 the X-inactivation centre. Nature **485:**381-385

MCB Accepts published online ahead of print


61. **Gibcus JH, Dekker J.** 2013. The hierarchy of the 3D genome. Mol. Cell **49:**773-782.
62. **Nora EP, Dekker J, Heard E.** 2013. Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? Bioessays. **35:**818-828.
63. **Schaaf CA, Kwak H, Koenig A, Misulovin Z, Gohara DW, Watson A, Zhou Y, Lis JT, Dorsett D.** 2013. Genome-wide control of RNA polymerase II activity by cohesin. **9:**e1003382. doi: 10.1371/journal.pgen.1003382.
64. **Li Y, Huang W, Niu L, Umbach DM, Covo S, Li L.** 2013. Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. BMC Genomics **14:**553. doi:10.1186/1471-2164-14-553
65. **DeMare LE, Leng J, Cotney J, Reilly SK, Yin J, Sarro R, Noonan JP.** 2013. The genomic landscape of cohesin-associated chromatin interactions. Genome Res. **23:**1224-1234.
66. **Sanyal A, Lajoie BR, Jain G, Dekker J.** 2012. The long-range interaction landscape of gene promoters. Nature. **489:**109-113.
67. **Zhang Y, Liang J, Li Y, Xuan C, Wang F, Wang D, Shi L, Zhang D, Shang Y.** 2010. CCCTC-binding factor acts upstream of FOXA1 and demarcates the genomic response to estrogen. J. Biol. Chem. **285:**28604-28613.
68. **Faure AJ, Schmidt D, Watt S, Schwalie PC, Wilson MD, , Xu H, Ramsay RG, Odom DT, Flicek P.** 2013. Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. Genome Res. **22:**2163-2175.
69. **Gavrilov AA, Gushchanskaya ES, Strelkova O, Zhironkina O, Kireev II, Iarovaia OV, Razin SV.** 2013. Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. Nucleic Acids Res. **41:**3563-3575.
70. **Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA.** 2010. Mediator and cohesin connect gene expression and chromatin architecture. Nature **467:**430-435.
71. **Saramäki A, Diermeier S, Kellner R, Laitinen H, Vaïsänen S, Carlberg C.** 2009. Cyclical chromatin looping and transcription factor association on the regulatory regions of the p21 (CDKN1A) gene in response to 1alpha,25-dihydroxyvitamin D3. J Biol Chem. 284:8073-8082.
72. **Wang J, Lan X, Hsu PY, Hsu HK, Huang K, Parvin J, Huang TH, Jin VX.** 2013. Genome-wide analysis uncovers high frequency, strong differential chromosomal interactions and their associated epigenetic patterns in E2-mediated gene regulation. BMC Genomics. 14: 70. doi: 10.1186/1471-2164-14-70.
73. **Eeckhoute J, Métivier R, Salbert G.** 2009. Defining specificity of transcription factor regulatory activities. J. Cell Sci. **122:**4027-4034.
74. **Attanasio C, Nord AS, Zhu Y, Blow MJ, Li Z, Liberton DK, Morrison H, Plajzer-Frick I, Holt A, Hosseini R, Phouanenavong S, Akiyama JA, Shoukry M, Afzal V, Rubin EM, FitzPatrick DR, Ren B, Hallgrímsson B, Pennacchio LA, Visel A.** 2013. Science 342:1241006. doi: 10.1126/science.1241006.
75. **Lupien M, Meyer CA, Bailey ST, Eeckhoute J, Cook J, Westerling T, Zhang X, Carroll JS, Rhodes DR, Liu XS,**

MCB Accepts published online ahead of print

**Brown M.** 2010. Growth factor stimulation induces a distinct ER(alpha) cistrome underlying breast cancer endocrine resistance. Genes Dev. **24:**2219-2227.

76. **Gotea V, Visel A, Westlund JM, Nobrega MA, Pennacchio LA, Ovcharenko I.** 2010. Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. Genome Res. **20:**565-577.
77. **Nielsen PE.** 2010. Gene targeting and expression modulation by peptide nucleic acids (PNA). Curr. Pharm. Des. **16:**3118-3123.
78. **Brunet E, Corgnali M, Cannata F, Perrouault L, Giovannangeli C.** 2006. Targeting chromosomal sites with locked nucleic acid-modified triplex-forming oligonucleotides: study of efficiency dependence on DNA nuclear environment. Nucleic Acids Res. **34:**4546-4553.
79. **Moreno PM, Geny S, Pabon YV, Bergquist H, Zaghloul EM, Rocha CS, Oprea II, Bestas B, Andaloussi SE, Jørgensen PT, Pedersen EB, Lundin KE, Zain R, Wengel J, Smith CI.** 2013. Development of bis-locked nucleic acid (bisLNA) oligonucleotides for efficient invasion of supercoiled duplex DNA. Nucleic Acids Res. **41:**3257-3273.
80. **Nicol JW, Helt GA, Blanchard SG Jr, Raja A, Loraine AE.** 2009. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics **25**: 2730-2731.

MCB Accepts published online ahead of print

852 **FIGURE LEGENDS**

853 **FIG 1** Cell-specific E2-sensitive genes and ER binding sites in a 2 Mb genomic region including the TFF locus. (A)
854 Venn diagram illustrating the overlap of identified E2-sensitive genes in MDA::ER and MCF-7 cells. (B) Heatmap
855 representation of RT-qPCR results obtained on RNAs prepared from MCF-7, MDA::ER and MDA-MB231 cells treated
856 for the indicated times with $10^{-8}$ M E2 and pre-treated for 36h with $10^{-6}$ M ICI164,384 where precised. Results are
857 the log2 of the fold inductions of gene expression levels by E2 obtained in two independent triplicate experiments.
858 (C) Integrated genome browser [IGB; (80)] illustration of the studied genomic region with RefSeq genes indicated.
859 ER binding signal obtained in an ER ChIP-chip analysis performed using chromatin of MDA::ER cells treated for 50
860 min with E2 is depicted in gray. MCF-7 data were obtained from published dataset (17). For the sake of clarity, only
861 the highest 5% signals are shown. Grey and red boxes delineate cell-specific ER binding sites (ERBSs) whilst
862 common ERBSs are in green. (D) Anti ER ChIP and FAIRE assays were conducted using chromatin prepared from
863 MCF-7, MDA-MB231 or MDA::ER cells treated with E2 or ethanol (EtOH) as vehicle control for 50 min. Results
864 shown within heatmaps are means from 6 to 9 values obtained in independent triplicate experiments. Values are
865 fold enrichments over control samples and a negative control region (promoter of the transcriptionally active *Rplp0*
866 gene). (E) Overlap of MDA::ER ERBSs with MCF-7 ones on regions spotted on the arrays. (F) Enrichment signals
867 obtained for anti-RNA Polymerase II (Pol II) ChIP-chip experiments performed in MDA::ER cells treated for 50 min
868 with E2 [MCF-7 data from (17)] were aligned on MCF-7 or MDA::ER ERBS identified within the regions spotted on
869 the arrays and located more than 10 Kbp away from the TSS of any annotated gene.
870
871 **FIG 2** E2 provokes a three-dimensional reorganization of the TFF cluster in MDA::ER cells. (A) Locations of the BACs
872 (B1 to B6 referring respectively to RP11-814F13, CTD-2337B13, RP11-35C4, CTD-260o11, RP11-113F1 and CTD-
873 1033M14) used to generate fluorescent probes for FISH experiments all along the genomic region of interest,
874 illustrated as in Figure 1. Positions of the ERBSs are also indicated. (B) Analysis of the 3D volume of the TFF loci by
875 DNA-FISH using a mix of all generated probes in cells treated with $10^{-8}$ M E2 or ethanol (EtOH) for 50 min.
876 Representative pictures of these assays are on the left side of the panel, with magnified views of the three TFF loci
877 present in MDA:ER cells. Quantitative measurements (n=416) of the 3D volume of the loci is shown at the right side
878 of the panel with indicated Fisher *t*-test *p*-value. (C) Quantitative analysis of the distribution of distances separating
879 indicated paired FISH probes. Values are shown within boxplots (top of the panel) or within quantile-quantile (Q-Q)
880 plots representation of conditionally ranked measured distances (bottom). The normal distribution expected from
881 non-variating distances is illustrated by the straight black line in each Q-Q plots. Non-parametric Fisher test *p*-
882 values are indicated when determined as significant, as well as the number of measurements made in at least three
883 independent experiments.
884
885 **FIG 3** ERBS-promoters interactomes. (A) IGB visualization of MCF-7, MDA-MB231 and MDA::ER interactomes linking
886 *Dpn*II fragments encompassing ERBS or promoter regions of indicated genes, as detected by 4C-qPCR on chromatin

