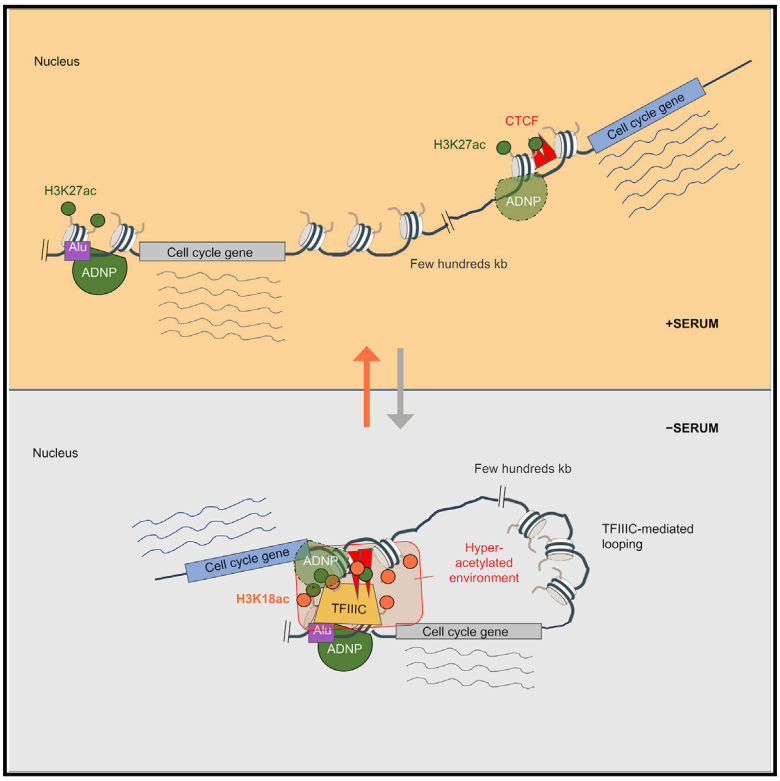
Article

Molecular Cell

TFIIIC Binding to Alu Elements Controls Gene Expression via Chromatin Looping and Histone Acetylation

Graphical Abstract



Authors

Roberto Ferrari,
Lara Isabel de Llobet Cucalon,
Chiara Di Vona, ..., Martin Teichmann,
Susana de la Luna, Miguel Beato

Correspondence

roberto.ferrari@crg.eu (R.F.),
miguel.beato@crg.eu (M.B.)

In Brief

Repetitive elements shape genome structure and function. Ferrari et al. find that cells respond to serum deprivation by redirecting the general transcription factor TFIIIC to acetylate ADNP-bound Alu elements in order to rewire the 3D genome architecture via CTCF looping, ultimately sustaining steady-state levels of cell-cycle-regulated gene expression.

Highlights

- Serum starvation recruits TFIIIC at ADNP-bound Alu Elements (AEs) near Pol II genes
- TFIIIC-associated histone acetylase activity acetylates H3K18 over the bound AEs
- TFIIIC-bound acetylated AEs loop to contact CTCF at distal cell-cycle genes' promoters
- CTCF-TFIIIC interaction ensures rapid cell-cycle genes' reactivation on serum exposure

Ferrari et al., 2020, Molecular Cell 77, 475–487
February 6, 2020
https://doi.org/10.1016/j.molcel.2019.10.020

CellPress

Molecular Cell
Article

TFIIIC Binding to Alu Elements Controls Gene Expression via Chromatin Looping and Histone Acetylation

Roberto Ferrari,[1,9,*] Lara Isabel de Llobet Cucalon,[1,9] Chiara Di Vona,[1,2] François Le Dilly,[1] Enrique Vidal,[1] Antonios Lioutas,[1] Javier Quilez Oliete,[1] Laura Jochem,[3] Erin Cutts,[3] Giorgio Dieci,[4] Alessandro Vannini,[3,8] Martin Teichmann,[5] Susana de la Luna,[1,2,7] and Miguel Beato[1,6,10,*]

[1]Center for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Dr. Aiguader 88, Barcelona 08003, Spain
[2]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain
[3]The Institute of Cancer Research (ICR), London, UK
[4]Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy
[5]Université de Bordeaux, INSERM U1212 CNRS UMR 5320 146, Bordeaux, France
[6]Universitat Pompeu Fabra (UPF), Barcelona, Spain
[7]ICREA, Pg. Lluis Companys 23, Barcelona 08010, Spain
[8]Human Technopole. Via Cristina Belgioioso, 171, 20157 Milano MI, Italy
[9]These authors contributed equally
[10]Lead Contact
*Correspondence: roberto.ferrari@crg.eu (R.F.), miguel.beato@crg.eu (M.B.)
https://doi.org/10.1016/j.molcel.2019.10.020

SUMMARY

How repetitive elements, epigenetic modifications, and architectural proteins interact ensuring proper genome expression remains poorly understood. Here, we report regulatory mechanisms unveiling a central role of Alu elements (AEs) and RNA polymerase III transcription factor C (TFIIIC) in structurally and functionally modulating the genome via chromatin looping and histone acetylation. Upon serum deprivation, a subset of AEs pre-marked by the activity-dependent neuroprotector homeobox Protein (ADNP) and located near cell-cycle genes recruits TFIIIC, which alters their chromatin accessibility by direct acetylation of histone H3 lysine-18 (H3K18). This facilitates the contacts of AEs with distant CTCF sites near promoter of other cell-cycle genes, which also become hyperacetylated at H3K18. These changes ensure basal transcription of cell-cycle genes and are critical for their re-activation upon serum re-exposure. Our study reveals how direct manipulation of the epigenetic state of AEs by a general transcription factor regulates 3D genome folding and expression.

INTRODUCTION

The mammalian genome is shaped by the expansion of repetitive elements that provide new regulatory networks for coordinated control of gene expression (Chuong et al., 2017) and genome folding (Cournac et al., 2016; Pombo and Dillon, 2015; van de Werken et al., 2017). In particular, Alu elements (AEs) are retained close to the transcription start site (TSS) of genes (Tsirigos and Rigoutsos, 2009), show proto-enhancer functions (Su et al., 2014), correlate with the level of chromatin interactions (Gu et al., 2016), and are recognized by the RNA polymerase III (Pol III) general transcription factor III (TFIIIC) (Dieci et al., 2007). TFIIIC recognizes AEs by binding to the so-called A- and B-boxes (Dieci et al., 2007). Besides acting in Pol III transcription, TFIIIC binds to so-called "extra TFIIIC" sites (ETCs) (Moqtaderi et al., 2010), which carry a non-canonical B-box and are devoid of the remaining Pol III machinery. It also acts as an insulator participating in 3D genome organization (Pascali and Teichmann, 2013; Van Bortle and Corces, 2012). In *Drosophila*, TFIIIC redistributes within the genome upon heat shock to rewire DNA looping within topologically associating domains (TADs) favoring proper gene expression (Li et al., 2015). In mouse, TFIIIC binding to short interspersed nuclear elements (SINEs) controls the relocation of the activity-dependent neuronal genes *Fos* and *Gadd45a* to transcription factories (Crepaldi et al., 2013). TFIIIC associates with promoters of N-MYC target genes, facilitates the recruitment of the Cohesin complex subunit RAD21, and is required for RNA polymerase II (Pol II) escape and pause release (Büchel et al., 2017). However, the precise role of human TFIIIC in 3D genome shaping during stress conditions remains unknown. Here, we use serum starvation (SS) to unveil a reversible mechanism by which AEs close to cell-cycle genes and marked by the transcription factor Activity-Dependent Neuroprotective Protein (ADNP) recruit TFIIIC to acetylate Histone 3 lysine-18 (H3K18ac). These acetylated AEs engage in long-range interactions with pre-bound CTCF sites within promoters of distal cell-cycle genes, which also become H3K18 acetylated. The hyperacetylated environment maintains basal levels of transcription and facilitates re-activation of cell-cycle genes transcription upon serum re-exposure. Thus, our work defines a precise architectural role for AEs and exposes novel roles for TFIIIC.

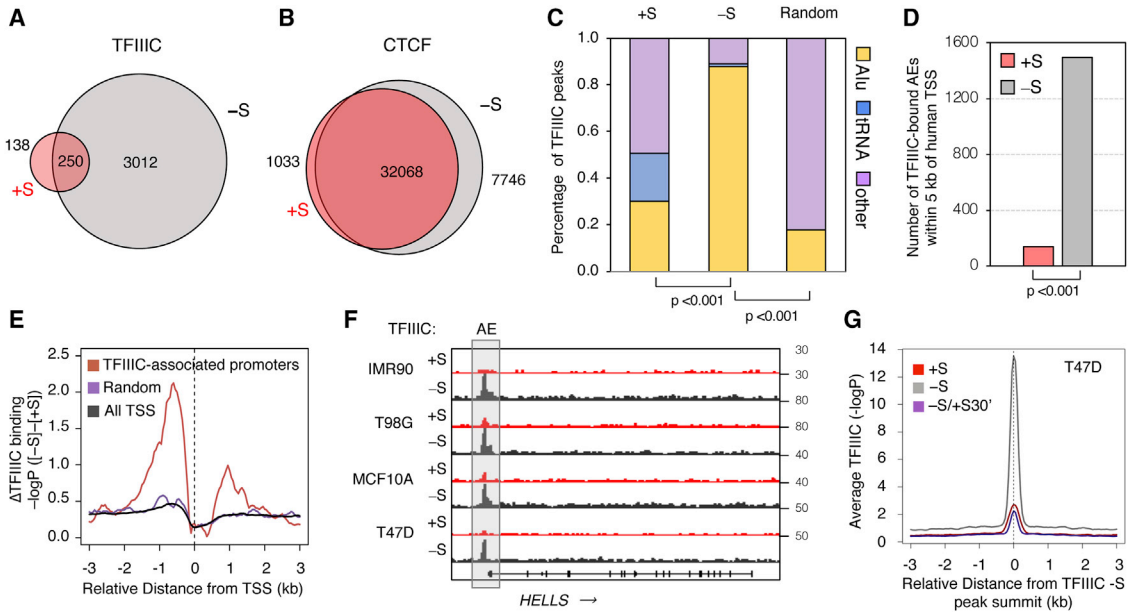


Figure 1. TFIIIC Occupancy at AEs Increases upon Serum Depletion in Tumor and Non-Tumor Cells

(A and B) Venn diagram of overlapping peaks for TFIIIC (GTF3C2) (A) and CTCF (B) in T47D in the presence (+S, red) or absence (–S, gray) of serum.

(C) Stacked plot for TFIIIC peaks over AEs, tDNA, or other loci in T47D in the presence (+S) or absence (–S) of serum. A comparison to a random set of genomic regions of same size as those in –S is also included (p value for the +S versus –S comparison, chi-square test due to differences in sample size; for the –S versus random comparison, two-tailed paired t test).

(D) Bar plot with the number of TFIIIC-bound AEs within a ±5 kb window around all human TSSs in T47D in the absence (–S) and presence of serum (+S) (p value, Kolmogorov-Smirnov test).

(E) *cis*-regulatory annotation system (CEAS) plots of the change in TFIIIC average binding in T47D in the presence (+S) or absence (–S) of serum (+S subtracted from –S) at TFIIIC-associated promoters (red). The profile of a random set of genes of the same size of the TFIIIC-associated promoters (purple), as well as the average for all human TSSs (black) are included.

(F) Genome browser view of the representative cell-cycle-regulated locus *HELLS* with ChIP-seq tag counts for TFIIIC in the presence (+S) or absence (–S) of serum. The AE bound by TFIIIC is highlighted with a gray rectangle. *HELLS'* genomic structure and direction of transcription (arrow) are shown.

(G) CEAS plots of TFIIIC enrichment in T47D cells in the presence (+S), absence (–S), or absence followed by serum addition for 30 min (–S/+S-30 min). The graphs are plotted over the summit of TFIIIC peaks in the –S condition (–log10 of the Poisson p value).

RESULTS

SS Provokes a Rapid and Reversible TFIIIC Increased Occupancy at AEs Close to Cell-Cycle Gene Promoters

First, we assessed the global occupancy of CTCF and TFIIIC by chromatin immunoprecipitation sequencing (ChIP-seq) in T47D breast cancer cells growing in normal conditions with serum (+S) and after 16 h of serum depletion (–S) (Figure S1A). Upon SS, a strong increase in the number of TFIIIC-bound sites was detected (Figure 1A, 92% increase), compared to a 24% increase in the total number of CTCF peaks occupancy (Figure 1B). We excluded that alterations of the cell-cycle profile were contributing to this effect, because SS did not induce strong changes in the profile (Figure S1B). Only ~30% (140) of the total TFIIIC peaks were located over AEs in the presence of serum, but this value increased to 89% (3,096) after SS (Figure 1C). This enrichment was statistically significant when compared with peaks detected in normal growth conditions or when using a random set of peaks as control (Figure 1C). Most of the new TFIIIC sites are bona fide AEs with the characteristic A- and B-box sequences, while the B-box consensus of the ETCs (Moqtaderi et al., 2010) was found in 14% of the new AEs bound by TFIIIC (Figure S1C). Notably, a large percentage of the new TFIIIC-bound AEs was in close proximity (within 5 kb) of annotated Pol II TSSs (Figures 1D and 1E) enriched in gene ontology (GO) terms associated with cell-cycle-related functions (Figure S1D). The TFIIIC enrichment at these sites was not simply reflecting an increase in AEs density, which was higher at Pol II promoters devoid of TFIIIC binding (Figure S1E). In contrast to tRNA genes (tDNAs), neither Pol III nor other components of the Pol III machinery were found at the TFIIIC-bound AEs (Figure S1F). The increased AEs occupancy by TFIIIC in response to SS was also observed in other cancer and normal cell lines, such as glioblastoma T98G, normal lung fibroblasts IMR90, and normal breast MCF10A (Figures 1F and S1G), where we observed the absence of the components of the Pol III machinery (Figure S1H). Finally, TFIIIC occupancy at AEs was reversed after just 30 min of serum re-addiction, indicating a rapidly reversible process and ruling out a cell-cycle direct role (Figures 1G and S1I). Thus, TFIIIC is reversibly recruited to AEs close to Pol II promoters of a subset of cell-cycle-related genes in response to SS.

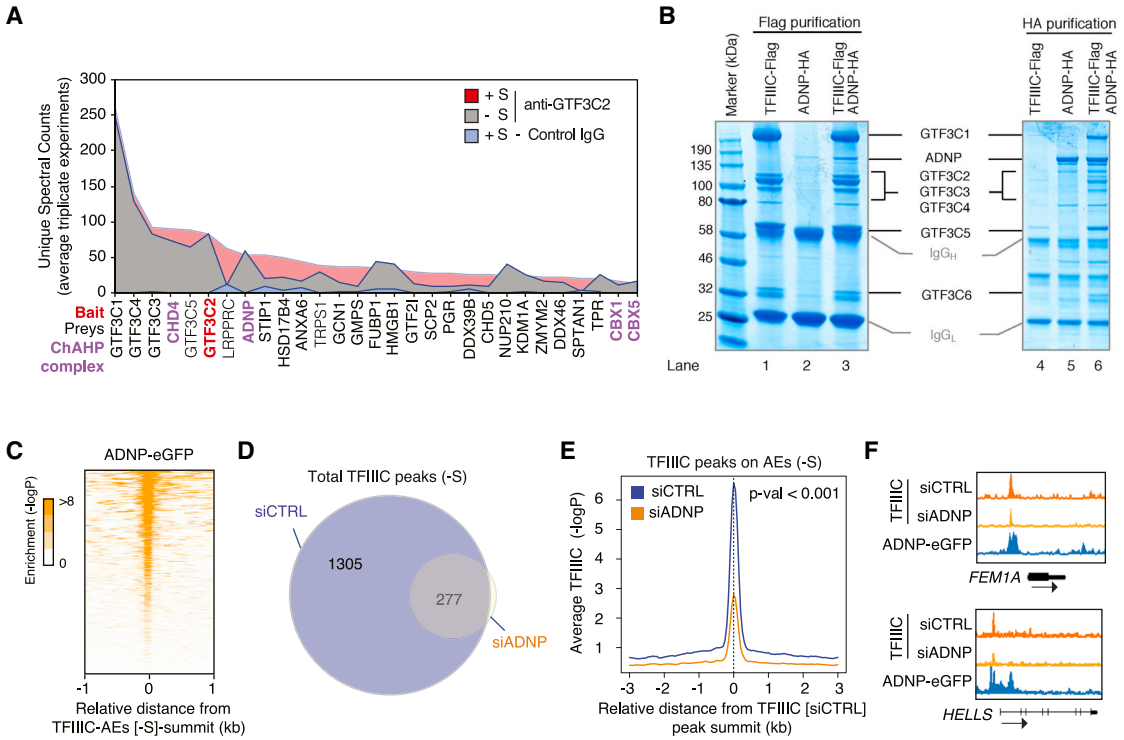


Figure 2. ADNP Guides TFIIIC Selective Recruitment to AEs and TFIIIC-Associated Promoters

(A) Area plot ranking unique spectral counts (average of three replicates) of proteins identified in a TFIIIC-RIME experiment using GTF3C2 as bait in T47D cells grown in the presence (+S) or absence of serum (–S). Data from control immunoprecipitations using normal IgG in the presence of serum are also included (blue). Prey proteins belonging to the ChAHP complex are shown in purple.

