PLOS GENETICS

RESEARCH ARTICLE

Ligand dependent gene regulation by transient ERα clustered enhancers

Bharath Saravanan[1,2], Deepanshu Soota[1], Zubairul Islam[1], Sudeshna Majumdar[3], Rajat Mann[1], Sweety Meel[1], Umer Farooq[1,4], Kaivalya Walavalkar[1], Srimonta Gayen[3], Anurag Kumar Singh[1], Sridhar Hannenhalli[5]*, Dimple Notani[1]*

1 Cellular Organization and Signalling, National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India, **2** School of Chemical & Biotechnology, SASTRA Deemed University, Thanjavur, Tamil Nadu, India, **3** Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India, **4** Centre for Functional Genomics and Bio-informatics, The University of Trans-Disciplinary Health Sciences and Technology, Bangalore, India, **5** Cancer Data Science Lab, National Cancer Institute, NIH, Bethesda, MD, United States of America

* sridhar.hannenhalli@nih.gov (SH); dnotani@ncbs.res.in (DN)

OPEN ACCESS

Citation: Saravanan B, Soota D, Islam Z, Majumdar S, Mann R, Meel S, et al. (2020) Ligand dependent gene regulation by transient ERα clustered enhancers. PLoS Genet 16(1): e1008516. https://doi.org/10.1371/journal.pgen.1008516

Editor: Ofir Hakim, Bar Ilan University, ISRAEL

Received: March 15, 2019

Accepted: November 12, 2019

Published: January 6, 2020

Data Availability Statement: The data has been deposited in GEO under accession number GSE136302, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136302.

Funding: DN is recipient of Wellcome/IA award (IA/14/2/501539) and was supported by the intramural funding from Tata Institute of Fundamental Research. Funder had no role in study design. SH is recipient of National Science Foundation award (NSF#1564785) and was partly supported by the Intramural Research Program of the National Cancer Institute, Center for Cancer Research, NIH. SM, UF and KW are supported by Council of

Abstract

Unliganded Estrogen receptor alpha (ERα) has been implicated in ligand-dependent gene regulation. Upon ligand exposure, ERα binds to several EREs relatively proximal to the pre-marked, unliganded ERα-bound sites and affects transient but robust gene expression. However, the underlying mechanisms are not fully understood. Here we demonstrate that upon ligand stimulation, persistent sites interact extensively, via chromatin looping, with the proximal transiently ERα-bound sites, forming Ligand Dependent ERα Enhancer Cluster in 3D (LDEC). The E2-target genes are regulated by these clustered enhancers but not by the H3K27Ac super-enhancers. Further, CRISPR-based deletion of *TFF1* persistent site disrupts the formation of its LDEC resulting in the loss of E2-dependent expression of TFF1 and its neighboring genes within the same TAD. The LDEC overlap with nuclear ERα condensates that coalesce in a ligand and persistent site dependent manner. Furthermore, formation of clustered enhancers, as well as condensates, coincide with the active phase of signaling and their later disappearance results in the loss of gene expression even though persistent sites remain bound by ERα. Our results establish, at TFF1 and NRIP1 locus, a direct link between ERα condensates, ERα enhancer clusters, and transient, but robust, gene expression in a ligand-dependent fashion.

Author summary

ERα occupies its cognate binding sites on the genome, upon estrogen stimulation, to bring about the signaling mediated response. However, ERα has been implicated in gene regulation even in the absence of hormone and its underlying mechanisms are not known. Here we show that, unliganded ERα-bound genomic sites are necessary for the emergence of active enhancers upon estrogen stimulation. Several ERα sites emerge in proximity to these unliganded ERα sites which loop with each other in three-dimensional nuclear space forming the enhancer:ERα clusters that we call Ligand dependent enhancer

Scientific and Industrial Research, India. DS, SM and RM are supported by the TIFR graduate program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No authors have competing interests

clusters (LDEC). These clusters are essential for expression of genes that are associated with the LDEC's as deletion of unliganded ERα-bound site completely perturbs the formation of clusters and expression of genes within the same TAD while clusters in neighboring TAD remain unaffected. At the mesoscale, these LDEC's appear as ERα condensates that emerge dynamically only upon addition of hormone. Further, the disappearance of clusters coincide with the loss of gene expression and condensate disappearance, marking the end of signaling response.

Introduction

Upon ligand stimulation ERα regulates gene expression by binding to distal regulatory elements [1,2,3]. Similarly, numerous studies have shown that unliganded nuclear ERα also binds to some specific sites in the genome [4,5,6,7,8] and these ERα-bound sites are specific as they exist even after prolonged deprivation (weeks) of ligand [8]. Knockdown of unliganded ERα inhibits subsequent ligand-dependent gene induction [8] confirming their role in gene regulation. Interestingly, unliganded ERα binding was recently shown to mark future functional enhancers, whereby these pre-bound ERα sites serve as seeds, around which multiple ERα-bound peaks emerge after ligand exposure [5]. Suggesting, the unliganded nuclear receptor recruitment on chromatin is required for the robust induction of ligand-dependent genes. However, mechanisms underlying these intriguing observations, and their role in gene regulation, are not entirely clear. For instance, it is not known whether unliganded pre-bound ERα sites are absolutely required for such ligand-dependent ERα-clustering at the genomic level and the downstream gene regulation. The estradiol (E2)-dependent gene expression peaks at 1 hour post stimulation and drops at 3 hours [3]. However, whether the transient response to signaling is driven by transient appearance of enhancers is not known.

Although it is known that ERα bound sites exhibit proximity, whether the proximity is ligand-dependent has been debated with evidences both for [1,9] and against [10], warranting additional studies. Furthermore, whether the constituent enhancers within clustered enhancers exhibit E2-dependent physical proximity to form a functional unit is not known.

Super-enhancers exhibit high density of coactivators caused by presence of multimeric enhancer units. Intrinsically disordered regions (IDR) in these coactivators mediate their molecular condensation on super enhancers [11]. ERα, with its co-activator MED1, was recently shown to form such condensates [12]. However, whether *bona fide* ERα clustered enhancers are tethered in such ERα condensates is not known. Importantly, whether enhancer condensates continue to exist during the entire interval when the ligand is present, or whether they exist only during the active signaling phase marked by robust transcription is not known.

Here we show that upon E2 treatment, as expected, ERα binds to numerous locations across the genome. As recently observed [5], the new ligand-dependent sites are significantly organized as clusters in relative proximity to pre-existing, or *persistent*, ligand-independent ERα-bound sites. However, we find that the ligand-dependent enhancer clusters (LDEC) are distinct from previously reported super-enhancers [13]. LDECs exhibit ligand induced extensive chromatin looping among constituent ERα sites within as well as across LDECs in 3D. Interestingly, LDECs may include multiple genes whose expressions are concomitant with the appearance of LDEC. We show via CRISPR deletion that persistent sites are absolutely required for ligand-induced binding of ERα to the neighboring EREs. Further, ERα forms condensates on ERα-clustered enhancers by coalescence upon E2 treatment as observed by ImmunoFISH experiments. Specifically, LDECs emerge robustly and transiently upon E2

treatment and their disappearance at 3 hours post-treatment coincides with the loss of eRNA at constituent enhancers and their cognate target gene expression as well as drop in ERα protein levels.

Overall, our results establish and clarify the role of unliganded ERα binding in priming enhancer clusters in 3D that drive transient, but robust, gene expression in a ligand-dependent fashion. Our work suggests a model of E2-induced gene regulation where during active phase, liganded ERα decorates the EREs closer to pre-marked unliganded ERα site. These pre-marked and new ERα-bound sites exhibit E2-dependent extensive chromatin looping, forming functional LDEC that drives gene expression. These LDECs correspond to ERα puncta in the nucleus that emerge upon E2 stimulation. Finally, upon ERα degradation at 3h or upon deletion of persistent site, these clusters disappear leaving persistent site behind still bound by ERα as bookmark, but that alone is unable to drive gene expression.

Results

ERα binds in clusters around pre-existing ERα-bound sites upon ligand stimulation

ERα binds predominantly in the intergenic regions of the genome upon E2 stimulation (S1A, S1B and S1C Fig). Further, we tested if ERα bound sites exhibit genomic clustering as has been reported recently[5]. Consecutive ERα peaks within 20kb were considered to be part of the same genomic cluster. We identified 1514 ERα clusters containing at least three ERα peaks in each cluster (see Methods). As a control, we repeated the clustering analysis based on hundred iterations of randomly selected 21,834 ERE motif, or DHS (Methods), and found that 7% of all EREs (482 clusters) and 22% of the DHS sites (1304 clusters) clustered compared with 30% of total ERα peaks (Fig 1A). These results strongly support the tendency of ERα peaks to cluster on the genome confirming the previous findings[5]. ERα clustering tendency was confirmed even at shorter distance of 5 and 10kb intervals. However, at distances longer than 20kb, we did not observe a consistent pattern of ERα clustering. ERα binding on specific sites in the genome has been described even in the absence of ligand [4,5,6,7,8]. To investigate if the existence of above found clusters is dependent on ligand stimulation, we compared the post-E2 ERα-bound peaks with those in non-stimulated MCF7 cells. We observed 6659 peaks in the untreated cells suggesting the unliganded nature of these peaks, of which 3779 peaks were also found in the post-E2 condition (Fig 1B); we refer to such sites as *persistent* sites, the 2880 lost sites as *-E2 unique* sites and the newly acquired 18,055 E2-dependent sites as *transient* sites.

Next, to test the authenticity of our observed -E2 specific sites we analysed publicly available ERα ChIP-seq data with and without ligand from different laboratories. Interestingly, although persistent sites were commonly found across datasets and exhibited robust gain of ERα upon ligand addition in these datasets, -E2 specific sites varied across the datasets; however, these sites consistently lost ERα binding upon ligand stimulation in all datasets (S1D Fig). However, persistent sites and -E2 unique sites exhibited comparable H3K27ac (S1E Fig). To further test the specificity of these peaks, we analyzed the ERα signal on persistent and–E2 specific 2880 sites upon ERα knockdown in unliganded condition [8]. The signals at persistent as well –E2 specific sites were significantly reduced upon knockdown of ERα (S1F and S1G Fig) suggesting that the ERα binding events in absence of ligand are not an artefact. Since, -E2 unique sites exhibited loss of ERα binding upon ligand stimulation, we focused only on persistent sites for their role in ligand dependent gene regulation. Next, we binned persistent sites into quantiles based on ERα binding strength post-E2 (Method). Persistent sites in general, and the 3rd quantile persistent sites in particular, exhibited stronger ERα binding as well as significantly greater DHS in both post and pre-E2 treated cells (Fig 1C; Methods). Further, to rule out insufficient

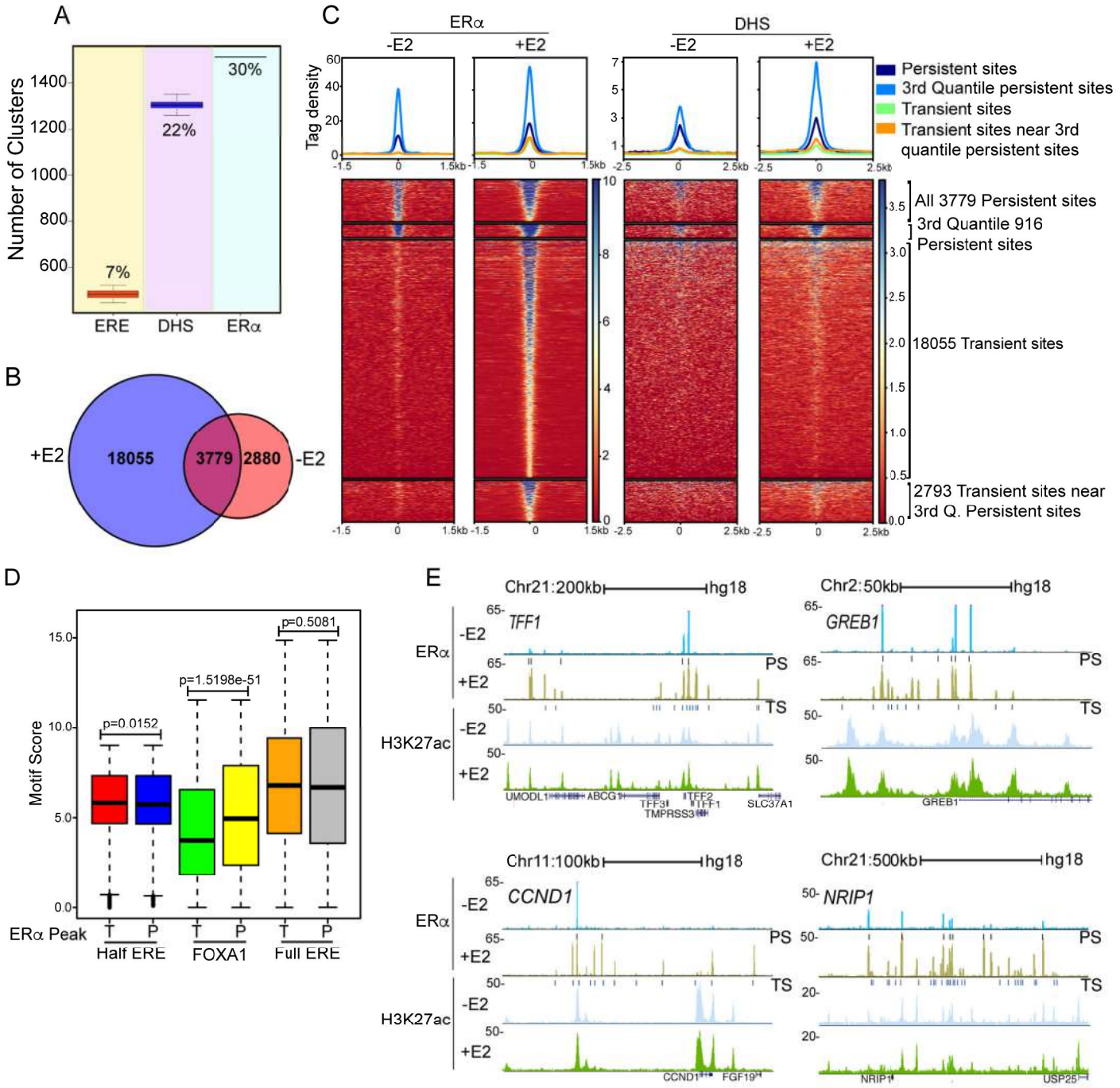

Fig 1. Estrogen receptor binds in clusters around pre-existing ERα-bound sites upon ligand stimulation. (A) ERα shows highest propensity for genomic clustering compared to ERE and DHS. Y-axis represents the number of clusters with at least 3 sites (ERα, ERE, or DHS individually) with less than 20kb between consecutive sites. (B) Venn diagram showing 3779 persistent ERα peaks being common in control and E2 treated cells. (C) Heat map exhibiting binding strength of ERα (left panel), and DHS (right panel) on all persistent, strongest persistent (3rd quantile), all transient, and transient sites clustered around a persistent site. (D) Motif enrichment plot showing relative scores of Half ERE, FOXA1 and Full ERE motifs in transient and persistent sites. (E) UCSC genome browser snapshots showing the binding of ERα and H3K27ac status on ERα clustered enhancers for four of the robustly E2-induced target genes–*TFF1*, *GREB1*, *NRIP1*, and *CCND1* in unliganded and in E2 conditions. The boxplots depict the minimum (Q1-1.5*IQR), first quartile, median, third quartile and maximum (Q3+1.5*IQR) without outliers.

