LETTER

doi:10.1038/nature12210

Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation

Wenbo Li[1]*, Dimple Notani[1]*, Qi Ma[1,2], Bogdan Tanasa[1,3], Esperanza Nunez[1], Aaron Yun Chen[1], Daria Merkurjev[1,2], Jie Zhang[1], Kenneth Ohgi[1], Xiaoyuan Song[1], Soohwan Oh[1,4], Hong-Sook Kim[1], Christopher K. Glass[5] & Michael G. Rosenfeld[1]

The functional importance of gene enhancers in regulated gene expression is well established[1–3]. In addition to widespread transcription of long non-coding RNAs (lncRNAs) in mammalian cells[4–6], bidirectional ncRNAs are transcribed on enhancers, and are thus referred to as enhancer RNAs (eRNAs)[7–9]. However, it has remained unclear whether these eRNAs are functional or merely a reflection of enhancer activation. Here we report that in human breast cancer cells 17β-oestradiol (E2)-bound oestrogen receptor α (ER-α) causes a global increase in eRNA transcription on enhancers adjacent to E2-upregulated coding genes. These induced eRNAs, as functional transcripts, seem to exert important roles for the observed ligand-dependent induction of target coding genes, increasing the strength of specific enhancer–promoter looping initiated by ER-α binding. Cohesin, present on many ER-α-regulated enhancers even before ligand treatment, apparently contributes to E2-dependent gene activation, at least in part by stabilizing E2/ER-α/eRNA-induced enhancer–promoter looping. Our data indicate that eRNAs are likely to have important functions in many regulated programs of gene transcription.

We performed ER-α chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-Seq) analysis using 1 h E2-treated (100 nM) MCF-7 human breast cancer cells and revealed 31,052 ER-α binding sites genome wide. This included only 902 on promoters (Supplementary Fig. 1a), in accordance with previously reported analyses[10–12], and 7,174 ER-α-bound potential enhancers based on the presence of H3K4me1 (refs 13, 14) and H3K27ac (ref. 15) (Supplementary Fig. 1b). Global run-on sequencing (GRO-seq) analysis of MCF-7 cells in similar conditions identified 1,309 E2-upregulated coding genes, of which 1,145 had an E2/ER-α-binding enhancer within 200 kilobases (kb) from their transcription start site (TSS) and were thus considered to be direct oestrogen-upregulated target genes (hereafter referred to as UP genes; Supplementary Fig. 1c). Of these, only 112 showed ER-α binding to their promoters (Supplementary Fig. 1c), consistent with suggestions[10,11] that ER-α occupancy on enhancers is a key strategy underlying E2-induced gene expression. Most E2-regulated enhancers showed a rapid bidirectional activation of eRNAs, exemplified by the *FOXC1* locus (Supplementary Fig. 1e), which is about ~1.5 kb long as identified by GRO-seq, although ~10% exhibited an apparent unidirectional eRNA transcription[8] (Fig. 1a and Supplementary Fig. 1f, g). These data suggest that eRNA induction in response to ER-α binding is a predictive mark of enhancer activity[9]. Binding of ER-α did not cause clear alterations in enhancer marks on ER-α-bound enhancers, such as H3K27ac (Supplementary Fig. 1h).

Approximately 83% of enhancers with detectable GRO-seq signals adjacent to UP genes exhibited E2-induced eRNA upregulation (Fig. 1a); for the remaining 17%, the tag count was not sufficient to assign upregulation bioinformatically. E2 induction of eRNA was not observed on non-ER-α-bound H3K27ac-marked enhancers (Supplementary Fig. 1i). The median distance between enhancers exhibiting E2-induced eRNAs ($n = 1{,}248$; referred to as UP enhancers) and their closest UP genes was ~52 kb, compared with a median distance of ~270 kb between enhancers exhibiting no E2 induction of

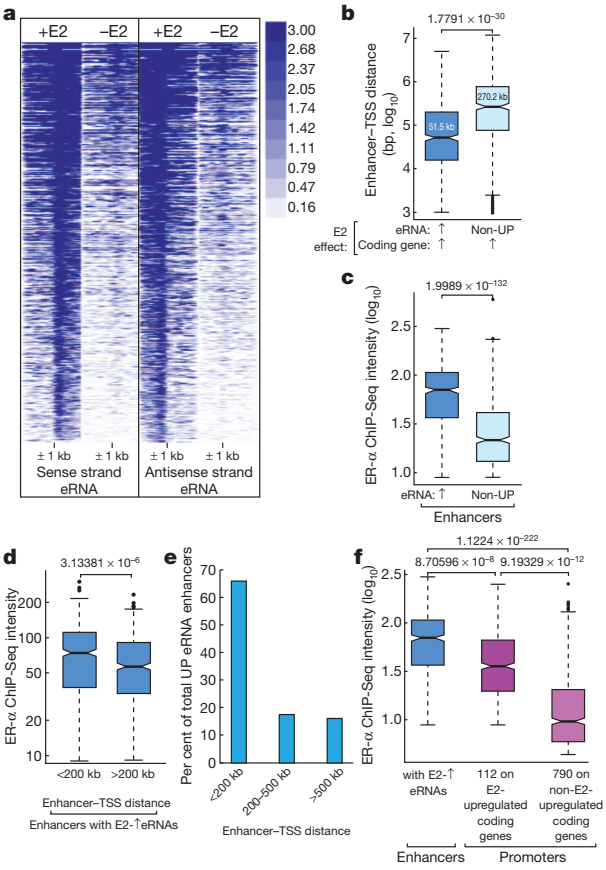

Figure 1 | E2 induction of eRNA in MCF-7 breast cancer cells. a, Heat map of GRO-seq showing bidirectional eRNA transcription at enhancers induced by E2. **b**, UP enhancers are closer to the UP genes (median ~52 kb) in comparison to enhancers with non-upregulated eRNAs (median ~270 kb). Up arrows indicate upregulation. **c**, ER-α binds more robustly to UP enhancers than to the enhancers with non-upregulated eRNA. **d**, Among the UP enhancers, the proximal ones within 200 kb from any E2-upregulated gene TSSs exhibit higher ER-α binding intensity than the distal cohort of UP enhancers located farther away. **e**, Most of the UP enhancers are in close proximity to E2-upregulated coding genes. **f**, ER-α binding intensity on UP enhancers is higher than on 112 promoters of E2-activated genes, which itself is higher than the 790 ER-α-bound promoters of genes did not show upregulation by E2. A $\log_{10}$ scale is used for panels **b**, **c** and **f**. *P* values are given at the top of graphs, and were calculated using the Student's *t*-test.

[1]Howard Hughes Medical Institute, Department of Medicine, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA. [2]Graduate Program in Bioinformatics, University of California, San Diego, La Jolla, California 92093, USA. [3]Graduate Program, Kellogg School of Science and Technology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA. [4]Graduate Program in Biological Sciences, University of California, San Diego, La Jolla, California 92093, USA. [5]Cellular and Molecular Medicine, Department of Medicine, University of California, San Diego, La Jolla, California 92093, USA.
*These authors contributed equally to this work.