(B) Direct interaction of recombinant ADNP and TFIIIC. For TFIIIC expression, insect cells were infected with a single baculovirus expressing GTF3C1-3XFLAG and the rest of the TFIIIC subunits (GTF3C2-6) untagged. Left panel: Coomassie staining of anti-FLAG immuno-purifications of insect cell lysates overexpressing FLAG-TFIIIC (lane 1), HA-ADNP (lane 2), or both simultaneously (lane 3). Note that GTF3C1 immunopurification via the FLAG-tag pulled down the whole TFIIIC complex (lane 1). ADNP co-purifies with TFIIIC (lane 3). GTF3C5 fractionates on top of the heavy IgG chain (IgG_H). Right panel: Coomassie staining of anti-HA immunopurification of lysates expressing FLAG-TFIIIC alone (lane 4), HA-ADNP alone (lane 5), or both together (lane 6). The presence of the TFIIIC subunits is detected in the ADNP immuno-complexes. The position of the IgG heavy and light chains (IgG_H and IgG_L, respectively) is indicated. The identity of the bands was confirmed by mass spectrometry and western blotting with HA and GTF3C2 antibodies (Figure S2A).

(C) Heatmap of ADNP-eGFP binding in K562 cells (GEO: GSE105573) spanning ±1 kb across all TFIIIC-bound AEs in T47D, ranked from high to low ADNP-eGFP enrichment. Color bar scale with increasing shades of color stands for higher enrichment (–log10 of the Poisson p value).

(D) Venn diagram with the number of TFIIIC peaks upon ADNP depletion (siADNP, yellow) compared to control cells (siCTRL, blue) in serum-starved T47D cells.

(E) Sitepro profile of TFIIIC occupancy in control (siCTRL) and ADNP-depleted cells (siADNP) in serum-starved T47D cells. The graphs are plotted over the summit of TFIIIC peaks in the siCTRL cells (p value, two-tailed unpaired t test). ADNP depletion levels are shown in Figures S2E and S2F.

(F) Genome browser view of representative TFIIIC-associated genes *FEM1A* and *HELLS* showing TFIIIC occupancy in ADNP-depleted (siADNP) and T47D control cells (siCTRL). ADNP binding from GEO: GSE105573 is represented by the blue track (eGFP-ADNP). The gene structure and the direction of transcription (arrow) are shown at the bottom.

ADNP Marks AEs to Favor Selective Recruitment of TFIIIC upon SS

To explore possible factors responsible for TFIIIC redistribution upon SS, we used rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) (Mohammed et al., 2016) using an antibody against the GTF3C2 subunit of TFIIIC (see STAR Methods). All six subunits of the TFIIIC holo-complex were identified (Figure 2A; Table S1), validating the approach. Furthermore, all subunits of the recently characterized ChAHP complex were also identified (Ostapcuk et al., 2018), with the chromodomain helicase DNA binding protein 4 (CHD4) and ADNP appearing as the top chromatin-bound TFIIIC-associated proteins (Figure 2A; Table S1). The binding of TFIIIC to ADNP was further validated *in vitro* using recombinant TFIIIC components and ADNP produced in insect cells (Figures 2B and S2A).

ADNP is a ubiquitously expressed transcription factor with a role in transcription of genes essential for embryogenesis (Ostapcuk et al., 2018). To gain insight into a functional cross-talk between ADNP and TFIIIC, we re-analyzed ADNP ChIP-seq data from mouse embryonic stem cells (mESCs) (Ostapcuk et al., 2018) and found that ADNP is mainly associated (95%) with repetitive elements (Figure S2B), over 50% of which are SINEs of the B1 and B2 families (Figure S2C), which are the corresponding murine AEs (Weiner, 2002). Motif enrichment analysis for ADNP binding in mESCs identified the B-box sequence as the second most represented motif (Ostapcuk et al., 2018).

Unfortunately, we could not generate ADNP ChIP-seq data due to the poor performance of all tested ADNP antibodies. Therefore, we analyzed ChIP-seq data from human cells expressing ADNP-eGFP (Consortium, 2012) and found that ADNP is strongly enriched at many TFIIIC-bound AEs (Figure 2C), including those close to Pol II promoters compared to a random set of promoters (Figure S2D). These findings point to ADNP as a candidate in helping TFIIIC selective recruitment to this subset of AEs. Indeed, small interfering RNA (siRNA)-mediated depletion of ADNP in T47D cells resulted in a 5-fold decrease in the total number of TFIIIC peaks upon SS (Figure 2D), as well as in impaired recruitment of TFIIIC to AEs (Figure 2E), in particular, to AEs associated with Pol II promoters (Figures 2F and S2G). Altogether, the data support a role for ADNP in promoting the recruitment of TFIIIC to a specific subset of AEs triggered by SS.

The Histone Acetyltransferase Activity of TFIIIC Acetylates H3K18 at AEs upon SS

The TFIIIC complex relieves chromatin-mediated repression in *in vitro* assays (Kundu et al., 1999), and we wondered whether TFIIIC binding to AEs induced by SS increased chromatin accessibility. We performed ATAC sequencing (ATAC-seq) in T47D cells and found that already in the presence of serum the regions around the TFIIIC-bound AEs were more accessible when compared with a random set of AEs of similar size that do not exhibit TFIIIC binding after SS (Figure 3A). These suggests that the chromatin over AEs that will bind TFIIIC upon SS is already more accessible, presumably due to binding of other associated factors. Upon SS the chromatin accessibility over these regions increased considerably (Figure 3A). Given that three TFIIIC subunits possess intrinsic histone acetyltransferase (HAT) activity (Hsieh et al., 1999; Kundu et al., 1999) and that TFIIIC interacts with p300/CREB-binding protein (CBP) (Mertens and Roeder, 2008) that specifically acetylates H3K18 and H3K27 *in vivo* (Horwitz et al., 2008; Jin et al., 2011), we hypothesized that the increase accessibility could be due to histone acetylation. Indeed, the AEs that will bind TFIIIC upon SS were already positive for H3K18ac and H3K27ac in cells grown in the presence of serum (Figures 3B and S3A). However, upon SS only H3K18ac was markedly increased at TFIIIC-bound AEs (Figure 3B), while no changes in H3K27ac or H3K9ac were observed at these loci (Figures S3A and S3B). Moreover, neither enrichment nor changes in the H3K18ac mark were found over tDNAs in any condition (Figure S3C).

Given the sharp peak of H3K18ac at the TFIIIC-bound sites, we wondered whether this could be due to the presence of nucleosomes positioned over the AE at the region occupied by TFIIIC. To answer this question, we used nucleosome position data derived from ATAC-seq by selecting ATAC fragments from 180 up to 250 bp in length as described (Buenrostro et al., 2013). This analysis showed that TFIIIC binding to AEs caused a decreased in the overall nucleosome occupancy or stability around TFIIIC-bound regions (Figure 3C). In addition, TFIIIC-bound AEs had a single or dimeric nucleosome positioned over the peak's center corresponding to the AE itself (Figure 3C), in agreement with a previous report (Tanaka et al., 2010). Notably, the central AE nucleosome signal was significantly reduced in serum-starved cells (Figure 3C and selected examples in Figure 3D), concurring with the increased H3K18ac signal at these sites (Figure 3B). All the results point to an *in vivo* role of TFIIIC in relieving chromatin-mediated repression by increasing H3K18ac and decreasing nucleosome stability at AEs.

To further support this idea, we interrogated the NucMap database (Zhao et al., 2019) and used published dataset of micrococcal nuclease (MNase) digestion in T47D to plot the nucleosome profile over the AEs bound by TFIIIC in serum-starved cells. In agreement with our ATAC-seq data, these regions showed a precise pattern of nucleosome occupancy with the center of the AE barring a mono- or di-nucleosome (Figure S3D). The difference in the relative nucleosome density of the AE's nucleosome(s) observed between ATAC-seq and MNase digestion could simply reflect the higher accessibility to the two nucleosome free regions surrounding the AE's nucleosome(s) of the MNase enzyme (~16 kDa) compared to the Tn5 transposase (~100 kDa). Despite the differences between the two experimental approaches, the detection of a nucleosome over the AE was clear, and therefore we conclude that the H3K18ac profile (Figure 3B) is compatible with the presence of a nucleosome at AEs bound by TFIIIC in T47D.

Around 80% of all TFIIIC-bound AEs and 70% of AEs at Pol II promoters bound by TFIIIC were found acetylated at H3K18 in response to serum depletion (Figures 3E and 3F). H3K18ac and H3K27ac are markers of p300/CBP function *in vivo* (Ferrari et al., 2008; Horwitz et al., 2008; Jin et al., 2011). However, low levels of p300 were detected at TFIIIC-bound AEs in both T47D (Figures 3G and S3E) and T98G cells (Figure S3F), and p300 levels decreased upon SS (Figures 3G, S3E, and S3F), suggesting that p300 could not account for H3K18ac at these loci. We reasoned that the effect could rely on the HAT activity of TFIIIC since the GTF3C1 subunit of TFIIIC robustly acetylates H3K18 *in vitro* and in HepG2 cells (Basu et al., 2019), and it was found enriched at TFIIIC-bound AEs in T47D cells upon SS (Figure S3G). Therefore, we used siRNAs against GTF3C1 to reduce its protein levels (Figure S3H) and found a dramatic reduction in H3K18ac total levels in the depleted cells (Figures 3H and S3I), indicating that the GTF3C1-associated HAT activity could be responsible of the increase in H3K18ac in serum-starved T47D cells and, in particular, of the increased H3K18ac signal in TFIIIC-bound AEs upon SS. To support this notion, we targeted the GTF3C5 subunit of TFIIIC, known to stabilize the interaction of the whole TFIIIC complex with the B-box (Jourdain et al., 2003), by siRNA, which drastically reduced H3K18ac at two TFIIIC-bound AEs in serum-starved T47D cells (Figure S3K). Although we cannot formally exclude the existence of a protein with HAT activity recruited to AEs by TFIIIC, the results support a TFIIIC-mediated H3K18 acetylation of nucleosomes over AEs upon SS and further point to TFIIIC-associated HAT activity as responsible for the increase in both H3K18ac and chromatin accessibility at TFIIIC-bound AEs.

TFIIIC Engages AEs in Long-Range Chromatin Interactions in Response to SS to Maintain Cell-Cycle Gene Expression

As AEs have evolved toward a proto-enhancer function and their epigenetic state participates in forging genome topology and gene expression (Su et al., 2014; van de Werken et al., 2017),

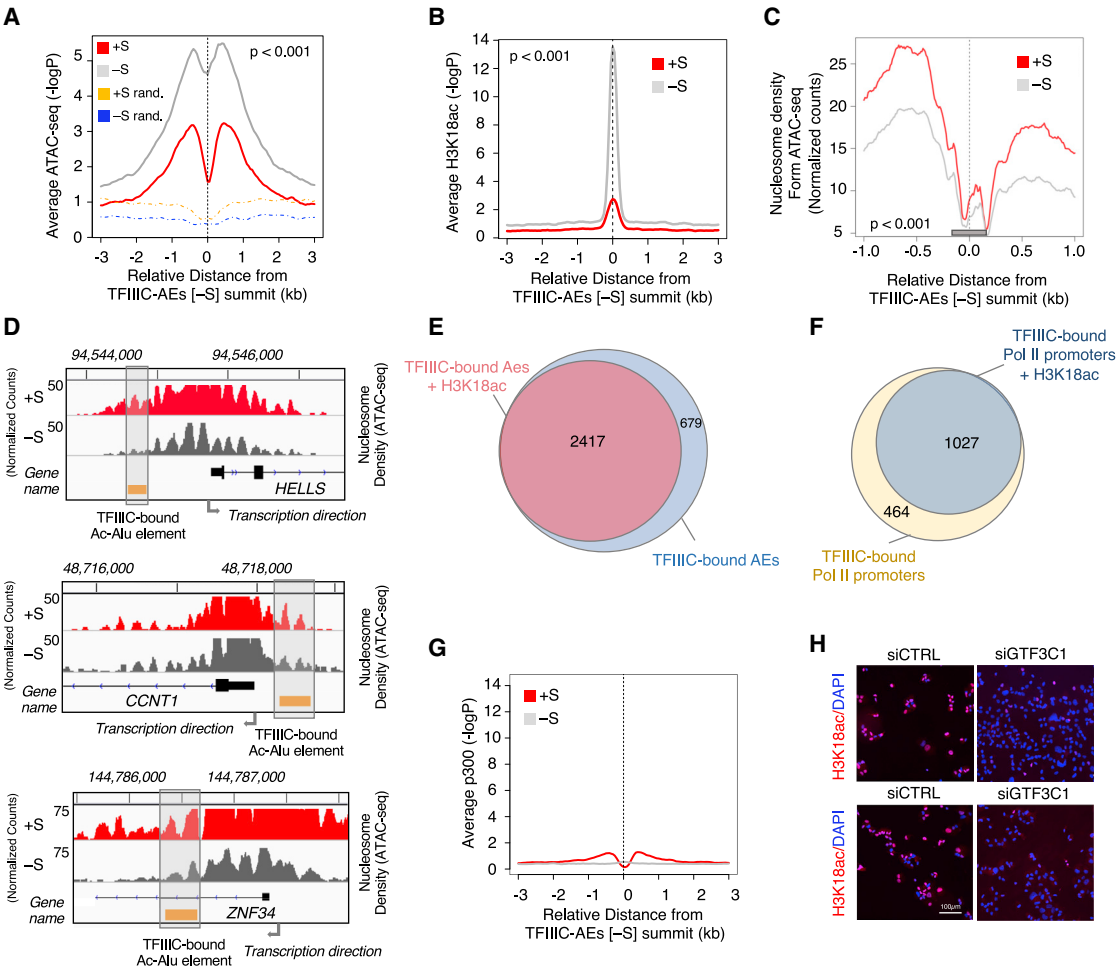


Figure 3. TFIIIC-Associated HAT Activity Acetylates H3K18 at Nucleosomes over AEs in Response to Serum Depletion

(A) ATAC-seq signal enrichment in T47D grown in the presence (+S) or in the absence of serum (–S) across TFIIIC-bound AEs detected in the absence of serum (–S) (–log10 of the Poisson p value). p value, two-tailed t test comparing each condition to its random control.

(B) Average profile of H3K18ac enrichment in T47D grown in the presence (+S,) or absence of serum (–S) across all TFIIIC-bound AEs (–log10 of the Poisson p value). p value, two-tailed paired t test.

(C) Sitepro average plot of nucleosome signals derived from ATAC-seq by selecting sequenced fragments of 180–250 bp (Buenrostro et al., 2013) spanning a 1 kb region over all the AEs bound by TFIIIC in T47D grown in the presence (+S) or absence (–S) of serum. Note the presence of a peak in the center of the region representing the AE (gray rectangle). SS causes a decrease in the nucleosome signal for both the AE-nucleosome and the surrounding ones (two-tailed unpaired t test).

(D) Genome browser view of representative cell-cycle-related TFIIIC-associated promoters with normalized counts of ATAC-derived nucleosome signal for T47D grown in the presence (+S) or absence (–S) of serum. The position of the AE bound by TFIIIC positive for H3K18ac is indicated by a yellow rectangle. AE nucleosomes reduced upon SS are highlighted by a gray rectangle.

(E) Venn diagram showing the total number of AEs bound by TFIIIC and those acetylated in H3K18 in serum-starved T47D.