https://doi.org/10.1371/journal.pgen.1008516.g001

ligand stripping as the potential cause for the 3779 persistent sites, we performed ERα ChIP-seq in cells stripped for seven days with daily media change. The persistent sites remained even in this stringent condition (S2A Fig). We also extracted chromatin fraction of seven days stripped cells in unstimulated and stimulated conditions (1, 3, and 24h) and found ERα bound to chromatin and the levels increased in response to E2 treatment (S2B Fig) confirming that these sites are bound by unliganded-ERα. After confirming the authenticity of unliganded ERα binding we pursued only persistent sites as they exhibited consistency across the data sets. We further probed the binding strength differences between persistent and transient sites via motif enrichment analyses as the motifs may guide these differences. As shown recently [5], both persistent and the transient sites were highly enriched for EREs (Figs 1D and S2C). Unlike what was reported in [5], we did not find a significant enrichment of full EREs in persistent sites relative to transient sites; however, there is a marginally significant but modest enrichment (ratio of means = 1.0119) of half ERE (Fig 1D). However, the only motifs enriched uniquely in the persistent sites were for FOX proteins where FOXA1 and FOXA2 motifs had the most significant p-values (Figs 1D and S2C). Intriguingly, -E2 unique sites are enriched only for FOXA1 motif but not for ERE motifs (S2C Fig). FOXA1 is a known pioneer *cis* co-factor for ERα binding [14,15,16] and is frequently mutated in breast tumors [17]. We confirmed the exclusive preference of FOXA1 binding on persistent sites (S2D Fig). These results indicate that FOXA1 bound sites along with classical ERE motifs are the hallmarks of persistent sites which may give rise to the clustered binding of ERα around these persistent sites on signaling. We found that about 50% of the above clusters had at least one persistent site, suggesting their potential role in seeding the genomic clusters of ERα sites upon E2 treatment. Further, clusters with persistent site exhibited high strength of ERα binding (S3A Fig) and a greater number of peaks (S3B Fig) as compared to the clusters without persistent site. We observed that most of the robust E2-induced genes are harbored around these clusters for example *CCND1*, *NRIP1*, *TFF1* and *GREB1* (Fig 1E). These clusters are marked by H3K27ac suggesting their potential enhancer activity (Fig 1E). Finally, relative to transient sites, persistent sites seem to be under a greater purifying selection as evidenced by higher cross-species conservation (Phast-cons scores, Methods) (S3C Fig). Furthermore, the ERα peaks within clusters exhibited an intriguing pattern, with the strongest binding at persistent sites showing a gradual decrease in ERα binding strength at consecutive ERE's as we go farther from the persistent site in either direction (S4A Fig). However, DHS signal on these transient sites remained unchanged (S4B Fig), suggesting a gradually decreasing sphere of influence of the persistent site.

ERα enhancer clusters but not the conventional super-enhancers regulate E2-dependent genes

To understand the role of identified ERα clusters in gene regulation, we compared the normalized GRO-seq tag counts of 4 closest genes from clusters with and without persistent site using publicly available GRO-seq in E2 time course [3]. Interestingly, genes closer to clusters with persistent site were robustly induced upon E2 treatment as compared to genes closer to the clusters lacking a persistent site (S5A Fig). Further, genes closer to persistent sites were in general highly expressed even in the absence of ligand (S5B Fig, first bar) however, we observed a slight down regulation of genes closer to -E2 unique sites (S5B Fig, last two bars). Note that ERα clusters share salient properties with super-enhancers, namely, target gene expression and genomically clustered constituent enhancers [13], raising the possibility that ERα clustered enhancers correspond to super-enhancers. Using the ROSE tool [13] to detect super-enhancers, we identified 858 H3K27ac-based super-enhancers and 390 ERα binding strength-based super-enhancers in E2 condition. Fig 2A shows a substantial overlap in ERα signal

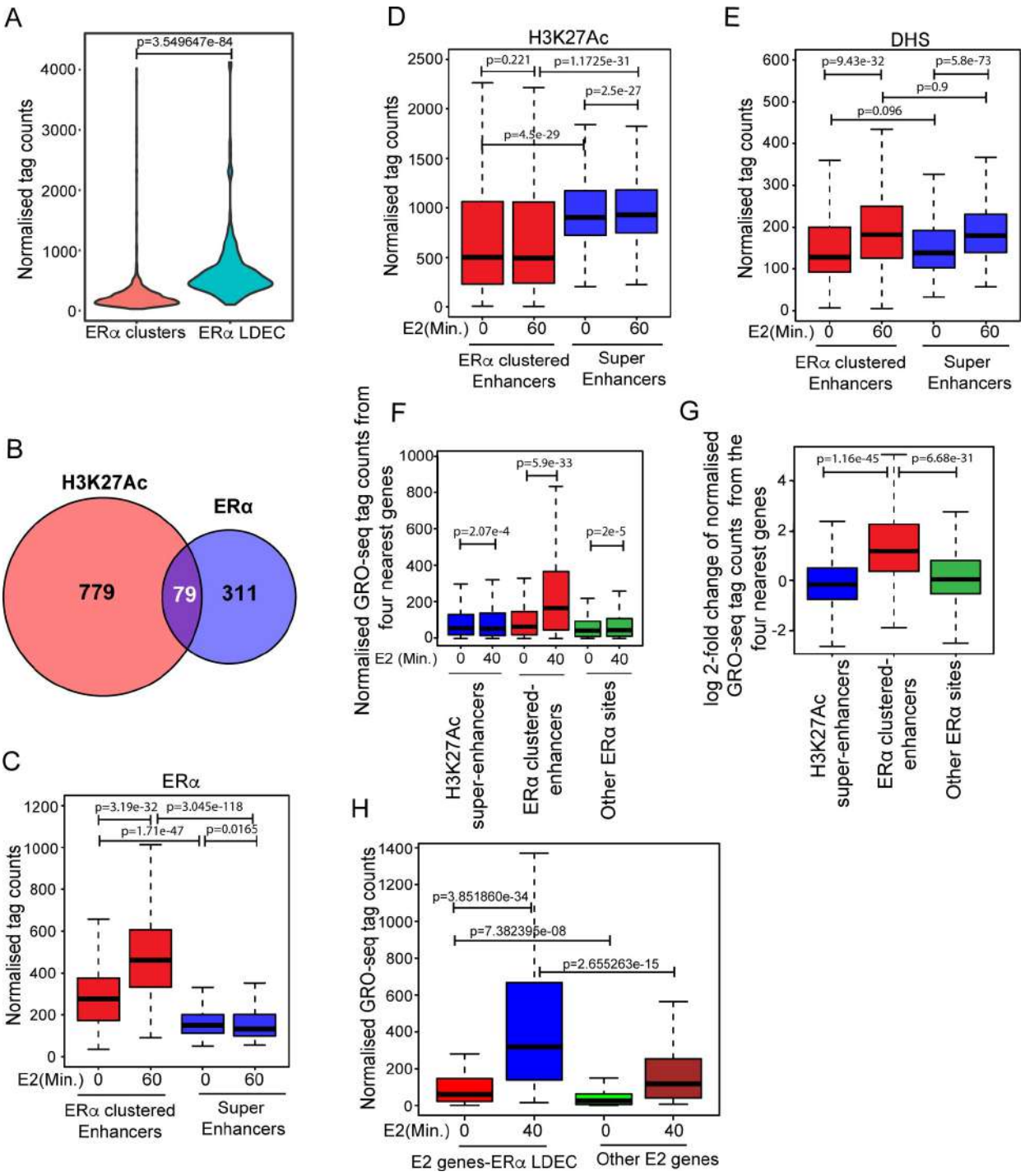

Fig 2. ERα super-enhancers but not conventional H3K27ac super-enhancers control E2 target genes. (A) Relative ERα tag counts on ERα clusters and LDECs. (B)Venn diagram showing overlap of 79, H3K27ac super-enhancers with ERα super-enhancers. (C-E) ERα binding strength (C), H3K27ac enrichment (D), and DHS signal (E) on H3K27ac super-enhancers and ERα super-enhancers in control and 60' post-E2 treated cells show higher levels of ERα induction on ERα super-enhancers, higher H3K27ac enrichment on H3K27ac super-enhancers, but similar DHS induction in both categories of enhancers. (F) Expression of genes closer to H3K27ac super-enhancers and other ERα sites not in cluster do not change upon E2 treatment but the genes that are closer to ERα super-enhancers are highly induced upon E2 treatments. (G) Pre-E2 to post-E2 log2 fold changes in expression of genes closer to different categories of enhancers as mentioned in panel C. (H) Expression of E2-responsive genes closer to LDEC vs. other ERα sites. All p-values were calculated by either Wilcoxon rank sum test or Wilcoxon signed rank test. The boxplots depict the minimum, first quartile, median, third quartile and maximum without outliers.

https://doi.org/10.1371/journal.pgen.1008516.g002

distribution between the ERα super enhancers and ERα density clusters. Of the 390 ERα super enhancers, 241 directly overlap an ERα density cluster and 368 are within 100 kb of one. In contrast, only 79 of the ERα super-enhancers overlapped with H3K27ac super-enhancers (Fig 2B). Hereon we will call these 390 ERα super-enhancers as LDEC. Although, as expected, H3K27ac super-enhancers exhibited higher enrichment of H3K27ac marks as compared to LDEC (Fig 2D), importantly, upon E2 stimulation H3K27ac and ERα occupancy did not change on H3K27ac super-enhancers, whereas ERα occupancy but not the H3K27ac increased significantly on LDEC (Fig 2C and 2D). Interestingly, DHS signal was strengthened on both classes of super-enhancers (Fig 2E) suggesting that modulation of H3K27ac super-enhancers by E2 is independent of direct ERα binding. Super-enhancers have been shown to activate their target genes robustly compared to typical enhancers. To compare the strength of gene activation by the two classes of super-enhancers, we compared the normalized GRO-seq tag counts of 4 closest genes to each enhancer in the two classes before and after E2 induction. We observed significant E2-dependent upregulation of genes closer to LDEC compared to the H3K27ac super-enhancers and other ERα sites, suggesting that ERα super-enhancers but not the H3K27ac super-enhancers robustly activate their target genes in an E2-dependent manner (Fig 2F and 2G). Further, even among the E2-responsive genes, the genes closer to LDEC exhibited robust induction as opposed to the E2-target genes elsewhere. Interestingly, genes closer to LDEC show higher basal expression even in absence of ligand, suggesting that persistent sites have a regulatory role in gene expression even in the absence of E2 induction; however, the expression of target genes is several fold higher upon LDEC appearance post-E2 (Fig 2H).

Genomically clustered ERα sites exhibit 3D proximity with each other and with target gene promoter(s)

Genes that are induced in a ligand-dependent manner exhibit induced physical proximity to their enhancers in a ligand-dependent manner [1,18,19]. Given our observation that the genes near LDECs exhibit robust E2-dependent upregulation, we hypothesized that for the LDECs to act as a regulatory unit in a way similar to the conventional super-enhancers, the constituent persistent and transient ERα sites within a LDEC should physically interact with each other and also with the target gene promoter. Since persistent ERα sites exhibit greater levels of ERα binding, we tested if these sites show relatively higher degree of physical interactions with other ERα-bound sites. Using publicly available ChIA-PET data on ERα in MCF7 cells [20], we found that overall the persistent sites exhibit a higher degree of interactions compared to the transient sites, and this tendency is even greater on highest occupancy persistent sites (3^{rd} quantile) (Fig 3A). Further, to assess whether LDECs are contained within a single Topologically Associated Domains (TAD) or span over multiple TADs, we interrogated HiC-inferred TADs from E2-treated and untreated MCF7 cells [9]. We found that 70% of LDECs are within a TAD and the mean length of TADs with LDEC is shorter (53kb) than the mean length of TADs without LDEC (61kb). Recently, strengthening of short-range interactions upon ligand stimulation was noted genome-wide using Shannon entropy measurements in MCF-7 cells [9] where, lower entropy means strengthened cis-interactions. Recapitulating these observations, we found that the 100 kb bins with LDEC have lower entropy relative to bins without LDEC but with ERα peaks and other bins in -E2 condition. However, in E2 condition, bins with LDEC show even greater drop in entropy as compared to the bins without LDEC (Fig 3B). These data suggest that regions with ERα LDEC gain short range interactions upon E2 stimulation. The strengthened looping potentially is between persistent and transient sites, and with the target gene promoters, within the LDEC upon ligand addition. To capture such

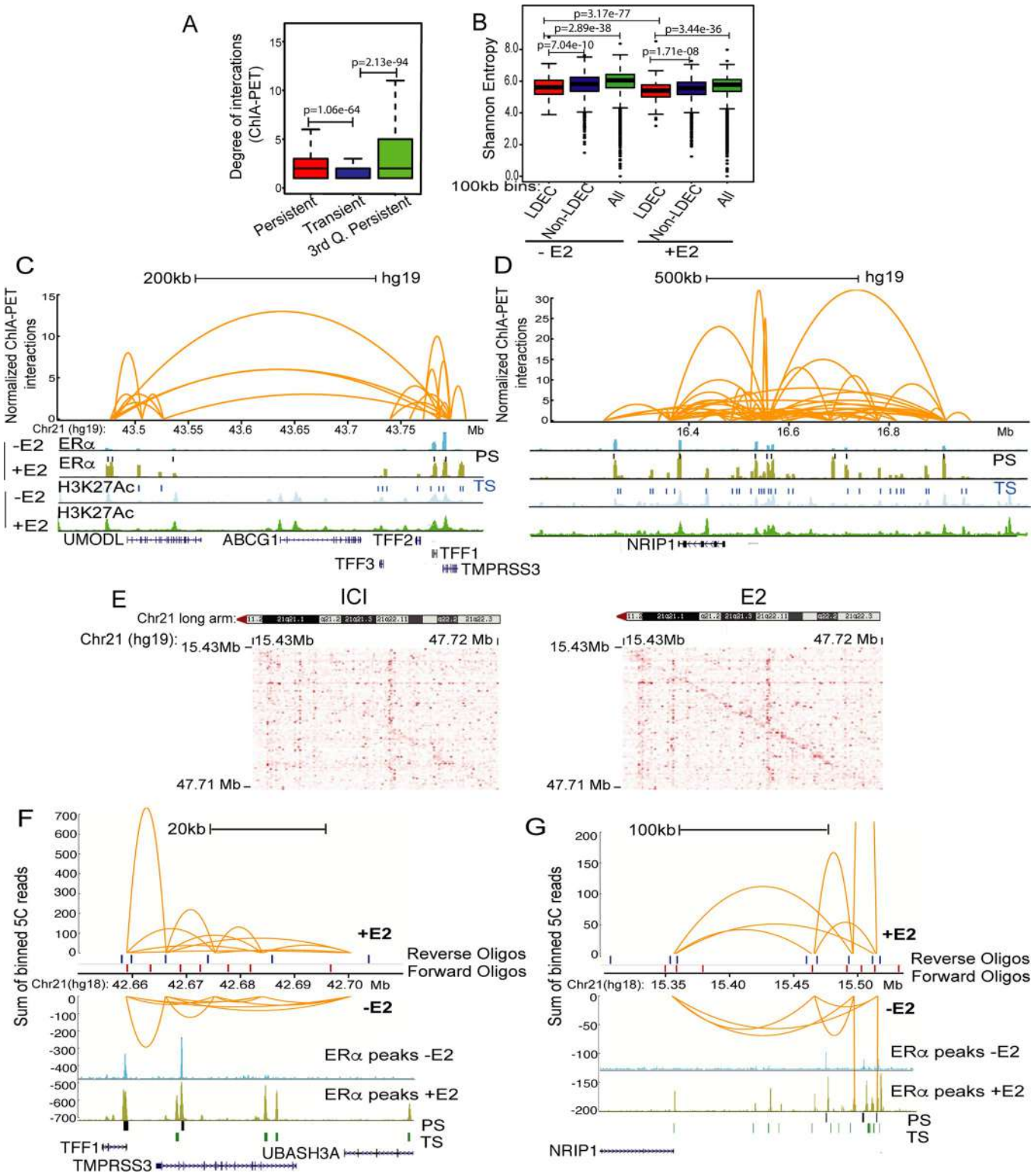

Fig 3. Genomically clustered ERα sites exhibit 3D proximity with each other and with target gene promoter(s). (A) Distributions of the degree of interactions emanating from the persistent, transient and 3rd quantile persistent sites, derived from ERα ChIA-PET data. (B) Shannon entropy plots taking 100kb genomic bins showing the drop of entropy upon E2 stimulation. The bins with LDEC exhibit highest drop compared to Non-LDEC and all bins. (C) ERα ChIA-PET data plotted for *TFF1* and (D) *NRIP1* LDEC. Orange arches depict the interaction pairs derived from one replicate, height of loops represent the strength of interaction. The panels are overlaid on ERα and H3K27ac ChIP-seq tracks in plus and minus E2 conditions from MCF-7 cells. (E) 5C matrix from