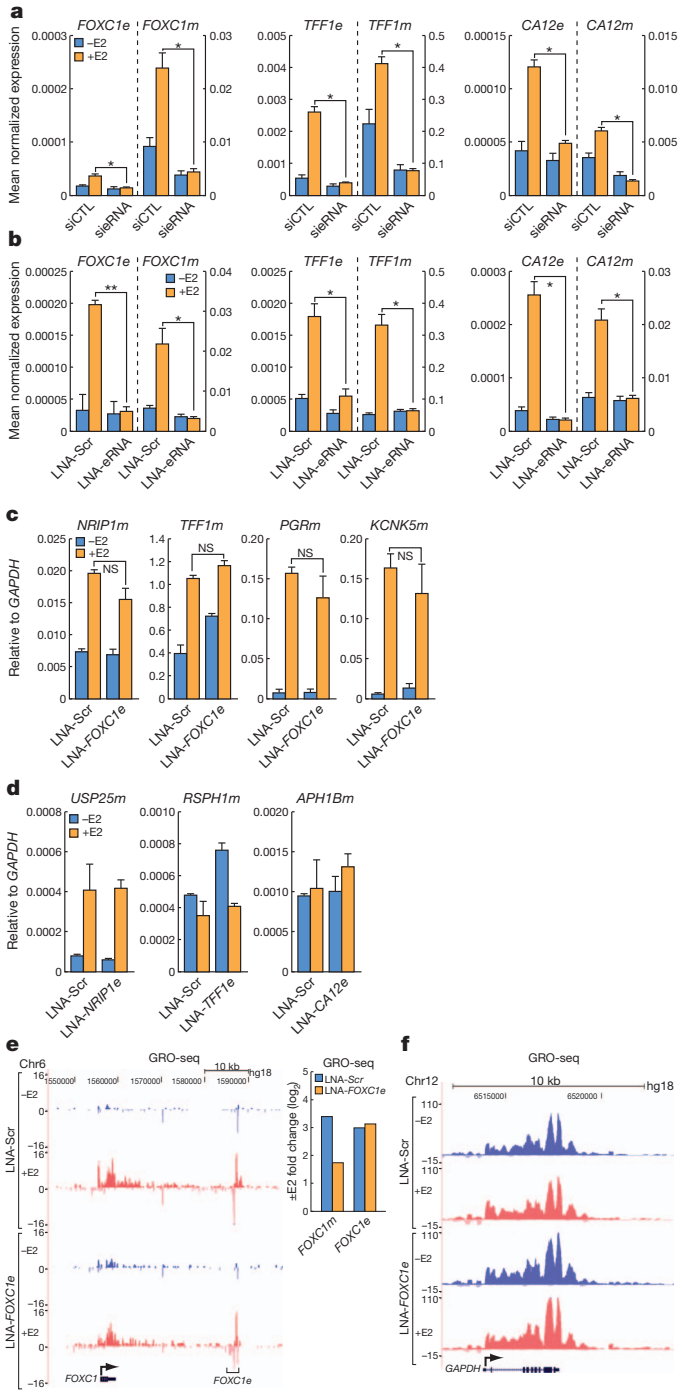

Figure 2 | Importance of eRNA for target gene activation. **a**, **b**, siRNA/LNA knockdown of eRNAs. Efficacy and effects on coding gene transcription were assessed by qPCR for the *TFF1*, *FOXC1* and *CA12* eRNAs and corresponding coding transcription units. Lower case 'e' and 'm' after gene names denote eRNA and gene mRNA, respectively. CTL, control; Scr, scramble. **c**, qPCR analysis showing no significant change of several E2 target coding genes when *FOXC1* eRNA was knocked down using LNA. NS, not significant. **d**, Lack of effect of *NRIP1*, *TFF1* or *CA12* eRNA knockdown on expression of other coding genes located distally, including *USP25* (520 kb from *NRIPe*), *RSPH1*, (120 kb from *TFF1e*) and *APH1B* (110 kb from *CA12e*). **e**, GRO-seq data from MCF-7 cells treated with LNA against *FOXC1* eRNA (*FOXC1e*) showing LNA's inhibitory effect on the transcription of the *FOXC1* coding locus, but not on the targeted enhancer region. The bar graph (right) shows that the LNA against *FOXC1e* knocked down E2 induction of *FOXC1* messenger RNA (tag counts over the whole gene length), but not transcription of the enhancer region. hg18, human genome 18. **f**, A similar GRO-seq snapshot as in **e**, showing the lack of effect from LNA against *FOXC1e* on *GAPDH* transcription. Data represent mean ± standard error of the mean (s.e.m.) (**a**, **b**) and mean ± standard deviation (s.d.) (**c**, **d**); ($n = 3$). $^*P < 0.05$, $^{**}P < 0.01$.

eRNAs with UP genes (Fig. 1b). ChIP-Seq analysis revealed that UP enhancers showed significantly stronger binding of ER-α than enhancers not exhibiting eRNA upregulation (Fig. 1c). Proximal (<200 kb) UP enhancers constituted a majority of all UP enhancers and had a higher affinity for ER-α than did distal UP enhancers (Fig. 1d, e). The strength of ER-α binding was much higher on UP enhancers than on 112 ER-α-bound promoters of coding genes that showed E2 induction, whereas the remaining 790 ER-α-bound promoters of genes with no E2 upregulation exhibited the weakest ER-α binding (Fig. 1f).

On the basis of GRO-seq analyses, we selected ten highly upregulated transcription units for further experimentation, each associated with adjacent UP enhancers exhibiting ~2.5–5-fold E2 induction of eRNAs (Supplementary Fig. 1j). Despite increasing evidence for crucial nuclear functions of lncRNAs[4–6], it remains an unresolved question whether eRNAs are merely a by-product of enhancer activation or whether they might serve as key regulators of coding gene transcription[7–9]. To investigate the potential roles of ligand-induced eRNAs on gene activation events, both specific short interfering RNAs (siRNAs)[16] and locked nucleic acid antisense oligonucleotides (LNAs)[17] directed against each eRNA transcript were designed on the basis of the peaks of eRNA exhibited by GRO-seq. To exclude off-target effects, experiments were performed with two different LNAs or siRNAs targeting each eRNA.

With a high efficiency of transfection (Supplementary Fig. 2a), both siRNA and LNA-mediated knockdown of the *TFF1*, *FOXC1* or *CA12* eRNAs revealed that, for each transcription unit, the induction of both the eRNA and of the adjacent coding gene, as assessed by quantitative polymerase chain reaction (qPCR) and GRO-seq, respectively, was significantly inhibited (Fig. 2a, b, e and Supplementary Fig. 2b, c). By contrast, these siRNAs/LNAs did not affect the housekeeping genes we tested (for example, *GAPDH*; Fig. 2f), or E2-regulated or non-E2-regulated transcription units more distal to the regulated enhancers (Fig. 2c, d). Ligand-induced increase of ER-α binding occurred even after eRNA knockdown (Supplementary Fig. 2d, e). Similar eRNA requirements for coding-gene induction by E2 were observed on the basis of knockdown of eRNAs adjacent to the *PGR*, *SIAH2*, *KCNK5*, *P2RY2*, *SMAD7*, *GREB1* and *NRIP1* genes using either siRNAs or LNAs (Fig. 3g, i and Supplementary Fig. 2b). GRO-seq data were consistent with the notion that LNA against eRNA reduces the levels of eRNA transcript post-transcriptionally, but not its nascent transcription (Fig. 2e, bar graph). Knockdown of an eRNA on an ER-α-bound distal enhancer (~222 kb from the *FOXC1* TSS) that did not exhibit E2-induced eRNA and with low ER-α-binding affinity did not affect neighbouring *FOXC1* gene induction (Supplementary Fig. 2f), further indicating that eRNA induction potentially marks E2-regulated functional enhancers. Although GRO-seq results (Fig. 2e) already indicate a lack of any LNA-mediated transgene silencing of the enhancer DNA, further assays—including methyl miner and enzyme digestion assays (Supplementary Fig. 3a–c)—confirmed unaltered enhancer methylation on the *FOXC1*, *P2RY2* or *NRIP1* enhancers. Additional supporting evidence was provided by using an LNA targeting the sense transcript from a regulatory region near the *GREB1* gene (GREB1-RR), which exhibits overlapping bidirectional transcription (Supplementary Fig. 3d, e); we observed no significant change in transcript level from the antisense strand by strand-specific qPCR. We also failed to observe any significant LNA effects on levels of total histone H3, H3K9me3 or H3K27me3 silencing marks on several targeted enhancers (Supplementary Fig. 3g). Together, these data suggest that siRNA/LNA-mediated knockdown of eRNAs does not elicit transgene silencing of the interrogated enhancers.

To validate independently that eRNAs *per se* are important for quantitative increases in target gene expression, we took advantage