(F) Venn diagram showing the total number of TFIIIC-associated Pol II promoters and those enriched in H3K18ac in T47D grown in the absence of serum.

(G) Plot for p300 occupancy across all TFIIIC-bound AEs spanning a 6 kb region in T47D grown in the presence (+S) or absence (–S) of serum.

(H) H3K18ac immunostaining (red) in serum-starved T47D cells knocked down for GTF3C1 (siGTF3C1) or control (siCTRL). DAPI was used to stain nuclei (blue). Two different fields are shown. Scale bar, 100 μm. GTF3C1 depletion levels are shown in Figure S3H and quantification in Figure S3I.

we postulated that TFIIIC might help in reorganizing the landscape of chromatin loops following SS. To explore this, we compared the transcript levels by mRNA sequencing (mRNA-seq) (Table S2) and the genomic contacts by *in nucleo* Hi-C (Rao et al., 2014) (Table S3) in T47D cells growing in in both conditions. Changes in chromatin compartments were observed in serum-deprived cells (Figure S4A), although without significant changes in the total number of TADs as detected by TADbit (Serra et al., 2017) (Figure S4B). SS also induced changes in the expression of Pol II genes (Figure S4C). Pol II genes with TFIIIC-bound AE at their promoters upon SS showed significantly higher expression than a random control set, but their expression levels were not significantly affected by the growth conditions (Figure S4D). The results correlated with the lack of changes in Pol II loaded at these promoters when the two conditions were compared (Figure S4E). We

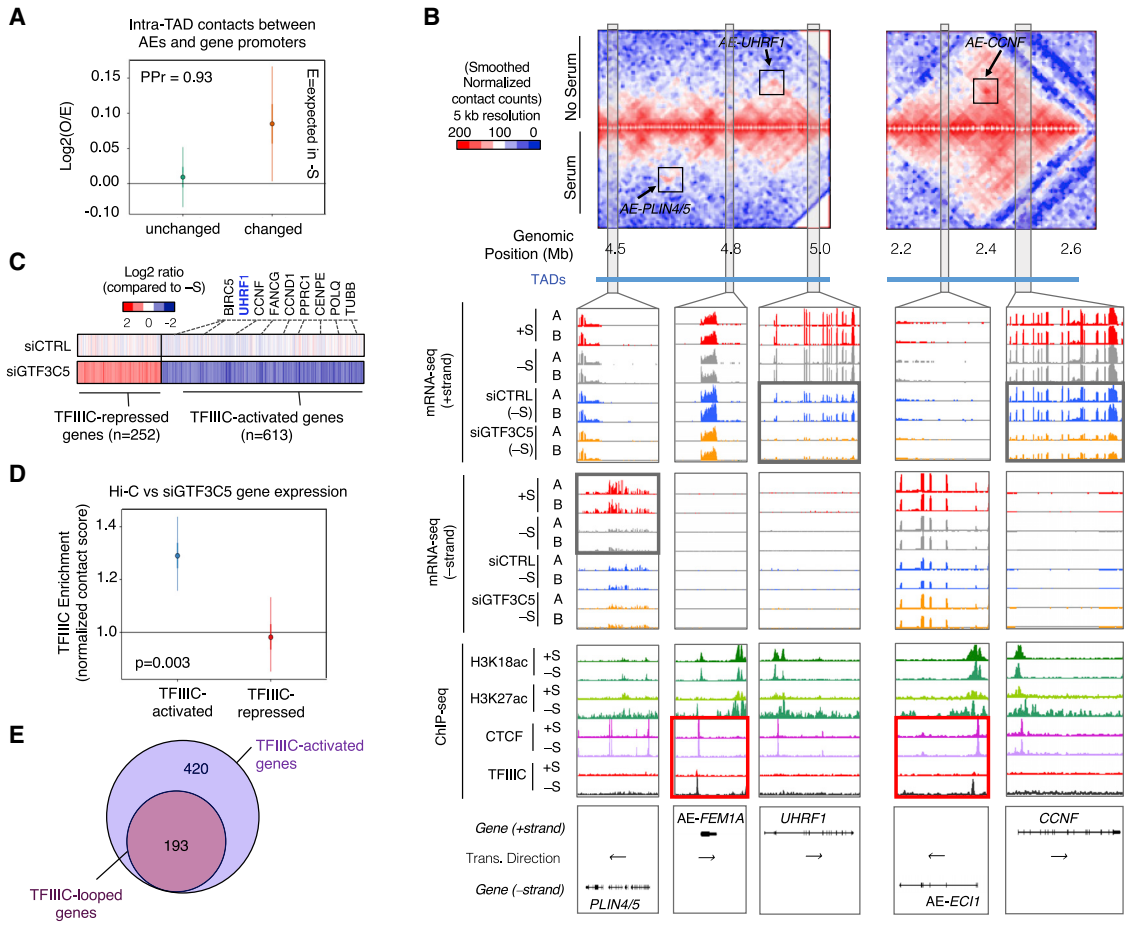


Figure 4. TFIIIC Controls Gene Expression in Response to SS via AEs-Mediated Long-Range Chromatin Looping

(A) Hi-C analysis of intra-TAD contacts represented as log2 fold change (Observed / Expected: O/E) and 95% confidence interval (CI) of the contact enrichment between TFIIIC-bound AEs and Pol II promoters for genes differentially expressed (changed) or not affected (unchanged) by SS in T47D cells. Posterior probability (PPr) = 0.93.

(B) Heatmap matrix of Hi-C interactions of *PLIN4/5-UHRF1* loci (left panel) and *CCNF* locus (right panel) for TFIIIC-bound AEs in T47D in the presence or absence of serum. Color scale bar is reported with red representing values from 200 and above, white values equal to 100, and blue values equal to 0. The loops representing the higher frequency of interactions of the AEs and the corresponding genes are framed (Hi-C data, top panel). TADs are indicated by blue rectangles. The regions with changes in their interactions upon SS are framed with gray rectangles and zoomed out to help the visualization of the genome browser tracks of mRNA-seq (A and B indicate biological replicates; the two strands are represented separately), and ChIP-seq data for the protein/histone marks are indicated. The gene structure and transcription direction are reported at the bottom of the figure. Gray bold rectangles highlight the mRNA-seq regions with significant changes for the *PLIN4/5*, *UHRF1*, and *CCNF* loci. Red bold rectangles highlight the position of TFIIIC binding to the AEs.

(C) Heatmaps of differential gene expression for siGTF3C5 and siCTRL T47D cells in the absence of serum compared to mock-transfected cells in the absence of serum. Color bar scale stands for log2FC of normalized RNA expression in each condition. Only the genes that changed their expression significantly in siGTF3C5 and not in siCTRL are shown. Two classes of genes were designated as TFIIIC-activated or TFIIIC-repressed genes. Cell-cycle-related genes whose expression was drastically reduced by siGTF3C5 are indicated, including *UHRF1* and *CCNF* further analyzed in (B).

(D) TFIIIC contact enrichment for the TFIIIC-activated and TFIIIC-repressed genes using Hi-C data in the absence of serum (p value for logistic regression comparing the two groups).

(E) Venn diagram showing the overlap between the TFIIIC-activated genes and those bound by TFIIIC either directly within a 10 kb region or via DNA looping (TFIIIC-looped genes).

conclude that the presence of TFIIIC does not alter transcription locally upon SS.

As the TFIIIC-bound AEs were positive for the active enhancer mark H3K27ac (Figure S3A), we speculated that TFIIIC could use these elements as enhancer modules to control gene expression at distance. Supporting this hypothesis, we found that TFIIIC-bound AEs interacted more frequently with genes whose expression was affected by serum (Figure 4A). Over 75% of the total TFIIIC interactions were detected within TADs, and, thus, we focused on intra-TADs interactions. As an example, we focused on the interaction patterns of the TFIIIC-bound AE near the cell-cycle-regulated gene *FEM1A* (Figure 4B). In the presence of serum, this region shows intra-TAD contacts with an ~200 kb upstream locus containing the *PLIN4/PLIN5* genes, whose

expression were significantly decreased upon SS (Figure 4B, left panel). SS induced binding of TFIIIC (Figure 4B, left panel) and a change in the intra-TAD contacts of the AE-containing region, which interacted more frequently with the *UHRF1* locus located ~150 kb downstream (Figure 4B, left panel). SS also induced both the binding of TFIIIC to the AE near the *ECI1* gene and changes in the long-rage interactions of this region with the ~200 kb downstream *CCNF* locus (Figure 4B, right panel). Indeed, when the interaction scores between all TFIIIC-bound AEs and Pol II TSSs were calculated, a significant increase was observed upon SS (Figure S4F), suggesting that TFIIIC binding to AEs participates in reshaping the genome topology in response to SS.

To gain insight into the functional meaning of TFIIIC-mediated looping induced by SS, we searched for transcript changes in cells depleted of one of the TFIIIC components (siGTF3C5). As shown by others (Büchel et al., 2017), no significant changes in tRNA expression were observed upon depletion (Figure S4G); however, *GTF3C5* depletion led to the dysregulation of a set of Pol II transcripts (Figures S4H–S4J). We focused on those differentially expressed genes that were affected by the siGTF3C5 treatment but did not exhibit significant changes in the siCTRL cells when compared to mock-transfected cells (Figure 4C; Table S2): 252 genes were upregulated and 613 were downregulated. We refer to these two sets of genes as TFIIIC-repressed and TFIIIC-activated genes, respectively. GO analysis of the TFIIIC-activated genes showed enrichment for cell-cycle-regulated activity (Figure S4K), in agreement with a recent report in glioblastoma cells (Büchel et al., 2017). This effect was reflected in a significant decrease of the S- and G2/M-phases in T47D (Figure S4L).

Given the dependence on TFIIIC for a set of genes in serum-starved cells and the fact that the Pol II genes with TFIIIC bound to their promoters did not show serum-dependent changes (Figure S4D and Figure 4B for *FEM1A* as an example), we wondered whether the TFIIIC-dependent formation of intra-TADs loops in response to SS was responsible for the transcriptional regulation. Thus, we analyzed Hi-C interactions of TFIIIC-occupied sites with genes whose expression was affected by TFIIIC depletion. We found that TFIIIC binding was significantly enriched in TADs containing TFIIIC-activated genes compared to those containing TFIIIC-repressed genes (Figure 4D), suggesting that TFIIIC-bound AEs could act as rescue modules to prevent drastic repression of these genes in the absence of serum. For instance, expression of *UHRF1*, encoding a E3 ubiquitin ligase necessary for cell-cycle progression (Tien et al., 2011), was not regulated by serum but was largely repressed upon TFIIIC depletion (Figure 4B, left panel: RNA sequencing [RNA-seq] tracks in +S versus –S and siCTRL versus siGTF3C5). We thus propose that the switch in looping of the AE-TFIIIC module from the *PLIN4/5* locus toward the *UHRF1* gene is the mechanism responsible for ensuring its proper steady-state levels in response to SS. Indeed, more than 30% of the TFIIIC-activated genes (193) contacted an AE bound by TFIIIC in serum-starved cells (Figure 4E); this subset was named TFIIIC-looped genes, and it represents the set of genes that are co-regulated by both local and long-range interactions through a TFIIIC-bound AE. Altogether, the results suggest that binding of TFIIIC to AEs is required to sustain basal transcription levels in the absence of serum of a subset of genes with cell-cycle-related functions.

Long-Range Interactions of TFIIIC with CTCF Mediate DNA Looping upon SS

We wondered how TFIIIC could promote chromatin looping. TFIIIC interacts with CTCF (Galli et al., 2013), which is enriched at promoters (Ruiz-Velasco et al., 2017). Notably, we found that CTCF occupancy at TFIIIC-looped genes was significantly higher than at a random set of promoters (Figures 5A and 5B). This effect was not general since the proportion of TFIIIC-bound sites co-occupied by CTCF in the absence of serum was very small (Figure S5A). CTCF binding to the TFIIIC-looped genes was likely direct because the CTCF-bound regions are enriched in the canonical CTCF motif (Figure S5B). Cohesin recruitment was detected at the CTCF-bound regions of the TFIIIC-looped genes, based on the presence of RAD21, and its recruitment was not different from a random set of CTCF peaks (Figure S5C), as expected since there is a 99% overlap between CTCF and RAD21 in serum-starved conditions (Figure S5D). Therefore, we hypothesized that the interaction of TFIIIC with CTCF could participate in long-range chromatin loops between the AEs and the promoters of TFIIIC-looped genes. To test this, we first assessed TFIIIC and CTCF interaction in co-immunoprecipitation experiments and found that SS induced a marked increase in the presence of TFIIIC, as measured by GTF3C2, in CTCF-containing complexes (Figure 5C), which was not due to changes in total TFIIIC protein levels (Figure S5E). This effect was also reflected in the Hi-C data with an increase of intra-TAD contacts between the two factors (Figures 5D, S5F, and S5G), which is visible in the Hi-C heatmap as a dot connecting the two regions forming the loop (Figure 5E; Hi-C heatmaps in Figure 4B). Moreover, as ADNP was identify as a putative looping factor (Weintraub et al., 2017), we wondered whether this protein could participate in the TFIIIC-mediated looping, To this aim, we calculated the level of occupancy of ADNP-eGFP to the TFIIIC-looped regions and found a very strong significant enrichment compared to all the TFIIIC-bound AEs or the AEs at TFIIIC-associated promoters (Figure S5H). Altogether, the data support a putative role for CTCF and ADNP in TFIIIC-mediated looping upon SS.

Long-Range Chromatin Looping Promotes H3K18ac Hyperacetylation of Distant Promoters Contacted by TFIIIC

Our data support that TFIIIC acetylates AEs and shapes chromatin looping landscape by contacting preloaded-CTCF at promoters of distal TFIIIC-looped genes in response to SS. This mechanism could favor the maintenance of steady-state expression levels of these genes by creating an acetylated "transcription-favorable" environment that favors H3K18ac. In support of this, we found that SS caused a drastic change in the overall profile of H3K18ac with a large fraction of promoters, including those for TFIIIC-looped genes, showing H3K18ac evenly distributed along a broader region around the TSS (Figures 5G and S5I). Indeed, almost 70% of the TFIIIC-looped

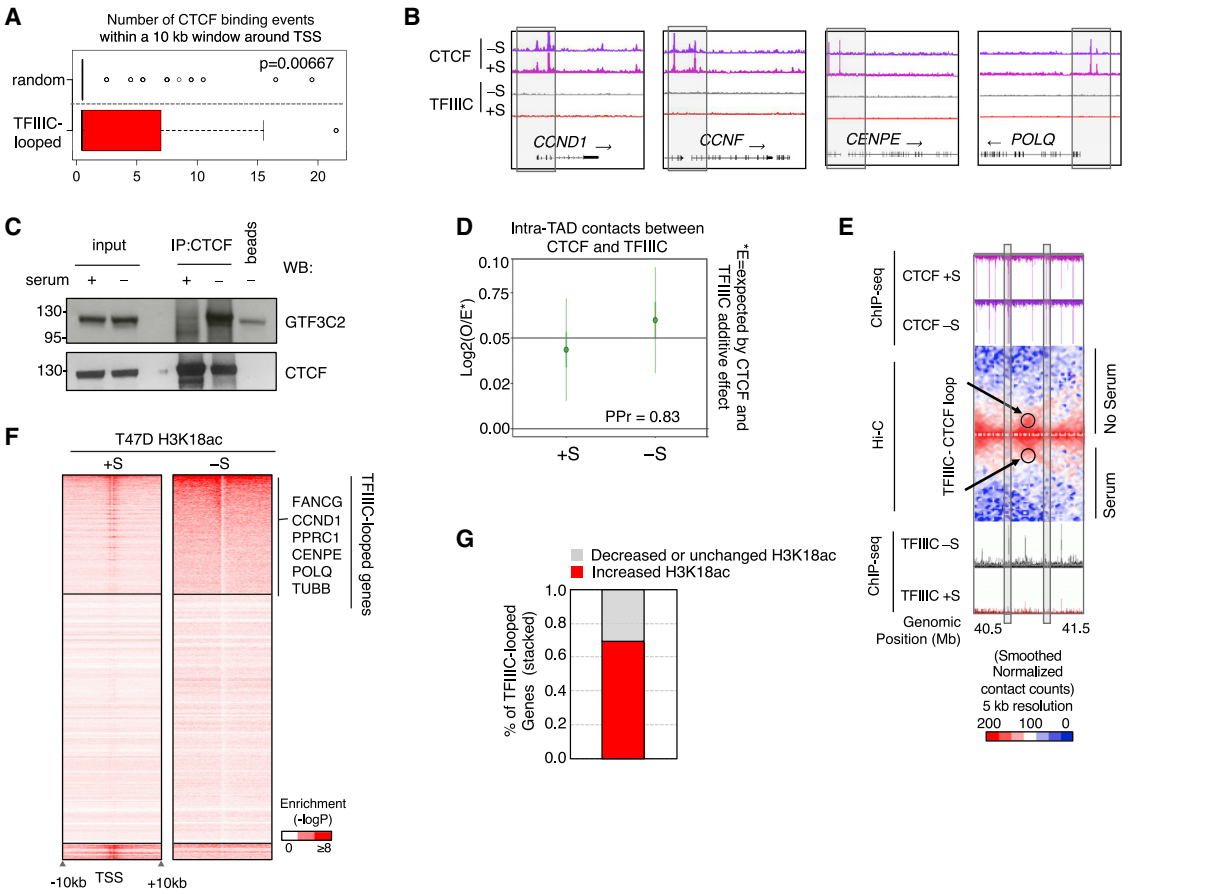


Figure 5. Serum Depletion Induces TFIIIC and CTCF Interaction to Allow Long-Range DNA Looping and H3K18 Hyperacetylation of Distal TFIIIC-Looped Genes

(A) Boxplot of CTCF binding events within a 10 kb region around the TSSs of TFIIIC-looped genes or of a random dataset of TSS of the same size (p value, Friedman X^2 test).