MCB Accepts published online ahead of print

prepared from cells treated for 50 min with $10^{-8}$ M E2. Shown are RefSeq genes coordinates along chr21, as well as the positions of ERBS with grey, red and green boxes delineating MDA::ER, MCF-7 or common ERBS. 4C data are represented as lines linking one ERBS to its target promoters. MCF-7 interactions are in red, MDA::ER ones in grey and those shared between MDA::ER and MDA-MB231 in orange. (B) Venn diagrams depicting the overlapping interactions characterized in this study and to those identified in published ER ChiA-PET dataset (15) restricted to loops involving the gene promoters that served here as 4C baits. (C) Stacked histograms illustrating the overlap of the 4C-detected interactions for each tested promoters in the different cell lines. (D and E) Cytoscape (41) circular layouts of the networks of interactions that link E2-regulated genes to ERBS in MCF-7 (D) and in MDA::ER (E) cells. The sizes of the nodes are directly related to the number of interactions they direct.

**FIG 4** CTCF and cohesin recruitment on ERBSs. (A) MDA::ER anti-ER, CTCF or RAD21 ChIP-Seq or ChIP-chip signals visualized under IGB as in Figure 1. All data were obtained from cells treated with $10^{-8}$ M E2 for 50 min. (B) Overlaps on array-spotted regions between RAD21 or CTCF binding sites determined in this study in MDA::ER cells and those previously determined in MCF-7 (53). (C) Overlap of RAD21 and CTCF binding sites with ERBSs. (D) Repartition of CTCF and RAD21 BS within the interactions between ERBS and gene promoters identified by 4C. (E) CTCF and RAD21 ChIP-qPCR were performed on chromatin prepared from MCF-7 or MDA::ER cells treated for 4h with $10^{-8}$M E2 or ethanol (EtOH) as vehicle control. Fold enrichment of the precipitated proteins on tested sequences was normalized over control ChIP and negative region (*PKNOX1* promoter) values. Data shown are mean values ± SD obtained in three independent triplicate experiments. The line depicts the significance threshold applied (Fold enrichment>2). (F) Boxplots of MACS normalized reads of RAD21 ChIP-seq experiments performed in the absence or presence of E2 in MCF-7 cells [dataset from (53)]. Values shown are mean counts measured in a 500 bp window centered on indicated binding sites located within the 2 Mb genome region studied. Calculations were also made on random sites of similar mean size and in equal number than the shared CTCF/RAD21 BSs.

**FIG 5** CTCF and cohesin organize the E2-responsiveness of the TFF cluster. (A) Amounts of indicated mRNA were evaluated in untreated MDA::ER cells following 72h of transfection with either control siRNA (directed against the Luciferase gene, siLuc) or siRNAs targeting mRNAs of interest. Results shown are mean data ± SEM of values obtained in at least 5 different experiments (n ranges from 15 to 18), normalized to the expression of control *PKNOX1* gene. Wilcoxon test was used to identify statistically relevant variations from control siLuc/EtOH condition: p-value <0.001 (***), <0.01 (**). (B) Western blots assaying the expression of RAD21, CTCF, ER and hCAPD2 (loading control) in cells transfected with the indicated siRNAs. (C) 3C-qPCR experiments following the impact of the reduction in CTCF and RAD21 contents on the interactions between indicated ERBS and gene promoters. Results shown originate from one representative experiment out of two and are expressed as the log2 of values normalized to those obtained in the control siLuc condition. (D) Quantitative DNA-FISH analysis of the distribution of distances separating paired fluorescent probes in transfected MDA::ER cells. Results are illustrated as in Figure 3.

MCB Accepts published online ahead of print

(E) Heatmaps summarizing RT-qPCR results obtained in at least 3 independent triplicate experiments following the transfection of control siLuc or siRNAs directed against CTCF or RAD21 expression and a 4h treatment of the cells with $10^{-8}$ M E2 or EtOH as vehicle control. Expression levels were normalized to the *PKNOX1* internal control and reported to those calculated for the siLuc EtOH condition. The fold inductions of these genes expression by E2 are also shown on the right side of the panel as relative to those measured in cells transfected by the siLuc. Results originate from at least 4 independent triplicate experiments.

**FIG 6** Dynamic three-dimensional reorganization of the studied genomic region in MCF-7 cells. Summary of one 3C experiment representative from two, performed on chromatin sampled from MCF-7 cells treated from 0 to 80 minutes with $10^{-8}$ M E2. As indicated, the size of the bubble that corresponds to one interaction is proportional to the fold changes in frequencies of interaction as compared to the basal (t0) situation. The location of the gene promoters that served as anchors is illustrated on the top of each subpanel and the ERBSs on the left. Distance scale is accurate (2 Mb between ticks) but had to be broken in some instances for sake of figure size and clarity. Bubbles highlighted in orange are those commented in the main text and those in yellow correspond to interactions made by the ERBS located within the *TFF1* promoter (comERBS2).

**FIG 7** Dynamic three-dimensional reorganization of the studied genomic region in MDA::ER cells. Summary of one 3C experiment representative from two, performed on chromatin sampled from MDA::ER cells treated from 0 to 80 minutes with $10^{-8}$ M E2. Results are illustrated as in Figure 6.