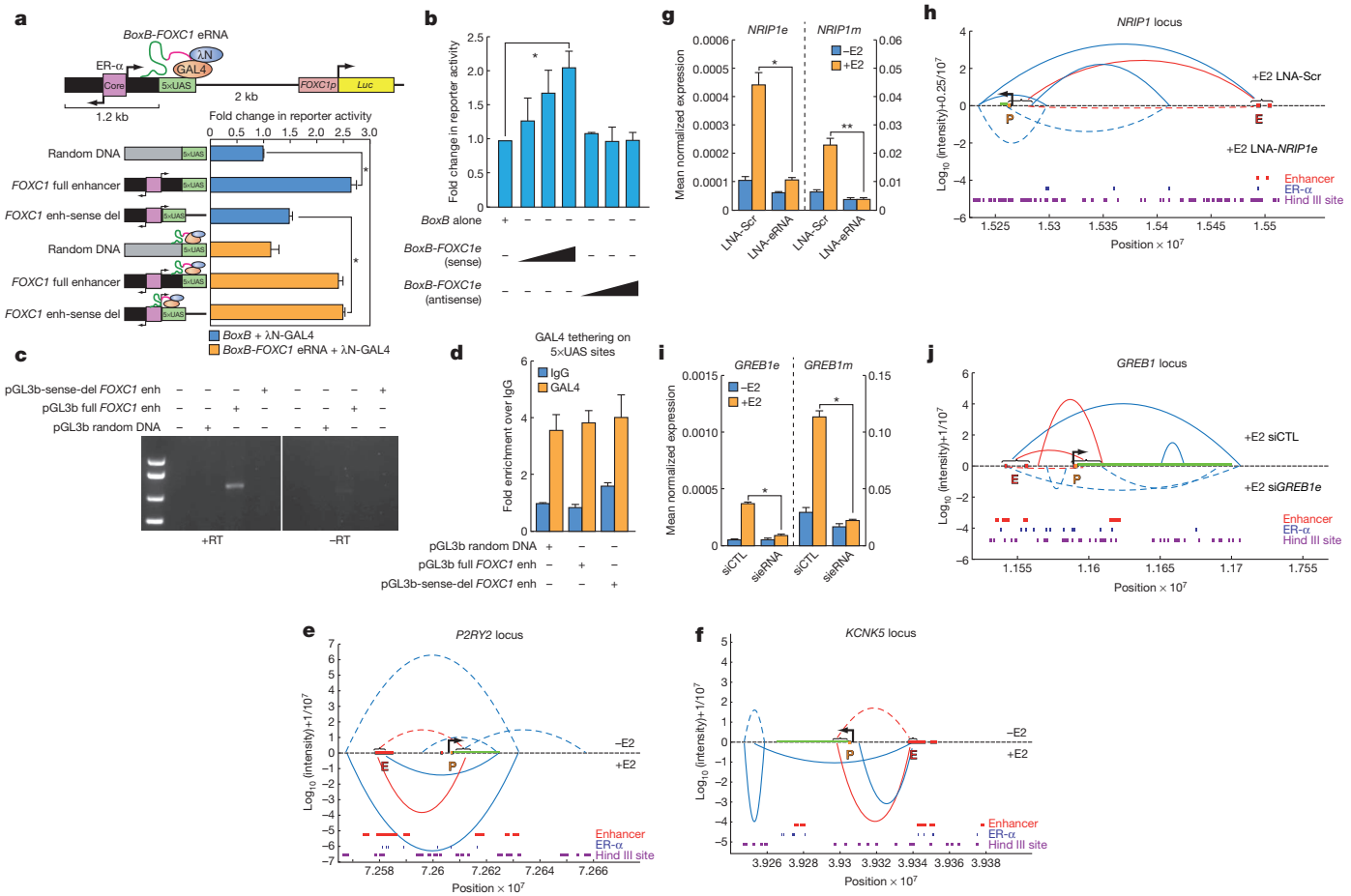

Figure 3 | Ligand-induced eRNA is functionally important. **a**, Schematic diagram of the *BoxB*-λN tethering system on the *FOXC1* enhancer, which is upstream of a *FOXC1* native-promoter-linked luciferase (*Luc*). *FOXC1p*, *FOXC1* promoter. 5×UAS sites are fused downstream to the *FOXC1* enhancer. GAL4–λN fusion protein tethers *BoxB-FOXC1* eRNA to the 5×UAS sites. Bar graph shows the effects of the *FOXC1* eRNA on *FOXC1* promoter-driven *Luc* activity in the presence of E2 (24 h). Blue bars show how the activating function of the native full-length enhancer (bar 2) over random DNA (bar 1) is lost when the sense eRNA cassette is substituted with the 5×UAS site (bar 3). Orange bars show how this loss is largely rescued upon *FOXC1* eRNA tethering to the sense-eRNA-deleted enhancer cassette (enh-sense del; bar 6). **b**, eRNA function is sequence specific: *FOXC1e* sense eRNA but not antisense strand RNA could rescue the activity of sense-eRNA-deleted enhancer in the tethering assay. **c**, Gel picture showing plasmid-based eRNA expression from full-length enhancer (enh) but not from the sense-eRNA-deleted enhancer construct (pGL3b-sense-del *FOXC1* enh). RT, reverse transcriptase. **d**, Bar graph showing efficiency of GAL4 tethering on various pGL3b constructs. **e**, 3D-DSL data for the *P2RY2* locus, revealing strengthened promoter–enhancer interactions over basal conditions after 1 h E2 treatment. For all 3D-DSL data, the $\log_{10}$ intensities of interaction counts plus 1 or 0.25 for presentation purposes are on the *y*-axis, and the *x*-axis depicts coordinates from the University of California, Santa Cruz (UCSC) genome browser. Interaction data are overlaid with positions of the enhancer, ER-α-binding sites and HindIII sites on the regions interrogated. The pertinent promoter–enhancer interaction is shown in red and other interactions are shown in blue. E, enhancer; P, promoter. **f**, 3D-DSL data for the *KCNK5* locus after 1 h E2 treatment. **g**, LNA knockdown of *NRIP1* eRNA effectively reduced the levels of both eRNA and associated coding gene transcripts. **h**, 3D-DSL data demonstrating significant reduction in promoter–enhancer interaction upon treatment of LNA against *NRIP1* eRNA. **i**, *GREB1e* siRNA knockdown diminished the levels of eRNA and associated coding gene transcript. CTL, control. **j**, 3D-DSL data for the *GREB1* locus showing significantly reduced enhancer–promoter looping as well as other genomic interactions after *GREB1e*-specific siRNA treatment. Dotted lines in panels **e** and **f** represent −E2 condition, but knockdown situation in panels **h** and **j**. Data show mean ± s.d.; ($n = 3$). $^*P < 0.05$, $^{**}P < 0.01$.

of a GAL4–*BoxB*-tethering-based reporter assay[18]. For this we engineered a chimaeric RNA by fusing *FOXC1* sense eRNA to *BoxB* viral RNA, permitting *BoxB-FOXC1* eRNA to be recruited by the RNA binding domain of λN protein fused with the GAL4 DNA binding domain (λN–GAL4). Thus eRNA can be artificially tethered to 5×UAS sites just downstream of the *FOXC1* enhancer in the reporter plasmid, in which luciferase (*Luc*) is under the control of the native *FOXC1* promoter (Supplementary Fig. 8). We observed that the presence of full-length *FOXC1* enhancer increased *Luc* expression to ~2.5 fold when compared to random DNA in place of the enhancer (Fig. 3a, blue bars). This effect was abolished when the sense eRNA sequence was deleted and substituted with 5×UAS sites, generating a non-functional 'missense' eRNA (Fig. 3a, blue bars, and Supplementary Fig. 8). Tethering of *BoxB-FOXC1* eRNA, but not *BoxB* alone, could fully rescue the activity loss of sense-eRNA-deleted enhancer (Fig. 3a, orange bars), whereas the antisense *FOXC1* eRNA could not (Fig. 3b). We confirmed the loss of plasmid-driven native *FOXC1* eRNA expression from the sense-eRNA-deleted reporter construct, and showed that GAL4 tethering was not altered (Fig. 3c, d). These data further support the suggestion that the sequence-specific eRNA transcript *per se*, rather than merely the process of enhancer transcription, is required for the actions of the eRNA on enhancer-dependent coding-gene activation events. This observation is consistent with recent studies of the role of ncRNAs in p53-dependent gene activation[19] and in regulation of the *SNAI1* gene[20].