(B) Genome browser view of representative cell-cycle-related TFIIIC-activated genes *CCND1, CCNF*, *CENPE*, and *POLQ*. ChIP-seq data of CTCF and TFIIIC in T47D in the presence (+S) or absence (–S) of serum. Note that multiple CTCF binding sites (gray boxes) are present at the 5′ end of these genes but no TFIIIC is detected. Transcription directionality is indicated with an arrow.

(C) Immunoprecipitation of CTCF in soluble extracts from T47D cell grown in the presence (+) or absence (–) of serum ("beads" are used as a specificity control). Input lysates (10%) and immunoprecipitates were probed with CTCF and GTF3C2 antibodies (as a marker for TFIIIC).

(D) Hi-C analysis of TFIIIC and CTCF contacts represented as log2 FC (O / E*; 95% CI) of the specific CTCF-TFIIIC contacts (O = observed) compared to the expected (E*) CTCF and TFIIIC additive effect for both +S and –S conditions. The PPr value indicates a high probability of an increase in TFIIIC/CTCF Hi-C contacts in the absence versus presence of serum.

(E) Genomic representation of a 1 Mb region of chromosome 17 containing an AE-TFIIIC-CTCF loop induced by SS (framed with a circle in the Hi-C heatmaps). Color scale bar is reported with red representing values from 200 and above, white values equal to 100, and blue values equal to 0. ChIP-seq data for CTCF and TFIIIC in the presence (+S) or absence (–S) of serum is included at the top and bottom, respectively. Matrix heatmaps with Hi-C data are also included and aligned with the ChIP-seq data. Gray rectangles highlight the position of the AE-TFIIIC and the CTCF interacting regions.

(F) Heatmap representation of H3K18ac spanning a 20 kb region of all human promoters in T47D cells grown in the presence or absence of serum. Biased clustering shows promoters with increased H3K18ac in the absence of serum, and the cluster contains several TFIIIC-looped genes (selected examples are listed).

(G) Stacked plot representing changes in H3K18ac induced by SS in TFIIIC-looped genes. Note that around 70% of them display increased acetylation.

genes exhibited increase H3K18ac upon SS (Figure 5G). As for the TFIIIC-bound AEs (Figures 3B and S3A), the effect was specific for H3K18ac, and H3K27ac did not show significant changes at promoters upon SS (Figure S5J). One interpretation for these results is that the TFIIIC-dependent chromatin reorganization facilitates the placement of its HAT-associated activity to the distal looped genes promoting H3K18 acetylation to favor transcription. Transcription activation might in turn promote further acetylation by TFIIIC and/or other HATs at the looped genes.

TFIIIC Depletion Impairs DNA Looping and Abrogates the Reactivation of Gene Expression upon Serum Re-exposure

An implication of our model is that TFIIIC depletion should decrease the frequency of genomic interactions upon SS.

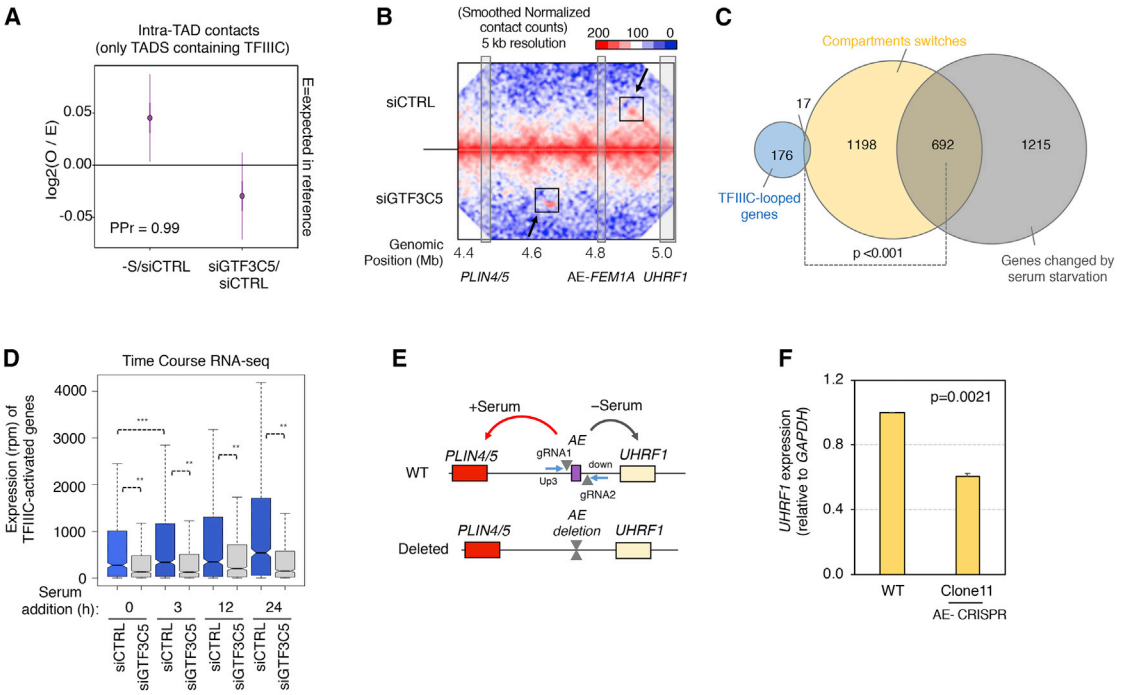


Figure 6. Impaired DNA Looping at TFIIIC-Looped Genes by TFIIIC Depletion or AEs Deletion Abrogates the Reactivation of Gene Expression upon Serum Re-exposure

(A) Changes in specific intra-TADs contacts made by TFIIIC in T47D treated with siGTF3C5 or siCTRL in the absence of serum. Data are the log2 FC of observed versus expected (95% CI) of Hi-C data. The changes in siGTF3C5 versus siCTRL show significant (PPr = 0.99) decrease of total intra-TADs contacts compared to mock-transfected (–S) versus siCTRL.

(B) Hi-C matrix heatmaps of the *UHRF1* locus for siCTRL and siGTF3C5 cells in the absence of serum. Color scale bar is reported with red representing values from 200 and above, white values equal to 100, and blue values equal to 0. The looping between the AE bound by TFIIIC and its respective targets (*PLIN4/5* and *UHRF1*) is indicated by a box when detected and an arrow when lost. Note that even in the absence of serum the Alu-*FEM1A* loops back to the *PLIN4/5* locus upon siGTF3C5 (as reported for the condition of presence of serum in Figure 4B).

(C) Venn diagram showing the overlap between the TFIIIC-looped and the serum-dependent genes compared to the regions showing compartment switch (Figure S4B). Note the higher overlap of the regions showing compartment changes with the serum-dependent genes (Fisher's exact test).

(D) RNA-seq expression analysis in T47D cells of TFIIIC-activated genes after serum re-exposure for the indicated times in conditions of siCTRL and siGTF3C5. Expression data are reported as RPM (**$p < 1x10^{-12}$, two-tailed paired t test in comparisons for each time-point). The comparison between 0 to 3 h in siCTRL cells is shown to highlight the rapidity of gene activation after serum addition (***$p < 1x10^{-20}$). *GTF3C5*-depleted levels were maintained during the time course (Figure S6B).

(E) Schematic representation of the CRISPR-Cas9 approach to delete the TFIIIC-bound AE located between the *PLIN4/5* and *UHRF1* loci in chromosome 17. The wild-type (WT) and the deleted alleles are shown. The targeted AE is shown as a purple box, the position of the guide RNAs (gRNA1 and gRNA2) is marked with triangles, and the primers used for the screen (Figure S6C) are indicated with blue arrows. Arrows indicate the chromatin interactions in +S (red) and –S (black) conditions based on Figure 4B.

(F) *UHRF1* expression by qRT-PCR in T47D parental cells (WT) and AE-deleted (Clone11) in the absence of serum (mean ± SEM of biological duplicates with data in WT cells normalized to 1; one-tailed t test).

Indeed, *in nucleo* Hi-C in GTF3C5-depleted cells in the absence of serum showed a significant reduction in the overall intra-TAD contacts (Figure 6A) and at distant TFIIIC-looped genes (Figure S6A) compared to siCTRL cells. The reduction was observed for the AE-*FEM1A* Hi-C contacts with *UHRF1*, while the contacts were increased at the *PLIN4/5* upon TFIIIC depletion (Figure 6B). These results resemble the situation in the presence of serum (Figure 4B) and suggest that TFIIIC is involved in the genomic re-wiring induced by SS. They also indicate that the AE-TFIIIC interaction with the *UHRF1* locus induced by SS is not the result of changes in gene transcription at the *PLIN4/5* loci, because the reversion of the looping by depletion of GTF3C5 was observed in the absence of serum (Figure 6B), when the transcription of the locus is abrogated (Figure 4B). Moreover, compartment switches overlapped more with changes in genes transcription depending on serum than with the TFIIIC-looped genes (Figure 6C). Therefore, the data support a causality link for TFIIIC-induced looping and the transcription changes.

One possible scenario is that TFIIIC might act by depositing an epigenetic bookmark (H3K18ac) necessary to enable a quick response to serum exposure of the subset of TFIIIC-activated genes (Figure 4C). To test this, we performed RNA-seq analysis of serum-starved cells at various time points after serum re-exposure. Whereas in control cells TFIIIC-activated genes rapidly increased their expression within 3 h post-addition, depletion of TFIIIC completely abrogated this transcriptional response (Figure 6D). These results support a positive role of TFIIIC-bound AEs in the regulating

transcriptional levels of cell-cycle-related genes in response to serum.

Genomic Deletion of a TFIIIC-Bound AE Decreases Transcription of the Associated Gene in Serum-Starved Cells

To show that the AEs are needed for the TFIIIC response to SS, we deleted the AE bound by TFIIIC between the *PLIN4/5* and *UHRF1* loci by CRISPR-Cas9 technology (Wang et al., 2016) (Figure 6E). This AE contacts the *PLIN4/5* locus in normal growth conditions and shifts to the *UHRF1* locus upon SS (Figure 4B), and TFIIIC depletion causes a drastic decrease in *UHRF1* expression (Figure 4B, left panel). We were only able to obtain clones with one allele modified, and we chose one of the heterozygous clones for further analysis (Figure S6C). Remarkably, deletion of the AE in just one allele caused almost 50% decrease in the expression of *UHRF1* upon SS, compared to the parental cell line (Figure 6F). This result concurs with that from RNA-seq of TFIIIC depletion in serum-starved cells (Figure 4C) and with the Hi-C data (Figure 6B) and supports the requirement of the AE to maintain steady-state levels of *UHRF1* transcripts during stress conditions such as deprivation of serum.

DISCUSSION

Here, we uncover chromatin-associated regulatory mechanisms of cell adaptation to serum withdrawal, as the trigger of stress, which involve the cooperation of previously unconnected *trans*- and *cis*-elements, and further propose that expansion of several families of repetitive elements during evolution might have served to generate new genomic *cis*-regulatory networks enabling the coordinated regulation of a large set of genes relevant for cellular stress survival.

Pioneering work dedicated to elucidate the genome-wide occupancy of Pol III machinery unveiled the presence of TFIIIC close to the TSS of Pol II-transcribed genes (Canella et al., 2010; Carrière et al., 2012; Moqtaderi et al., 2010; Oler et al., 2010). Albeit these studies clearly established a connection between the Pol III and Pol II machinery, no mechanistic insight was provided for such observation. More recently, such connection has been proposed to rely on AE transcripts for heat shock (Allen et al., 2004; Mariner et al., 2008; Yakovchuk et al., 2009) or for neuronal depolarization (Policarpi et al., 2017), with TFIIIC regulating the production and nuclear localization of AE transcripts, which directly regulates Pol II loci (Crepaldi et al., 2013; Policarpi et al., 2017). However, the possibility of a direct regulatory function of TFIIIC on the transcriptional response to stress was not considered in previous reports. The novelty of our results thus resides in the identification of a function of TFIIIC independent of Pol III but dependent on changes in chromatin 3D structure. In response to SS, TFIIIC harnesses a subset of AEs pre-marked by the ChAHP complex, at least by its component ADNP, to establish long-distance contacts with CTCF bound to promoters of cell-cycle genes and uses its H3K18 acetylation activity to maintain these genes ready to respond to serum re-exposure (Figure S6D).

One relevant finding is the description of a mechanism underlying the selective binding of TFIIIC to a specific subset of the more than 1.1 million AEs in the human genome. Our study identifies ADNP, a subunit of the ChAHP complex, as one of the major chromatin-associated TFIIIC interactors and proves its requirement for TFIIIC loading to a specific subset of AEs, in agreement with the suggested putative role for ADNP in shaping genome topology (Kaaij et al., 2019; Weintraub et al., 2017). Interestingly, such mechanism could be evolutionary conserved as murine *Adnp*, which occupies SINEs, has been shown to interact with TFIIIC (Kaaij et al., 2019; Ostapcuk et al., 2018). In the context of the selectivity of the interactions, our analysis has characterized these regions as CTCF-loaded sites enriched in Cohesin. Moreover, these TFIIIC-looped regions show much higher levels of occupancy of ADNP than all the TFIIIC-bound AEs or all the AEs at the TFIIIC-bound promoters. Therefore, ADNP, and maybe the whole ChAHP complex, could be involved in the TFIIIC-mediated process of long distal looping, together with CTC*F.* This might appear to conflict recent findings regarding the ability of the ChAHP complex to serve as a defense mechanism against genome architecture rewiring upon transposon-mediated CTCF motif spreading in mESCs (Kaaij et al., 2019). Neither TFIIIC occupancy nor the effects of serum deprivation on ChAHP complex activity were considered in that study. In addition, as indicated by Kaaij and colleagues, there is no evidence of recent expansion of tRNA-derived SINEs in the primate lineage, and, accordingly, CTCF association to AEs has not been detected (Schmidt et al., 2012). Therefore, our study, which was conducted in human cells, might just add a new layer of complexity to this issue suggesting that TFIIIC might help ADNP-containing complexes to interact with distant CTCF sites.

TFIIIC relieves chromatin-mediated repression *in vitro* (Kundu et al., 1999) and mediates p300/CBP recruitment to SINEs in mouse neuronal cells to locally increase histone acetylation (Crepaldi et al., 2013). Our results show that H3K18ac, but not H3K27ac or H3K9ac, responds to serum depletion. In addition, we show a strong dependence on TFIIIC for the increase of this mark at AEs, as knockdown of TFIIIC reduces H3K18ac at these sites. Finally, the increase in H3K18ac does not depend on known H3K18 HATs such as p300, which is not detected at these sites. Together with the known TFIIIC HAT activity *in vitro* (Basu et al., 2019; Hsieh et al., 1999; Kundu et al., 1999), these findings lead us to propose that one of the TFIIIC components is directly responsible for H3K18ac. Published data support a role of H3K18ac in the last steps preceding activation of transcription (Ferrari et al., 2012, 2014). However, what guarantees its specificity *in vivo* and whether H3K18 acetylation occurs before or after transcription initiation remain as open questions worth of further investigation.