**FIG 8** Functionalization of MDA::ER and MCF-7 interactomes. (A) Triplex forming oligonucleotides (TFOs) were designed to interfere with ER binding and thus to identify the roles of ER on specific BSs for the regulation of E2-sensitive genes. (B) Formation of DNA triplex as analyzed by gel-shift. Increasing amounts of TFO (25 to 1,500 pmol) were added to 25 pmoles of target DNA duplexes and incubated for 16 h at 37 °C. Control was made using an unspecific TFO (Ctrl TFO) at the highest concentration. Complexes were separated by electrophoresis and stained with methylene blue. (C) MCF-7 cells were transfected for 36 h with 10 µmol of TFO or biotinylated (Biot-) TFO directed against the ERBS1, subjected to cross-linking and sonicated chromatin was then incubated with streptavidin-coated beads. Amounts of captured DNA were analyzed by qPCR. Values are mean ± SD of two independent duplicates, and are expressed as % of captured DNA relative to input DNA normalized to the amounts of recovered negative control region (*Rplp0* promoter). (D) Anti-ER ChIP-qPCR performed on MCF-7 cells transfected as previously and treated for 50 min with $10^{-8}$ M E2. Results are mean values ± SD of three independent triplicates expressed as relative enrichment towards the *PKNOX1* promoter. (E) Fold changes in ER mobilization measured by ChIP-qPCR on each tested ERBS following the transfection of corresponding TFO was plotted against the distance separating the sequence targeted by the TFO from the center of the ERBS defined from ChIP-chip data. (F) Amounts of streptavidin-captured DNA following the transfection of Biot-TFOs are plotted against the relative

MCB Accepts published online ahead of print

957 chromatin accessibility of their target regions as measured by FAIRE experiments in control conditions (ethanol
958 vehicle control, EtOH). (G) RT-qPCR experiments performed on MCF-7 and MDA::ER cells transfected with the
959 indicated TFOs. Boxes at the left of each heatmap indicate identified interactions by 4C. Experimental values were
960 normalized to those obtained in untransfected cells and expressed as log2. Data originate from at least three
961 independent duplicate experiments.

MCB Accepts published online ahead of print

**TABLE 1.** Sequences and characteristics of TFOs.

| Target | sequence | Distance from ERBS center | Off-target *vs.* specific e-values[a] | 35 bp probe sequence[b] |
|---|---|---|---|---|
| MCF-7 BS1 | UGGUGTUUGGUUUGGUUGG | 276 | -1.95 [1] | CACAGACGT*GG**AAGGAAAGGAATGAGGA*TGATATT |
| MCF-7 BS4 | UGGGUGUGUUGGGUCUUGGUUU | 479 | -5.74 [1] | TGACCCT***AAAGGAACAGGGAAGAGAGGGA***TTCAGC |
| MCF-7 BS6 | UUUUGUUUGGGUUGGUCUG | 2 | -1.95 [1] | TGGGCACT***GACAGGAAGGGAAAGAAAA***CAGCCTGC |
| MCF-7 BS10 | GUUGGUUCGUUGGUGUUU | 248 | -1.98 [1] | GATCTTCCT***AAAGAGGAAGCAAGGAAG***CCAGCCTC |
| MCF-7 BS14 | UUGUUCGGUUUGUGGUGG | 511 | -2.02 [2] | GATCTGAT***GGAGGAGAAAGGCAAGAA***CATGTGCGA |
| MCF-7 BS16 | GGGGUTGUGGUGGUGGUGUUU | 355 | -5.28 [2] | AGGTGTCC***AAAGAGGAGGAGGAGTAGGGG***CAACAG |
| MDA::ER BS1 | UGUUGUUUUGUGGTGGGUG | 371 | -1.98 [1] | CTAGGATGT***GAGGGTGGAGAAAAGAAGA***CGTGAGG |
| MDA::ER BS10 | UUUGGGUCGUGUUGGUGG | 273 | -1 [1] | GCAGGAGATG***GAGGAAGAGCAGGGAAA***TAGAAGCT |
| MDA::ER BS12 | UGGGUGGTUUUGUGGUGGGGUU | 301 | -5.74 [1] | AGGTGACC***AAGGGGAGGAGAAATGGAGGGA***CATTC |
| Common BS1 | UGGUUUGUUGGGUGUGUGUGUG | 185 | -1.74 [2] | GGAGTTA***GAGAGAGAGGGAAGAAAGGA***GGGAGGGA |
| Common BS2 | UUUUUUUGUGGGUGGUCGGG | 450 | -3.93 [1] | GGCTGG***GGGCAGGAGGGAGAAAAAAA***TAGTATATA |

[a] *Expressed as log(Inv(off target/specific target)). The number in brackets indicates the number of off-targets in the top5 hits determined from BLAST (http://blast.ncbi.nlm.nih.gov/).* [b] *Nucleotides targeted by the TFOs are in bold italic. Only the sense oligonucleotide is indicated.*

MCB Accepts published online ahead of print

**TABLE 2.** Genomic regions spotted on microarrays.

| Chromosome | Start [a] | Stop [a] | Cluster | Regulated in | Top gene [a] |
|---|---|---|---|---|---|
| 1 | 16,920,000 | 18,300,000 | Yes | MDA::ER | *PADI1* |
| 1 | 150,450,000 | 151,200,000 | No | MCF-7 | *LCE3B* |
| 2 | 118,600,000 | 123,280,000 | Yes | MDA::ER | *INHBB* |
| 3 | 49,890,000 | 51,600,000 | Yes | Both | *SEMA3B* |
| 7 | 72,600,000 | 73,200,000 | Yes | MDA::ER | *CLDN4* |
| 8 | 67,060,000 | 68,460,000 | Yes | MCF-7 | *MYBL1* |
| 9 | 138,882,000 | 139,620,000 | Yes | MCF-7 | *ENTPD2* |
| 10 | 43,500,000 | 44,600,000 | No | MCF-7 | *CXCL12* |
| 10 | 99,800,000 | 100,130,000 | No | MDA::ER | *LOXL4* |
| 11 | 1,500,000 | 3,300,000 | Yes | MDA::ER | *TH3* |
| 11 | 93,800,000 | 94,930,000 | Yes | MCF-7 | *FUT4* |
| 12 | 9,600,000 | 10,920,000 | Yes | MDA::ER | *KLRC3* |
| 12 | 14,100,000 | 16,300,000 | Yes | MCF-7 | *ART4* |
| 14 | 91,800,000 | 96,600,000 | Not in MCF-7 | Both | *SerpinA1* |
| 16 | 21,500,000 | 24,420,000 | Yes | MCF-7 | *SCCN1C1* |
| 16 | 54,600,000 | 55,900,000 | Yes | MDA::ER | *MT1X* |
| 16 | 65,600,000 | 66,240,000 | Yes | MDA::ER | *HSF4* |
| 17 | 35,530,000 | 37,440,000 | Yes | Both | *KRT9* |
| 19 | 54,150,000 | 54,480,000 | Yes | MCF-7 | *LHB* |
| 20 | 21,000,000 | 26,800,000 | Yes | MDA::ER | *ACSS1* |
| 21 | 36,200,000 | 47,000,000 | Yes | Both | *TFF1* |
| 22 | 48,720,000 | 49,380,000 | Yes | MDA::ER | *MAPK12* |
| X | 36,600,000 | 39,600,000 | Yes | Both | *SYTL5* |

*[a] Genomic coordinates are given from the hg18 assembly of the human genome. [b] Gene exhibiting the most important fold-change in expression upon E2 treatment in either MDA::ER or MCF-7 cells.*