We next investigated whether enhancer–promoter looping is induced in the E2-activation events[21], using a strategy analogous to chromosome conformation capture carbon copy (5C), which is named three-dimensional DNA selection and ligation (3D-DSL)[22], to study the spatial organization of genomes[23,24]. We first examined two E2-regulated transcription units: *P2RY2* and *KCNK5*. For *P2RY2*, E2 treatment significantly increased the specific promoter–enhancer interaction (Fig. 3e), and

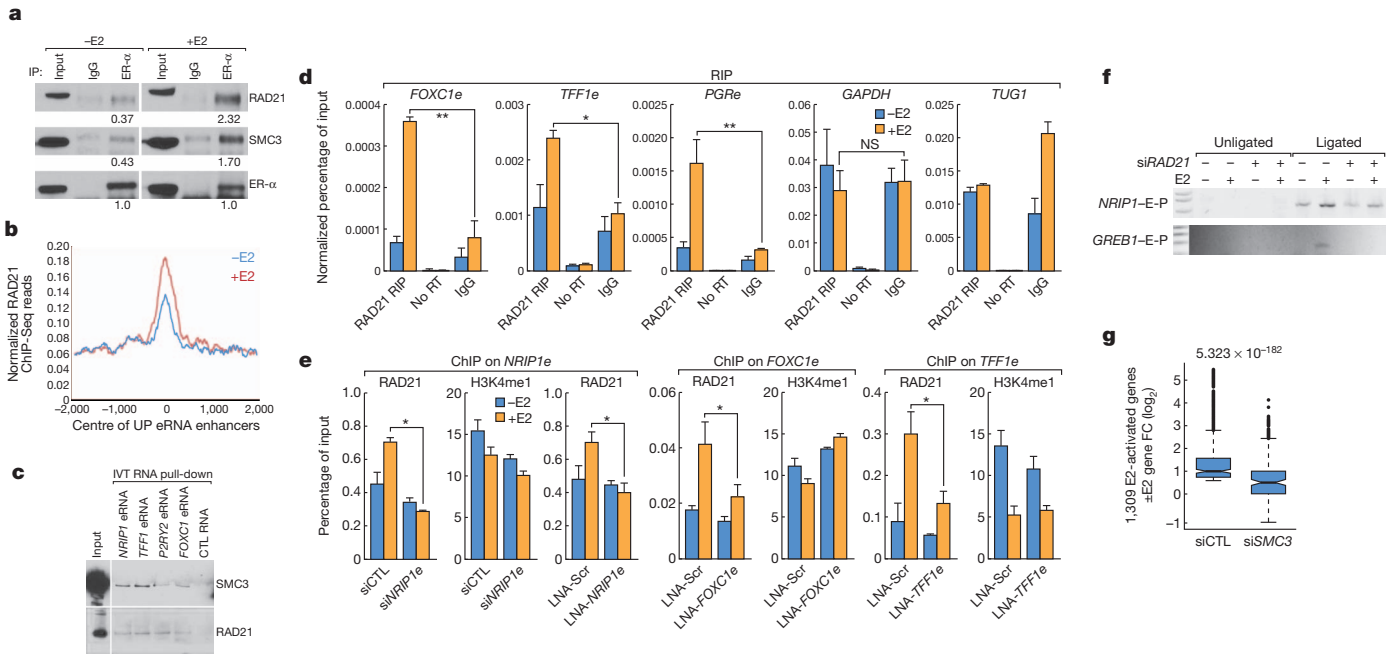

Figure 4 | Role of eRNA in cohesin-dependent gene activation. **a**, Co-immunoprecipitation (IP) of RAD21 and SMC3 with ER-α from E2 or ethanol-treated MCF-7 whole-cell extracts showing the physical interaction between ER-α and cohesin subunits, which is enhanced by E2 treatment. Numbers below blots indicate the band density (Image J, http://rsbweb.nih.gov/ij/) relative to that of the corresponding density of ER-α. **b**, RAD21 enrichment centred at UP enhancers as determined by ChIP-Seq, which shows moderate E2-induced increase. **c**, IVT RNA pull-down assay showing the interaction between cohesin subunits and eRNAs, but not a control RNA (RNA fragment of *Xenopus* elongation factor α; CTL). **d**, RIP-qPCR showing binding of RAD21 to selected regulated eRNAs but not to *GAPDH* or *TUG1*. RT, reverse transcriptase. **e**, ChIP-qPCR analyses represent the inhibitory effect from knockdown of *NRIP1e* (siRNA and LNA), *FOXC1e* or *TFF1e* on E2-induced RAD21 additional recruitment, but not on H3K4me1 binding. **f**, Effect of RAD21 depletion on the physical interaction between promoter–enhancer for the *GREB1* and *NRIP1* genes, assessed by 3C assay. E, enhancer; P, promoter. **g**, 1,309 E2-induced coding genes were defined by GRO-seq from the control siRNA (*siCTL*; ± E2) group, and then their fold changes ($\log_2$ FC) in siCTL- versus si*SMC3*-transfected MCF-7 cells were plotted. Data show mean ± s.d.; ($n = 3$). $^*P < 0.05$, $^{**}P < 0.01$.

also induced a new E2-dependent interaction between the enhancer and the gene terminus region. Similarly, for the *KCNK5* locus, E2 treatment caused a clear increase in loops from enhancer to promoter, as well as to other regions near the terminator and promoter (Fig. 3f). These observations indicate that a major effect of ligand is to enhance specific promoter–enhancer interactions in parallel to induction of eRNA.

We next investigated whether E2-induced enhancer–promoter interactions are affected by eRNAs. The *NRIP1* locus exhibited specific enhancer–promoter and promoter–gene-terminus loops, whereas treatment with LNA against *NRIP1* eRNA caused a marked inhibition of these interactions (Fig. 3h) and E2 activation of the *NRIP1* gene (Fig. 3g). siRNA-mediated *GREB1* eRNA knockdown also coordinately inhibited *GREB1* gene induction and the two specific enhancer–promoter interactions induced by E2 and two additional non-enhancer loops (Fig. 3i, j). Together these experiments indicate that oestrogen causes quantitative, as well as some qualitative, alterations in the interactions between enhancers and coding-gene promoters, and that eRNAs are of functional importance, at least under the experimental conditions here, for enhancer–promoter interactions.

To address the possibility that eRNAs might also work *in trans*, we first estimated the absolute expression levels of the eRNAs, finding that most of the eRNAs we investigated were present at levels of <5–15 copies per cell, although several, including *TFF1* eRNA, were present at ~70–95 molecules per cell (Supplementary Fig. 4a, b), suggesting that these eRNAs were likely to function primarily *in cis*. Furthermore, we used chromatin isolation by RNA purification (ChIRP)[25] to identify potential sites where *FOXC1* eRNA localizes in the genome; despite robust detection of *FOXC1* eRNA from its transcribing site—establishing the efficacy of the biotin-labelled probes used (Supplementary Fig. 4c)—only 15 peaks could be confidently called, and for none was the nearest gene E2 regulated (Supplementary Fig. 4c). In addition, qPCR analysis after knockdown of *FOXC1* eRNA revealed no significant effects on E2 activation of *NRIP1*, *TFF1*, *PGR* or *KCNK5* genes (Fig. 2c). GRO-seq after LNA transfection against *FOXC1* eRNA revealed that a large majority (>95%) of the E2-upregulated coding genes continued to exhibit clear E2-dependent upregulation. Therefore, any *trans* effects of eRNAs are likely to be relatively infrequent or quantitatively small. Of course, there are inevitably indirect effects observed after knockdown of any eRNA that downregulates a functional coding gene. However, at least for a few gene areas, there may be effects of enhancer-based long-range interactions. We identified at least one such example, between the *NRIP1* and *TFF1* loci, separated by ~27 Mb on chromosome 21, exhibiting an E2-induced increase of colocalization by fluorescence *in situ* hybridization (FISH; Supplementary Fig. 5a–c). Surprisingly, knockdown of *NRIP1e* eRNA by LNA caused a clear decrease in the interactions between these two genomic loci (Supplementary Fig. 5b, c), suggesting that such E2-induced colocalization was eRNA dependent.

Because several studies have implicated a role for cohesin in chromosomal interactions and enhancer–promoter looping events[26–28], we investigated whether cohesin was involved in the observed eRNA functions. First, co-immunoprecipitation showed that ER-α can interact with cohesin subunits (Fig. 4a). ChIP-Seq revealed that ~30–40% of RAD21 (a subunit of cohesin) binding sites overlap with putative H3K4me1/H3K27ac-marked enhancers in MCF-7 cells[28] (Supplementary Fig. 7h). After E2 treatment, both ChIP-Seq (Fig. 4b) and ChIP-qPCR data revealed a reproducible, but modest (50–200%), increased occupancy of RAD21 and SMC3 on the interrogated enhancers, as exemplified by *FOXC1e*, *NRIP1e* and *TFF1e* (Fig. 4e and Supplementary Fig. 6a). By *in vitro* transcribed (IVT) RNA pull-down, the investigated eRNAs could pull-down SMC3 and RAD21 from MCF-7 nuclear extracts (Fig. 4c and Supplementary Fig. 6b). RIP-qPCR confirmed the interaction between cohesin and several eRNAs,

but not with *GAPDH* or another nuclear RNA, *TUG1* (Fig. 4d). To test possible direct or indirect involvement of RNAs in cohesin recruitment to enhancers, we found that RNase treatment caused some decrease of the cohesin level in the chromatin-bound fraction of cells (Supplementary Fig. 6c). Knockdown of specific eRNAs by LNA or siRNA resulted in a decrease of cohesin recruitment (Fig. 4e) to enhancers in response to E2, with no significant alteration of the H3K4me1 mark (Fig. 4e), or ligand-dependent increase of ER-α recruitment (Supplementary Fig. 2d, e). Expression levels of cohesin subunits were not affected by knockdown of eRNAs (Supplementary Fig. 7a). siRNA-mediated depletion of RAD21 caused loss of enhancer–promoter interactions, both basal and E2 induced, when assessed by chromatin conformation capture (3C) assay for the *NRIP1* and *GREB1* loci (Fig. 4f). When we tested the role of cohesin in the oestrogen transcription program by GRO-seq, we noted that si*SMC3* caused a broad inhibition of coding gene activation by E2 (Fig. 4g and Supplementary Fig. 7e, f), with only ~34% of E2-upregulated genes remaining induced (Supplementary Fig. 7g). Similarly, RAD21 knockdown inhibited E2 induction of genes, as revealed by the five targets evaluated (Supplementary Fig. 7d). We excluded alterations in levels of ER-α as the cause for these marked effects of cohesin depletion (Supplementary Fig. 7b, c). On the basis of these results, we speculate that many regulatory genomic regions, such as enhancers, harbour the cohesin complex, which 'poises' the enhancer for the stable eRNA-induced looping necessary for gene activation events. However, we cannot exclude the possibility that the role of cohesin could also reflect non-enhancer-based regulation.