ICI (Left panel) and E2 (Right Panel) conditions showing normalized raw reads. (F and G) Yellow arches on y-axis represent the sum of normalized binned reads from 5C as shown for two LDEC, namely, *TFF1* (F) and *NRIP1* (G). Positive y-axis shows the interactions upon E2 treatment whereas negative y-axis corresponds to the interactions in ICI treated cells. The positions of forward and reverse oligos used to derive the plots are shown in red and blue vertical points along the x-axis. The plots are overlaid (bottom) with ERα peaks from -E2 and +E2 ChIP-seq data and BED peaks of persistent (Black) and transient sites (Green). All p-values were calculated by either Wilcoxon rank sum test or Wilcoxon signed rank test. The boxplots depict the minimum, first quartile, median, third quartile and maximum without outliers.

https://doi.org/10.1371/journal.pgen.1008516.g003

interactions, we investigated in depth specific LDECs using the available ERα ChIA-PET data. Interestingly, for the LDEC in the TFF1 gene locus, the persistent and its neighboring transient sites exhibited strong interactions with each other and with TFF1 promoter (Fig 3C). Interrogating long-range interactions, we found interactions between TFF1 LDEC and another LDEC near *UMODL1* gene 200 kb away (Fig 3C). Interestingly, TFF1 and UMODL1 LDECs are present in two adjacent TADs (S5D Fig) indicating the presence of extensive looping within LDEC and also between neighboring LDECs. Intra-LDEC and LDEC-promoter interactions were also observed for other tested genes such as *NRIP1* (Fig 3D) and *GREB1* (S5C Fig). Because ChIA-PET was performed in presence of E2, we next aimed to determine whether the observed interactions were E2-dependent. Due to relative genomic proximity of ERα sites within a cluster, requiring greater sensitivity in spatial proximity determination, we performed 5C in two replicates on all clustered enhancers on Chr21, as it harbors the most robustly E2-induced genes, viz., *TFF1*, *NRIP1*, and the lncRNA *DSCAM-as*. These genes span the entire length of the long arm of acrocentric Chr21. The 5C library contained forward and reverse oligos for all active enhancers [2] and the respective promoters. Oligos derived from ENCODE desert region on Chr16 were used to normalize the digestion and ligation efficiency biases between the libraries as reported earlier [21]. The 5C was performed in ICI and 1hr E2-treated MCF-7 cells, number of reads and contacts are mentioned in (S1 Table). The replicates showed 98% correlation with each other and the data in E2 treated and untreated cells clearly showed an overall comparable 5C matrix ruling out any bias. Interestingly, the overall normalized 5C data showed a strong diagonal upon E2 induction reflecting enriched cis-interactions (Fig 3E) as seen in Fig 3B, 3C and 3D. Similar enrichment in cis-interactions upon E2 treatments has been reported recently [9]. To interrogate the observed E2-dependent strengthening of interactions within LDECs (Fig 3B), we plotted the Chr21 normalized reads arising from the binned reads overlapping genomic regions within the cluster (Method and S2 Table). A snapshot of *TFF1* and *NRIP1* loci confirms the strengthening of *cis*-interactions between persistent and transient sites within the clustered enhancer as well as with target *TFF1* and *NRIP1* promoters (Fig 3F and 3G). These gained or strengthened interactions overlap with the transiently gained ERα peaks within *TFF1* and *NRIP1* LDECs. Overall, these results demonstrate an induced and specific 3D interaction between the transient and the persistent sites, as well as with target gene promoter, in an E2-dependent manner.

Persistent sites are required for the binding of ERα at neighboring transient sites

Given that the transient sites cluster around persistent sites, we assessed whether persistent sites play a direct role in the emergence of clustered enhancers. We either deleted or blocked persistent site (PS) from LDECs of the most E2 inducible genes *TFF1* on Chr21 and *GREB1* on Chr2 using CRISPR-Cas9 strategy (S6A to S6D Fig for *TFF1* PS deletion, S6E and F for *GREB1* PS blocking). ERα occupancy at transient sites around *TFF1* persistent site was completely lost in ΔPS–TFF1 MCF7 cells as compared to the wild-type cells as seen by ChIP-seq of ERα in these cells (Fig 4A, pink highlighted region). Similar observations were made in a different homozygous deleted line (S7A, S7B and S7C Fig). Importantly, ERα occupancy at distal sites

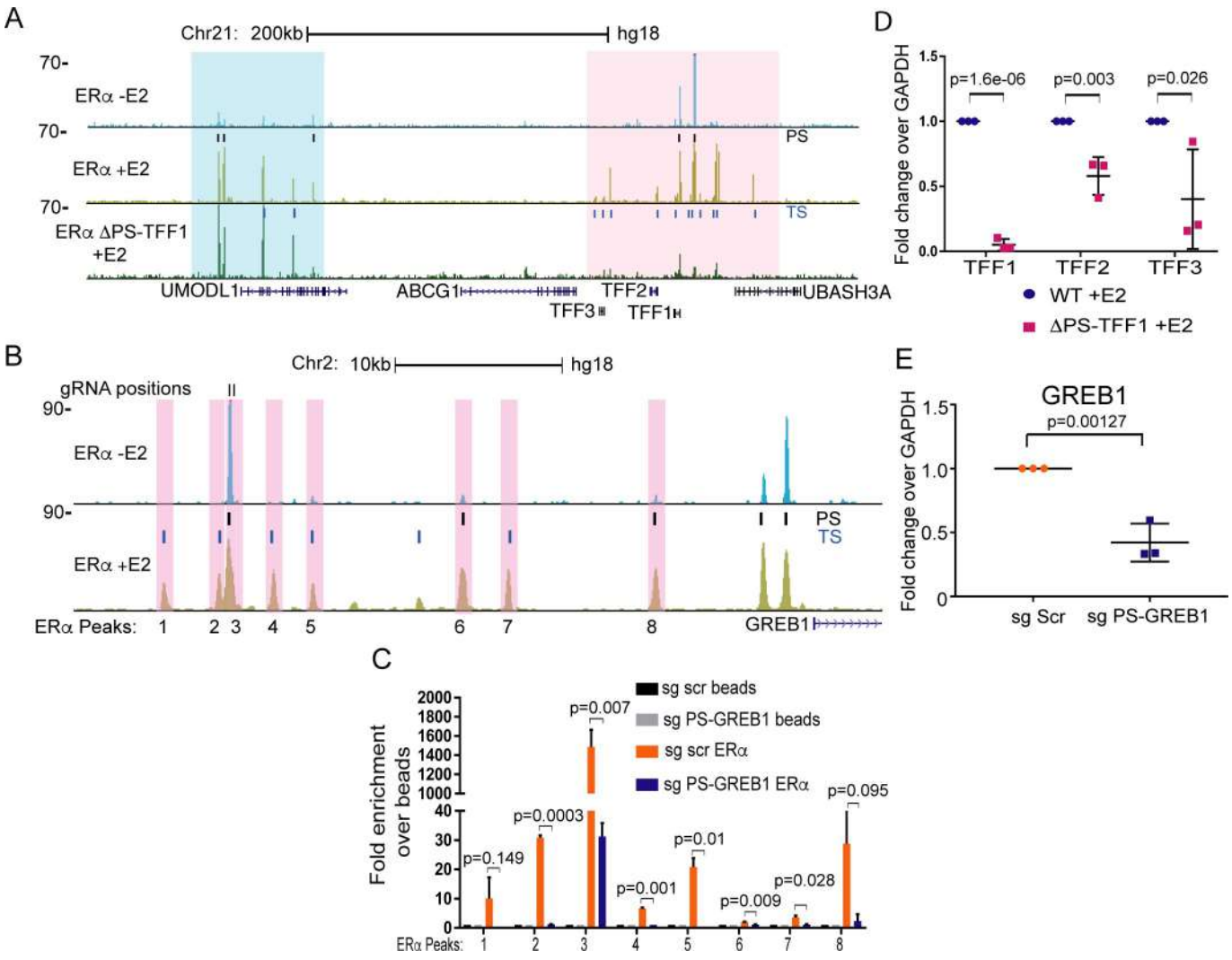

Fig 4. Persistent sites are required for the binding of ERα at neighboring transient sites. (A) Genome browser snapshot of *TFF1* region depicting ERα binding in WT and ΔPS-TFF1 lines. First and second ERα ChIP-seq tracks are in untreated and E2 treated WT cells, and the third track represents ERα ChIP-seq in the E2-treated ΔPS-TFF1 MCF7 line. Pink and blue highlighted regions represent the *TFF1* and *UMODL1* LDEC respectively. (B) Genome browser snapshot of *GREB1* LDEC exhibiting ERα peaks in untreated and E2 conditions. The vertical highlighted boxes show the regions selected for the measurement of ERα occupancy upon CRISPR blocking of persistent site using gRNAs designed at positions indicated as vertical lines on peak 2/3. (C) ERα occupancy on highlighted regions in panel C upon CRISPR blocking of persistent site shows significant loss of ERα binding on not only the persistent site but also on transient sites. (D) qRT-PCRs for TFF1, TFF2, and TFF3 genes in E2 treated WT and ΔPS-tff1 lines (E) qRT-PCR exhibiting the loss of GREB1 induction in the cells where *GREB1* persistent site is blocked as compared to the cells transfected with scr gRNAs. qPCR Plots represent data from two biological replicates and each replicate had three technical repeats. Data are plotted as average, SD and p values are indicated. All p-values were calculated by unpaired student's t-test.

https://doi.org/10.1371/journal.pgen.1008516.g004

was unchanged (Fig 4A, Blue highlighted region). The distal sites are in fact the LDEC for *UMODL1* gene which is in the neighboring TAD (S5D Fig). The ERα occupancy on E2 responsive regions located on other chromosomes was unaffected (S8 Fig). Similarly, we blocked persistent site at GREB1 LDEC using specific gRNA and measured ERα binding at transient sites. Again, not only persistent sites but transient sites also showed loss of ERα binding (Fig 4B and 4C) but the distal site in the neighboring TAD exhibited modest loss of ERα (S7D Fig), supporting essentiality of persistent sites for ERα recruitment at transient sites. Interestingly, blocking of another persistent site near the promoter of GREB1 (Fig 4B) also perturbed ERα binding on enhancer cluster suggesting that persistent sites in general play an

important role in the LDEC formation (S7E Fig). We further quantified the effects of persistent site deletion (TFF1 PS) and blocking (GREB1 PS) on the expression of the neighboring genes. The loss of E2-induced expression of *TFF1* was noted (Fig 4D). Interestingly, other genes within the LDEC such as *TFF2 and TFF3* also showed reduced response to E2 signaling in the ΔPS–TFF1 cells (Fig 4D); these genes have been shown to physically interact with each other and are regulated by E2 [22,23]. Similarly, GREB1 expression was also reduced upon blocking of GREB1 persistent site (Fig 4E). These data strongly suggest the specific and circumscribed effects of persistent site in regulating one (*GREB1* LDEC) or more (*TFF1* LDEC) genes present within LDECs.

ERα puncta are formed on LDECs by coalescing

Our results so far point towards ligand-dependent binding of ERα on the EREs near persistent sites and concomitant physical interaction among persistent and transient sites, suggesting spatial crowding of both DNA and ERα in these clusters. We asked if such clusters can be visualized by fluorescence microscopy. Toward this we performed live cell imaging of GFP-ERα upon E2 treatment. Interestingly, the ERα distribution was found to be uniform in untreated cells (Fig 5A) but we observed robust formation of ERα punctate pattern as early as 12–15' post-E2, formed by coalescing (Fig 5A, frames from S1 and S2 Movies respectively), the puncta continued to increase in number and size till 40' post signaling. Further, to test if the genomic LDECs that emerge in 3D upon ligand stimulation are indeed the dynamic ERα puncta observed in immunofluorescence, we conducted immunoFISH studies on *TFF1* LDEC. BAC clone overlapping LDEC was used. ERα intensity on *TFF1* LDEC increased significantly upon ligand stimulation as compared to random regions (Fig 5B). Interestingly, the deletion of persistent site on *TFF1* LDEC weakens high intensity puncta on this LDEC as seen by loss of ERα intensity and colocalization of FISH spots with ERα puncta (Fig 5C). Similar gain of ERα intensity was also noticed on *NRIP1* LDEC (S9A and S9B Fig). These data suggest a model where binding of liganded ERα to transient sites strengthens 3D proximity within LDEC, which can be observed as ERα puncta in the nucleus. The persistent sites mediate the crowding of ERα molecules on these puncta. To understand the ERα exchange dynamics in these E2 induced puncta, we performed FRAP experiments on GFP-ERα foci in untreated cells and cells that were treated with E2 for 1h. Although some puncta were observed even in untreated cells but the recovery of puncta was so robust that a complete bleaching could not be achieved due to rapid exchange of ERα in the puncta (Fig 5D). The recovery of ERα puncta in E2 treated cells was 60%, suggesting a stable ERα puncta formation post-E2 treatment and the loss of fluorescence recovery to this extent has been shown for other punctate patterns [11].