MCB Accepts published online ahead of print

Figure 1

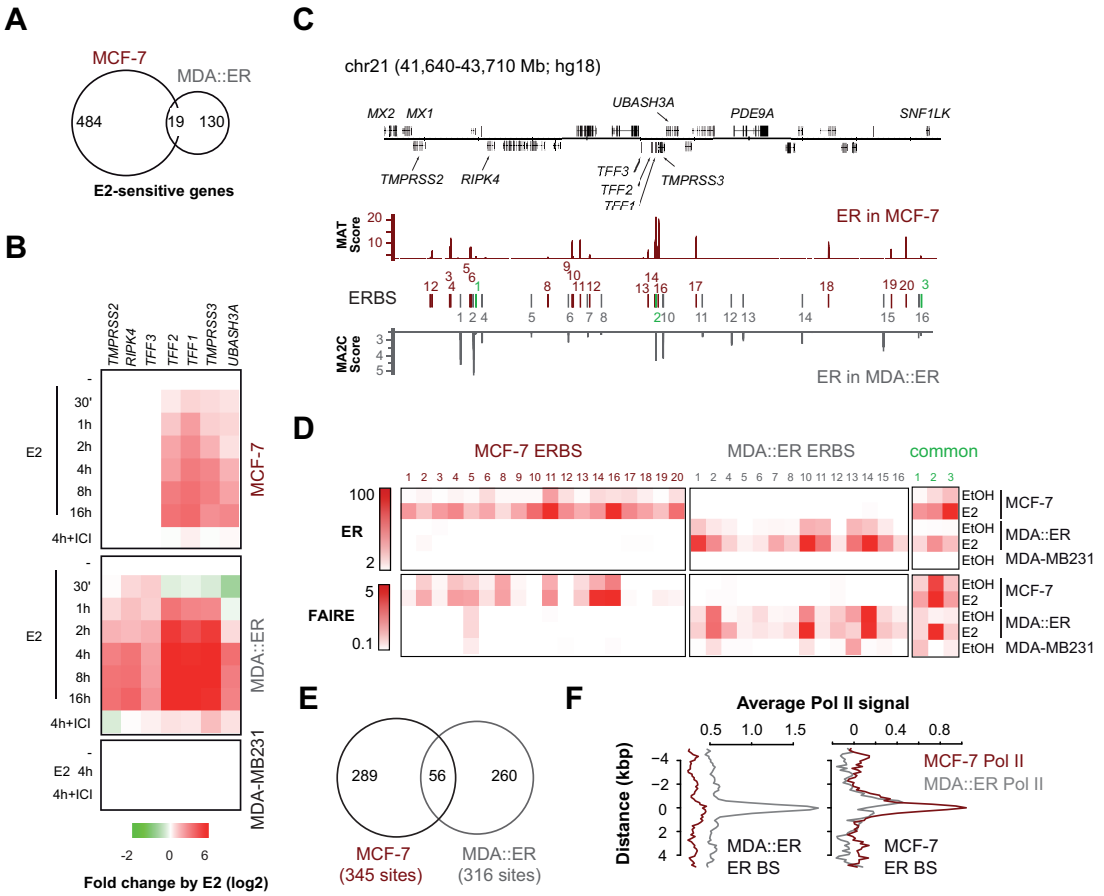


**FIG 1** Cell-specific E2-sensitive genes and ER binding sites in a 2 Mb genomic region including the TFF locus. (A) Venn diagram illustrating the overlap of identified E2-sensitive genes in MDA::ER and MCF-7 cells. (B) Heatmap representation of RT-qPCR results obtained on RNAs prepared from MCF-7, MDA::ER and MDA-MB231 cells treated for the indicated times with 10-8 M E2 and pre-treated for 36h with 10-6 M ICI164,384 where precised. Results are the log2 of the fold inductions of gene expression levels by E2 obtained in two independent triplicate experiments. (C) Integrated genome browser [IGB; (80)] illustration of the studied genomic region with RefSeq genes indicated. ER binding signal obtained in an ER ChIP-chip analysis performed using chromatin of MDA::ER cells treated for 50 min with E2 is depicted in gray. MCF-7 data were obtained from published dataset (17). For the sake of clarity, only the highest 5% signals are shown. Grey and red boxes delineate cell-specific ER binding sites (ERBSs) whilst common ERBSs are in green. (D) Anti ER ChIP and FAIRE assays were conducted using chromatin prepared from MCF-7, MDA-MB231 or MDA::ER cells treated with E2 or ethanol (EtOH) as vehicle control for 50 min. Results shown within heatmaps are means from 6 to 9 values obtained in independent triplicate experiments. Values are fold enrichments over control samples and a negative control region (promoter of the transcriptionally active *Rplp0* gene). (E) Overlap of MDA::ER ERBSs with MCF-7 ones on regions spotted on the arrays. (F) Enrichment signals obtained for anti-RNA Polymerase II (Pol II) ChIP-chip experiments performed in MDA::ER cells treated for 50 min with E2 [MCF-7 data from (17)] were aligned on MCF-7 or MDA::ER ERBS identified within the regions spotted on the arrays and located more than 10 Kbp away from the TSS of any annotated gene.

MCB Accepts published online ahead of print

Figure 2

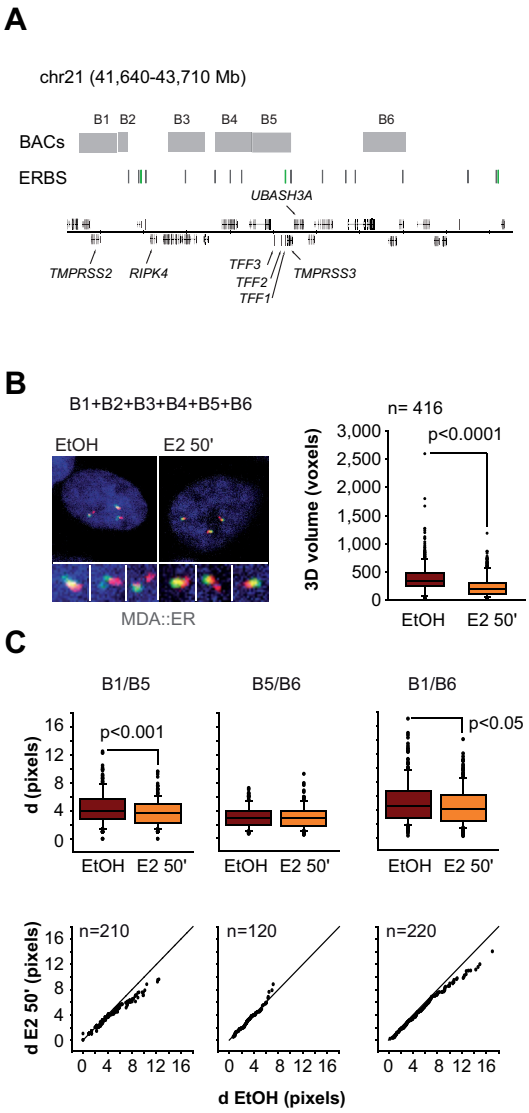


**FIG 2** E2 provokes a three-dimensional reorganization of the TFF cluster in MDA::ER cells. (A) Locations of the BACs (B1 to B6 referring respectively to RP11-814F13, CTD-2337B13, RP11-35C4, CTD-260o11, RP11-113F1 and CTD-1033M14) used to generate fluorescent probes for FISH experiments all along the genomic region of interest, illustrated as in Figure 1. Positions of the ERBSs are also indicated. (B) Analysis of the 3D volume of the TFF loci by DNA-FISH using a mix of all generated probes in cells treated with 10-8 M E2 or ethanol (EtOH) for 50 min. Representative pictures of these assays are on the left side of the panel, with magnified views of the three TFF loci present in MDA:ER cells. Quantitative measurements (n=416) of the 3D volume of the loci is shown at the right side of the panel with indicated Fisher t-test p-value. (C) Quantitative analysis of the distribution of distances separating indicated paired FISH probes. Values are shown within boxplots (top of the panel) or within quantile-quantile (Q-Q) plots representation of conditionally ranked measured distances (bottom). The normal distribution expected from non-variating distances is illustrated by the straight black line in each Q-Q plots. Non-parametric Fisher test p-values are indicated when determined as significant, as well as the number of measurements made in at least three independent experiments.