Despite the discovery of enhancers more than 35 years ago[1,2], a full understanding of the mechanisms by which they regulate gene expression has been difficult to achieve. We have provided several lines of evidence that induced eRNA transcripts are functionally important for the actions of oestrogen-regulated gene enhancers, at least in part by contributing to the dynamic generation or stabilization of enhancer–promoter looping between the regulated coding transcription units and these ER-α-bound enhancers.

METHODS SUMMARY

MCF-7 cells were initially obtained from ATCC, maintained in culture and treated as described previously[12]. They were hormone stripped for 3 days and treated or untreated with 100 nM oestradiol for 1 h to induce oestrogen target gene expression. Custom siRNAs were designed in-house and chemically synthesized by Bioneer and Sigma-Aldrich, whereas LNAs were designed and synthesized by Exiqon. Knockdown experiments with either siRNAs or LNAs were performed as transient transfections using Lipofectamine 2000, as per the manufacturer's instructions (Invitrogen). For siRNAs, hormone-stripped cells were subjected to two rounds of transfection and then treated with either vehicle or E2 for 1 h; for LNAs, cells were cultured for 24 or 48 h after transfection and exposed to the same treatment as described earlier. Real-time qPCR was carried out as previously described[9,12], normalized to either *ACTD* or *GAPDH*. FISH experiments were carried out as detailed previously[29], using commercially available bacterial artificial chromosome (BAC) probes obtained from Empire Genomics (Supplementary Table 1). The high-throughput sequencing libraries were prepared as per Illumina's HiSeq 2000 library reagent kit. Global run-on sequencing experiments (GRO-seq) were performed as previously reported[9,30]. 3D-DSL used a conventional 3C step for ~200 kb surrounding each interrogated E2-regulated gene TSS, and was performed as previously described[22], with all donor and acceptor probes designed using HindIII restriction sites. ChIP-Seq for ER-α, H3K4me1, H3K27ac and RAD21 were performed as described[9]. Detailed descriptions of bioinformatic analyses are provided in Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 21 May 2012; accepted 22 April 2013.
Published online 2 June 2013.

1. Newman, J. J. & Young, R. A. Connecting transcriptional control to chromosome structure and human disease. *Cold Spring Harb. Symp. Quant. Biol.* **75,** 227–235 (2010).
2. Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144,** 327–339 (2011).
3. Ong, C. T. & Corces, V. G. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature Rev. Genet.* **12,** 283–293 (2011).
4. Guttman, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482,** 339–346 (2012).
5. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43,** 904–914 (2011).
6. Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nature Rev. Genet.* **10,** 155–159 (2009).
7. Kim, T.-K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465,** 182–187 (2010).
8. Hah, N. *et al.* A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* **145,** 622–634 (2011).
9. Wang, D. *et al.* Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474,** 390–394 (2011).
10. Welboren, W. J. *et al.* ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* **28,** 1418–1428 (2009).
11. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nature Genet.* **38,** 1289–1297 (2006).
12. Kwon, Y. S. *et al.* Sensitive ChIP-DSL technology reveals an extensive estrogen receptor α-binding program on human gene promoters. *Proc. Natl Acad. Sci. USA* **104,** 4852–4857 (2007).
13. Heintzman, N. D. & Ren, B. Finding distal regulatory elements in the human genome. *Curr. Opin. Genet. Dev.* **19,** 541–549 (2009).
14. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459,** 108–112 (2009).
15. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107,** 21931–21936 (2010).
16. Ahlenstiel, C. L. *et al.* Direct evidence of nuclear Argonaute distribution during transcriptional silencing links the actin cytoskeleton to nuclear RNAi machinery in human cells. *Nucleic Acids Res.* **40,** 1579–1595 (2012).
17. Mayer, C., Schmitz, K. M., Li, J., Grummt, I. & Santoro, R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell* **5,** 351–361 (2006).
18. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472,** 120–124 (2011).
19. Melo, C. A. *et al.* eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell* **49,** 524–535 (2013).
20. Lai, F. *et al.* Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494,** 497–501 (2013).
21. Fullwood, M. J. *et al.* An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462,** 58–64 (2009).
22. Harismendy, O. *et al.* 9p21 DNA variants associated with coronary artery disease impair interferon-γ signalling response. *Nature* **470,** 264–268 (2011).
23. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489,** 109–113 (2012).
24. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–293 (2009).
25. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44,** 667–678 (2011).
26. Hadjur, S. *et al.* Cohesins form chromosomal *cis*-interactions at the developmentally regulated *IFNG* locus. *Nature* **460,** 410–413 (2009).
27. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467,** 430–435 (2010).
28. Schmidt, D. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20,** 578–588 (2010).
29. Cai, S. & Kohwi-Shigematsu, T. Intranuclear relocalization of matrix binding sites during T cell activation detected by amplified fluorescence *in situ* hybridization. *Methods* **19,** 394–402 (1999).
30. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322,** 1845–1848 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank K. Hutt for help with statistical analyses; M. Ghassemian from the University of California, San Diego (UCSD) Biomolecular/Proteomics Mass Spectrometry Facility for assistance with mass spectrometry; C. Nelson for cell culture assistance; J. Hightower for assistance with figure and manuscript preparation. We thank H. Chang for providing the *BoxB*, λN–GAL4 constructs. We acknowledge the UCSD Cancer Center Specialized Support Grant P30 CA23100 for confocal microscopy. W.L. and D.N. are supported by Department of Defense (DoD) postdoctoral fellowships, BC110381 and BC103858, respectively. M.G.R. is an investigator with the Howard Hughes Medical Institute. This work was supported by grants DK 039949, DK018477, NS034934, HL065445, CA173903 to C.K.G., and from the DoD.

Author Contributions M.G.R., W.L., D.N., E.N. and C.K.G. conceived the project. W.L. and D.N. performed most of the experiments reported, with contributions from E.N. and A.Y.C. (FISH). Q.M., B.T. and D.M. performed bioinformatic analyses. Q.M., E.N. and B.T. made equivalent contributions to this study. Additional experiments/methods were contributed by X.S., S.O. and H.-S.K. J.Z. and K.O. assisted in deep-sequencing library preparations and sequencing. W.L., D.N. and M.G.R. wrote the final paper with input from C.K.G.

Author Information The sequencing data sets are deposited in the Gene Expression Omnibus database under accession GSE45822 Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G.R. (mrosenfeld@ucsd.edu).

METHODS

Antibodies. The antibodies used in this study were: anti-ER-α (HC-20, Santa Cruz); anti-H3K4me3 (07-473, Santa Cruz); anti-H3K4me1 (ab8899, Abcam); anti-H3K27ac (ab4729, Active Motif); anti-RAD21 (ab992, Abcam); anti-SMC3 (ab9263, Abcam); anti-α-tubulin (T5168, Sigma), anti-GAL4 (DNA-binding domain (DBD)) (06-262, Millipore); and anti-IgG (I5006, Sigma).