Finally, although SS is a technique widely used to synchronize cultured cells, it has been shown to act as a major cellular stress triggering a plethora of distinct responses (Pirkmajer and Chibalin, 2011). In particular, solid tumors with poor blood supply are exposed to serum deprivation (Anastasiou, 2017), and they might require the TFIIIC-mediated response for their survival, given the link to the regulation of proliferation-associated genes. Remarkably, the analysis of published cancer data shows that expression of the genes regulated by TFIIIC-directed H3K18ac in response to SS (TFIIIC-associated promoters and TFIIIC-activated genes) is predictive of the clinical outcome of

breast cancer patients (Figures S6E and S6G) and that H3K18ac levels also correlate with cancer prognosis (Seligson et al., 2005; Tasselli et al., 2016). Thus, our study points to TFIIIC and its HAT activity as potential new targets for cancer management.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- METHODS DETAILS
  - Cell Lines and Treatments
  - Antibodies
  - TFIIIC and ADNP Protein Expression and Pull Down
  - ChIP-seq
  - ATAC-seq
  - ChIP-seq and ATAC-seq Peak Calling
  - ChIP-seq Downstream Analysis
  - ATAC-seq Downstream Analysis
  - Micrococcal Nuclease (MNase) Analysis of AEs
  - RNA-seq
  - RNA-seq Pipeline and Differential Gene Expression Analysis
  - Co-immunoprecipitation Assay
  - siRNA Knockdowns
  - Cellular Biochemical Fractionation
  - Gene Ontology (GO) Analysis
  - External Data Sources
  - Bedtools
  - Hi-C
  - *In Situ* Hi-C Data Processing and Normalization
  - Identification of Sub-nuclear Compartments and TADs
  - Identification of Intra-TAD contacts
  - MICROSCOPY
  - CRISPR/CASP9
  - PROTEOMICS
  - Mass Spectrometry Analysis
  - Analysis of Breast Cancer Tumor Samples
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.10.020.

ACKNOWLEDGMENTS

We thank all the members of the Beato lab, the CRG Gene Regulation, Stem Cells, and Cancer Program, the CRG Genome and Proteomic Facilities, the 4D Genome Unit (Synergy Program), Jose Luis Villanueva (CRG), Yasmina Cuartero (CRG), and Prof. Simone Ottonello (University of Parma, Italy) for invaluable insight and help. This work was supported by the Spanish Ministry of Economy and Competitiveness Centro de Excelencia Severo Ochoa 2013–2017 SEV-2012-0208 (to CRG) and BFU2016-76141-P (to S.d.l.L.); ACER (to CRG); Italian Association for Cancer Research (AIRC, grant IG16877 to G.D.); the Cancer Research UK Programme Foundation (CR-UK C47547/A21536 to A.V.); a Wellcome Trust Investigator Award (200818/Z/16/Z to A.V.); and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013/ERC Synergy grant agreement 609989-4DGenome to M.B.). We acknowledge the support of the Spanish Ministry of Science, Innovation, and Universities to the EMBL partnership, the CERCA Programme/Generalitat de Catalunya, and the Centro de Excelencia Severo Ochoa. The ERC provided funding for the open access charge. The proteomics analyses were performed at the CRG/UPF Proteomics Unit (part of the of Proteored, PRB3; supported by grant PT17/0019 [ISCIII and ERDF]). This work was also supported by the Ligue Contre le Cancer (to M.T.).

AUTHOR CONTRIBUTIONS

R.F., L.I.d.L.C., A.V., and M.B. designed the experiments. R.F., L.I.d.L.C., C.D.V., L.J., E.C., and F.L.D. performed the experiments. R.F., E.V., J.Q.O., and A.L. carried out the biostatistics analysis. F.L.D. also assisted with the sequencing. R.F. and M.B. wrote the manuscript in consultation with G.D., M.T., A.V., and S.d.l.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 12, 2019
Revised: August 20, 2019
Accepted: October 13, 2019
Published: November 20, 2019

REFERENCES

Allen, T.A., Von Kaenel, S., Goodrich, J.A., and Kugel, J.F. (2004). The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. Nat. Struct. Mol. Biol. *11*, 816–821.

Anastasiou, D. (2017). Tumour microenvironment factors shaping the cancer metabolism landscape. Br. J. Cancer *116*, 277–286.

Aparicio-Prat, E., Arnan, C., Sala, I., Bosch, N., Guigó, R., and Johnson, R. (2015). DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. BMC Genomics *16*, 846.

Ay, F., Vu, T.H., Zeitz, M.J., Varoquaux, N., Carette, J.E., Vert, J.P., Hoffman, A.R., and Noble, W.S. (2015). Identifying multi-locus chromatin contacts in human cells using tethered multiple 3C. BMC Genomics *16*, 121.

Basu, M.B.R., Das, S., and Kundu, T.K. (2019). The Largest Subunit of Human TFIIIC Complex, TFIIIC220, a Lysine Acetyltransferase Targets Histone H3K18. bioRxiv. https://doi.org/10.1101/513127.

Bates, D., Mächler, M., Bolker, B., M., and Walker, S., C. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. *67*, 1–48.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Büchel, G., Carstensen, A., Mak, K.Y., Roeschert, I., Leen, E., Sumara, O., Hofstetter, J., Herold, S., Kalb, J., Baluapuri, A., et al. (2017). Association with Aurora-A Controls N-MYC-Dependent Promoter Escape and Pause Release of RNA Polymerase II during the Cell Cycle. Cell Rep. *21*, 3483–3497.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Canella, D., Praz, V., Reina, J.H., Cousin, P., and Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. Genome Res. *20*, 710–721.

Carrière, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., Jounier, S., Aude, J.C., Keime, C., Murvai, J., et al. (2012). Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. Nucleic Acids Res. *40*, 270–283.

Choi, H., Larsen, B., Lin, Z.Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods *8*, 70–73.

Chuong, E.B., Elde, N.C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. Genet. *18*, 71–86.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Conti, A., Carnevali, D., Bollati, V., Fustinoni, S., Pellegrini, M., and Dieci, G. (2015). Identification of RNA polymerase III-transcribed Alu loci by computational screening of RNA-Seq data. Nucleic Acids Res. *43*, 817–835.

Cournac, A., Koszul, R., and Mozziconacci, J. (2016). The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. Nucleic Acids Res. *44*, 245–255.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240–244.

Crepaldi, L., Policarpi, C., Coatti, A., Sherlock, W.T., Jongbloets, B.C., Down, T.A., and Riccio, A. (2013). Binding of TFIIIC to sine elements controls the relocation of activity-dependent neuronal genes to transcription factories. PLoS Genet. *9*, e1003699.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, 3.

Di Vona, C., Bezdan, D., Islam, A.B., Salichs, E., López-Bigas, N., Ossowski, S., and de la Luna, S. (2015). Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. Mol. Cell *57*, 506–520.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. Trends Genet. *23*, 614–622.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Ferrari, R., Pellegrini, M., Horwitz, G.A., Xie, W., Berk, A.J., and Kurdistani, S.K. (2008). Epigenetic reprogramming by adenovirus e1a. Science *321*, 1086–1088.

Ferrari, R., Su, T., Li, B., Bonora, G., Oberai, A., Chan, Y., Sasidharan, R., Berk, A.J., Pellegrini, M., and Kurdistani, S.K. (2012). Reorganization of the host epigenome by a viral oncogene. Genome Res. *22*, 1212–1221.

Ferrari, R., Gou, D., Jawdekar, G., Johnson, S.A., Nava, M., Su, T., Yousef, A.F., Zemke, N.R., Pellegrini, M., Kurdistani, S.K., and Berk, A.J. (2014). Adenovirus small E1A employs the lysine acetylases p300/CBP and tumor suppressor Rb to repress select host genes and promote productive virus infection. Cell Host Microbe *16*, 663–676.

Galli, G.G., Carrara, M., Francavilla, C., de Lichtenberg, K.H., Olsen, J.V., Calogero, R.A., and Lund, A.H. (2013). Genomic and proteomic analyses of Prdm5 reveal interactions with insulator binding proteins in embryonic stem cells. Mol. Cell. Biol. *33*, 4504–4516.

Giorgetti, L., Lajoie, B.R., Carter, A.C., Attia, M., Zhan, Y., Xu, J., Chen, C.J., Kaplan, N., Chang, H.Y., Heard, E., and Dekker, J. (2016). Structural organization of the inactive X chromosome in the mouse. Nature *535*, 575–579.

Gu, Z., Jin, K., Crabbe, M.J.C., Zhang, Y., Liu, X., Huang, Y., Hua, M., Nan, P., Zhang, Z., and Zhong, Y. (2016). Enrichment analysis of Alu elements with different spatial chromatin proximity in the human genome. Protein Cell *7*, 250–266.

Győrffy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q., and Szallasi, Z. (2010). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res. Treat. *123*, 725–731.

Horwitz, G.A., Zhang, K., McBrian, M.A., Grunstein, M., Kurdistani, S.K., and Berk, A.J. (2008). Adenovirus small e1a alters global patterns of histone modification. Science *321*, 1084–1085.

Hsieh, Y.J., Kundu, T.K., Wang, Z., Kovelman, R., and Roeder, R.G. (1999). The TFIIIC90 subunit of TFIIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. Mol. Cell. Biol. *19*, 7697–7704.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods *9*, 999–1003.

Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. EMBO J. *30*, 249–262.

Jourdain, S., Acker, J., Ducrot, C., Sentenac, A., and Lefebvre, O. (2003). The tau95 subunit of yeast TFIIIC influences upstream and downstream functions of TFIIIC.DNA complexes. J. Biol. Chem. *278*, 10450–10457.

Kaaij, L.J.T., Mohn, F., van der Weide, R.H., de Wit, E., and Buhler, M. (2019). The ChAHP Complex Counteracts Chromatin Looping at CTCF Sites that Emerged from SINE Expansions in Mouse. Cell *178*, 1437–1451.

Kundu, T.K., Wang, Z., and Roeder, R.G. (1999). Human TFIIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity. Mol. Cell. Biol. *19*, 1605–1615.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al.; International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

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

Le Dily, F., Baù, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H., Ballare, C., Filion, G., et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev. *28*, 2151–2162.

Li, H., Handsaker, B., Wysoker, A., Fennel, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.T., Cubeñas-Potts, C., Hu, M., Lei, E.P., Bosco, G., et al. (2015). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol. Cell *58*, 216–231.

Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., and Goodrich, J.A. (2008). Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Mol. Cell *29*, 499–509.

Mertens, C., and Roeder, R.G. (2008). Different functional modes of p300 in activation of RNA polymerase III transcription from chromatin templates. Mol. Cell. Biol. *28*, 5764–5776.

Mohammed, H., Taylor, C., Brown, G.D., Papachristou, E.K., Carroll, J.S., and D'Santos, C.S. (2016). Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nat. Protoc. *11*, 316–326.

Moqtaderi, Z., Wang, J., Raha, D., White, R.J., Snyder, M., Weng, Z., and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat. Struct. Mol. Biol. *17*, 635–640.

Oler, A.J., Alla, R.K., Roberts, D.N., Wong, A., Hollenhorst, P.C., Chandler, K.J., Cassiday, P.A., Nelson, C.A., Hagedorn, C.H., Graves, B.J., and Cairns, B.R. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat. Struct. Mol. Biol. *17*, 620–628.

Ostapcuk, V., Mohn, F., Carl, S.H., Basters, A., Hess, D., Iesmantavicius, V., Lampersberger, L., Flemr, M., Pandey, A., Thomä, N.H., et al. (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature *557*, 739–743.

Pascali, C., and Teichmann, M. (2013). RNA polymerase III transcription - regulated by chromatin structure and regulator of nuclear chromatin organization. Subcell. Biochem. *61*, 261–287.

Pellegrini, M., and Ferrari, R. (2012). Epigenetic analysis: ChIP-chip and ChIP-seq. Methods Mol. Biol. *802*, 377–387.

Pirkmajer, S., and Chibalin, A.V. (2011). Serum starvation: caveat emptor. Am. J. Physiol. Cell Physiol. *301*, C272–C279.

Policarpi, C., Crepaldi, L., Brookes, E., Nitarska, J., French, S.M., Coatti, A., and Riccio, A. (2017). Enhancer SINEs Link Pol III to Pol II Transcription in Neurons. Cell Rep. *21*, 2879–2894.

Pombo, A., and Dillon, N. (2015). Three-dimensional genome architecture: players and mechanisms. Nat. Rev. Mol. Cell Biol. *16*, 245–257.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Ramos, Y.F., Hestand, M.S., Verlaan, M., Krabbendam, E., Ariyurek, Y., van Galen, M., van Dam, H., van Ommen, G.J., den Dunnen, J.T., Zantema, A., and 't Hoen, P.A. (2010). Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. Nucleic Acids Res. *38*, 5396–5408.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Ringnér, M., Fredlund, E., Häkkinen, J., Borg, Å., and Staaf, J. (2011). GOBO: gene expression-based outcome for breast cancer online. PLoS ONE *6*, e17911.

Rueden, C., T., Schindelin, J., Hiner, M., C., DeZonia, B., E., Walter, A., E., Arena, E., T., and Eliceiri, K., W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics *18*, https://doi.org/10.1186/s12859-017-1934-z.

Ruiz-Velasco, M., Kumar, M., Lai, M.C., Bhat, P., Solis-Pinson, A.B., Reyes, A., Kleinsorg, S., Noh, K.M., Gibson, T.J., and Zaugg, J.B. (2017). CTCF-Mediated Chromatin Loops between Promoter and Gene Body Regulate Alternative Splicing across Individuals. Cell Syst. *5*, 628–637.

Schmidt, D., Schwalie, P.C., Wilson, M.D., Ballester, B., Gonçalves, A., Kutter, C., Brown, G.D., Marshall, A., Flicek, P., and Odom, D.T. (2012). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell *148*, 335–348.

Seligson, D.B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., and Kurdistani, S.K. (2005). Global histone modification patterns predict risk of prostate cancer recurrence. Nature *435*, 1262–1266.

Serra, F., Baù, D., Goodstadt, M., Castillo, D., Filion, G.J., and Marti-Renom, M.A. (2017). Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. PLoS Comput. Biol. *13*, e1005665.

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. Bioinformatics *25*, 2605–2606.

Su, A.G.Y.-S. (2016). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.9-3 (Cambridge University Press).

Su, M., Han, D., Boyd-Kirkup, J., Yu, X., and Han, J.J. (2014). Evolution of Alu elements toward enhancers. Cell Rep. *7*, 376–385.

Tanaka, Y., Yamashita, R., Suzuki, Y., and Nakai, K. (2010). Effects of Alu elements on global nucleosome positioning in the human genome. BMC Genomics *11*, 309.

Tasselli, L., Xi, Y., Zheng, W., Tennen, R.I., Odrowaz, Z., Simeoni, F., Li, W., and Chua, K.F. (2016). SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. Nat. Struct. Mol. Biol. *23*, 434–440.

Tien, A.L., Senbanerjee, S., Kulkarni, A., Mudbhary, R., Goudreau, B., Ganesan, S., Sadler, K.C., and Ukomadu, C. (2011). UHRF1 depletion causes a G2/M arrest, activation of DNA damage response and apoptosis. Biochem. J. *435*, 175–185.

Toska, E., Osmanbeyoglu, H.U., Castel, P., Chan, C., Hendrickson, R.C., Elkabets, M., Dickler, M.N., Scaltriti, M., Leslie, C.S., Armstrong, S.A., and Baselga, J. (2017). PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. Science *355*, 1324–1330.

Tsirigos, A., and Rigoutsos, I. (2009). Alu and b1 repeats have been selectively retained in the upstream and intronic regions of genes of specific functional classes. PLoS Comput. Biol. *5*, e1000610.

Van Bortle, K., and Corces, V.G. (2012). tDNA insulators and the emerging role of TFIIIC in genome organization. Transcription *3*, 277–284.

van de Werken, H.J.G., Haan, J.C., Feodorova, Y., Bijos, D., Weuts, A., Theunis, K., Holwerda, S.J.B., Meuleman, W., Pagie, L., Thanisch, K., et al. (2017). Small chromosomal regions position themselves autonomously according to their chromatin class. Genome Res. *27*, 922–933.

Vicent, G.P., Nacht, A.S., Ballaré, C., Zaurin, R., Soronellas, D., and Beato, M. (2014). Progesterone receptor interaction with chromatin. Methods Mol. Biol. *1204*, 1–14.