MCB Accepts published online ahead of print

Figure 3

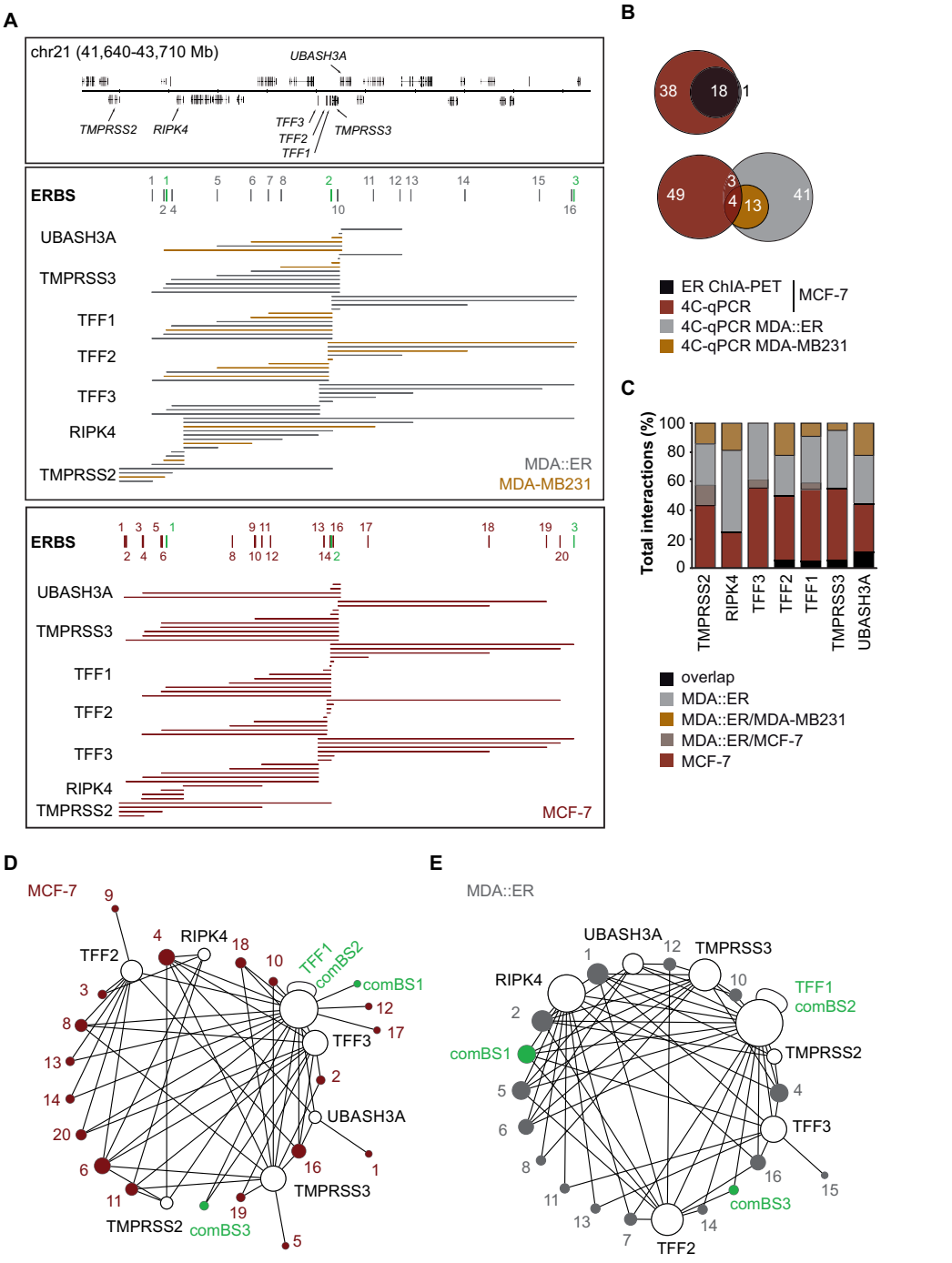


**FIG 3** ERBS-promoters interactomes. (A) IGB visualization of MCF-7, MDA-MB231 and MDA::ER interactomes linking *DpnII* fragments encompassing ERBS or promoter regions of indicated genes, as detected by 4C-qPCR on chromatin prepared from cells treated for 50 min with 10-8 M E2. Shown are RefSeq genes coordinates along chr21, as well as the positions of ERBS with grey, red and green boxes delineating MDA::ER, MCF-7 or common ERBS. 4C data are represented as lines linking one ERBS to its target promoters. MCF-7 interactions are in red, MDA::ER ones in grey and those shared between MDA::ER and MDA-MB231 in orange. (B) Venn diagrams depicting the overlapping interactions characterized in this study and to those identified in published ER ChiA-PET dataset (14) restricted to loops involving the gene promoters that served here as 4C baits. (C) Stacked histograms illustrating the overlap of the 4C-detected interactions for each tested promoters in the different cell lines. (D and E) Cytoscape (41) circular layouts of the networks of interactions that link E2-regulated genes to ERBS in MCF-7 (D) and in MDA::ER (E) cells. The sizes of the nodes are directly related to the number of interactions they direct.