Cell culture. MCF-7 cells obtained from ATCC were cultured in DMEM media supplemented with 10% FBS in a 5% CO_2 humidified incubator. They were hormone stripped for 3 days in phenol-free media with charcoal-stripped FBS before receiving 100 nM E2 (Sigma) or ethanol treatment for 1 h for oestrogen signalling induction. MCF10A cells were a gift from B. H. Park and were essentially grown as described previously[31]. For E2 induction of MCF10A, the culture media was stripped of EGF.

siRNA and LNA transfections. LNAs were obtained from Exiqon; siRNAs were from Bioneer and Sigma-Aldrich (Supplementary Tables 2 and 3). For transfection of both siRNA and LNAs, cells were first hormone stripped for 1 day followed by siRNA/LNA (both at 40 nM) transfection using Lipofectamine 2000. After 2 days they were then treated with ethanol or E2 for 1 h. For some experiments, transfections were performed twice to achieve higher efficiency. Similarly, LNA transfections were performed 2 days after starvation in stripped media, and thus the LNA treatment lasted 24 h in some experiments.

RT–qPCR. RNA was isolated using Trizol (Invitrogen) or RNeasy column (Qiagen), and total RNA was reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen). Quantitative PCRs were performed mostly with StepOne Plus (Applied Biosystem). For normalization, ΔC_t values were calculated relative to the levels of *ACTB*/*GAPDH* transcripts. The experiments were repeated at least three times, and one representative plot is shown in figures; most *P* values were obtained using a two-tailed Student's *t*-test. Primers are listed in Supplementary Table 4.

ChIP-Seq. ChIP was performed as previously described[9]. Briefly, approximately 10^7 cells were cross-linked with 1% formaldehyde at room temperature (~25 °C) for 10 min and neutralized with 0.125 M glycine. After sonication, ~75 μg soluble chromatin was incubated with 1–5 μg of antibody at 4 °C overnight. Immunoprecipitated complexes were collected using Dynabeads A/G (Invitrogen). Subsequently, immuno-complexes were washed, and DNA was extracted and purified by QIAquick Spin columns (Qiagen). For ChIP-Seq, the extracted DNA was ligated to specific adaptors followed by deep sequencing with the Illumina's HiSeq 2000 system according to the manufacturer's instructions. Usually, the first 48 bp for each sequence tag returned by the Illumina Pipeline was aligned to the hg18 assembly using BFAST or Bowtie2. Only uniquely mapped tags were selected for further analysis. The data was visualized by preparing custom tracks on the UCSC genome browser using HOMER[32] (http://biowhat.ucsd.edu/homer). The total number of mapable reads was normalized to 10^7 for each experiment presented in this study.

Identification of ChIP-Seq peaks. The ChIP-Seq peaks were identified by HOMER. Given different binding patterns of transcription factors and histones, parameters were optimized for the narrow tag distribution characteristic of transcription factors by searching for high read-enrichment regions within a 200-bp sliding window. Regions of maximal density exceeding a given threshold were called as peaks, and adjacent peaks were set to be >500 bp away to avoid redundant detection. The common artefacts from clonal amplification were circumvented by considering only one tag from each unique genomic position. The threshold was set at a false discovery rate (FDR) of 0.001 determined by peak finding using randomized tag positions in a genome with an effective size of 2×10^9 bp. For ChIP-Seq of histone marks, seed regions were initially found using a peak size of 500 bp (FDR <0.001) to identify enriched loci. Enriched regions separated by <1 kb were merged and considered as blocks of variable lengths. All called peaks were then associated with genes by cross-referencing with the RefSeq TSS database. Peaks from individual experiments were considered overlapping if their peak centres were located within 200 bp (for some analyses the distance between them could extend to 1 kb). The peaks within ±1 kb apart from the RefSeq gene TSS site were considered to be promoter bound.

GRO-seq. GRO-seq experiments were performed as previously reported[9,30,33]. Briefly, MCF-7 cells were swelled in swelling buffer (10 mM Tris-Cl pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$) for 5 min on ice and then lysed in lysis buffer (swelling buffer with 0.5% IGEPAL and 10% glycerol) before being finally re-suspended in 100 μl of freezing buffer (50 mM Tris-Cl pH 8.3, 40% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA). For the run-on assay, re-suspended nuclei were mixed with an equal volume of reaction buffer (10 mM Tris-Cl pH 8.0, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 300 mM KCl, 20 units of Superase-In, 1% sarkosyl, 500 μM ATP, GTP, Br-UTP and 2 μM CTP) and incubated for 5 min at 30 °C. The nuclear-run-on RNA (NRO-RNA) was then extracted with TRIzol LS reagent (Invitogen) following the manufacturer's instructions. After base hydrolysis on ice for 40 min and followed by treatment with DNase I and antarctic phosphatase, the Br-UTP-labelled NRO-RNA was purified by anti-BrdU argarose beads (Santa Cruz Biotech) in binding buffer (0.5× SSPE, 1 mM EDTA, 0.05% Tween) for 3 h at 4 °C while rotating. Then T4 PNK (NEB) was used to repair the end of NRO-RNA. Subsequently, complementary DNA synthesis was performed as reported[9,33] with few modifications. The RNA fragments were subjected to the poly-A-tailing reaction by poly-A polymerase (NEB) for 30 min at 37 °C. Reverse transcription was then performed using superscript III (Invitrogen) with oNTI223 primer (for sequence see Supplementary Table 5). The cDNA products were separated on a 10% polyacrylamide TBE-urea gel with the right product (~100–500 bp) being excised and recovered by gel extraction. After that, the first-strand cDNA was circularized by CircLigase (Epicentre) and re-linearized by Ape1 (NEB). Re-linearized single-strand cDNA was separated by TBE gel and the products of the desired size were excised (~120–320 bp) for gel extraction. Finally, the cDNA template was amplified by PCR using the Phusion High-Fidelity enzyme (NEB) with primers oNTI200 and oNTI201 for deep sequencing (primers listed in Supplementary Table 5).

Computational analysis of GRO-seq. The sequencing reads were aligned to hg18 using Bowtie2. For analysing oestrogen effects on gene transcription, we counted the reads from the first 30 kb (assuming an RNA polymerase speed of ~0.5 kb min^{-1} during 1 h E2 treatment) of the entire gene body, excluding the promoter-proximal region on the sense strand with respect to the gene orientation by using BED Tools or HOMER. EdgeR (http://www.bioconductor.org/) was used to compute the significance of the differential gene expression (FC ≥ 1.5, FDR ≤ 0.01). Additionally, a read density threshold (that is, normalized GRO-seq read counts per kb) was used to exclude lowly expressed genes.

***De novo* identification of GRO-seq transcripts.** GRO-seq read densities were analysed in a similar manner to ChIP-Seq. Provided that GRO-seq generated strand-specific data, separate tracks were uploaded onto the UCSC genome browser; tag-enriched sites were identified using a sliding window of 250 bp. Transcript initiation sites were identified as regions where the GRO-seq read density was increased threefold relative to the preceding 1 kb region. Transcript termination sites were defined by either a reduction in reads below 10% as compared to that of the TSS or when another transcript's start was identified on the same strand. Individual high-density peaks spanning a region less than 250 bp were considered artefacts and removed from the analysis. Transcripts were defined as putative eRNAs if their *de novo* called start sites was located distal to the RefSeq TSS (≥3 kb) and were associated with ER-α and H3K27ac co-bound regions.

Bioinformatics characterization of ER-α enhancers. The ER-α-H3K27ac co-bound regions are defined as those in which the distance from the centre of an ER-α peak to the H3K27ac peak-occupied region is ≤1 kb. Overall, two methods were used to assign the ER-α-bound enhancers to E2-upregulated genes: (1) identifying the E2-upregulated coding genes from GRO-seq first and then coupling each of them to their closest ER-α-H3K27ac co-bound enhancer within a certain distance (200 kb) (a 'gene-centric' view); and (2) characterizing the ER-α-H3K27ac co-bound enhancers first and then coupling each of them to their closest TSS that belongs to 1,309 E2-upregulated coding genes (an 'enhancer-centric' view). The comparison of ChIP-Seq tag intensity, GRO-seq transcription levels or distances between different categories (Figs 1 and 4 and Supplementary Fig. 1) are presented as boxplots by using either log or normal scales. The *P* values were determined by two-tailed Student's *t*-test.

ChIRP-Seq. The ChIRP experiment was performed essentially as per the original protocol[25], except for a few modifications. First, we designed antisense DNA probes targeting *FOXC1* eRNA ('odd' and 'even') (~40-base oligonucleotide) based on high oligonucleotide specificity (using BLAST and BLAT), moderate GC content (40–60%) and a T_m around 65 °C, with probes for *lacZ* RNA as control (all probes listed in Supplementary Table 8). All DNA probes were biotinylated and purified using the Label IT Nucleic Acid Labelling Kit (Mirus Bio). The sequencing reads were aligned to hg18 by Bowtie2 and the peaks were called by HOMER if they fulfilled three criteria: (1) they were consistently called in both the 'odd' and 'even' ChIRP-Seqs; (2) they did not intersect with the peaks in the ChIRP-Seq for *lacZ* RNA sample; and (3) they did not intersect with the satellite repeats or retrotransposon sequences. The remaining ChIRP peaks were divided into two categories: (1) highly confident peaks (peak score >8); and (2) weak peaks (peak score ≤ 8). The peaks were extended with 1 kb for intersection analysis by using BedTools, and the peak annotation was carried out in HOMER.