Recent reports have shown that crowding of protein-occupied regulatory units involve hydrophobic interactions, such as HP1-mediated heterochromatin formation and co-activator mediated condensate formation on enhancers [24, 11, 12]. To test whether hydrophobic interactions play a role in our observed puncta formation, we treated the E2-pretreated cells with 1,6-hexanediol (1,6-HD) that disturbs hydrophobic interactions. We found that ERα punctate pattern was lost upon 1,6-HD application (S9C Fig). To recapitulate the loss of puncta at the genomic level, we monitored the occupancy of ERα in cells treated with 1,6-HD. Loss of ERα occupancy was seen both at *TFF1* transient as well as persistent sites in presence of ligand (Fig 5E). Similarly, persistent sites exhibited significant loss of ERα binding even in the absence of ligand (S9D and S9E Fig), suggesting the loss of puncta in E2 condition is a result of 1,6-HD mediated perturbations in the hydrophobic interactions of ERα without affecting its binding with E2. The expression of most E2-inducible genes driven by LDECs such as TFF1, GREB1 and NRIP1 was significantly reduced upon administration of 1,6-HD confirming the loss of

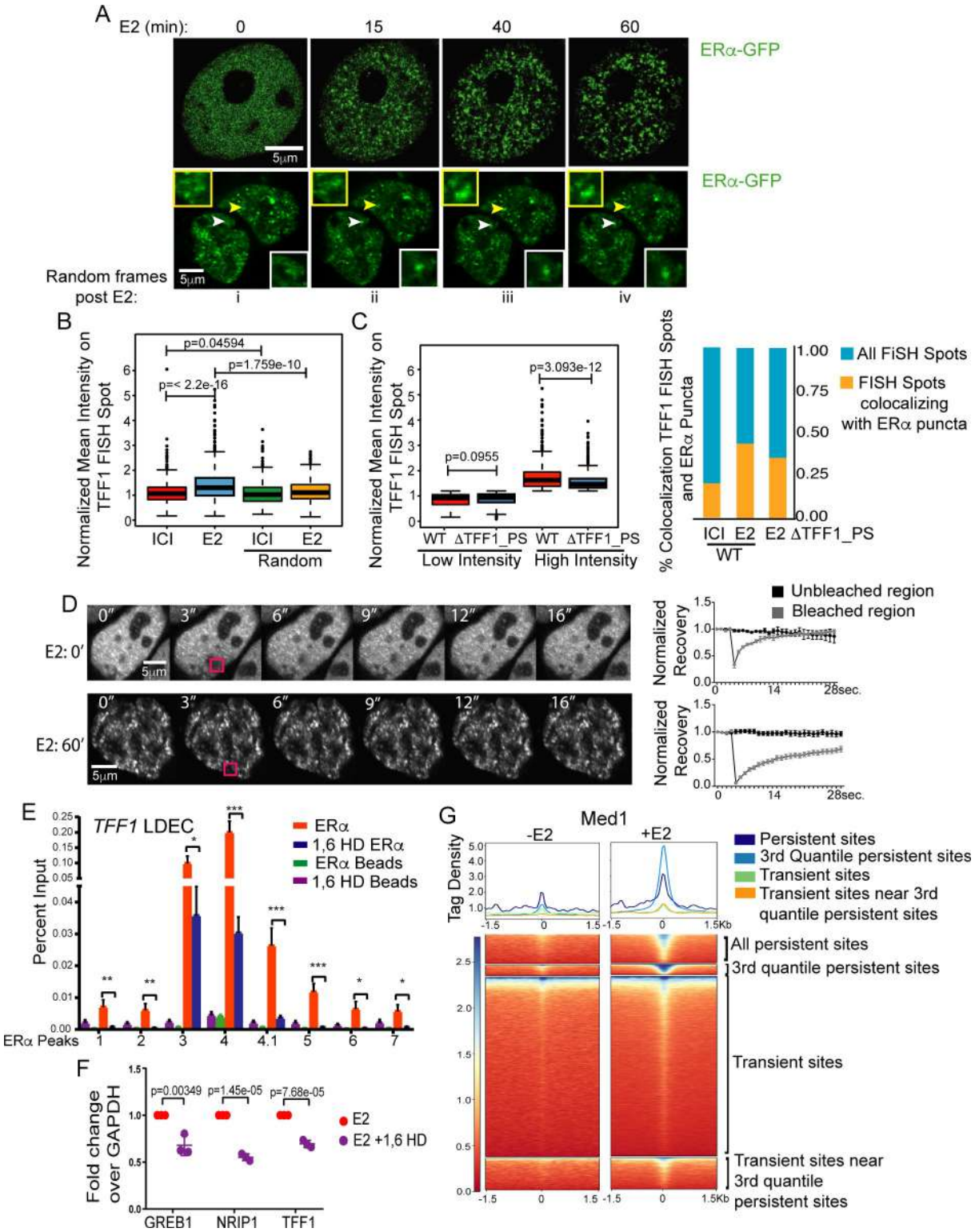

Fig 5. ERα puncta are formed on LDECs by coalescing. (A) GFP-ERα frames from movie showing ERα puncta appear upon E2 signaling during the mentioned time, post-E2 treatment (upper panel). ERα puncta appear upon E2 treatment and puncta of interest are marked by arrows (lower panel). Insets show the zoomed in images of marked puncta. (B) Plot showing the relative mean ERα intensity on TFF1 FISH spots and random regions upon ICI and E2 treatments. (C) Boxplot shows the mean intensities of low and high intensity ERα puncta on TFF1 FISH spots in WT and ΔTFF1_PS cells. p-values were calculated by Wilcoxon rank sum test. Right panel depicts the percent colocalization of

ERα puncta and TFF1 FISH signal before and after E2 treatment and in WT cells and TFF1_PS delete line. (D) FRAP recovery of puncta in control cells (upper panel of images and graph) and at 60' post-E2 treatment (lower panel and graph). Data are shown as mean +/− SEM. (E) Chromatin Immunoprecipitation shows the levels of ERα at persistent as well as at transient peaks in the *TFF1* LDEC before and after 1,6-HD treatments. (F) E2-dependent expressions of GREB1, NRIP1 and TFF1 is perturbed upon 1,6-HD treatments. (G) Heatmaps represent the MED1 binding strength at different categories of ERα peaks in untreated and E2 treated conditions. Tag density was measured on 1.5 Kb upstream and downstream region from center of ERα peaks. qPCR Plots represent data from three biological replicates and each replicate had three technical repeats. Data are plotted as average, SD and p values are indicated. All p-values were calculated by unpaired student's t-test. The boxplots depict the minimum, first quartile, median, third quartile and maximum.

https://doi.org/10.1371/journal.pgen.1008516.g005

ERα-clustered enhancers and their functions (Fig 5F). The data suggests that ERα genomic clusters identified indeed correspond to ERα puncta that involve hydrophobic interactions and are formed in a ligand dependent manner.

MED1 and ERα cooperate in gene regulation and were recently shown to cooperate in creating condensates on DNA [12]. Consistently, we observed E2-dependent binding of MED1 on the clustered enhancers with highest binding on 3rd quantile persistent sites (Fig 5G). The data points towards the ERα crowding upon ligand stimulation which corroborates with the 3D genomic interactions that take place within and perhaps between LDECs in E2 dependent manner.

LDECs exist transiently only during the active phase of signaling

Expression of E2-upregulated genes is known to peak at 45 min to 1 hour before declining at 3 hours post-E2 stimulation [3]. Further, ERα-bound enhancers have been shown to control most E2-inducible genes [1,2,3]. We therefore directly assessed whether E2-induced changes in gene expression over the course of signaling are directly driven by E2-induced LDECs, and further, whether the binding dynamics of ERα is different for persistent and transient sites. Towards this, we measured the genome-wide ERα occupancy on various time points post-E2 (5 to 1280 min) using publicly available data [4]. From their data, we partitioned ERα sites into persistent (present in 0 min as well as latter time points) and transient (absent in 0 min but present at latter time points) sites as mentioned before (Fig 1B) to monitor their binding dynamics separately. In general, binding at transient sites is relatively weak compared to persistent (Fig 6A). However, in all categories of sites, a pattern becomes evident, although to different degrees due to varying ERα binding strengths (Fig 6A). We observed an increase in ERα binding strength starting at 5', peaking at 40', and then gradually declining to minimal levels at 160', followed by an unexpected increase in binding strength at later time points. Most evident in the 3rd quantile (strongest) persistent sites, the binding strength recovers almost to the maximum levels ~24 hours post-E2. These results suggest that, LDEC emerge transiently around persistent sites to drive active phase of signaling in terms of target gene expression but they disappear during later phase of signaling which is recapitulated by loss of E2-target gene expression. However, the persistent sites remain ERα-bound as bookmarks (Fig 6B, 0 vs 1280 min). These alterations in ERα binding pattern occur even though, estradiol remains present in the medium at these later hours as the half-life of estradiol is between 8-36h. Given that LDECs correspond to ERα puncta (Fig 5B and 5C), we tested if these puncta also follow the similar transient pattern as genomic clusters. We performed ERα immunostaining in cells treated with either ICI for 24h or with E2 for 10', 60', 180', and 24h. Although ERα forms small puncta even in absence of ligand, the overall intensity, size, and the number of puncta increased significantly at 10', stayed high at 60' followed by a significant drop at 180', reaching a minimal level at 24h (Fig 6C). This temporal change in ERα puncta intensity is similar to ERα binding pattern in genome as observed in Fig 6A and 6B. Further, it is evident collectively from Fig 6A, 6B and 6C that even though the persistent sites reappear at 24h (Fig 6A), lack of transient ERα

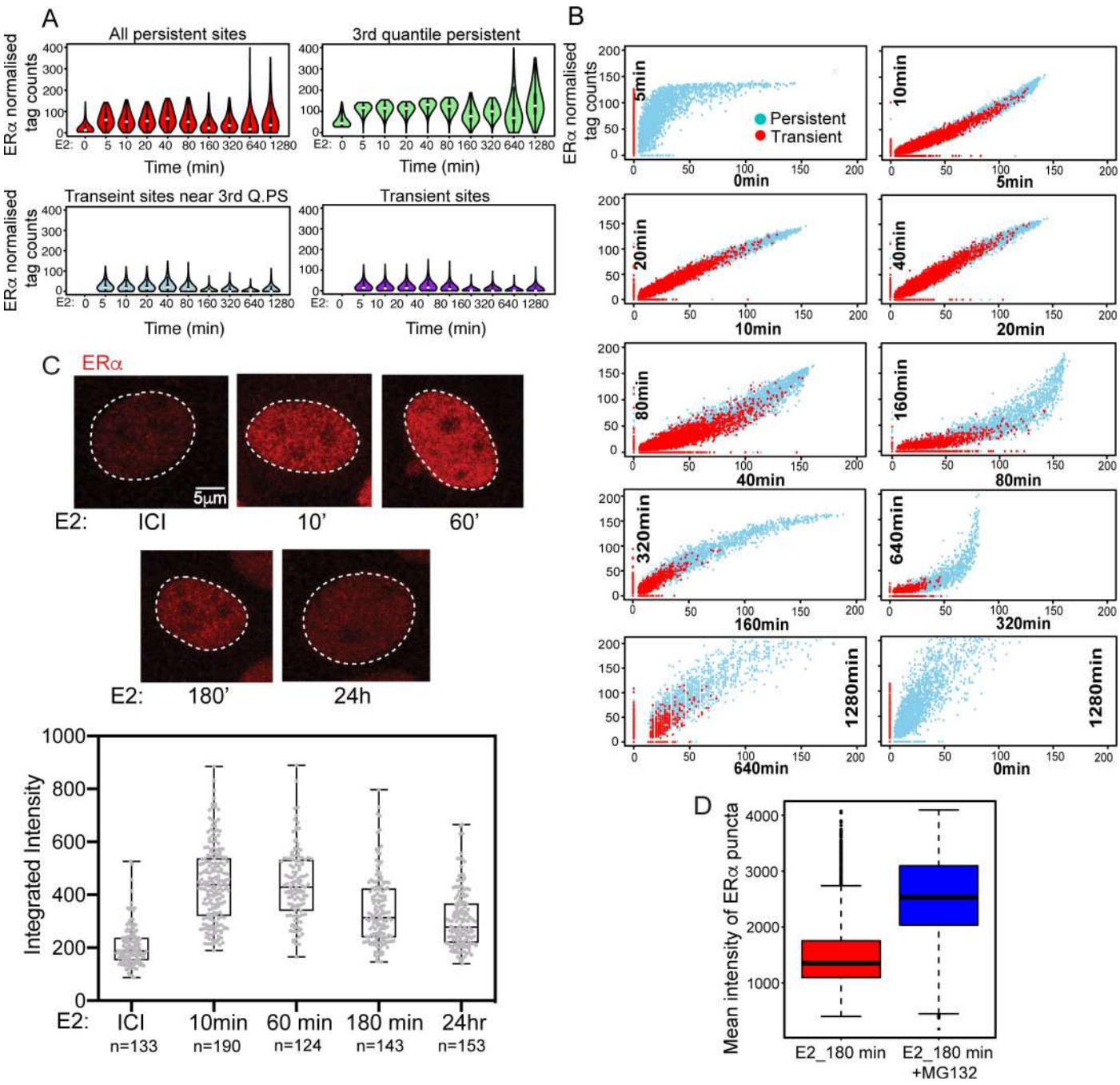

Fig 6. Clustered enhancers exist transiently only during active phase of signaling. (A) ERα binding strength is shown for different categories of ERα peaks at various time points post-E2. The four plots represent All persistent (Top left), Transient near 3rd quantile persistent (bottom left), 3rd quantile persistent (Top Right), and transient (Bottom right) ERα peaks. (B) Comparison of binding strengths at consecutive time points for persistent and transient sites reveal an initial surge in binding strength at persistent sites, followed by stably high binding at both persistent and transient sites, followed by loss of transient sites. (C) ERα immunostaining at ICI, 10', 60', 180', and 24h post-E2. Bottom panel shows the quantification of ERα integrated intensity at these time points. (D) Mean intensity of ERα puncta in E2 treated cells with and without MG132 pretreatments. The box plot depicts the minimum, first quartile, median, third quartile and maximum without outliers.

https://doi.org/10.1371/journal.pgen.1008516.g006

sites around these persistent sites do not allow persistent sites to form the genomic cluster, thereby punctate pattern of ERα disappear dramatically at 3h post stimulation and remains low even at 24h when persistent sites show significant ERα binding (Fig 6A). Next, inhibition

of proteasome-mediated protein degradation by MG132 rescued the intensity of ERα puncta 3h post-E2 signaling, suggesting that the ERα turnover is responsible for the loss of punctate pattern at 3h (Fig 6D). These data suggest that signaling response is at the peak around 1h however, it drops significantly at 3h post stimulation.

Emergence of LDECs is correlated with robust target gene expression

Robust eRNA induction at enhancers is directly linked to robust target gene activation [1,3]. Given that ERα clustered enhancers and punctate pattern decline 3h post-E2 treatment, we tested whether expression of eRNA at clustered enhancers undergoes a similar pattern and whether the persistent and transient sites exhibit varying levels of eRNA induction given their differential ERα binding affinity. Toward this, we analyzed time course GRO-seq data [3] and found a higher level of eRNAs at strong persistent sites relative to transient and weak persistent sites. Further, the fluctuations in ERα binding at different time points post-E2 stimulation were also reflected in the relative eRNA expression (Figs 7A, S10A). Not surprisingly, the persistent sites were the strongest enhancers in the cluster, and the target genes exhibited similar pattern of expression peaking at 40' and dropping to a minimal level at 180' post-E2 stimulation (Fig 7B). Conversely, the genes that are away from clustered enhancers but closer to random ERα sites did not show E2-dependent induction. Even though the level of nascent transcripts drop at 3h post-E2, mRNAs levels exhibit such a drop between 3-4h post-E2 as observed in a publicly available time-course mRNA-seq upon E2 addition [25] (S10B Fig). These signaling responses recapitulate the predicted ERα half-life of 3–4 hours in the presence of E2 [26,27], and the role of ERα degradation kinetics in E2-target gene expression [28,29,30]. Similarly, we found that the chromatin-bound fraction of ERα increases from untreated cells to 1h post-E2 treatment but goes down by 3h (Fig 7C), consistent with eRNA, E2-target gene expression, existence of LDECs, and the punctate pattern. Interestingly, protein levels of MED1 and FOXA1 do not exhibit alterations over the same E2 time course (S10C Fig). This strongly suggests that when ERα protein degrades, the transient sites lose their ERα resulting in loss of clusters reflected by a loss of puncta 3h post-E2. However, ERα binds to persistent sites due to presence of FOXA1 motif at these sites and as expected, FOXA1 knockdown clearly showed the loss of ERα binding at persistent sites (Fig 7D), without affecting the ERα mRNA and protein levels (S10D Fig) [15]. Interestingly, FOXA1 knockdown also reduced ERα binding at transient sites lacking FOXA1 motif, supporting a potential crosstalk between the persistent and the transient sites within the cluster (Fig 7D). Our data again suggests that although persistent sites remain bound with ERα, without the ERα binding at the neighboring transient sites they are incapable of driving the target gene expression.