MCB Accepts published online ahead of print

Figure 4

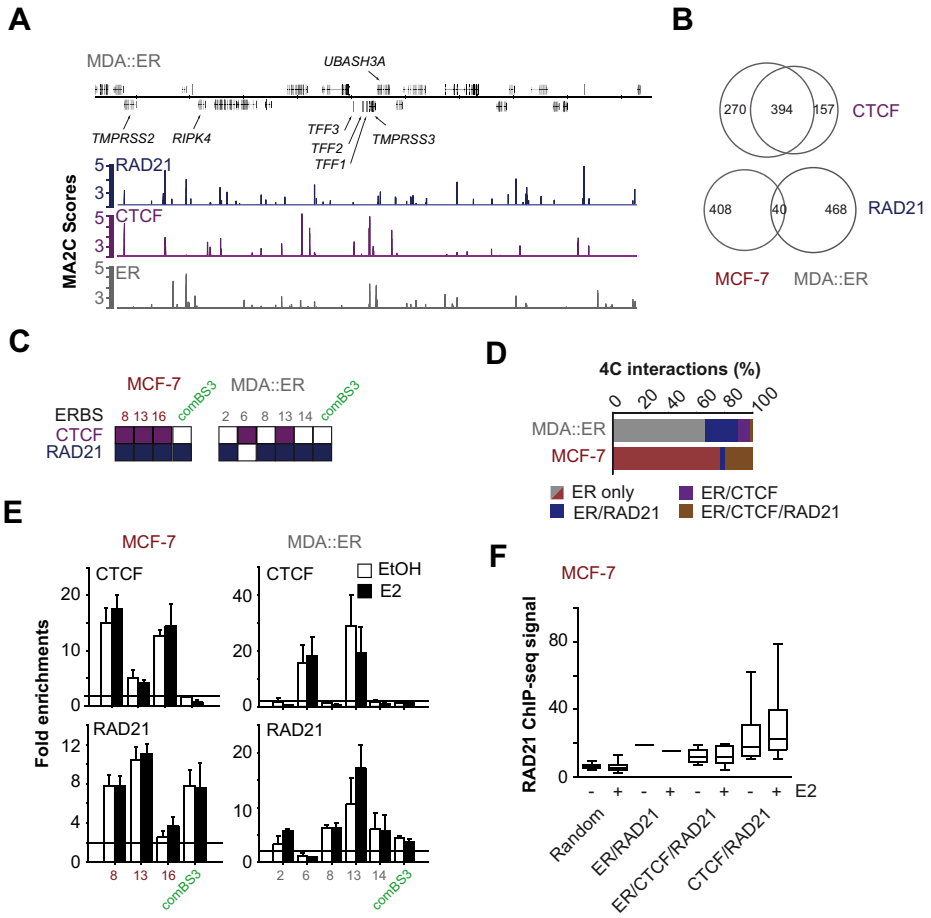


**FIG 4** CTCF and cohesin recruitment on ERBSs. (A) MDA::ER anti-ER, CTCF or RAD21 ChIP-Seq or ChIP-chip signals visualized under IGB as in Figure 1. All data were obtained from cells treated with 10-8 M E2 for 50 min. (B) Overlaps on array-spotted regions between RAD21 or CTCF binding sites determined in this study in MDA::ER cells and those previously determined in MCF-7 (53). (C) Overlap of RAD21 and CTCF binding sites with ERBSs. (D) Repartition of CTCF and RAD21 BS within the interactions between ERBS and gene promoters identified by 4C. (E) CTCF and RAD21 ChIP-qPCR were performed on chromatin prepared from MCF-7 or MDA::ER cells treated for 4h with 10-8M E2 or ethanol (EtOH) as vehicle control. Fold enrichment of the precipitated proteins on tested sequences was normalized over control ChIP and negative region (*PKNOX1* promoter) values. Data shown are mean values ± SD obtained in three independent triplicate experiments. The line depicts the significance threshold applied (Fold enrichment>2). (F) Boxplots of MACS normalized reads of RAD21 ChIP-seq experiments performed in the absence or presence of E2 in MCF-7 cells [dataset from (53)]. Values shown are mean counts measured in a 500 bp window centered on indicated binding sites located within the 2 Mb genome region studied. Calculations were also made on random sites of similar mean size and in equal number than the shared CTCF/RAD21 BSs.

MCB Accepts published online ahead of print

Figure 5

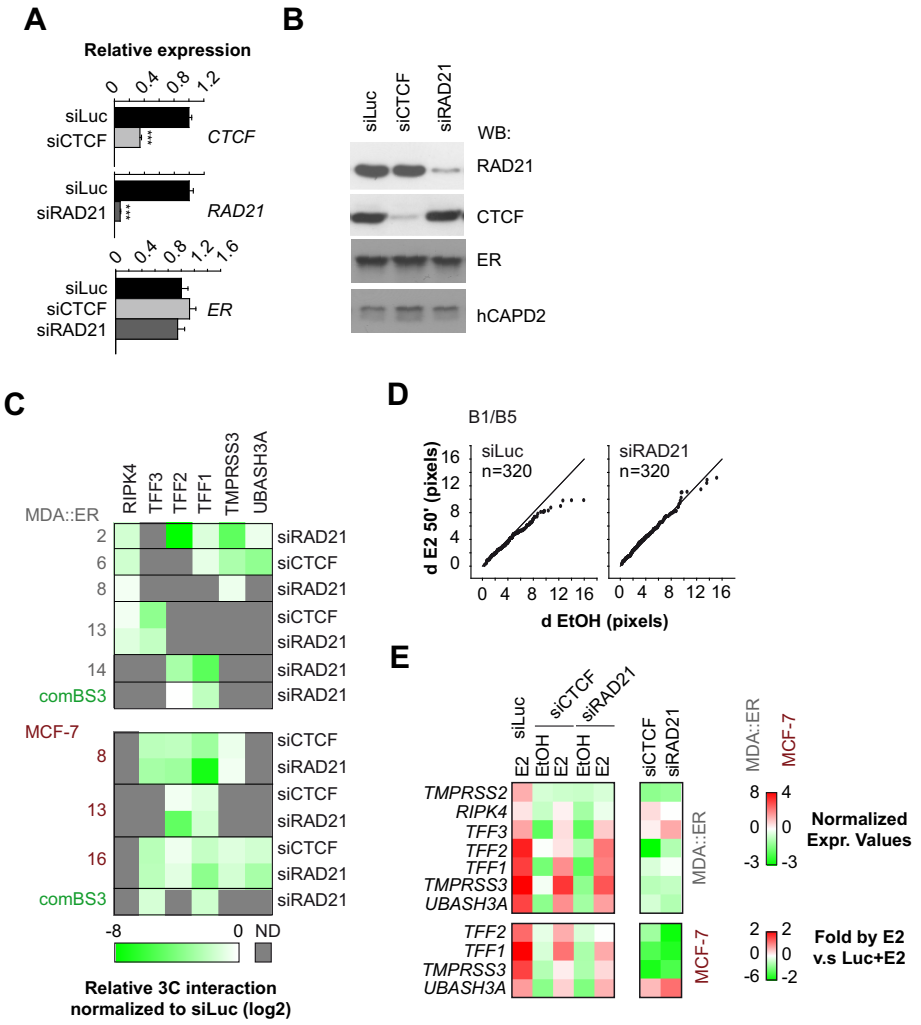


**FIG 5** CTCF and cohesin organize the E2-responsiveness of the TFF cluster. (A) Amounts of indicated mRNA were evaluated in untreated MDA::ER cells following 72h of transfection with either control siRNA (directed against the Luciferase gene, siLuc) or siRNAs targeting mRNAs of interest. Results shown are mean data ± SEM of values obtained in at least 5 different experiments (n ranges from 15 to 18), normalized to the expression of control PKNOX1 gene. Wilcoxon test was used to identify statistically relevant variations from control siLuc/EtOH condition: p-value <0.001 (***), <0.01 (**). (B) Western blots assaying the expression of RAD21, CTCF, ER and hCAPD2 (loading control) in cells transfected with the indicated siRNAs. (C) 3C-qPCR experiments following the impact of the reduction in CTCF and RAD21 contents on the interactions between indicated ERBS and gene promoters. Results shown originate from one representative experiment out of two and are expressed as the log2 of values normalized to those obtained in the control siLuc condition. (D) Quantitative DNA-FISH analysis of the distribution of distances separating paired fluorescent probes in transfected MDA::ER cells. Results are illustrated as in Figure 3. (E) Heatmaps summarizing RT-qPCR results obtained in at least 3 independent triplicate experiments following the transfection of control siLuc or siRNAs directed against CTCF or RAD21 expression and a 4h treatment of the cells with 10-8 M E2 or EtOH as vehicle control. Expression levels were normalized to the *PKNOX1* internal control and reported to those calculated for the siLuc EtOH condition. The fold inductions of these genes expression by E2 are also shown on the right side of the panel as relative to those measured in cells transfected by the siLuc. Results originate from at least 4 independent triplicate experiments.