Vidal, E., le Dily, F., Quilez, J., Stadhouders, R., Cuartero, Y., Graf, T., Marti-Renom, M.A., Beato, M., and Filion, G.J. (2018). OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes. Nucleic Acids Res. *46*, e49.

Wang, Z., and Roeder, R.G. (1997). Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. Genes Dev. *11*, 1315–1326.

Wang, X., Spandidos, A., Wang, H., and Seed, B. (2012). PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. Nucleic Acids Res. *40*, D1144–D1149.

Wang, H., La Russa, M., and Qi, L.S. (2016). CRISPR/Cas9 in Genome Editing and Beyond. Annu. Rev. Biochem. *85*, 227–264.

Weiner, A.M. (2002). SINEs and LINEs: the art of biting the hand that feeds you. Curr. Opin. Cell Biol. *14*, 343–350.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573–1588.

Weissmann, F., Petzold, G., VanderLinden, R., Huis In 't Veld, P.J., Brown, N.G., Lampert, F., Westermann, S., Stark, H., Schulman, B.A., and Peters, J.M. (2016). biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. Proc. Natl. Acad. Sci. USA *113*, E2564–E2569.

Welch, R.P., Lee, C., Imbriano, P.M., Patil, S., Weymouth, T.E., Smith, R.A., Scott, L.J., and Sartor, M.A. (2014). ChIP-Enrich: gene set enrichment testing for ChIP-seq data. Nucleic Acids Res. *42*, e105.

Weser, S., Gruber, C., Hafner, H.M., Teichmann, M., Roeder, R.G., Seifart, K.H., and Meissner, W. (2004). Transcription factor (TF)-like nuclear regulator, the 250-kDa form of Homo sapiens TFIIIB", is an essential component of human TFIIIC1 activity. J. Biol. Chem. *279*, 27022–27029.

Wysocka, J., Reilly, P.T., and Herr, W. (2001). Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. Mol. Cell. Biol. *21*, 3820–3829.

Yakovchuk, P., Goodrich, J.A., and Kugel, J.F. (2009). B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. Proc. Natl. Acad. Sci. USA *106*, 5569–5574.

Zhao, Y., Wang, J., Liang, F., Liu, Y., Wang, Q., Zhang, H., Jiang, M., Zhang, Z., Zhao, W., Bao, Y., et al. (2019). NucMap: a database of genome-wide nucleosome positioning map across species. Nucleic Acids Res. *47* (D1), D163–D169.

Ziyatdinov, A., Vázquez-Santiago, M., Brunel, H., Martinez-Perez, A., Aschard, H., and Soria, J.M. (2018). lme4qtl: linear mixed models with flexible covariance structure for genetic studies of related individuals. BMC Bioinformatics *19*, 68.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit polyclonal antibody anti-GTF3C2	This paper	NA
mouse monoclonal antibody anti-GTF3C2	Abcam	RRID:AB_2043034
rabbit polyclonal antibody anti-BDP1	Wang and Roeder, 1997; Weser et al., 2004	NA
rabbit polyclonal antibody anti-RPC39	Wang and Roeder, 1997; Weser et al., 2004	NA
rabbit polyclonal antibody anti-H3K18ac	Active motif	RRID:AB_2793308
rabbit polyclonal antibody anti-CTCF	Millipore	RRID:AB_441965
rabbit monoclonal antibody anti-CTCF	Abcam	ab128873
rabbit polyclonal antibody anti-Pol II	Cell Signaling	(NTD) (D8L4Y, Cat# 14958 RRID:AB_2687876)
rabbit affinity-purified anti-GTF3C5/TFIIIC63	Bethyl	RRID:AB_890669
rabbit affinity-purified anti-GTF3C1/TFIIIC220	Novus Biologicals	RRID:AB_922616
rabbit affinity-purified anti-GTF3C4/TFIIIC90	Abcam	RRID:AB_1269036
mouse Monoclonal antibody anti-tubulin	Sigma	RRID:AB_477593
rabbit anti-GTF3C1/TFIIIC220 Antibody, Affinity Purified ChIP	Bethyl	RRID:AB_938038
rabbit anti-GTF3C1/TFIIIC220 Antibody, Affinity Purified WB	Bethyl	RRID:AB_938042
mouse monoclonal [102C1a] to ADNP	Abcam	RRID:AB_879525
Rabbit polyclonal HA tag antibody	Abcam	RRID:AB_307019
rabbit affinity-purified anti-EP300	Santa Cruz	RRID:AB_2231120-Lot: I2815
rabbit affinity-purified anti-EP300	Santa Cruz	RRID:AB_2293429-Lot: J0915
Rabbit affinity-purified anti-H3K9ac	Active Motif	RRID:AB_2561017-Lot16918002
Mouse Monoclonal anti-H3K27ac	Millipore	RRID:AB_1977529-Lot.2626032
Rabbit polyclonal to Histone H1.2	Abcam	RRID:AB_2117983
Horseradish peroxidase conjugated anti-mouse	Amersham	NA931V
Horseradish peroxidase conjugated anti-rabbit	Amersham	NA934V
Anti-DYKDDDDK resin	Genscript	L00432
Anti-HA Agarose	Pierce	26181
Bacterial and Virus Strains		
One Shot Stbl3 Chemically Competent *E. coli*	Invitrogen	C7373-03
Chemicals, Peptides, and Recombinant Proteins		
RPMI Red phenol	GIBCO	42401-018
RPMI No Phenol Red	GIBCO	32404-014
DMEM Phenol Red	GIBCO	41965-039
DMEM Phenol Red	GIBCO	21063-029
DMEM/F12 Phenol Red	GIBCO	11330-032
DMEM/F12 No Phenol Red	GIBCO	11039-021
EMEM Phenol Red	GIBCO	31095-029
EMEM No Phenol Red	GIBCO	51200-038
Fetal Bovine Serum	GIBCO	10270-106
Fetal Bovine Serum, charcoal stripped	GIBCO	12676029
0.5% Trypsin-EDTA 1x	GIBCO	25300-054

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
L-Glutamine 200 mM 100x	GIBCO	25030-024
Penicillin-Streptomycin	GIBCO	15140-122
Human Insulin (Humulin regular)	Lilly	U100
EGF	SIGMA	E-9644
Hydrocortisone	SIGMA	H-0888
Horse Serum	Life technologies	16050122
Cholera Toxin	SIGMA	C8052
Trizol Reagent	Ambion	15596018
Lipofectamine 3000	Invitrogen	11668-019
Proteinase K	ThermoFisher Scientific	AM2546
Protein G Plus / Protein A Agarose	Millipore	IP05
protease inhibitor cocktail, cOmplete EDTA-free	Roche	05 892 791 001
Na_3VO_4		
MboI	New England BioLabs	r0147-mboi
Herculase II Fusion DNA Polymerase	Agilent	600675
AMPure XP beads	Beckman Coulter	A63881
Endopeptidase LysC,	Wako	125-05061
Sequencing grade Trypsin	Promega	V5111
Critical Commercial Assays		
Ovation Ultralow DR Multiplex System 9-16 kit	NUGEN	0535-32
Micro BCA Protein Assay Kit	ThermoFisher Scientific	23235
qScript cDNA Synthesis kit	Quanta Biosciences	95047-025
LightCycler FastStart DNA Master SYBR Green I kit	Roche	03 515 885 001
Illumina TruSeq kit	Illumina	20020594
TruSeq Stranded Total RNA Library Prep Human/ Mouse/Rat	Illumina	20020596
Pierce Coomassie (Bradford) kit	Thermo Fisher	23200
NEBNext DNA Library Prep Kit	New England BioLabs	NEB #E7645
Qubit HS kit	Thermo Fisher	Q32854
FISH Tag RNA Multicolor Kit	Thermo Fisher	mp32956
Gibson Assembly Master	New England Biolabs	E2611S
BsmbI	Thermo Fisher	ER0451
Anti-DYKDDDDK G1 Affinity Resin	Genscript	L00432
Anti-HA Agarose	Thermo Fisher Scientific	26181
NuPAGE MOPS SDS Running Buffer	Novex by Life Technology	NP0001
NuPAGE 4-12% Bis Tris Gels	Invitrogen	NP0335box
Skim Milk Powder	Sigma	70166-500G
XCell SureLock Mini-Cell Electrophoresis System	Thermo Fisher	EI0001
Deposited Data		
ChIP-seq, RNA-seq, Hi-C data	This study	GSE120162
Human reference genome NCBI build 38, GRCh38	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
CTCF ChIP-seq of T47D in -S	Le Dily et al., 2014	GSE53463
ChIP-seq of eGFP-ADNP in K562 cells	Consortium, 2012	GSE105573
ChIP-seq of mouse Adnp	Ostapcuk et al., 2018	GSE97945
Gene expression data of breast cancer samples	Ringnér et al., 2011	http://co.bmc.lu.se/gobo/
EP300 ChIP-seq in T98G	Ramos et al., 2010	GSE21026

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Proteomic data: PRIDE	This study	PXD011250
Mendeley	This study	https://doi.org/10.17632/mzjf96t3gc.3
Experimental Models: Cell Lines		
T47D	ATCC	CRL-2865
T98G	ATCC	CRL-1690
IMR90	ATCC	CCL-186
MCF10A	ATCC	CRL-10317
Oligonucleotides		
UHRF1-associated AE H3K18ac ChIP Forward: ATTGTAAT CCCGGTCGTTTG	This study	NA
UHRF1-associated AE H3K18ac ChIP Reverse: CGGGTTC AAGTGATTCTCGT	This study	NA
UHRF1-expression Forward: GCCATACCCTCTTTCGACTACG	This study	NA
UHRF1-expression Reverse: GCCCCAATTCCGTCTCATCC	This study	NA
HELLS-associated AE H3K18ac ChIP:Forward: TAGCCTGGAATGGGCTAAT	This study	NA
HELLS-associated AE H3K18ac ChIP: Reverse: TCAGTTGATCCTCCCACCTC	This study	NA
PPIA primers: For – GCCGAGGAAAACCGTGTACT	This study	NA
PPIA primers: Rev- GTCTTTGGGACCTTGTCTGC	This study	NA
siGENOME against human *GTF3C5* (9328) siRNA	Dharmacon	D-020031-02
siGENOME Non-Targeting siRNA Pool #2	Dharmacon	D-001206-14
siGENOME against human *GTF3C1* siRNA	Dharmacon	LQ-012581-00-0002
siGENOME against human *ADNP* siRNA	Dharmacon	LQ-012857-01-0002
GTF3C5 expression Forward B: GCGGCAAGCATACGTCAATG	This study	NA
GTF3C5 expression Rev B: TGGTCGGTAGAAGTAGTCCAC	This study	NA
GAPDH expression forward: GACTCAACGGATTTGGTCGT	This study	NA
GAPDH expression reverse: TTGATTTTGGAGGGATCTCG	This study	NA
Oligo1 reverse: TGGGATTCCTGGATCCGGTGTTTCGTCCTTT CCACAAGAT	This study	NA
Oligo2 forward: CACCGGATCCAGGAATCCCACTTCTGTTTT AGAGCTAGAAGAGAC	This study	NA
Oligo3 reverse: GAGACGGGATCCTAGGAATTCCGTCTCTTC TAGCTCTAAAAC	This study	NA
Oligo4 forward: TTCCTAGGATCCCGTCTCTCTGTATGAGAC CACTCTTTCCC	This study	NA
Oligo5 reverse: AACAAACCAAAAAACCCCACAAAGGGAA AGAGTGGTCTCAT	This study	NA
Oligo6 forward: GGGGTTTTTTGGTTTGTTTTAGAGCTAGAA ATAGCAAGTT	This study	NA
C557F: GTTTTAGAGCTAGAAATAGCAAG	This study	NA
C557R: GTGGTCTCATACAGAACTTATAAG	This study	NA
C542F: GTACAAAATACGTGACGTAG	This study	NA
C542R: ATGTCTACTATTCTTTCCCC	This study	NA
gRNA1: GATCCAGGAATCCCACTTCT	This study	NA
gRNA2: TTTGTGGGGTTTTTTGGTTT	This study	NA
Alu_up3: CCGAAGGCTAAAAGCGACTA	This study	NA
Alu_down: ACGTTGGCAAGGATTTGAAG	This study	NA
GTF3C1_expression_for: GGGAAGCTGCACTATCACAGA	This study	NA
GTF3C1_expression_rev: GGTAATCGGATCACATGGGACT	This study	NA

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
ADNP_expression_for: CATGGGAGGATGTAGGACTGT	This study	NA
ADNP_expression_rev: ATGGACATTGCGGAAATGACT	This study	NA
Recombinant DNA		
pDECKO_mCherry	Addgene	78534
pBIG2abc vector	Addgene	80617
Software and Algorithms		
Samtools	Li et al., 2009	http://samtools.sourceforge.net/
FACS Diva Software v6.1.2	Becton Dickinson	NA
ModFit v3.2	Verity Software	NA
cis-egulatory annotation system (CEAS)	Shin et al., 2009	NA
Trimmomatic	Bolger et al., 2014	https://goo.gl/VzoqQq
STAR	Dobin et al., 2013	NA
ChIP-seq Enrich	Welch et al., 2014	http://chip-enrich.med.umich.edu/
DAVID	Dennis et al., 2003	https://david.ncifcrf.gov/
ImageJ	Rueden et al., 2017	https://imagej.nih.gov/ij
Xcalibur software v2.2	Thermo Fisher Scientific	OPTON-30487
Proteome Discoverer software suite v1.4	Thermo Fisher Scientific	OPTON-30812
Mascot search engine v2.5	Matrix Science	NA
SAINT software	Choi et al., 2011	http://saint-apms.sourceforge.net/Main.html
NucMap	Zhao et al., 2019	http://bigd.big.ac.cn/nucmap
Other		
Gene expression analysis script	This paper	https://doi.org/10.17632/mzjf96t3gc.3
Hi-C analysis pipeline	Serra et al., 2017; Vidal et al., 2018	

LEAD CONTACT AND MATERIALS AVAILABILITY

Requests for reagents, cell lines generated in this study and resource sharing should be addressed to and will be fulfilled by the lead contact, Miguel Beato (miguel.beato@crg.eu).

All unique/stable reagents generated in this study are available from the Lead Contact without restriction.

METHODS DETAILS

Cell Lines and Treatments

Human T47D cells (American Type Culture Collection [ATCC]: CRL-2865) were grown in RPMI supplemented with 10% fetal bovine serum (FBS) (referred as +S condition in the Figures); for serum starvation experiments, cells were treated with RPMI supplemented with 10% charcoal-treated FBS for 48 h and starvation was achieved by culturing cells in the absence of FBS for 16 h (referred as –S condition in the Figures). T98G cells (ATCC: CRL-1690) were grown in DMEM with 10% FBS (referred as +S in the Figures); for serum starvation experiments, cells were cultured in DMEM with 0.1% FBS for 48 h (referred as –S in the Figures). IMR90 fibroblasts (ATCC: CCL-186) were grown in EMEM with 10% FBS (referred as +S in the Figures); for serum starvation experiments, cells were cultured in the absence of FBS for 16 h (referred as –S in the Figures). MCF10A (ATCC: CRL-10317) were cultured in DMEM/F12 supplied with 20 ng/ml epidermal growth factor, 0.5 μg/ml hydrocortisone, 10 μg/ml insulin, and 5% horse serum (with the addition of 100 ng/ml cholera toxin) (referred as +S in the Figures); for serum starvation experiments, cells were grown in RPMI supplemented with 10% charcoal-treated FBS for 48 h and in the absence of serum for 16 h (referred as –S in the Figures). All cultures were maintained with antibiotics (100 u/ml penicillin, 100 μg/ml streptomycin).

Experiments of serum re-exposure in Figures 1G and S1I were performed by adding DMEM supplemented with 10% FBS to serum-starved cells for 30 min. For siRNA-treated cells, serum re-exposure was performed by adding RPMI supplemented with 10% charcoal-treated FBS to the serum-starved cells at the indicated times.

For cell cycle profiling, cells were fixed with ethanol and DNA was stained with propidium iodide. Labeled cells were analyzed with a LSF II flow cytometer (Becton Dickinson) using the FACS Diva Software v6.1.2 (Becton Dickinson). The cell cycle profile was determined with the program ModFit v3.2 (BD Bioscience).