Discussion

Signaling response is pre-established by an unliganded receptor

Steroid receptors mediate cellular response to hormonal cues by binding to specific enhancers to induce target genes. Our results suggest that, in the case of ERα, a timely and robust response is ensured by pre-marked enhancers bound by unliganded ERα, which nevertheless remain in an inactivated form. However, upon signaling cues, additional liganded ERα binding at EREs in relative proximity to persistent sites is triggered, transiently creating an active cluster of enhancers capable of driving robust gene expression. We and others [5] have found that the presence of FOXA1 motif is a striking feature of pre-marked enhancers (Fig 1D). FOXA1, as a pioneering factor stabilizes the binding of ERα on persistent sites in absence of ligand. Upon E2 treatment, when the liganded nuclear ERα levels rise in the nucleus, ERα binds to transient sites albeit with low affinity (Fig 1C). However, at later time points when the

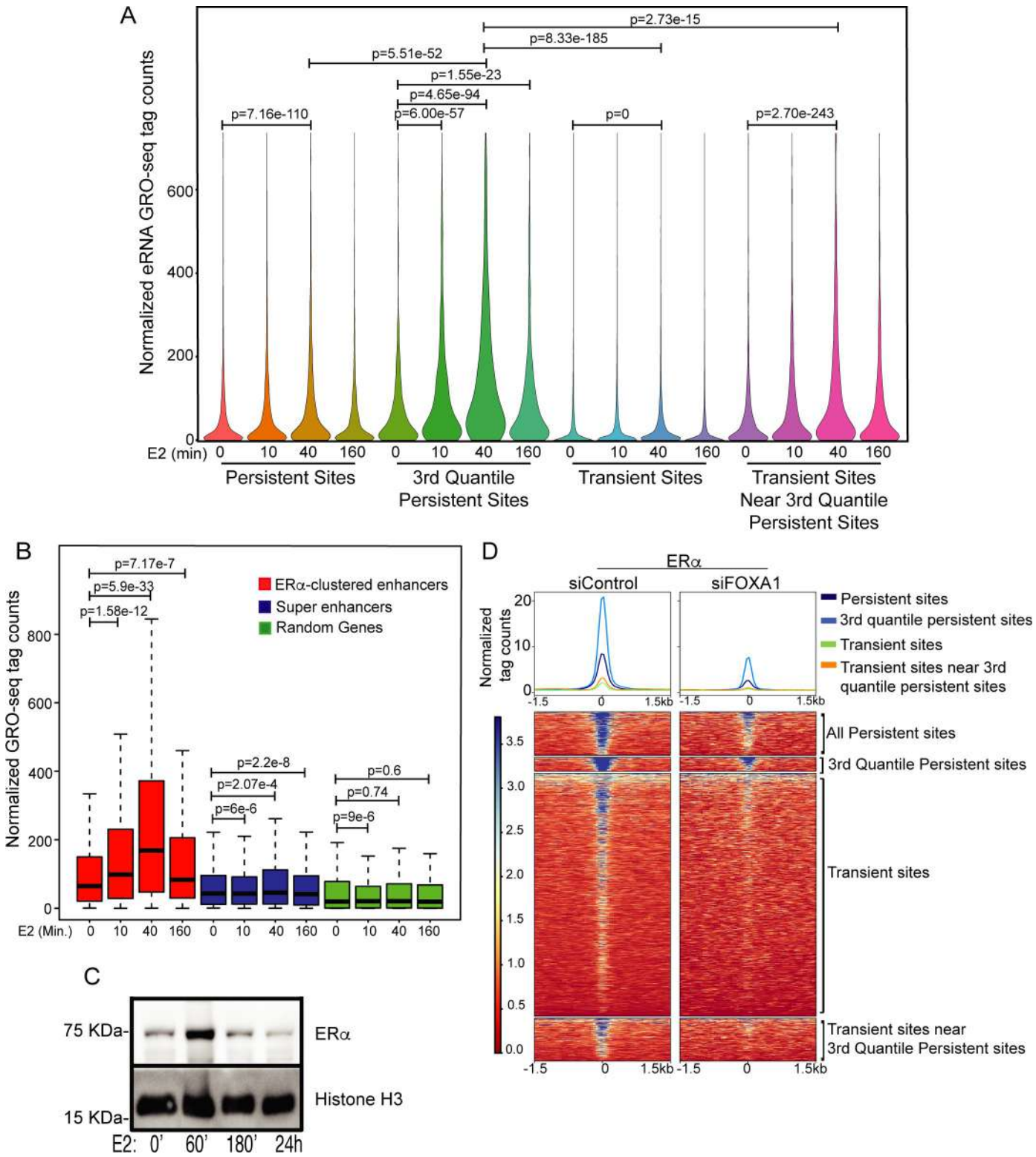

Fig 7. Emergence of ERα LDEC is correlated with robust target gene expression. (A) Combined eRNA levels from minus and plus strand show a robust increase in transcription at persistent sites at 40' post-E2, much more so for the 3rd quantile persistent sites. This is also true to a much smaller extent for transient sites near persistent sites, but not for distal transient sites that are not a part of LDECs. eRNA expression drops to basal level at all sites at 160', except the strongest persistent sites still show relatively higher levels of eRNAs. (B) Expression of genes closer to all LDEC, H3K27ac super-enhancers and 300 random genes at 0', 10', 40', and 160' post-E2. (C) Chromatin-bound fractions of ERα at 0, 60', 180', and 24hr post-E2 shows the drop in ERα chromatin bound fraction at 180'. (D) Heatmaps

showing the binding strength of ERα at various categories of ERα peaks in sicontrol and siFOXA1 transfected cells. All p-values were calculated by either Wilcoxon rank sum test or Wilcoxon signed rank test. The boxplots depict the minimum, first quartile, median, third quartile and maximum without outliers.

https://doi.org/10.1371/journal.pgen.1008516.g007

level of nuclear liganded-ERα decline by proteolysis again (Fig 7C), FOXA1 still assists binding of ERα at persistent sites but transient sites lose their ERα, restoring the binding pattern of ERα to the pre-treated state.

Despite harboring EREs, unliganded ERα does not bind to transient sites due to low chromatin accessibility (DHS), even when the levels of ERα are raised by ERα overexpression (Fig 5A, 0 min) (movie1). Liganded ERα binds to transient sites upon stimulation despite the fact that DHS does not change significantly at these sites as seen for thyroid receptor (TR) and glucocorticoid receptors (GR), it may do so by acting as a pioneering factor or by binding with cooperative factors and chromatin remodelers as in the case of GR [31,32,33,34,35]. However, in the absence of persistent site or upon FOXA1 knockdown, transient sites fail to recruit ERα, suggesting a role of FOXA1 in stabilizing the ERα binding not only at persistent sites but also at transient sites within LDECs, likely aided by 3D proximity with persistent sites. Collectively, these results suggest that a framework for signaling is setup by persistent sites via bookmarking the future functional enhancers with FOXA1 and unliganded ERα, which act as a nucleating point for the subsequent ERα binding on the neighboring EREs. Furthermore, chromatin remodeling around persistent sites do not appear to occur as the intermittent H3K27ac sites within *TFF1* LDEC that are not occupied by ERα remain unaffected by E2 signaling (Fig 1E), suggesting that the transient sites within LDECs respond specifically to E2 signaling. Furthermore, ERα clustered enhancers form a hierarchy where only some of the enhancers control the target gene expression while others seem redundant [36], which is very similar to STAT5-driven super-enhancers [37]. Our data suggest that perhaps the enhancers that are pre-marked are the functional enhancers. Lastly, while the pre-seeding of enhancer with unliganded receptor has been previously suggested to be crucial [5,8], our study experimentally demonstrates the absolute requirement of pre-marked enhancers for the emergence of enhancer clusters where the existence of entire cluster spanning over long linear distances depends on single enhancer.

LDECs exhibit extensive chromatin looping in ligand-dependent manner

Using 5C on chr 21 and HiC data, we observe the ligand-dependent alterations in chromatin proximity which has been debated with contrasting evidences [1, 9, 10]. Importantly, we observe that persistent sites exhibit an overall higher degree of interactions (Fig 3A). On the other hand, bins that contain LDECs exhibit E2-induced increase in short range interactions relative to other regions (Fig 3B). This may suggest that in absence of signalling, persistent sites may interact with other regulatory sites across different TADs, but upon E2 induction, as transient sites appear bound with liganded ERα, persistent sites redirect their interactions with transient sites favouring more short range cis-interactions. This E2-dependent shift in short range interactions is a prominent feature of LDEC. These observations also suggest that LDEC formation may not require major changes in chromatin architecture as most alterations are contained within the TADs. Such E2-dependent alterations in interactions favouring more short-range interactions from long range was also reported recently [9, 23]. However, our study teases apart the specific features of regions that exhibit such behaviour.

In some instances, LDEC from two neighbouring TADs also exhibit interactions, for example, *TFF1* LDEC shows interaction with *UMODL1* LDEC (Fig 3C), but ERα occupancy at *UMODL1* locus was unaffected by the perturbation of persistent site in *TFF1* LDEC (Fig 4A), suggesting a degree of autonomy and circumscribed influence of each LDEC. Related to this

observation, there is a recent interest in understanding whether a fraction of chromatin interactions may have functional roles other than transcriptional regulation [38].

ERα puncta are formed on LDECs by coalescing

While ERα binds to the DNA mostly as a monomer or dimer and dimerization-deficient or DNA-binding mutant of ERα do not form the punctate pattern in the nucleus [39]. The interactions among multiple ERα sites within a LDEC by the virtue of ERα interacting proteins and RNA could result in ERα-tethered 3D genomic clusters appearing as ERα puncta. Looking closely at Fig 3C, one of the central enhancers in *TFF1* LDEC that is not occupied by ERα does not participate in ERα-bound chromatin network, suggesting that the intervening DNA that is not bound by ERα may have been looped out allowing the regulation of only tethered DNA in such structures. Along similar lines, intrinsically disordered proteins have been shown to mediate protein-protein interactions while being bound on DNA [40]. They create genomic clusters by tethering protein bound DNA which results in mechanical exclusion of intervening chromatin fibre. These mechanical principles could define most genomic clusters created in 3D.

We observed MED1 occupancy on LDEC is entirely ligand dependent (Fig 5G) which recapitulates the established interaction between Mediator complex and ligand binding domain of ERα [41,42]. It has been shown that deletion of N-terminal region in ERα exhibits complete loss of punctate pattern in the nucleus upon E2 treatment [43]. Further, PONDR analyses suggest that this region of ERα is unstructured. Since, IDRs in transcription factors allow them to phase separate [44,45], thus IDR of ERα at its N-terminus and MED1 may allow them to form multivalent homotypic or heterotypic complexes on LDECs. The high density of ERα along with MED1 on persistent sites may allow them to interact more with other genomic regions in the same LDEC to form liquid condensates as shown recently [12]. Furthermore, persistent sites show robust induction of eRNAs and they interact with Mediator complexes [46,47] thus, it is likely that eRNA, MED1, and ERα form RNA protein complexes (RNP) strictly upon E2 stimulation, as shown for other RNA binding proteins [48,49].

Loss in signaling response recapitulates the loss of ERα protein

The ERα level is significantly reduced at 3h which is the half-life of ERα in presence of ligand (Fig 7C). The reduced levels of ERα at 3h overlaps with disappearance of LDEC (Fig 6A, 6B and 6C) and concomitant loss of eRNA and gene expression (Fig 7A and 7B) suggesting that upon loss of ERα levels, transient sites lose their ERα. However, persistent sites regain the binding back due to the presence of FOXA1 at these sites. Interestingly, MG132 pre-treatment of E2-treated cells which stabilizes the ERα levels even at later hours post signaling, allows longer and persistent expression of E2-regulated genes [50] conforming that the 26S-proteosome mediated degradation of ERα in the nucleus exerts enhancer-mediated decline in the signaling response. This suggests a simple model of regulation where protein degradation brings transcription factor levels below saturation point for the condensate to remain phase separated. This might be a universal mechanism driving the decline of response to signaling mediated by such phase-separated enhancer condensates.