MCB Accepts published online ahead of print

**Figure 6**

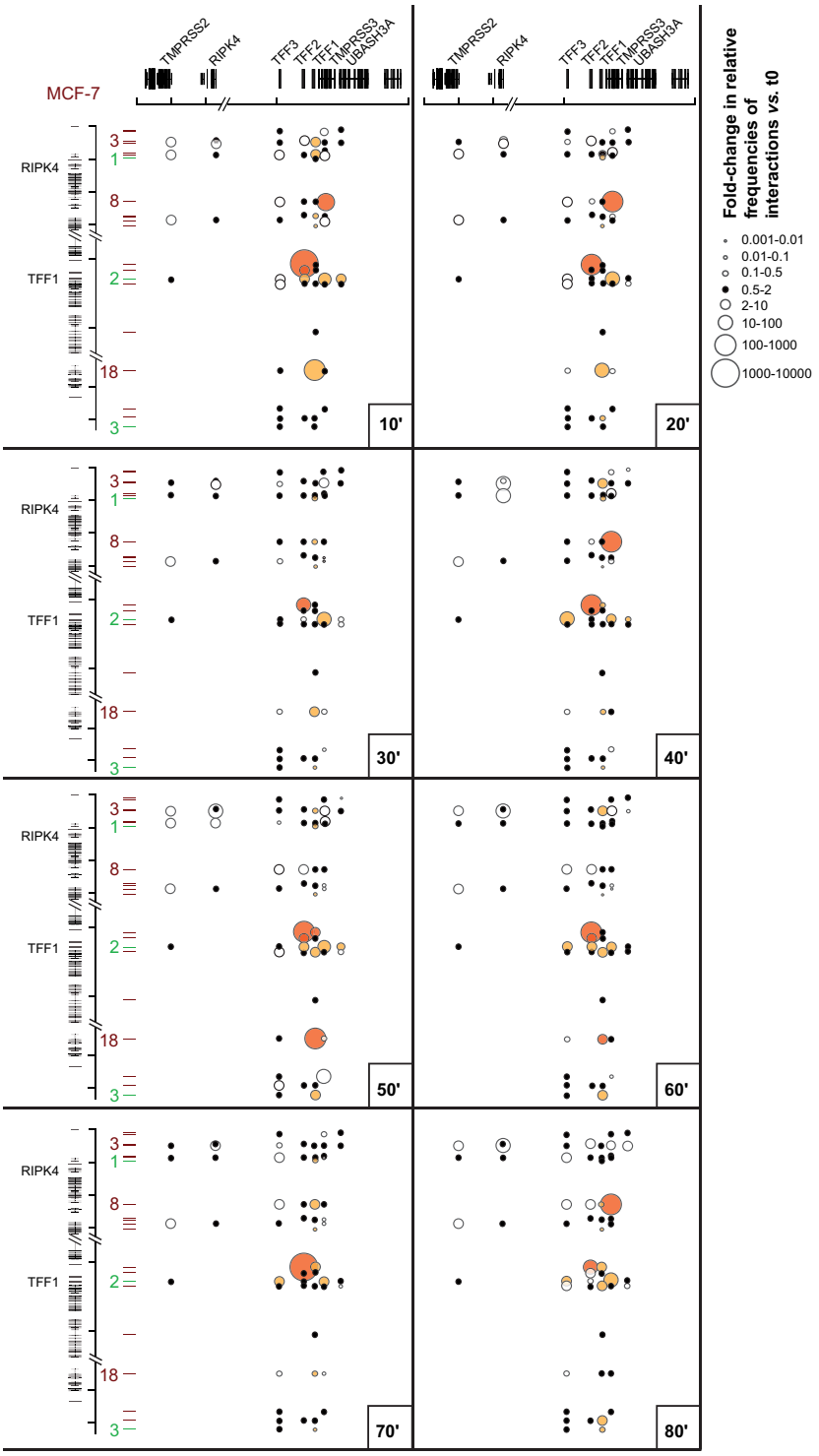


**FIG 6** DDynamic three-dimensional reorganization of the studied genomic region in MCF-7 cells. Summary of one 3C experiment representative from two, performed on chromatin sampled from MCF-7 cells treated from 0 to 80 minutes with 10-8 M E2. As indicated, the size of the bubble that corresponds to one interaction is proportional to the fold changes in frequencies of interaction as compared to the basal (t0) situation. The location of the gene promoters that served as anchors is illustrated on the top of each subpanel and the ERBSs on the left. Distance scale is accurate (2 Mb between ticks) but had to be broken in some instances for sake of figure size and clarity. Bubbles highlighted in orange are those commented in the main text and those in yellow correspond to interactions made by the ERBS located within the *TFF1* promoter (comERBS2).