RNA copy number quantification. To quantify each transcript, the PCR product using the qPCR primers for the transcript was purified and the concentration was measured. The absolute copy numbers of the PCR product were calculated as per the following. For example, for *GAPDH*, the number of single-stranded (ss)DNA molecules from 1 μl of the 17 ng μl^{-1} PCR product of the *GAPDH* fragment (142 bp) with 87,788.56 Da molecular weight is about $2 \times (17 \times 10^{-9} \times 6.023 \times 10^{23})/87{,}788.56 = 2.32 \times 10^{11}$. The number of ssDNA molecules from 1 μl of

the 16 ng μl^{-1} PCR product of the *TFF1e* fragment (82 bp) with 50,696.92 Da molecular weight is about $2 \times (16 \times 10^{-9} \times 6.023 \times 10^{23})/50{,}696.92 = 3.8 \times 10^{11}$. Using these PCR products with known molecule copy numbers, standard curves can be generated by qPCR, which forms the basis of the quantification of the number of copies of eRNAs from cDNA samples. For cDNA samples, 3 μg of total RNA (which is, according to the QIAgen manual, $\sim 2 \times 10^5$ cells) were converted into 20 μl cDNA. During multiple qPCR experiments using cells from different batches, the cycle number of target eRNA being amplified will vary within 2–3 cycles (~4–8 fold). The copy number of *GAPDH* mRNA is largely consistent with previous reports[34]. Considering that the efficiency of reverse transcriptase on *GAPDH* mRNA is estimated to be ~50% (ref. 35), which might be even lower for eRNAs, the real numbers of eRNA copies could be higher than the estimation.

Immunoprecipitation. Cells were collected with cold PBS and lysed with RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate, 0.5 mM DTT, protease inhibitor). The lysate was diluted 2–4 times with dilution buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.1% NP-40 and 10% glycerol, protease inhibitor). 2–5 μg of antibodies were added into the diluted cell lysate and incubated overnight at 4 °C. The next day, the protein complexes were collected by magnetic Dynabeads G for 2 h at 4 °C with rotation. The beads–antibody–protein complexes were then washed four times with wash buffer (50 mM Tris pH 7.4, 125 mM NaCl, 1 mM EDTA and 0.1% NP-40) and boiled for western blot analysis.

RIP and IVT RNA pull-down. The RIP experiment was done largely as per a previous protocol[36]. Briefly, cells were cross-linked with 0.3% formaldehyde for 10 min at 37 °C. 2.5 M glycine was added (1/20 of the medium volume) to neutralize for 10 min at room temperature. The cell pellet was re-suspended in 0.6 ml of RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate, 0.5 mM DTT, protease inhibitor and Superase-In 40 units per ml), sonicated once and incubated on ice with frequent vortex for 25 min. Subsequently, the supernatant was diluted with RIP dilution buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP40) and pre-cleared with ~25 μl protein A sepharose slurry for 30 min at 4 °C. Antibodies were added and incubated overnight at 4 °C with rotation. The next day, the RNA–protein complex was collected using pre-washed ~60 μl protein A sepharose beads for 1.5–2.5 h at 4 °C. After washing in RIPA buffer and RIPA-500 buffer (RIPA with higher salt: 500 mM NaCl), the beads were re-suspended in 150 μl of RIPA buffer with proteinase K at 45 °C for 45 min. RNA was extracted with TRIzol followed by DNaseI digestion. Reverse transcription was performed with SuperScript III RT kit (Invitrogen). For RIP-qPCRs, the amount of RNA in pull-down samples was calculated as the percentage of input *GAPDH* RNA of its respective group. The assay was repeated at least two times but was presented as a representative plot. *P* values were obtained using two-tailed Student's *t*-test.

For IVT RNA pull-down, first, plasmids carrying DNA sequences of the eRNA being investigated were linearized and *in vitro* transcribed using MEGAtranscript kit (Ambion) with 25% of UTP being replaced by biotinylated UTP (Ambion). About 10 μg of biotinylated RNA was heated to 90 °C for 3 min, put on ice for 2 min and added into RNA structure buffer (10 mM Tris pH 7.2, 0.1 M KCl, 10 mM $MgCl_2$, 1 μl Superase-In) for 20 min to form a structure. The biotinylated RNA was then mixed with pre-washed Streptoavidin magnetic beads and incubated at room temperature for 30 min to conjugate the RNA with the beads, following the manufacturer's protocol. After that, ~10 mg nuclear extract in RIP buffer (20 mM Tris pH 7.4, 1 mM EDTA, 150 mM NaCl, 0.5 mM DTT, 0.1% NP-40, 5% glycerol, protease inhibitor and Superase-In 40 units per ml) was then mixed with biotinylated RNA and incubated at 4 °C for 4 h. After being washed four times in high salt buffer (20 mM Tris pH 7.4, 500 mM NaCl, 0.05% Triton X-100, 10 units per ml Superase-In), the beads were boiled for western blots.

DNA enzyme digestion methylation assay. The protocol largely follows a previous paper[17]. Briefly, genomic DNA was extracted from MCF-7 cells (AccuPrep, Bioneer) 24 h after transfection with LNA/siRNA against eRNAs, and was digested with HpaII (*NRIP1e*) or HhaI (*FOXC1e*), both from NEB, before qPCR amplification using primers (Supplementary Table 4) that spanned the enzyme digestion sites (Supplementary Fig. 3a). The relative resistance to restriction digestion was calculated by dividing the amount of DNA that remained after digestion by the amount before digestion.

Methyl miner assay. MCF-7 cells were transfected with LNA or siRNA for 24 h, after which they were subjected to DNA isolation using a QIAgen DNA isolation column. 1 μg of each DNA was used for Biotin-tagged-MBD peptide pull-down as per the manufacturer's protocol (Invitrogen), after which unmethylated and methylated DNA fractions were collected, purified and subjected to qPCR analysis using the primers specified (Supplementary Table 4).

FISH and imaging. The cells were processed for DNA Immuno-FISH essentially as described previously[29], with BAC probes from Empire Genomics (Supplementary Table 1). MCF-7 cells were treated with ethanol or E2, grown on acid-washed poly-lysine-coated coverslips, were washed with 1× PBS and immediately fixed with freshly made 4% paraformaldehyde/PBS for 10 min. Permeabilization was achieved by incubating in PBS containing 0.5% Triton X-100 for 15 min. FISH pre-hybridization treatments include incubating the coverslips in 0.1 N HCl for 5 min at room temperature, followed by digestion with 0.01 N HCl/0.002% pepsin for 5 min at 37 °C, stopped by 50 mM $MgCl_2$/PBS and equilibrated in 50% formamide/2× SSC 2 h before hybridization. 5 μl of probe/hybridization buffer mix was used per coverslip, with a hybridization programme of 76 °C for 3 min followed by overnight hybridization at 37 °C in a humidified dark chamber. The coverslips were then washed with pre-warmed WS1 (0.4× SSC/0.3% NP-40), WS2 (2× SSC/0.1% NP-40) and PBS, before being finally mounted with prolong gold-DAPI anti-fade mounting reagent (Invitrogen).

For FISH Image acquisition and data analysis[37], images were acquired using the Leica SP5 II confocal microscopy (×63 objective lens) with a resonance scanner. Z-stack data acquisition was set up across 3.2 μm thickness at 0.4 μm each step (9 steps for each three-dimensional image set). The three-dimensional images were then generated in Volocity (v.6.0.1). The FISH-positive gene loci were identified using the "Find Object Using % Intensity" (generally >20%) function in combination with "Exclude Objects by Size" (generally >0.1 μm^3). The overlap between two FISH-positive gene loci was calculated by the function "Intersect" with size exclusion (>0.03 μm^3). The cells counted ($n > 100$ for each group; Supplementary Fig. 5) were from eight images/fields; the percentage of overlapping events from each one was calculated separately, which together generates the mean and s.d.