Antibodies

The TFIIIC antibody is a rabbit polyclonal antibody raised against the N-terminal 477 amino acids of TFIIIC110 (*GTF3C2*) (Q8WUA4) generated by Martin Teichmann. The BDP1 and RPC39 antibodies were already described (Wang and Roeder, 1997; Weser et al., 2004). H3K18ac antibody was from Active Motif (39693) and its use was already described (Ferrari et al., 2014; Ferrari et al., 2008; Ferrari et al., 2012). Commercial antibodies were: CTCF, Millipore (07-729); Pol II (NTD, D8L4Y), Cell Signaling (14958); TFIIIC63 (*GTF3C5*), Bethyl (A301-242A); TFIIIC220 (*GTF3C1*), Novus Biologicals (NB100-60657), Bethyl (A301-293A) for western blotting and Bethyl (A301-291A) for ChIP; TFIIIC90 (*GTF3C4*), Abcam (ab74229); α-Tubulin (*TUBA4A*), Sigma (T9026); ADNP, Abcam (ab54402); histone H1.2, Abcam (ab4086); H3K27ac, Millipore (17-683); p300 (*EP300*), Santa Cruz (sc584-Lot: J0915 and sc585-Lot: I2815). Rabbit polyclonal anti-HA Abcam (ab9110) and mouse monoclonal anti-GTF3C2 Abcam (ab89113).

TFIIIC and ADNP Protein Expression and Pull Down

The open reading frames of human TFIIIC components were cloned into the pBIG2abc vector (Weissmann et al., 2016) (Addgene #80617); GTF3C1-was cloned with a C-terminal 3xFLAG-tag and the rest of the TFIIIC subunits (GTF3C2-6) were cloned untagged. The open reading frame of human ADNP with a C-terminal HA-tag was cloned into the pLIB vector (Weissmann et al., 2016) (Addgene #80610). The constructs were transposed into *E. coli* DH10 Multi-emBassY cells to generate a Bacmid, which was subsequently used to infect Sf9 cells to generate a virus for each of the construct. For protein expression, Hi5 cells were infected either separately with the viruses for TFIIIC or ADNP or co-infected with both viruses. Hi5 cells were grown for 4 days at 27°C. Cells were harvested by centrifugation, resuspended in lysis buffer (500 mM NaCl, 20 mM HEPES pH 8, 1 mM $MgCl_2$, 10% glycerol, 5 mM β-mercaptoethanol) and sonicated for 10 s. The lysates were centrifuged at 16,000x*g* for 30 min and the supernatant was incubated with anti-DYKDDDDK G1 affinity resin (Genscript) or anti-HA agarose (Thermo Fisher Scientific) for 3 h at 4°C on a rolling plate. The samples were centrifuged at 1,200x*g* for 3 min, the supernatant was removed and the beads were washed 2 times with 20 column volumes of lysis buffer by centrifugation. The washed resin was directly mixed with Laemmli sample buffer, heated for 5 min at 100°C and ran on a 4%–12% NuPAGE Bis-Tris Gel (Thermo Fisher Scientific) in MES buffer (Invitrogen) for 40 min at 200 V. One aliquot of the immunocomplexes was detected by Coomassie staining of the gel and another one by immunoblotting with anti-HA and anti-GTF3C2 antibodies.

ChIP-seq

Chromatin Immunoprecipitation (ChIP): chromatin purification and library preparation

Preparation of cross-linked chromatin free of RNA, sonication, and immunoprecipitation was performed as previously described (Vicent et al., 2014). For Pol II, ChIP was performed as described (Di Vona et al., 2015). DNA was quantified with the Qubit HS kit (Invitrogen).

Single-end (SE) sequencing libraries were constructed from 1 ng of immunoprecipitated and input DNA using the Ovation Ultralow DR Multiplex System 1-8 and 9-16 kit (NuGen). To minimize false positives calling, several input libraries were sequenced to reach saturation with a coverage of 4 reads/bp of the human genome for each condition. The T47D ChIPs for TFIIIC, BDP1, RPC39, and H3K18ac were done in two separate biological repeats and pulled together. The T47D ChIP for CTCF+S was done in a single biological experiment and compared with previously published CTCF ChIP data in absence of serum (Le Dily et al., 2014). The Pol II ChIP was performed in a single biological experiment.

For ChIP-qPCR, 0.2 ng of chromatin were used in reactions with the Light Cycler FastStart DNA Master SYBR Green I kit (Roche) and specific primers (see Oligonucleotides at Key Resources Table). The ChIP for p300 was carried out mixing the two Santa Cruz antibodies at ratio 1:1 (25 μl of each).

ATAC-seq

ATAC-seq reaction and library preparation

For nuclei preparation, $5x10^6$ cells were washed with cold phosphate-buffered saline (PBS), harvested in cold PBS supplemented with protease inhibitor cocktail (PIC, Roche), and cell pellets obtained by centrifugation at 900x*g* for 5 min at 4°C. The cell pellet was gently resuspended in 50 μl of RBS buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$) supplemented with PIC, and then 1.3 mL of RBS buffer - 0.1% Igepal CA-630 was added, followed by centrifugation at 500xg for 10 min at 4°C. The supernatant was carefully discarded and 1 mL of RBS buffer was added to resuspend the cell nuclei.

For the transposition reaction, 50,000 nuclei were resuspended in 25 μl 2X TD Buffer (Illumina 121-1030), 2.5 μl Tn5 Transposase (Illumina 121-1030) in a final volume of 50 μl. The reaction mix was incubated for 30 min at 37°C C and DNA purified using the QIAGEN Mini-Elute Kit in 10 μl Tris buffer 10mM, pH8.0. Eluate was subjected to PCR amplification in a volume of 50 μl as follow:

10 μl Transposed DNA
10 μl Nuclease free water
2.5 μl Customized Nextera PCR Primer1
2.5 μl Customized Nextera PCR Primer2 702 for sample -S (replicate 1) and 702 for sample -S (replicate 2).
25 μl NEB Next High-Fidelity 2x PCR Master Mix

PCR Cycle as follows:

(1) 72°C, 5 min
(2) 98°C, 30 s
(3) 98°C, 10 s
(4) 63°C, 30 s
(5) 72°C, 1 min
(6) Repeat steps 3-5, 4x
(7) Hold at 4°C

In order to reduce GC and size bias in the PCR, the reaction is monitored using a side qPCR to avoid amplification to saturation. qPCR of a total volume of 20 μl is assemble as following:

5 μl PCR amplified DNA
3.88 μl H2O
0.5 μl Primer 1
0.5 μl Primer2 701 for sample -S (replicate 1) and 701 for sample -S (replicate 2).
0.12 μl Sybr Green 100x
10 μl NEBNext High Fidelity 2x PCR Master Mix

qPCR cycle as follow:

(1) 98°C, 30 s
(2) 98°C, 10 s
(3) 63°C, 30 s
(4) 72°C, 1 min
(5) Repeat steps 2-4, 25x
(7) Hold at 4°C

The additional calculated number of cycles needed for the remaining 45 μl PCR reaction:

Sample -S replicate 1 = 7 cycles
Sample -S replicate 2 = 7 cycles

For the remaining 45 μl a PCR reaction is setup with the corrected number of cycles as follow:

(1) 98°C, 30 s
(2) 98°C, 10 s
(3) 63°C, 30 s
(4) 72°C, 1 min
(5) Repeat steps 2-4, (7x sample -S replicate 1 and 2)
(7) Hold at 4°C

Purify amplified library using QIAGEN PCR Cleanup Kit. Elute the purified library in 20 μl Elution Buffer (10 mM Tris Buffer, pH8). Ensure to dry the column before adding elution buffer

ChIP-seq and ATAC-seq Peak Calling

Analysis of sequence data was carried out as previously described (Ferrari et al., 2014) with minor modifications. Reads were aligned to the hg38 human genome reference sequence (GRCh38) using Bowtie (Langmead et al., 2009) and aligning parameters of uniqueness (-S –m1 –v2 –t –q). p values for the significance of ChIP-seq counts compared to input DNA were calculated as described (Pellegrini and Ferrari, 2012) using a threshold of 10^{-8} and a false discovery rate (FDR) < 1%.

ChIP-seq Downstream Analysis

Average ChIP-seq signals of 50 bp windows around 3 kb (or 5 kb) upstream and downstream of annotated TSSs were calculated using the *cis*-regulatory annotation system (CEAS) (Shin et al., 2009). Boxplots for ChIP-seq data or RNA-seq data were generated with R, and show median and the interquartile range; the whiskers indicate the minimum and maximum.

For the selection of the Pol II genes closed to TFIIIC peaks, we selected all the promoters of the human protein-coding gene version 4 (V4) and sorted the genes based on higher occupancy of TFIIIC measured by the sum of significant counts within 50 bp bins spanning a 10 kb region for each TSS.

ATAC-seq Downstream Analysis

Nucleosome-derived signals were extrapolated from ATAC-seq data by selecting sequenced fragments of 180-250 bp as reported (Buenrostro et al., 2013). From the BAM-aligned files two wiggle files (corresponding to the +S and the -S conditions) were extracted using BedTools GenomeCoverage function (Quinlan and Hall, 2010). Read counts were normalized through the GenomeCoverage function of factor correction. Wiggle were then used with Sitepro (Shin et al., 2009) to calculate the average nucleosome signal across genome coordinates of selected BED file as reported in Figure 3C. Wiggles were also converted to bigwig format for visualization on IGV genome browser (Figure 3D).

Micrococcal Nuclease (MNase) Analysis of AEs

Profiles of MNase digestion were obtained using the online tool NucMap (Zhao et al., 2019). Selected genomic coordinates were arranged in the format suitable for the NucMap. The calculation of nucleosome positioning was carried out using Nucleosome peaks analyzed by iNPS (as reported in Figure S3D).

RNA-seq

RNA extraction, RNA-seq library preparation, and qRT-PCR

RNA was isolated from cells with TRIzol reagent (Ambion), ethanol precipitated, and dissolved in sterile water. RNA concentration was measured with a Qubit fluorometer and RNA subjected to Bioanalyzer for quality control. Libraries were prepared using 1 μg of polyA+ RNA by PCR amplification of cDNA with bar-coded primers using the Illumina TruSeq kit at the CRG Genomic Facility. Libraries were sequenced using Illumina HIseq-2500 to obtain pair-ended (PE) 100-base-long reads.

For gene expression analysis, RNA (250 ng) was subjected to cDNA synthesis using the qScript cDNA Synthesis kit (Quanta Biosciences). qPCR was carried out using the LightCycler FastStart DNA Master SYBR Green I kit (Roche), and specific primers selected from the list of available designed primers at Primer Bank (https://pga.mgh.harvard.edu/primerbank) (Wang et al., 2012) (see Oligonucleotides at Key Resources Table). As reference gene, *GAPDH* was used.

RNA-seq Pipeline and Differential Gene Expression Analysis

Sequencing adapters and low-quality ends were trimmed from the reads using Trimmomatic, using the parameters values recommended (Bolger et al., 2014) and elsewhere (https://goo.gl/VzoqQq) (trimmomatic PE raw_fastq trimmed_fastq ILLUMINACLIP:TruSeq3-PE.fa:2:30:12:1:true LEADING:3 TRAILING:3 MAXINFO:50:0.999 MINLEN:36). The trimmed reads were aligned to GRCh38 (Lander et al., 2001) using STAR (Dobin et al., 2013).

First, the genome index files for STAR were generated with: star–runMode genomeGenerate–genomeDir GENOME_DIR–genomeFastaFiles genome_fasta–runThreadN slots–sjdbOverhang read_length–sjdbGTFfile sjdb–outFileNamePrefix GENOME_DIR/

Where genome_fasta is the FASTA file containing the GRCh38 sequence downloaded from the University of California Santa Cruz (UCSC) Genome Browser, excluding the random scaffolds and the alternative haplotypes; and sjdb is the GTF file with the GENCODE's V24 annotation.

Second, trimmed reads were aligned to the indexed genome with: star–genomeDir GENOME_DIR/–genomeLoad NoSharedMemory–runThreadN slots–outFilterType "BySJout"–outFilterMultimapNmax 20–alignSJoverhangMin 8–alignSJDBoverhangMin 1–outFilterMismatchNmax 999–outFilterMismatchNoverLmax 0.04–alignIntronMin 20–alignIntronMax 1000000–alignMatesGapMax 1000000–readFilesIn read1 read2–outSAMtype BAM SortedByCoordinate–outTmpDir TMP_DIR/–outFileNamePrefix ODIR1/$sample_id.–outWigType bedGraph–readFilesCommand zcat

(https://docs.google.com/document/d/1yRZevDdjxkEmda9WF5-qoaRjIROZmicndPl3xetFftY/edit?usp=sharing)

For the analysis of AE expression, a published algorithm was used (Conti et al., 2015). To avoid miss-alignment, we only accepted uniquely mapped reads and we focused on AE placed within intergenic regions (at least 5 kb away from any human TSS) or on the opposite strand of a known annotated transcript.

Differences in gene expression were calculated by using a DESeq.R script for RNA analysis. The script is provided as a downloadable link with the manuscript (https://www.dropbox.com/s/026pc48kfuqr88g/RNA_analysis_deseq.2.R?dl=0). Genes with fold change (FC) ± 1.5 (p value < 0.05; FDR < 0.01) were considered as significantly regulated. Sitepro profiles were generated with the script, provided in the CEAS package.

Co-immunoprecipitation Assay

Cells were lysed with lysis buffer (50 mM Tris-HCl pH 7.4, 130 mM NaCl, 1 mM EDTA, 1 mM EGTA, 5 mM $MgCl_2$, 1% Triton X-100 and 0.2 mg/ml bovine serum albumin [BSA]). A protease inhibitor cocktail (Roche), 25 mM α-glycerophosphate and 10 μM Na_3VO_4 were all added to the lysis buffer. The lysate was incubated for 30 min at 4°C in rotation, and then centrifuged at 13,000 rpm for 20 min at 4°C. Proteins were quantified with the Pierce Coomassie (Bradford) kit (Thermo Scientific). Soluble cell extracts (1 mg protein) were incubated with protein A/G agarose beads (Millipore) previously coupled with 3 μg of the corresponding antibodies or control beads (Millipore) at 4°C for 16 h on rotation. The beads were washed 10 times with 1 mL lysis buffer (with protease inhibitors) and the immunoprecipitated proteins (IPs) were eluted by boiling the beads in SDS sample buffer (1% SDS and 2X loading buffer). Both the lysate (10%) and the IPs were analyzed by western blot using specific antibodies (Key Resources Table).

siRNA Knockdowns

siRNA knockdown of GTF3C5, GTF3C1 and ADNP

Dharmacon D-020031-02, LQ-012581-00-0002, LQ-012857-01-0002 siGENOME against human *GTF3C5*, *GTF3C1*, *ADNP* siRNA and D-001206-14 siGENOME Non-Targeting siRNA Pool #2 were used to carry out TFIIIC and ADNP knockdown in T47D cells. Cells were seeded in the absence of antibiotics and culture for 16 h prior to transfection with lipofectamine (Lipofectamine 2000, Invitrogen). siRNAs were used at 12.5 nM and cells were left in culture for 48 h in the presence of the siRNA. When required, cells were subjected to serum starvation for 16 h prior to further processing. The knockdown efficiency was evaluated by qRT-PCR (Primers in Key Resources Table), with efficiencies around 80% depletion in different experiments.