Together, these observations suggest a mechanism by which unliganded receptor acts as a nucleating point for the new ERα binding in its proximity. Many transient sites along with persistent sites loop with each other to form a 3D complex. These clusters appear to be formed by coalescing of ERα as seen by emerging puncta in E2-dependent manner. Further, the clusters are transient and disappear 3h post signaling coinciding with the loss of eRNA expression and target gene inducibility (Fig 8). Our work establishes the direct regulatory link between

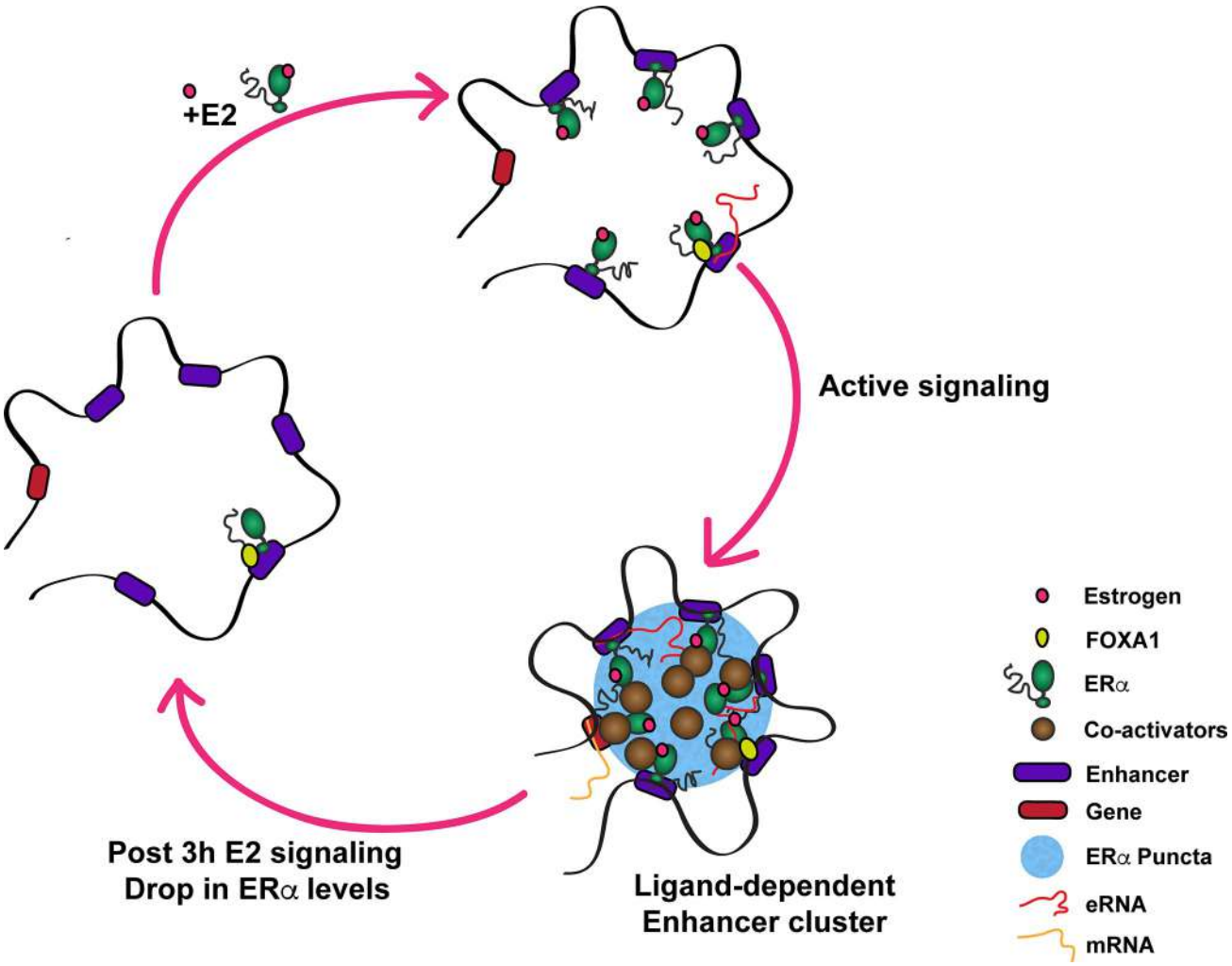

Fig 8. Proposed Model of E2-mediated transcriptional regulation. Model depicts the active and inactive phases of estrogen signaling. During active phase, liganded ERα decorates the EREs closer to premarked-unliganded ERα site. Together, these persistent and transient sites form LDEC in 3D manifesting as ERα puncta, resulting in robust expression of target genes. Upon ERα degradation at 3h or upon deletion of persistent site, these clusters disappear leaving persistent site behind still bound by ERα as bookmark, for next round of ligand stimulation.

https://doi.org/10.1371/journal.pgen.1008516.g008

appearance of ERα condensates, ERα-genomic clusters and robust gene activation over the time course of signaling response, clearly depicting the inability of persistent sites in triggering the gene expression on their own unless the transient sites form a functional unit.

Materials and methods

Cell culture

MCF7 cells were obtained from ATCC. They were cultured in High glucose DMEM media (Invitrogen) at 37˚C and 5% CO_2 conditions in humidified chamber. For ligand stimulation, cells were grown in stripping media containing high glucose DMEM without phenol red and 10% charcoal stripped FBS for three days for all the experiments except for the ChIP-seq and biochemical fractionation studies where stripping was performed for seven days with daily media change. On the third day cells were treated with β-estradiol (E2758, Sigma-Aldrich) at 10nM concentration for various periods as mentioned in the respective figures. For untreated

control, cells were either treated with equal microliters of ethanol on the third day or with ERα inhibitor ICI182780 (1047, Tocris Biosciences) at 10nM concentration for 24 h after two days of stripping.

1,6 Hexanediol treatments

Cells pretreated with either E2 or vehicle for 30 min were then treated with 5% 1,6-HD for 30 min in the same E2/vehicle containing media. Cells were then cross-linked to be processed for ChIP or were directly subjected to RNA isolation using Trizol method (Invitrogen) and subsequently cDNA synthesis was performed.

Chromatin immunoprecipitation and qPCRs

ChIP protocol was followed as described in [1]. Briefly, 5 million MCF-7 cells were cross-linked using 1% formaldehyde for 10 min at room temperature. Formaldehyde was quenched by 125μM Glycine for 5 min at RT with gentle rotation. Cells were scrapped in cold PBS, pelleted at 2000 rpm for 5 min at 4˚C. Pellet was dissolved in lysis buffer (50μM Tris HCl pH 7.4,1%SDS, 10μM EDTA and protease inhibitors), sonicated using Bioruptor Pico (Diagenode) to obtain fragments of 500 bp. Sonicated lysate was spun to remove the insoluble debris. Supernatant was diluted 2.5 times with dilution buffer (20mM Tris HCl pH7.4, 100μM NaCl, 2μM EDTA, 0.5% TritonX100 and protease inhibitors). 100 μg chromatin was taken for each IP, 1 μg ERα sc-8002 (Santa Cruz Biotechnology) and ab32063 (Abcam Inc.), antibody was added to bind with complexes overnight at 4˚C with gentle rocking motion. 15 μl 50% slurry of Protein G dynabeads (10004D, Invitrogen) were added for 4 hours and complexes were captured on magnetic rack, and the complexes were washed three times. The eluted complexes were reverse cross-linked overnight at 65˚C followed by purification with phenol:chloroform: isoamylalcohol. Purified chromatin was dissolved in 100 μl of TE (pH.8) buffer and 4 μl of chromatin was used per q-PCR reaction on CFX96 touch real time PCR (Bio-rad Laboratories). q-PCRs were carried out in three technical replicates and at least three biological replicates were performed. The fold changes were calculated by $2^{-\Delta\Delta ct}$ method or percent over input was calculated as described in https://www.thermofisher.com/in/en/home/life-science/epigenetics-noncoding-rna-research/chromatin-remodeling/chromatin-immunoprecipitation-chip/chip-analysis.html.

ChIP-seq library preparation and data analysis

Library preparation was performed as per manufacturer instruction for NEB Next ChIP-seq library preparation Kit (New England Biolabs). Briefly, at least 10 ng chromatin was subjected to end repair, T tailing and adapters compatible with Illumina sequencing platform were ligated. The amplified library was size-selected using AMPure XP beads to enrich fragments from 300 to 500 bp in size. 12 pico moles of library was used for cluster generation and sequencing was performed in 1X50 bp format on Hiseq 2500 (Illumina Inc.).

The sequenced reads were aligned to hg18 assembly using default Bowtie2 [51] options. Tag directories were made from the aligned reads to identify ChIP-seq peaks using HOMER [52]. A 200bp sliding window was used to identify narrow peaks which are characteristic of transcription factor peaks. The common artifacts from clonal amplification were neglected as only one tag from each unique genomic position was considered. 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 500bp (FDR<0.001) to identify enriched loci. Enriched regions separated by <1kb were merged and considered as blocks of variable lengths.

The read densities as bed graph files were calculated across the genome and this track was uploaded to UCSC genome browser.

ChIP-seq data analysis for persistent/ transient sites classification

ERα ChIP-seq peaks for vehicle and E2 condition were obtained from [2]. Coordinates overlapping with UCSC blacklisted regions, chromosome arm end, and peaks called in MCF7 Input ChIP-seq were removed from the ERα peaks. Common sites called as peaks in both Vehicle and E2 condition were classified as persistent sites and those called as peaks in only E2 condition were classified as transient sites. Genomic clustering analysis was performed on persistent sites. ROSE algorithm [13] was used on the same dataset to call ERα LDECs.

For the temporal data analysis we used time course ChIP-seq data from [4]. We have defined the unique and persistent sites based on the temporal dataset (Sites with no signal in 0 min in the temporal dataset is classified as transient and those with signal in 0 min classified as persistent). The persistent sites were divided into 4 quantiles based on their signal in the–E2 condition, defined based on signal intensity range as 0–0.25, 0.25–0.5, 0.5–0.75, and 0.75–1. We combined the bottom 2 quantiles into a signal quantile as the temporal data showed no signal in the bottom-most quantile, thus resulting in 3 quantiles, with 1961, 908, and 910 sites in the first, second, and the third quantile respectively.

ERα-clustered enhancers

ChIP-seq peaks called (described above) were selected if they were within 20kb distance of each other. Clustered enhancers were considered as clusters of three or more such peaks.

Super-enhancer calling

Super-enhancers were identified using the ROSE (Rank Ordering of Super-enhancers) algorithm (https://bitbucket.org/young_computation/) using the aligned ChIP-seq reads as input with parameters -s 15000.

Phast-Cons analysis

Phast-Cons scores (scores for multiple alignments of 99 vertebrate genomes to the human genome) across the human genome as a bigwig file was obtained from UCSC genome browser. This was then used to plot Phast-Cons scores across different regions.

GRO-seq analysis

The reads were aligned to hg18 assembly using default Bowtie2 options. The reads were counted from the region between +1kb of gene promoter proximal end to 13kb of the gene body using HOMER for each gene. The read densities were calculated similar to that of the ChIP-seq analysis in a strand specific format.

FOXA1 knockdown

Cells were transfected with FOXA1 specific siRNAs (L-010319-00-0005). Cells were harvested 24h post transfections to measure the RNA and protein levels of ERα and FOXA1.

5C library preparation and data analysis

5C oligos were designed in alternate orientation [21] on ERα peaks overlapping with very strong enhancers on chr21 as identified in [2]. For normalizing the digestion and ligation

efficiency biases between the samples, the 5C oligos were designed on ENCODE chr16 gene desert region [21], sequences of all 5C oligos are mentioned in S1 Table. 5C was performed in duplicates using BamHI restriction endonuclease in ICI and 1h E2 treated cells as reported [53,54]. Number of reads for each 5C in ICI are shown in the S2 Table and the Pearson correlation was more than 0.95 for ICI as well as E2 libraries. Raw sequencing reads were mapped to an artificial genome of all possible combinations of forward and reverse primers using Bowtie2. The uniquely mapped reads were then used to build a matrix of interactions between all the reverse and forward primers. The intra-chromosome 21 primers interactions were then normalized to their interactions with chr16 gene dessert for each interaction combination. The normalized read counts for the forward and the reverse oligos falling in a given region were binned and sum of the reads was plotted as arches using Sushi (http://www.bioconductor.org/) in Fig 3F and 3G.

ChIA-PET data analysis

ERα ChIA-PET in MCF-7 interaction files were obtained from ENCODE. All three replicates were merged to obtain common ends in all replicates using pairtopair tool from BED tools. Interaction ends were annotated as the different classes of ERα sites using pairtobed BEDtool. The number of interacting partners from each annotated interaction end was defined as the degree of the ERα site and the degree was then plotted for the different classes.

Shannon entropy calculations

Entropy values for every 100 kb bins before and after E2 treatmnets were directly taken from [9] and were plotted for the bins mentioned in Fig 3B.

Immunostaining

Immunostaining was performed as mentioned in [55]. Briefly, MCF-7 cells were grown on TC coated coverslips in phenol free media containing charcoal stripped serum for 24 hrs. Cell were treated with 10nM E2 or ethanol as vehicle for various time points before fixing them with 4% paraformaldehyde or methanol for 10 minutes followed by permeabilization for 10 min with 0.1% Triton-X-100 in PBS followed by blocking with 1% BSA for 15 min. Coverslips were incubated with ERα antibody (sc-8002, Sigma-Aldrich Corporation) followed by incubation with secondary antibodies conjugated with fluorophores (Life sciences Inc.). Coverslips were treated with 1nM DAPI for 2 min and mounted using Vectashield antifade mounting media (H-1200, Vector Laboratories). Immunofluorescence slides using similar settings were imaged using PLAPON 60x/1.42 oil objective of Olympus FV3000 microscope. DAPI signal was used to identify surfaces denoting the nuclear volume. ERα immunofluorescence signal was used to identify spherical spots of 0.8um in both the treated and untreated cells using similar intensity thresholds. The mean intensity of ERα signal at the identified ERα spots was calculated and plotted.

Cell imaging

Cells were imaged using PLAPON 60x/1.42 oil objective of Olympus FV3000 microscope. Images were taken with intervals of 1.1 seconds.

Immunofluorescence (IF) coupled with DNA-FISH and the analysis

IF/DNA-FISH was performed following the protocol as described in [56]. Briefly, Cells were permeabilized through treatment with cytoskeletal extraction buffer (CSK:100 mM NaCl, 300

mM sucrose, 3 Mm MgCl2, and 10 mM PIPES buffer, pH 6. 8) containing 0.4% Triton X-100 (SRL, #64518) and fixed with 4% paraformaldehyde. For IF, cells were washed 3X in PBS for 3 min each and then incubated in blocking buffer (0.5 mg/mL BSA, in 1X PBS with 0.2% Tween-20) for 30 min at 37˚C in a humid chamber. Following blocking, cells were next incubated with ERα antibody (sc-8002, Sigma-Aldrich Corporation) for 1 hr. The samples were then washed 3X in PBS/0.2% Tween 20 for 3 min. and incubated with fluorescently-conjugated secondary antibody (Alexa Fluor, Invitrogen) for 30 min. After three washes in PBS/0.2% Tween-20 for 3 min each, cells were processed for DNA FISH. For DNA FISH, cells were refixed with 1% paraformaldehyde containing 0.5% Tergitol and 0.5% Triton X-100. Next, dehydration was done through wash with ethanol series (70%, 85%, and 100% ethanol, 2 min each) and air dried for 15 mins. The cells were then treated with RNase A (1.25 ug/μl) at 37˚C for 30 min. The cells were again dehydrated and air dried as described above. The samples were then denatured in a prewarmed solution of 70% formamide in 2X SSC on a glass slide stationed on top of a heat block set at 95˚C for 11 min followed immediately by dehydration through a -20˚C chilled ethanol series (70%, 85%, 95%, and 100% ethanol, 2 min each). The cells were then air dried and hybridized with probe for overnight at 37˚C. The next day, the samples were washed twice with prewarmed 50% formamide/2X SSC solution at 39˚C and 2X with 2X SSC, 7 min each. The dsDNA FISH probes were made by randomly-priming DNA templates using BioPrime DNA Labeling System (18094011, Invitrogen). Probes were labelled with Fluorescein-12-dUTP (Invitrogen) and Cy3-dUTP (ENZO life science).

Following BAC probe (BACPAC Resources) was used:

NRIP1: RP11-213G23 (hg18, chr21:15245033–15430733)

TFF1: RP11-619I15 (hg18, chr21:42514453–42703548)

ImmunoFISH intensity analysis

The Immunofluorescence coupled DNA-FISH slides were imaged using PLAPON 60x/1.42 oil objective of Olympus FV3000 microscope with the same settings. IMARIS was used for the colocalisation analysis. ERα immunofluorescence and FISH signal was used to identify spherical spots of 0.8um and 1um diameter respectively. Spots were identified using the same intensity thresholds. Spots colocalising with each other were identified with distance threshold of 1 pixel and classified as colocalised and non-colocalised spots. The intensity of ERα signal at the ERα and NRIP1 spots was calculated and plotted.

For the TFF1 FISH analysis we have used normalised intensity values to correct for any differences in the total ERα intensities between wildtype and TFF1 delete cells. These values were calculated by first identifying TFF1 FISH spots as described above. The nucleus is then identified based on DAPI. The mean intensity of ERα was calculated on the FISH spot. This value is then divided by the mean intensity of ERα in the corresponding nucleus to obtain the normalised intensity of ERα on the FISH spot. This value is a measure of relative enrichment of ERα signal on the FISH spot compared to the corresponding nucleus.

$$Normalised\ Intensity = \frac{Mean\ Intensity\ of\ ERa\ Within\ FISH\ spot}{Mean\ Intensity\ of\ ERa\ Within\ Corresponding\ Nucleus}$$

We used an arbitrary normalised intensity (enrichment ratio) of 1.4 to define co-localisation of ERα with the the FISH spot.