***BoxB*-λN tethering assay.** Similar to the previous method[18], as described in Fig. 3a and Supplementary Fig. 8, the *BoxB* tethering system uses viral RNA–protein interactions, in which *BoxB* is a viral RNA that can be recognized and bound by viral anti-terminator protein λN. Fusion of *FOXC1* eRNA with *BoxB* enables the fused *BoxB-FOXC1e* to be bound by λN. Subsequently, λN protein was fused to the DNA-binding domain (DBD) of GAL4, which then recognizes 5×UAS sites on the reporter plasmid DNA. Using this technique, *BoxB*-eRNA can be tethered to the 5×UAS sites on a reporter plasmid with the help of the λN–GAL4 fusion protein[18]. Full-length *FOXC1* eRNA was cloned in pCDNA3.1 downstream to five copies of *BoxB*. This construct was co-transfected along with the reporter plasmids and λN–GAL4 vector (Supplementary Fig. 8), which is also based on a pCDNA3.1 vector with CMV promoter.

FOXC1 promoter was cloned in KpnI and BglII sites in pGL3-basic vector, 5×UAS sites were cloned at upstream SalI site in pGL3-basic vector, *FOXC1* full-length enhancer (1.2 kb) was placed just upstream to 5×UAS sites at the BamHI site. For deletion of the sense eRNA, the enhancer region was amplified including the full antisense transcript, the core region and 20 nucleotides from the sense eRNA of the *FOXC1* enhancer (thus called *FOXC1* enh-sense del enhancer, Fig. 3c) was also cloned at BamHI site upstream to 5×UAS site (Supplementary Fig. 8).

Luciferase reporter assay. Tethered plasmids alone or in combination were transfected along with Renilla-TK plasmid into MCF-7 cells that had been hormone stripped for 3 days. Six hours post-transfection, they were treated with 10 nM E2 for 24 h further, and then they were subjected to the luciferase assay using the Dual-Lucifersae reporter assay kit (Promega); plates were read in Veritas Microplate Luminometer (Turner Biosystems).

3C. 3C was performed as previously described[22,24]. Briefly 25×10^6 MCF-7 cells were fixed by adding 1% formaldehyde at room temperature for 10 min, and the reaction was stopped by glycine. Lysis buffer (500 μl 10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% Igepal CA630, protease inhibitors) was added and cells were incubated on ice. Next, cells were lysed with a Dounce homogenizer, and the suspension was spun down at 5,000 r.p.m. at 4 °C. The supernatant was discarded and the pellet was washed twice with 500 μl ice-cold 1× NEBuffer 2 (NEB). The pellet was then re-suspended in 1× NEBuffer 2 and split into five separate 50 μl aliquots. The extracted chromatin was then digested overnight with 400 units HindIII (NEB). Each digested chromatin mixture was ligated by T4 DNA Ligase (800 units) in 20 times the initial volume for 4 h at 16 °C. The ligase step was omitted in one chromatin aliquot from the five mentioned above as the unligated control. The chromatin was subsequently de-cross-linked overnight at 65 °C and purified twice with phenol and then with a mixture of phenol, chloroform and IAA (at a ratio of 25:24:1). DNA was precipitated and pellets were air-dried before re-suspending in 250 μl 1× TE buffer. To degrade any carryover RNA, 1 μl RNase A (1 mg ml^{-1}) was added to each tube and incubated at 37 °C for 15 min. DNA was further purified using phenol/chloroform/IAA and precipitated. The digestion and ligation efficiencies were checked and normalized before 3D-DSL.

Probe design for 3C. Donor and acceptor probes were designed on HindIII sites covering the enhancers and the gene body of the following genes: *GREB1*, *NRIP1*,

KCNK5 and *P2RY2*; by using custom Perl scripts (available upon request). The chosen regions for probe design covered the most prominent ER-α binding sites as well as enhancers. The uniqueness of the probe sequences was verified by Bowtie alignment to the human genome hg18 assembly. Universal adaptor sequences that are compatible with HiSeq 2000 flow cell design were added to the probe ends for bridge amplification of the ligation products and for direct sequencing. Acceptors were phosphorylated and both acceptors and donors were pooled individually in equimolar amounts for 3D-DSL (Supplementary Table 6).

3D-DSL. The DSL ligation products were prepared as described previously[12,22]. 3D-DSL was performed as described previously[22]. Briefly, after 3C efficiency estimation, equal amount of 3C chromatin was biotinylated using the Photoprobe Kit (Vector Lab). Donor and acceptor probe pools (20 fmol per probe) were annealed to the biotinylated 3C samples at 45 °C for 2 h followed by 10 min at 95 °C. The biotinylated DNA was immunoprecipitated with magnetic beads conjugated to streptavidin, and during this process unbound oligonucleotides were removed by stringent washes. The 5′-phosphate of acceptor probes and the 3′-OH of donor probes were ligated using Taq DNA ligase at 45 °C for 1 h. These ligated products were washed and eluted from beads and then amplified by PCR using primers A and B-AD (or primer B-BC1 and B-BC2 if bar coding was used) for deep sequencing on the Illumina HiSeq 2000, using primer A as the sequencing primer.

3D-DSL data analysis. After removing the adaptor sequences, the reads are aligned to a custom library that includes all the combinations of donors and acceptors. The alignment was performed with Novoalign, and the reads were counted for every possible interaction by using custom Perl scripts (available upon request). The reads that were generated by donor–acceptor ligations on the same restriction site were removed: the remaining number of reads included both intra- and inter-chromosomal interactions. We used the median value (~6 million) of all the samples from the same sequencing run for normalization; the reads accounting for ligation products in unligated controls were subtracted. In addition to standard tools, such as my5C[38] and HiTC[39], we used an intensity-based method to characterize the set of interactions. A related method was also used in 3C-seq procedures[40]. A *P* value is assigned to an interaction based on the Poisson probability distribution function, $p(x) = e^{-\lambda} \times \lambda^x/x!$, where $p(x)$ is the probability of an interaction, x is the interaction intensity, and λ is the average interaction intensity considering all the potential interactions in the library; that is, the ratio (total number of usable reads)/(total number of all possible interactions given the set of acceptors and the set of donors). The *P* values were corrected for multiple testing by using the Bonferroni correction method. In addition, for each interaction, we define supplementary parameters, such as (1) fold enrichment over Poisson's λ, and (2) a fold enrichment over background (where the background represents the average intensity of the ligations between the probes on the neighbouring restriction sites). We consider the interactions that meet the following criteria significant for downstream analyses: (1) a (corrected) *P* value <0.01; and (2) a fold enrichment over background >2 (although for display purposes in the figure plots we may also show the weak interactions).

To generate 3D-DSL plots Matlab was used; a 10-kb window was used to bundle the interactions, except for a 20-kb window for *NRIP1*. The interactions were plotted using a Bezier curve between the two positions with the third point in the middle of the positions with the *y*-axis corresponding to the $\log_{10}$ intensity. For example, if the two *x*-axis positions are 1 and 2, and the intensity is 4, a Bezier curve is drawn between (1,0),(1.5,4), and (2,0). The peak locations were then added on the bottom of the plot as stated in the legend. Interactions at distances generally <10 kb were not plotted for the *NRIP1* locus.

31. Abukhdeir, A. M. *et al.* Physiologic estrogen receptor α signaling in non-tumorigenic human mammary epithelial cells. *Breast Cancer Res. Treat.* **99,** 23–33 (2006).
32. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38,** 576–589 (2010).
33. Ingolia, N. T. *et al.* Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324,** 218–223 (2009).
34. White, A. K. *et al.* High-throughput microfluidic single-cell RT-qPCR. *Proc. Natl Acad. Sci. USA* **108,** 13999–14004 (2011).
35. Zhong, J. F. *et al.* A microfluidic processor for gene expression profiling of single human embryonic stem cells. *Lab Chip* **8,** 68–74 (2008).
36. Tsai, M. C. *et al.* Long non-coding RNA as modular scaffold of histone modification complexes. *Science* **329,** 689–693 (2010).
37. Rueden, C. T. *et al.* Visualization approaches for multidimensional biological image data. *Biotechniques* **43,** 31–36 (2007).
38. Lajoie, B. R. *et al.* My5C: web tools for chromosome conformation capture studies. *Nature Methods* **6,** 690–691 (2009).
39. Servant, N. *et al.* HiTC: exploration of high-throughput 'C' experiments. *Bioinformatics* **28,** 2843–2844 (2012).
40. Stadhouders, R. *et al.* Dynamic long-range chromatin interactions control *Myb* proto-oncogene transcription during erythroid development. *EMBO J.* **31,** 986–999 (2011).