MCB Accepts published online ahead of print

Figure 7

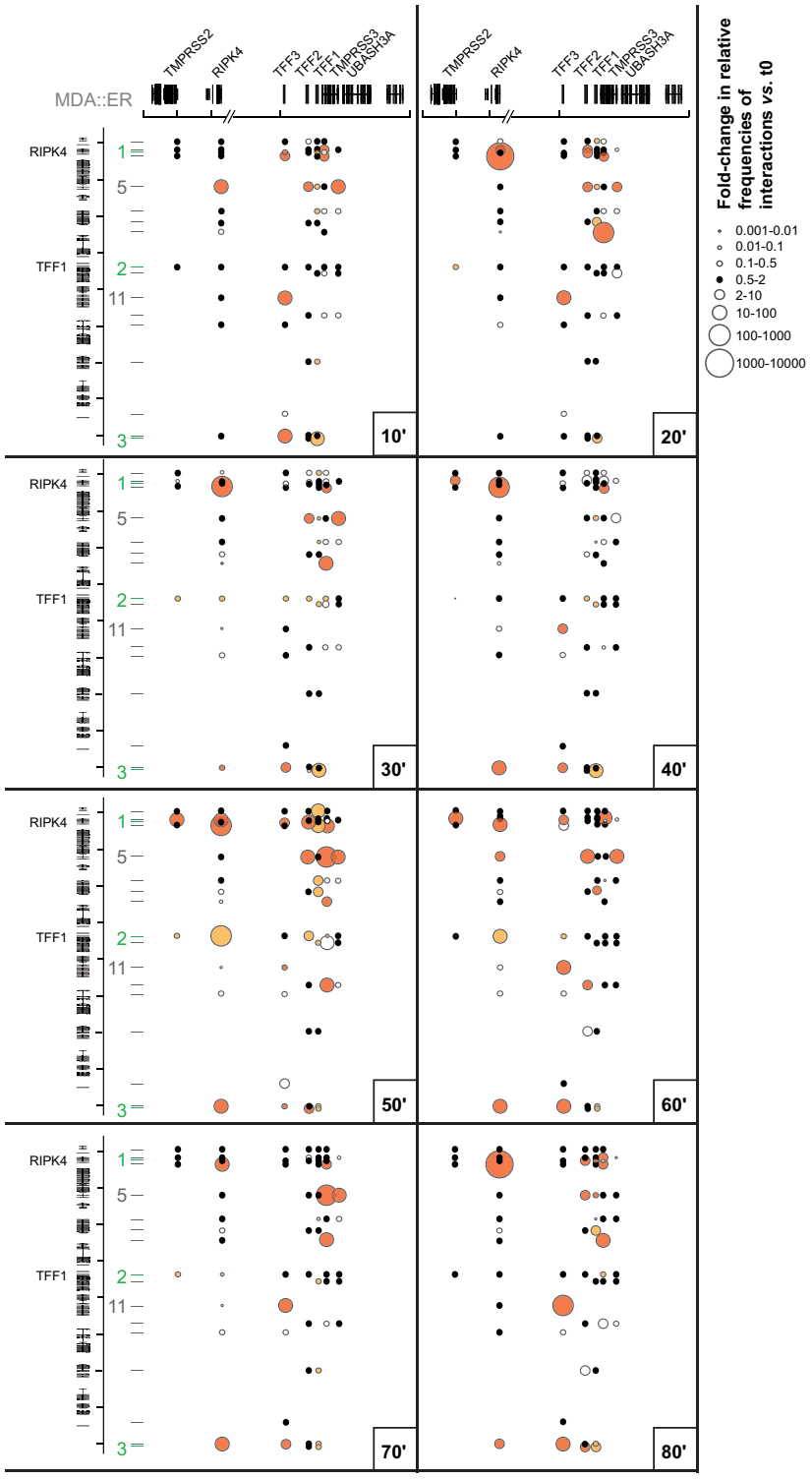


**FIG 7** Dynamic three-dimensional reorganization of the studied genomic region in MDA::ER cells. Summary of one 3C experiment representative from two, performed on chromatin sampled from MDA::ER cells treated from 0 to 80 minutes with $10^{-8}$ M E2. Results are illustrated as in Figure 6.

MCB Accepts published online ahead of print

Figure 8

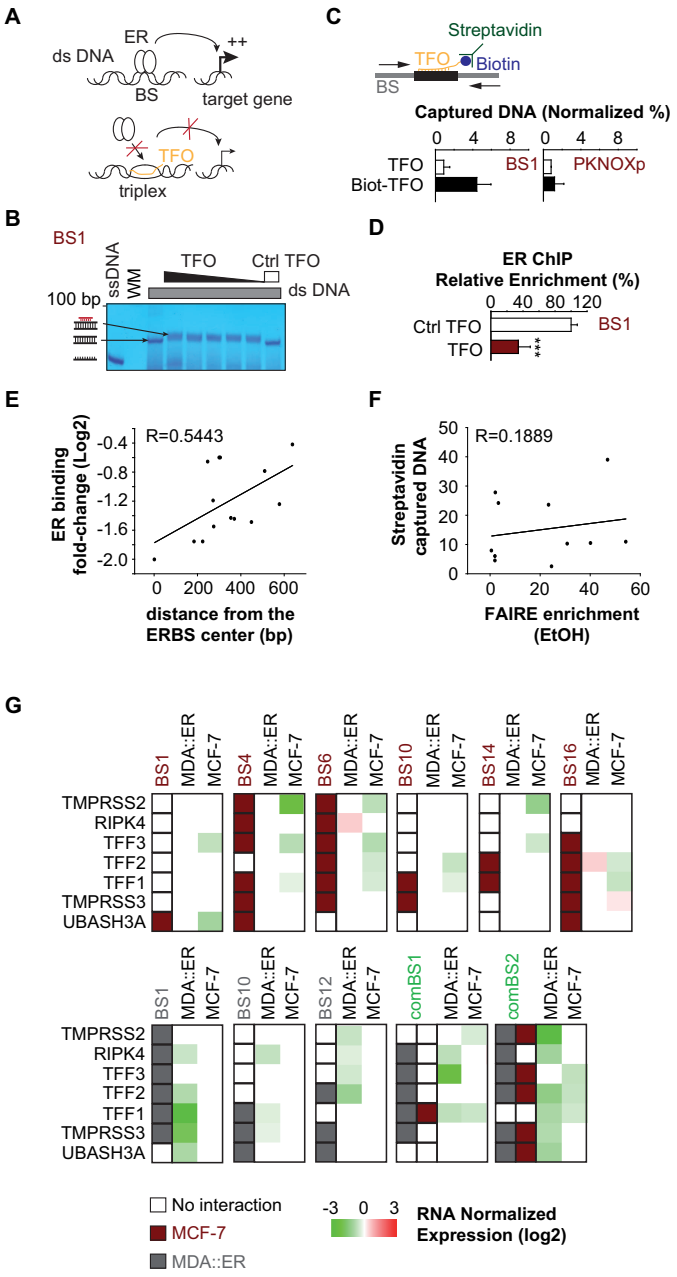


**FIG 8** Functionalization of MDA::ER and MCF-7 interactomes. (A) Triplex forming oligonucleotides (TFOs) were designed to interfere with ER binding and thus to identify the roles of ER on specific BSs for the regulation of E2-sensitive genes. (B) Formation of DNA triplex as analyzed by gel-shift. Increasing amounts of TFO (25 to 1,500 pmol) were added to 25 pmoles of target DNA duplexes and incubated for 16 h at 37 °C. Control was made using an unspecific TFO (Ctrl TFO) at the highest concentration. Complexes were separated by electrophoresis and stained with methylene blue. (C) MCF-7 cells were transfected for 36 h with 10 µmol of TFO or biotinylated (Biot-) TFO directed against the ERBS1, subjected to cross-linking and sonicated chromatin was then incubated with streptavidin-coated beads. Amounts of captured DNA were analyzed by qPCR. Values are mean ± SD of two independent duplicates, and are expressed as % of captured DNA relative to input DNA normalized to the amounts of recovered negative control region (*Rplp0* promoter). (D) Anti-ER ChIP-qPCR performed on MCF-7 cells transfected as previously and treated for 50 min with 10-8 M E2. Results are mean values ± SD of three independent triplicates expressed as relative enrichment towards the PKNOX1 promoter. (E) Fold changes in ER mobilization measured by ChIP-qPCR on each tested ERBS following the transfection of corresponding TFO was plotted against the distance separating the sequence targeted by the TFO from the center of the ERBS defined from ChIP-chip data. (F) Amounts of streptavidin-captured DNA following the transfection of Biot-TFOs are plotted against the relative chromatin accessibility of their target regions as measured by FAIRE experiments in control conditions (ethanol vehicle control, EtOH). (G) RT-qPCR experiments performed on MCF-7 and MDA::ER cells transfected with the indicated TFOs. Boxes at the left of each heatmap indicate identified interactions by 4C. Experimental values were normalized to those obtained in untransfected cells and expressed as log2. Data originate from at least three independent duplicate experiments.