To test the specificity of ERa signal on TFF1 FISH spot, we generated random ROIs (region of interest) within the nuclei of the different fields using matlab. We then identified spots and quantified the normalised intensity within random nuclear regions as described earlier.

FRAP quantification

GFP-ERα (Addgene #Plasmid 28230) was overexpressed in MCF7 cells in phenol free media with charcoal stripped serum. Cells were treated with Ethanol (Vehicle) or with E2 (10 nM) for one hour post 24h of transfections. For 1,6-HD treatments, 5% 1,6-HD was given to transfected cells that were pretreated with E2 for 30 min to continue other 30 min but along with 1,6-HD. FRAP was performed on Olympus FV3000 microscope with 488nm laser. Bleaching was performed over an area of 1μm using 100% laser power and images were collected every two seconds. The intensity of the photo-bleached ROI was calculated across all the frames in Fiji [57]. Background intensity was subtracted from the value of intensity at the ROI at each frame. This value was then normalized to the whole cell intensity and plotted. The same value is calculated for an unbleached ROI in the same way as well and plotted to control for bleaching due to image acquisition [58].

RNA isolation and cDNA synthesis

RNA was isolated using Trizol (Invitrogen) as per the manufacturer protocol. 1 μg of RNA was taken for cDNA synthesis using random hexamer as per manufacturer guideline (Superscript IV RT-PCR kit, Life Technologies). cDNA was subjected to real time PCR in triplicates in CFX96 touch real time PCR (Bio-rad laboratories) using oligos mentioned in S3 Table. Fold changes were calculated by $2^{-\Delta\Delta ct}$ formula. GAPDH or b-actin expression was taken as control for fold changes in gene expression.

HiC-analysis

The raw reads were mapped to hg18 assembly using bowtie 2. The HOMER program makeTagDirectory was first used to create tag directories with tbp 1. Data was further processed by HOMER in order to remove small fragment and self-ligations using makeTagDirectory with the following options: -removePEbg -removeSpikes 10000 5. Next, findTADsAndLoops.pl was used to obtain overlapping TADs, produced at a 20kb resolution with 40kb windows. Smallest overlapping TADs were selected and intersected with ERα clusters. Strength of these TADs were obtained as Inclusion Ratios (the ratio of intra-TAD interactions relative to interactions from the TAD to the surrounding region (both upstream and downstream of the TAD to regions the same size as the TAD) across the different conditions using findTADsAndLoops.pl score option.

Nuclear lysate fractionation and immunoblotting

Cells were washed twice in phosphate buffered saline (1XPBS) and cell pellet was carefully suspended in 200ul of SFI Buffer (100mM NaCl, 300mM Sucrose, 3mM MgCl2, 10mM PIPES [pH 6.8], 1mM EGTA, 0.2% Triton X-100,) with PIC. Cells were incubated for 30 mins at 4˚C and then pelleted down at 2900 rpm for 5 min at 4˚C, and the supernatant was collected (Soluble fraction). The pellet was carefully washed twice in SFI buffer and chromatin bound fraction was extracted by adding 2X protein loading dye to the pellet. Fractions were loaded on 15% SDS PAGE and western was done for ERα (sc-8002 Santa Cruz Biotechnology), GAPDH (sc-32233 Santa Cruz Biotechnology), Histone H3 (H0164, Sigma-Alderich). Cellular lysates prepared were separated on 15% SDS-PAGE gels. The protein transfer was carried out in Tris-glycine buffer at 30V for 1 hour on ice using 0.45μ PVDF membrane (Millipore). Membrane was blocked in 5% non-fat dry milk made in TBST and further incubated for three hours with primary antibodies followed by stringent washings. Membranes were probed with HRP-conjugated secondary antibody (Bio-rad Laboratories). Signal amplification was performed using

ECL substrate (RPN2106, GE Healthcare). Images were captured on Image Quant LAS4000 with CCD camera. Anti-FOXA1 (ab5089, Abcam) and MED1 (A3007938, Bethyl) were used for total protein levels.

gRNA design and cloning

We used gRNA design tool at crispr.mit.edu to design the gRNAs for ERα peak in the GREB1 and TFF1 persistent and transient sites (S3 Table). The gRNAs were cloned in a customized PX459 vector (pSpCas9(BB)-2A-Puro V2.0, Addgene # 62988) [59] was a gift from Feng Zhang. The Cas9 enzyme cassette was removed from PX459 by digesting it with Xba1 and Not1 followed by blunt end ligation. The cloning of the gRNAs in the customized PX459 plasmid was performed as per the Zhang Lab general cloning protocol.

Deletion of persistent sites

Cells were transfected with cas9 plasmid and gRNAs constructs described above. Transfected cells were kept under puromycin selection for 48h post 24h of transfection. After two days of selection, cells were diluted at the density of 0.5 cell/100 μl and 100 μl was dispensed into each well of a 96 plate. Wells containing single cells were identified under microscope and marked. Media was changed every 5 days till the colony appeared. Colonies from single cells were screened for the homozygous deletion. Second round of gRNA transfection was done on heterozygous line that was obtained from the first round of cas9 experiment.

CRISPR blocking and surveyor assays on GREB1 persistent site

For CRISPR blocking, dcas9 plasmid along with scr gRNA or two gRNA specific to GREB1 persistent site were transfected in 24h hormone stripped MCF-7 cells. As a control for gRNA specificity, surveyor assays were performed for which, active cas9 along with scr or GREB1 PS gRNA's were transfected in a separate dish. 24h post transfections, CRISPRi cells were harvested for ChIP and gene expression analysis whereas, cells transfected with active cas9 were analysed for cut products using primers that flank the gRNA binding sites of GREB1 persistent sites.

Supporting information

S1 Fig. ERα clusters around persistent sites in an E2-dependent manner. (A) Venn diagram shows distribution of ERα on various genomic regions. (B) Strength of ERα occupancy, DHS, and H3K27ac levels at ERα peaks before and after E2 treatments. (C) Heatmaps of ERα, DHS, and H3K27ac 2.5 kb upstream and downstream regions from the center of the ERα peaks in untreated and E2 treated condition. (D) Heatmaps depicting the relative binding of ERα on -E2 specific unique sites, persistent sites and transient sites across different datasets. (E) H3K27ac enrichment plot in minus and plus E2 on 2880 -E2 unique, persistent, transient closer to 3rd quantile, and transient sites. (F) ERα binding strength is affected upon siRNA mediated knockdown of ERα in untreated cells (data from Caizzi et al., 2014.) (G) ERα binding strength on 2880 -E2 unique sites is affected upon siRNA mediated knockdown of ERα in untreated cells (data from Caizzi et al., 2014).
(TIF)

S2 Fig. Persistent sites are bound by ERα in ligand independent manner. (A) Heatmap showing the relative ERα binding strength on various classes of ERα binding sites after 7 days of stripping and frequent change of media. (B) Immunoblot for ERα, GAPDH and Histone H3 in nucleoplasmic (soluble) and chromatin bound biochemical fractions in cells stripped for

7 days followed by 60 and 180 min E2 treatments. (C) (top panel) Known Motif enrichment analysis identifies full ERE in both persistent ($p = 10^{-1136}$) and transient sites ($p = 10^{-5402}$) whereas FOXA1 is enriched uniquely in persistent with $p = 10^{-226}$ and in -E2 unique sites $p = 10^{-174}$. (D) Heatmaps representing the strength of FOXA1 binding in different categories of ERα peaks in treated and untreated cells. Strength was measured at 1.5 kb upstream and downstream of center of ERα peak.
(TIF)

S3 Fig. ERα binding in genomic clusters. (A) ERα binding strength in clusters with and without persistent sites in E2 untreated and treated conditions. (B) Number of ERα peaks are greater in clusters with persistent site as compared to clusters without persistent sites. (C) Phast-cons score of persistent, 3rd quantile persistent, transient, and transient near persistent sites.
(TIF)

S4 Fig. ERα and DHS in genomic clusters. (A) Heatmaps exhibit the loss of ERα binding strength at every 2 consecutive EREs from persistent site (B) Heatmap shows DHS signal on sites in panel A.
(TIF)

S5 Fig. ERα clustered enhancers but not conventional super enhancers control E2 target genes. (A) GRO-seq tag count shows the relative higher expression of genes near clusters with a persistent site. (B) GRO-seq tag count shows higher expression of genes closer to persistent vs. transient, random and -E2 unique sites. Note: relative higher expression of these genes (first bar) even in untreated cells. (C) ChIA-PET data plotted from one ERα ChIA-PET replicates on *GREB1* LDEC as shown in Fig 3D. (D) TAD structure around *TFF1* and *UMODL1* genes in MCF-7 cells.
(TIF)

S6 Fig. Deletion and blocking of persistent sites. (A) UCSC genome browser snap shot of *TFF1* region showing blue highlighted persistent sites. Dashed line box marks the deleted regions. (B) Surveyor assay using the oligos specific for the region outside the deleted PS. Wt genomic DNA exhibits the larger molecular weight amplicon compared to the amplicon from ΔPS-Tff1 genomic DNA. (C) Sanger sequencing chromatogram shows the fusion of yellow and blue highlighted regions in TFF1 delete line, whereas these regions are 1611 nucleotide apart in wild type cells. (D) UCSC genome browser snapshot on TFF1 PS region shows the loss of ERα ChIP-seq peaks in delete cells as compared to wild-type cells (Upper track). Browser snap shot on the wider region around deleted TFF1 site, highlighted region depicts the deleted region (Lower track). (E) UCSC genome browser snap shot of *GREB1* region showing blue highlighted persistent site which was blocked by specific gRNAs. (F) gRNA's cut the specific region within the enhancer as shown by surveyor assay using oligos outside of blocked region, PCR was performed on population of cells after transfection so larger and smaller both amplicons are seen.
(TIF)

S7 Fig. Persistent sites are required for the emergence of transient ERα sites within clustered enhancers. (A) Conformation of TFF1 persistent site deletion in second CRISPR clone (Left panel) and conformation of deletion by sequencing of genomic DNA (Right Panel) (B) UCSC genome bowser track on TFF1 cluster showing ERα ChIP-seq peaks in untreated and treated cells. Highlighted regions depict the deleted region. (C) ChIP-qPCRs shows the loss of ERα binding at different regions of TFF1 cluster upon persistent site deletion in second

CRIPSR clone. (D) ChIP-qPCRs confirm the loss of ERα binding at P3 site where gRNAs were targeted in the GREB1 cluster (Upper panel); a modest loss of ERα binding was seen on distal ERα binding site in neighboring TAD (Lower panel). (E) ChIP-qPCRs shows effect in ERα binding strength at GREB1 enhancer region when persistent site near promoter was blocked with gRNAs.
(TIF)

S8 Fig. ERα occupancy on E2-target regions. UCSC Genome Browser shots show the occupancy of ERα in WT and ΔTFF1-PS line at various E2 target genes.
(TIF)

S9 Fig. ERα puncta are formed on LDECs by coalescing. (A) Representative immunoFISH image for NRIP1 (Left panel) and mean ERα intensity on NRIP1 FISH spots (Right panel). (B) Comparison of ERα intensity on all puncta vs. the puncta that overlap with *NRIP1* loci by immunoFISH upon ICI and E2 treatment for 1h. (C) Time lapse microscopy images of GFP-ERα upon 1,6-HD treatment in cells treated with E2 for one hour. 0 vs. 440 sec shows significant loss of ERα punctate pattern (D) UCSC genome browser snapshot of *TFF1* region showing the ERα ChIP-seq peaks in untreated (Top track) and E2 treated (Bottom track) conditions. Dashed boxes mark the regions on which ERα occupancy was measured upon 1,6-HD treatments. (E) ChIP-qPCRs for ERα on R3 and R4 peaks upon 1,6-HD treatments in E2 treated and untreated conditions.
(TIF)

S10 Fig. Emergence of ERα clustering is correlated with robust target gene expression. (A) GRO-seq tag counts from plus and minus strands at different categories of ERα peaks during E2 time course. (B) Relative mRNA expression of genes closer to LDEC at different hours of E2 signaling. (C) Immunoblots probed for MED1, FOXA1 and GAPDH in ICI and 60, 180min and 24h post E2 exposure. (D) Levels of ESR1 and FOXA1 by qRT-PCR (upper Panel) and immunoblot for ERα and GAPDH (lower Panel).
(JPG)

S1 Movie. Live imaging of GFP-ERα upon E2 stimulation. Rest period 4.8sec, exposure 3.2 sec. total interval 8 sec. Three movies taken for different periods were stitched together.
(ZIP)

S2 Movie. Live imaging of GFP-ERα post E2 stimulation. Rest period 4.8sec, exposure 3.2 sec. total interval 8 sec.
(ZIP)

S1 Table. The table lists the sequences of forward and reverse oligo used in 5C experiments.
(PDF)

S2 Table. The table shows the number of reads from 5C experiments.
(PDF)

S3 Table. The table lists the sequences of primers used in qRT-PCRs and CRISPR studies.
(PDF)

S4 Table. The table shows the accession numbers of publicly available and our unpublished data used in the study.
(DOCX)

S5 Table. The table shows the called peaks for all ERα peaks, Persistent peaks, transient peaks, LEDC and TADs from processed NGS data.
(XLSX)

Acknowledgments

We thank Sakshi Gorey for help with CRISPR experiments. Authors would like to thank Dan Larson and Amanjot Singh for their inputs on an earlier draft of the manuscript. We also thank Awadhesh Pandit from Genomics core facility at NCBS for his help in sequencing.

Author Contributions

Conceptualization: Bharath Saravanan, Sridhar Hannenhalli, Dimple Notani.

Data curation: Bharath Saravanan, Deepanshu Soota, Sridhar Hannenhalli, Dimple Notani.

Formal analysis: Bharath Saravanan, Deepanshu Soota, Zubairul Islam, Rajat Mann, Sweety Meel, Umer Farooq, Kaivalya Walavalkar, Anurag Kumar Singh, Dimple Notani.

Funding acquisition: Dimple Notani.

Investigation: Bharath Saravanan, Deepanshu Soota, Zubairul Islam, Rajat Mann, Sweety Meel, Umer Farooq, Dimple Notani.

Methodology: Bharath Saravanan, Deepanshu Soota, Zubairul Islam, Sudeshna Majumdar, Rajat Mann, Sweety Meel, Umer Farooq, Dimple Notani.

Project administration: Dimple Notani.

Resources: Srimonta Gayen, Dimple Notani.

Software: Bharath Saravanan.

Supervision: Sridhar Hannenhalli, Dimple Notani.

Validation: Bharath Saravanan, Sridhar Hannenhalli, Dimple Notani.

Visualization: Bharath Saravanan, Kaivalya Walavalkar.

Writing – original draft: Bharath Saravanan, Sridhar Hannenhalli, Dimple Notani.

Writing – review & editing: Deepanshu Soota, Zubairul Islam, Rajat Mann, Sweety Meel, Umer Farooq, Kaivalya Walavalkar, Srimonta Gayen, Anurag Kumar Singh, Dimple Notani.

References

1. 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 CK, Rosenfeld MG. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013. 498(7455):516–20. https://doi.org/10.1038/nature12210 PMID: 23728302
2. Liu Z, Merkurjev D, Yang F, Li W, Oh S, Friedman MJ, Song X, Zhang F, Ma Q, Ohgi KA, Krones A, Rosenfeld MG. Enhancer activation requires trans-recruitment of a mega transcription factor complex. Cell. 2014. 159(2):358–73. https://doi.org/10.1016/j.cell.2014.08.027 PMID: 25303530
3. Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL. A Rapid, Extensive, and Transient Transcriptional Response to Estrogen Signaling in Breast Cancer Cells. Cell. 2011. 145(4): 622–634. https://doi.org/10.1016/j.cell.2011.03.042 PMID: 21549415
4. Dzida T, Iqbal M, Charapitsa I, Reid G, Stunnenberg H, Matarese F, Grote K, Honkela A, Rattray M. Predicting stimulation-dependent enhancer-promoter interactions from ChIP-Seq time course data. PeerJ. 2017. 5:e3742. https://doi.org/10.7717/peerj.3742 PMID: 28970965

5. Bojcsuk D, Nagy G, Balint BL. Inducible super enhancers are organized based on canonical signal-specific transcription factor binding elements. Nucleic Acids Res. 2017. 45(7): 3693–3706. https://doi.org/10.1093/nar/gkw1283 PMID: 27994038
6. Holding AN, Cullen AE, Markowetz F. Genome-wide Estrogen Receptor-α activation is sustained, not cyclical. Elife. 2018. 7. pii: e40854. https://doi.org/10.7554/eLife.40854 PMID: 30457555
7. Welboren WJ1, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG. ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J. 2009. 28(10):1418–28. https://doi.org/10.1038/emboj.2009.88 PMID: 19339991
8. Caizzi L1, Ferrero G, Cutrupi S, Cordero F, Ballaré C, Miano V, Reineri S, Ricci L, Friard O, Testori A, Corà D, Caselle M, Di Croce L, De Bortoli M. Genome-wide activity of unliganded estrogen receptor-α in breast cancer cells. Proc Natl Acad Sci U S A. 2014. 111(13):4892–7. https://doi.org/10.1073/pnas.1315445111 PMID: 24639548
9. Rodriguez J, Ren G, Day CR, Zhao K, Chow CC, Larson DR. Intrinsic Dynamics of a Human Gene Reveal the Basis of Expression Heterogeneity. Cell. 2018. pii: S0092-8674(18)31518-6.
10. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B. A high-resolution map of the three dimensional chromatin interactome in human cells. Nature. 2013 Nov 14; 503 (7475):290–4. https://doi.org/10.1038/nature12644 PMID: 24141950
11. Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, Li CH, Guo YE, Day DS, Schuijers J, Vasile E, Malik S, Hnisz D, Lee TI, Cisse II, Roeder RG, Sharp PA, Chakraborty AK, Young RA. Coactivator condensation at super-enhancers links phase separation and gene control. Science. 2018. 361(6400).
12. Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, Abraham BJ, Afeyan LK, Guo YE, Rimel JK, Fant CB, Schuijers J, Lee TI, Taatjes DJ, Young RA. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell. 2018. 175(7):1842–1855.e16. https://doi.org/10.1016/j.cell.2018.10.042 PMID: 30449618
13. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke H, Young RA. Transcriptional super-enhancers connected to cell identity and disease. Cell. 2013. 155(4):10.
14. Swinstead EE, Miranda TB, Paakinaho V, Baek S, Goldstein I, Hawkins M, Karpova TS, Ball D, Mazza D, Lavis LD, Grimm JB, Morisaki T, Grøntved L, Presman DM, Hager GL. Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. Cell. 2016. 165(3):593–605. https://doi.org/10.1016/j.cell.2016.02.067 PMID: 27062924
15. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll, JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat. Genet. 2011. 43:27–33. https://doi.org/10.1038/ng.730 PMID: 21151129
16. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell. 2005. 122:33–43. https://doi.org/10.1016/j.cell.2005.05.008 PMID: 16009131
17. Fu X, Jeselsohn R, Pereira R, Hollingsworth EF, Creighton CJ, Li F, Shea M, Nardone A, De Angelis C, Heiser LM, Anur P, Wang N, Grasso CS, Spellman PT, Griffith OL, Tsimelzon A, Gutierrez C, Huang S, Edwards DP, Trivedi MV, Rimawi MF, Lopez-Terrada D, Hilsenbeck SG, Gray JW, Brown M, Osborne CK, Schiff R. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8expression in ER-positive breast cancer. Proc Natl Acad Sci U S A. 2016. 25; 113(43):E6600–E6609. https://doi.org/10.1073/pnas.1612835113 PMID: 27791031
18. Hsieh CL, Fei T, Chen Y, Li T, Gao Y, Wang X, Sun T, Sweeney CJ, Lee GS, Chen S, Balk SP, Liu XS, Brown M, Kantoff PW. Enhancer RNAs participate in androgen receptordriven looping that selectively enhances gene activation. Proc Natl Acad Sci U S A. 2014. 111(20):7319–24. https://doi.org/10.1073/pnas.1324151111 PMID: 24778216
19. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat Rev Genet. 2016. 17(4):207–23. https://doi.org/10.1038/nrg.2016.4 PMID: 26948815
20. Fullwood MJ, Liu MH, Pan YF, Liu J, Han X, Mohamed Y, Orlov YL, Velkov S, Ho A, Mei PH. An Oestrogen Receptor α-bound Human Chromatin Interactome. Nature. 2009. 462(7269): 58–64. https://doi.org/10.1038/nature08497 PMID: 19890323
21. Sanyal A, Lajoie B, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012. 489(7414): 109–113. https://doi.org/10.1038/nature11279 PMID: 22955621
22. Quintin J, Le Péron C, Palierne G, Bizot M, Cunha S, Sérandour AA, Avner S, Henry C, Percevault F, Belaud-Rotureau MA, Huet S, Watrin E, Eeckhoute J, Legagneux V, Salbert G, Métivier R. Dynamic estrogen receptor interactomes control estrogen-responsive trefoil Factor (TFF) locus cell-specific

activities. Mol Cell Biol. 2014. 34(13):2418–36. https://doi.org/10.1128/MCB.00918-13 PMID: 24752895

23. Rafique S, Thomas JS, Sproul D, Bickmore WA. Estrogen-induced chromatin decondensation and nuclear re-organization linked to regional epigenetic regulation in breast cancer. Genome Biol. 2015. 16:145. https://doi.org/10.1186/s13059-015-0719-9 PMID: 26235388

24. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature. 2017. 547(7662): 236–240. https://doi.org/10.1038/nature22822 PMID: 28636604

25. Baran-Gale J, Purvis JE, Sethupathy P. An integrative transcriptomics approach identifies miR-503 as a candidate master regulator of the estrogen response in MCF-7 breast cancer cells. RNA. 2016. 22 (10):1592–603. https://doi.org/10.1261/rna.056895.116 PMID: 27539783

26. Valley CC, Solodin NM, Powers GL, Ellison SJ, Alarid ET. Temporal variation in estrogen receptor-alpha protein turnover in the presence of estrogen. J Mol Endocrinol. 2008. 40(1):23–34. https://doi.org/10.1677/JME-07-0067 PMID: 18096994

27. Reid G, Hübner MR, Métivier R, Brand H, Denger S, Manu D, Beaudouin J, Ellenberg J, Gannon F. Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling. Mol Cell. 2003. 11(3):695–707. https://doi.org/10.1016/s1097-2765(03)00090-x PMID: 12667452

28. Nawaz Z, Lonard DM, Dennis AP, Smith CL, O'Malley BW. Proteasome-dependent degradation of the human estrogen receptor. Proc Natl Acad Sci U S A.1999. 96:1858–62. https://doi.org/10.1073/pnas.96.5.1858 PMID: 10051559

29. Lonard DM, Nawaz Z, Smith CL, O'Malley BW. The 26S proteasome is required for estrogen receptor-α and coactivator turnover and for efficient estrogen receptor-α transactivation. Mol Cell. 2000. 5:939–48. https://doi.org/10.1016/s1097-2765(00)80259-2 PMID: 10911988

30. Callige M, Kieffer I, Richard-Foy H. CSN5/Jab1 is involved in ligand-dependent degradation of estrogen receptor {α} by the proteasome. Mol Cell Biol. 2005. 25:4349–58. https://doi.org/10.1128/MCB.25.11.4349-4358.2005 PMID: 15899841

31. Grøntved L, John S, Baek S, Liu Y, Buckley JR, Vinson C, Aguilera G et al. C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. The EMBO journal. 2013. 32(11):1568–83. https://doi.org/10.1038/emboj.2013.106 PMID: 23665916

32. Grøntved L, Waterfall JJ, Kim DW, Baek S, Sung MH, Zhao L, Park JW, Nielsen R, Walker RL, Zhu YJ, Meltzer PS, Hager GL, Cheng SY. Transcriptional activation by the thyroid hormone receptor through ligand-dependent receptorrecruitment and chromatin remodelling. Nat Commun. 2015. 28;6:7048.

33. Johnson TA, Chereji RV, Stavreva DA, Morris SA, Hager GL, Clark DJ. Conventional and pioneer modes of glucocorticoid receptor interaction with enhancer chromatin in vivo. Nucleic Acids Res. 2018 Jan 9; 46(1):203–214. https://doi.org/10.1093/nar/gkx1044 PMID: 29126175

34. Voss TC1, Schiltz RL, Sung MH, Yen PM, Stamatoyannopoulos JA, Biddie SC, Johnson TA, Miranda TB, John S, Hager GL. Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell. 2011 Aug 19; 146(4):544–54. https://doi.org/10.1016/j.cell.2011.07.006 PMID: 21835447

35. He HH, Meyer CA, Chen MW, Jordan VC, Brown M, Liu XS. Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. Genome Res. 2012. (22):1015–1025.

36. Carleton JB, Berrett KC, Gertz J. Multiplex Enhancer Interference Reveals Collaborative Control of Gene Regulation by Estrogen Receptor Alpha Bound Enhancers. Cell Syst. 2017. 5(4): 333–344.e5. https://doi.org/10.1016/j.cels.2017.08.011 PMID: 28964699

37. Shin HY, Willi M, HyunYoo K, Zeng X, Wang C, Metser G, Hennighausen L. Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nat Genet. 2016. 48(8): 904–911. https://doi.org/10.1038/ng.3606 PMID: 27376239

38. Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, Dostie J, Bickmore WA. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. Genes Dev. 2014. 28(24):2778–91. https://doi.org/10.1101/gad.251694.114 PMID: 25512564

39. Wang LH, Yang XY, Zhang X, An P, Kim HJ, Huang J, Clarke R, Osborne CK, Inman JK, Appella E, Farrar WL. Disruption of estrogen receptor DNA binding domain and related intramolecular communication restores tamoxifen sensitivity in resistant breast cancer. Cancer Cell. 2006. 10(6):487–99. https://doi.org/10.1016/j.ccr.2006.09.015 PMID: 17157789

40. Shin Y, Chang YC, Lee DSW, Berry J, Sanders DW, Ronceray P, Wingreen NS, Haataja M, Brangwynne CP.Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. Cell. 2018. 175(6):1481–1491. https://doi.org/10.1016/j.cell.2018.10.057 PMID: 30500535

41. Kang YK1, Guermah M, Yuan CX, Roeder RG. The TRAP/Mediator coactivator complex interacts directly with estrogen receptors alpha and beta through the TRAP220 subunit and directly enhances estrogen receptor function in vitro. Proc Natl Acad Sci U S A. 2002. 99(5):2642–7. https://doi.org/10.1073/pnas.261715899 PMID: 11867769

42. Malik S, Roeder RG. Nuclear Receptors. Elsevier Methods in Enzymology. 2003. 364257–284.

43. Tanida T, Matsuda KI, Yamada S, Hashimoto T, Kawata M. Estrogen-related Receptor β Reduces the Subnuclear Mobility of Estrogen Receptor α and Suppresses Estrogen-dependent Cellular Function. J Biol Chem. 2015. 290(19):12332–45. https://doi.org/10.1074/jbc.M114.619098 PMID: 25805499

44. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. Biochemistry. 2006 45(22): 6873–88. https://doi.org/10.1021/bi0602718 PMID: 16734424

45. Harmon TS, Holehouse AS, Rosen MK, Pappu RV. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. elife. 2017. 6, e30294. https://doi.org/10.7554/eLife.30294 PMID: 29091028

46. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature. 2013. 494(7438):497–50. https://doi.org/10.1038/nature11884 PMID: 23417068

47. Cheng D, Vemulapalli V, Lu Y, Shen J, Aoyagi S, Fry CJ, Yang Y, Foulds CE, Stossi F, Treviño LS, Mancini MA, O'Malley BW, Walker CL, Boyer TG, Bedford MT. CARM1 methylates MED12 to regulate its RNA-binding ability. Life Sci Alliance. 2018. 1(5).

48. Lin Y, Protter DS, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell. 2015. 60(2):208–19. https://doi.org/10.1016/j.molcel.2015.08.018 PMID: 26412307

49. Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, Rosen MK. Compositional Control of Phase-Separated Cellular Bodies. Cell. 2016. 166(3):651–663. https://doi.org/10.1016/j.cell.2016.06.010 PMID: 27374333

50. Fan M, Nakshatri H, Nephew KP. Inhibiting proteasomal proteolysis sustains estrogen receptor-alpha activation. Mol Endocrinol. 2004. (11):2603–15. https://doi.org/10.1210/me.2004-0164 PMID: 15284335

51. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Mar 4; 9 (4):357–9. https://doi.org/10.1038/nmeth.1923 PMID: 22388286

52. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010 May 28; 38(4):576–89. https://doi.org/10.1016/j.molcel.2010.05.004 PMID: 20513432

53. van Berkum NL, Dekker J. Determining spatial chromatin organization of large genomic regions using 5C technology. Methods Mol Biol. 2009. 567:189–213. https://doi.org/10.1007/978-1-60327-414-2_13 PMID: 19588094

54. Ferraiuolo MA, Sanyal A, Naumova N, Dekker J, and Dostie J. Mapping chromatin interactions with 5C technology. 5C; a quantitative approach to capturing chromatin conformation over large genomic distances. Methods. 2012. 58(3): 10.1016.

55. Notani D, Gottimukkala KP, Jayani RS, Limaye AS, Damle MV, Mehta S, Purbey PK, Joseph J, Galande S.Global regulator SATB1 recruits beta-catenin and regulates T(H)2 differentiation in Wnt-dependent manner. PLoS Biol. 2010. 8(1):e1000296. https://doi.org/10.1371/journal.pbio.1000296 PMID: 20126258

56. Gayen S, Maclary E, Buttigieg E, Hinten M, Kalantry S. A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation. Cell Rep. 2015. 11(8):1251–65. https://doi.org/10.1016/j.celrep.2015.04.039 PMID: 25981039

57. Schindelin J1, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012. 289(7):676–82.

58. Sprague BL, Pego RL, Stavreva DA, McNally JG. Analysis of binding reactions by fluorescence recovery after photobleaching. Biophys J. 2004. 86:3473–3495. https://doi.org/10.1529/biophysj.103.026765 PMID: 15189848

59. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013. 8(11):2281–308. https://doi.org/10.1038/nprot.2013.143 PMID: 24157548