Cellular Biochemical Fractionation

Protein fractions from cytoplasm, nucleoplasm, and chromatin (as reported in Figure S2F) were obtained as described in (Wysocka et al., 2001). A total of $1 \times 10^7 - 2 \times 10^7$ cells were washed on plate with cold PBS and harvest in 1ml PBS using a cell scraper; cells were spun down at 1000 rpm for 2 min and the supernatant discarded. The remaining cell pellet was washed twice with PBS. (For each wash cells were resuspended in 1ml cold PBS added of protease inhibitor cocktail cOmplete EDTA-free (Roche), and spun down at 1000 rpm for 2 min). Cell pellet was then resuspended in 200 μl of Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol, 1 mM DTT and protease inhibitor cocktail cOmplete EDTA-free). Triton X-100 was then added to the cell pellet to a final concentration of 0.1% and incubated on ice for 8 min. After incubation, the resuspended cell pellet was centrifuged at 1,300 x g at 4°C, for 5 min; The resulting supernatant (named fraction S1) was separate from the pellet (corresponded to nuclei and named P1). S1 fraction was then clarified by high-speed centrifugation at 20,000 x g at 4°C, for 5 min; the resulting supernatant was collected and named fraction S2 (soluble cytoplasmic extract). The pellet (named P2) was discarded. P1 fraction was then washed once with 200 μl of Buffer A, spun down at 1000 rpm at 4°C for 2 min and lysed for 30 min in 100 μl of Buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT and protease inhibitor cocktail cOmplete EDTA-free). The resulting lysate was then centrifuged at 1,700 x g at 4°C, for 5 min. The supernatant corresponding to the soluble nuclear extract was collected (and named fraction S3) and separated from the pellet (insoluble chromatin) named fraction P3. P3 was washed one more time with 100 μl of Buffer B, centrifuged at 1,700 x g at 4°C, for 5 min and the resulting pellet (after discarding the supernatant) was resuspend in 50 μl SDS sample buffer (62.5 mM Tris-HCl pH 6.8, 2.5% SDS, 0.002% Bromophenol Blue, 0.7135 M β-mercaptoethanol and 10% glycerol) and boiled for 10 min at 70°C. The Micro BCA Protein Assay Kit was used to quantify fractions S1 and S3. 30 μg of each S1 and S3 fractions and 5 μl of the P3 fraction were loaded for SDS-PAGE/Western Blot analysis (Figure S2F). Fraction S3 and P3 were combined and loaded as a unique fraction to check GTF3C1 and GTF3C5 knock down experiments (Figures S3H and S3J) and total levels of TFIIIC subunits (Figure S5E). Pre-casted gels (4%–12% polyacrylamide) were used for all SDS-PAGE/Western analysis at room temperature (100V for 90 min). Proteins transfer to nitrocellulose membranes 0.45 μm was carried out using MOPS transferring buffer (Key Resources Table) and X Cell SureLock caste system at 4°C (100V for 90 min). Membrane blocking was carried out using Skim Milk Powder (5% in TBS-T) at room temperature for 1 h. For western blots primary antibodies were used at 1:50 to 1:1000 dilution and incubated overnight at 4°C followed by 1 h incubation with horseradish peroxidase conjugated anti-mouse (NA931V) or anti-rabbit (NA934V, Amersham) and blots were developed using ECL prime western blotting detection reagent (RPN2232, GE Healthcare) according to the manufacturer instructions.

Gene Ontology (GO) Analysis

DAVID (https://david.ncifcrf.gov) (Dennis et al., 2003) and ChIP-seq Enrich (http://chip-enrich.med.umich.edu) (Welch et al., 2014) were both used with default parameters to detect GO terms enrichments and association of peaks with genes, respectively. Lists of genome coordinates derived from downstream analysis of ChIP-seq data were analyzed with ChIP-seq Enrich. Parameters used: locus definition (Nearest TSS), Enrichment Method (Chip-Enrich), filter (2000 genes), adjust for the mappability of the gene locus regions (False). Gene symbols lists resulted from ChIP-seq Enrich analysis and RNA-seq downstream analysis were used as input for DAVID GO (default parameters) to generate Figure S1D and Figure S4K. For all DAVID GO analysis the Bonferroni-corrected p values were reported. ChIP-seq Enrich also generated a GO analysis that was used to corroborate DAVID GO terms enrichment reported in Figure S1D (data not shown).

External Data Sources

ChIP-seq of CTCF in T47D in the absence of serum was taken from GEO: GSE53463 (Le Dily et al., 2014). ChIP-seq of eGFP-ADNP in K562 cells was from GEO: GSE105573 (Consortium, 2012). ChIP-seq of mouse Adnp was taken from GEO: GSE97945 (Ostapcuk et al., 2018). T47D ATAC-seq in normal growth conditions (+S) was taken from GEO: GSM2241147, GEO: GSM2241148 (Toska et al., 2017). ChIP-seq of p300 in T98G cells grown in normal conditions or in the absence of serum was from GEO: GSE21026 (Ramos et al., 2010).

Bedtools

Bed intersection was carried out using bedtools (Quinlan and Hall, 2010) "intersectBed" function with default parameters of 1-bp overlap. Graphic representation of Venn diagrams has been obtained with R graphic, using R-studio (https://www.rstudio.com).

Hi-C

In situ *Hi-C library preparation*

In situ Hi-C was performed as previously described (Rao et al., 2014) with the following modifications: (i) $2x10^6$ cells were used as starting material; (ii) chromatin was initially digested with 100 U MboI (New England BioLabs) for 2 h, and then another 100 U (2 h incubation) and a final 100 U were added before overnight incubation; (iii) before fill-in with bio-dATP, nuclei were pelleted and re-suspended in fresh 1 × NEB2 buffer; (iv) ligation was performed overnight at 24 °C with 10,000 cohesive end units per reaction; (v) de-cross-linked and purified DNA was sonicated to an average size of 300-400 bp with a Bioruptor Pico (Diagenode; seven cycles of 20 s on and 60 s off); (vi) DNA fragment-size selection was performed only after final library amplification; (vii) library preparation was performed with an NEBNext DNA Library Prep Kit (New England BioLabs) with 3 μL NEBNext adaptor in the ligation step; (viii) libraries were amplified for 8-12 cycles with Herculase II Fusion DNA Polymerase (Agilent) and were purified/size-selected with Agencourt AMPure XP beads (> 200 bp). Hi-C library quality was assessed through *Cla*I digestion and low-coverage sequencing on an Illumina NextSeq500 instrument, after which every technical replicate (n = 2) of each biological replicate (n = 2) was sequenced at high coverage on an Illumina HiSeq2500 instrument. Data from technical replicates were pooled for downstream analysis. We sequenced > 18 billion reads in total to obtain 0.78–1.21 billion valid interactions per time point per biological replicate.

In Situ Hi-C Data Processing and Normalization

Hi-C data were processed by using an in-house pipeline based on TADbit and OneD algorithms (Serra et al., 2017; Vidal et al., 2018). First, the quality of the reads was checked with FastQC to discard problematic samples and detect systematic artifacts. Trimmomatic (Bolger et al., 2014) with the recommended parameters for PE reads was used to remove adaptor sequences and poor-quality reads (ILLUMINACLIP: TruSeq3-PE.fa:2:30:12:1:true; LEADING:3; TRAILING:3; MAXINFO:targetLength:0.999; and MINLEN:36).

For mapping, a fragment-based strategy implemented in TADbit was used, which was similar to previously published protocols (Ay et al., 2015). Briefly, each side of the sequenced read was mapped in full length to GRCh38. After this step, if a read was not uniquely mapped, we assumed that the read was chimeric, owing to ligation of several DNA fragments. We next searched for ligation sites, discarding those reads in which no ligation site was found. The remaining reads were split as often as ligation sites were found. Individual split read fragments were then mapped independently. These steps were repeated for each read in the input FASTQ files. Multiple fragments from a single uniquely mapped read resulted in a number of contacts identical to the number of possible pairs between the fragments. For example, if a single read was mapped through three fragments, a total of three contacts (all-versus-all) was represented in the final contact matrix. We used the TADbit filtering module to remove non-informative contacts and to create contact matrices. The different categories of filtered reads applied were:

1. Self-circle: reads coming from a single restriction enzyme (REnz) fragment and pointing to the outside.
2. Dangling end: reads coming from a single REnz fragment and pointing to the inside.
3. Error: reads coming from a single REnz fragment and pointing in the same direction.
4. Extra dangling end: reads coming from different REnz fragments but that were sufficiently close and point to the inside; the distance threshold used was left to 500 bp (default), which was between percentiles 95 and 99 of average fragment lengths.
5. Duplicated: the combination of the start positions and directions of the reads was repeated, thus suggesting a PCR artifact; this filter removed only extra copies of the original pair.
6. Random breaks: the start position of one of the reads was too far from REnz cutting site, possibly because of non-canonical enzymatic activity or random physical breaks; the threshold was set to 750 bp (default), > percentile 99.9.

From the resulting contact matrices, low-quality bins (those presenting low contact numbers) were removed, as implemented in TADbit's 'filter columns' routine. A single round of ICE normalization (Imakaev et al., 2012), also known as 'vanilla' normalization (Rao et al., 2014), was performed. That is, each cell in the Hi-C matrix was divided by the product of the interactions in its columns and the interactions in its row. Finally, all matrices were corrected to achieve an average content of one interaction per cell.

Identification of Sub-nuclear Compartments and TADs

To segment the genome into A/B compartments, normalized Hi-C matrices at 100-kb resolution were corrected for decay as previously described, by grouping diagonals when the signal-to-noise ratio was below 0.05 (Rao et al., 2014). Corrected matrices were then split into chromosomal matrices and transformed into correlation matrices by using the Pearson product-moment correlation.

Normalized contacts matrices at 20-kb resolution were used to define TADs, and for visualization purposes, through a previously described method with default parameters (Crane et al., 2015; Giorgetti et al., 2016). First, for each bin, an insulation index was obtained on the basis of the number of contacts between bins on each side of a given bin. Differences in the insulation index between both sides of the bin were computed, and borders were called, searching for minima within the insulation index. The insulation score of each border was determined as previously described (Crane et al., 2015), by using the difference in the delta vector between the local maximum to the left and the local minimum to the right of the boundary bin. This procedure resulted in a set of borders for each time point and replicate. To obtain a set of consensus borders along the time course, we proceeded in two steps: (i) merging borders of replicates and overlapping merged borders (that is, for each pair of replicates, we expanded the borders one bin on each side and

kept only those borders present in both replicates as merged borders) and (ii) further expanding two extra bins (100 kb) on each side and determining the overlap to obtain a consensus set of borders common to any pair of time points.

Identification of Intra-TAD contacts

Using Hi-C matrices at 5 kb resolution, we focused on TADs containing TFIIIC. Each bin was labeled according to TFIIIC and CTCF occupancy as well as to gene promoter annotation. Genes were classified according to mRNA-seq results as ''changed'' or ''unchanged'' by serum starvation. We marked as ''others'' bins not overlapping any of the corresponding categories.

The non-observed contacts were gathered within the TADs from the different types of bins and computed expected contacts frequencies based on the genomic distance that separate each pair (the expected distance decay was calculated excluding entries outside TADs). We used the log2 of the ratio on observed over expected as a contact score, and summarized the results via linear mixed-effects models fitted using lme4 (Ziyatdinov et al., 2018). TADs were considered as a random effect and the aforementioned bin categories (and their interactions) as fixed effects.

We fitted linear mixed models using lmer function of lme4 R package (Bates et al., 2015) (and computed the posterior probabilities using sim function of arm R package (Su, 2016) for the corresponding interaction terms.

MICROSCOPY

Immunofluorescence

Cells were washed with PBS twice and fixated with 4% paraformaldehyde (PFA) in PBS for 10 min at room temperature and permeabilized with 0.2% Triton X-100 in PBS for 10 min at room temperature. After washing, the cells were blocked with 5% BSA - 0.1% Triton X-100 in PBS for 1 h at room temperature. The incubation with primary antibodies diluted in 5% BSA - 0.1% Triton X-100 in PBS proceeded for 1 h at room temperature, followed by incubation with fluorophore-labeled secondary antibodies (anti-rabbit Alexa-594). After repeated washes with PBS, the cells were incubated with 4′,6-diamidino-2-phenylindole (DAPI) in PBS for 30 s and mounted with Mowiol (Sigma).

Fluorescence images were acquired with a Leica (DM 6000B).

CRISPR/CASP9

Generation of Alu-deleted T47D cells by CRISPR/Casp9

The generation of the targeting vector was carried out by DECKO2 cloning (Aparicio-Prat et al., 2015), using the pDECKO-mCherry (Addgene #78534) as backbone, and the oligonucleotides 1-6 included in the Key Resources Table. Oligonucleotides C542F and C542R, were used for colony screen. For the amplification of the constant part oligonucleotides C557F and C557R were used. The design of the gRNAs was done with CRISPETA (crispeta.crg.eu).

For clone selection, T47D Cas9-expressing cells were transfected with the resulting plasmid of the Decko2 cloning using Lipofectamine 3000 in a ratio 1:3, following manufacturer's instructions. The day after the transfection, the BFP^{+}/$cherry^{+}$ cells were sorted using a FACSaria Cell Sorter at the UPF Cytometry Facility. Cells were single-plated in 96-well plates and allowed for growth till enough number for further processing. For the identification of the clones, genomic DNA was prepared and analyzed by PCR for the presence of the deletion with primers Alu_up3 and Alu_down (expected size of the WT allele = 509 nt; deleted allele = 156 nt).

PROTEOMICS

RIME (Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins) and mass spectrometry analysis

For the identification of affinity-purified proteins associated to chromatin, the RIME procedure was used. The protocol is an adaptation of previous publications (Mohammed et al., 2016) to our model system, and of previous lab experiments from using ChIP-seq, to be able to match results from both protocols. It was also adapted for obtaining a broader set of interactors, which thanks to several replicates and time points end up with a big amount of high confidence interactors.

For extract preparation, cells were cross-linked with PFA for 8 min (Vicent et al., 2014), and stopped by adding a final 200 mM glycine and incubating for 5 min at room temperature. Plates were kept on ice and washed twice with cold phosphate-buffered saline (PBS). Cells were scrapped in ice-cold PBS with protease inhibitors and collected in a 15 mL tube suitable for sonication (BD Polyestyrene, 352095). Cells were centrifuged at 4,000 rpm for 5 min and washed twice with cold PBS. All buffers contain freshly added inhibitors in the following concentration: cOmplete EDTA-free as recommended by manufacturer (1 tablet for 50mL), 10 μM phenylmethylsulfonyl fluoride and 10 μM Na_3VO_4. Cell pellets were resuspended in 10 mL of lysis buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100), and incubated on ice for 10 min. After centrifugation, the pellet was resuspended in 10 mL of lysis buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) and incubated on rotation for 5 min at 4°C and centrifuged again. The pellet was finally resuspended in 400 μl of lysis buffer 3 (10 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% N-lauroylsarcosine) by carefully pipetting up and down for ten times. The extracts were sonicated in a Bioruptor (Diagnode) at 4°C, for 9 cycles of 30 s on / 30 s off, at high output. After sonication, the sample was transferred to a 1.7 mL siliconized tube, and 10% Triton X-100 was added. Lysates were centrifuged and supernatant was added to the antibody-conjugated beads.

Antibody binding to the beads was done typically, for 10^7 cells, with 100 μl of Protein A magnetic beads washed once in PBS, resuspended in 500 μl of LB3, with the appropriate amount of antibody or IgG (6 μl of anti-GTF3C2, at 0.2 μg/μl, or 12 μL of rabbit IgG),

incubated for 3 h at 4°C and washed with LB3 twice, 500 μl each. After overnight incubation, the beads were washed 10 times with RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.5% Na-deoxycholate, 1% NP-40, 0.1% SDS) and 2 times with 100 mM ammonium hydrogen carbonate (AMBIC) solution. For the second wash, the beads were transferred to new 1.7 mL tubes.

Mass Spectrometry Analysis

The proteomics analyses were performed at the CRG/UPF Proteomics Unit. The immunoprecipitated proteins were reduced by adding 10 μl of 10 mM DTT in 100 mM ammonium bicarbonate (ABC) buffer (1 h, 37°C) and alkylated by adding 10 μl of 20 mM iodocetamide in 100 mM ABC (30 min, room temperature, in the dark). The digestion was done in two steps: first, with 1 μg of endopeptidase LysC, incubated over night at 37°; second, 1 μg of sequencing grade trypsin was added and incubated for 8 h at 37°C. The digestion reaction was stopped with formic acid (5% final concentration). The supernatant was taken and tryptic peptides were desalted with C18 columns, dried in a Speed-vac and re-suspended in 10 μl 0.1% formic acid.

From the resuspended sample, 4.5 μl of each peptide mixture was analyzed using a LTQ-Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific, San Jose, USA) coupled to a nano-LC (Proxeon, Odense, Denmark) equipped with a reversed-phase chromatography 2-cm C18 pre-column (Acclaim PepMap-100, Thermo; 100 μm i.d., 5 μm), and a 25-cm C18 analytical column (Nikkyo Technos, 75 μm i.d., 3 μm). Chromatographic gradients started at 3% buffer B with a flow rate of 300 nL/min and gradually increased to 7% buffer B in 1 min and to 35% buffer B in 60 min. After each analysis, the column was washed for 10 min with 90% buffer B (Buffer A: 0.1% formic acid in water; Buffer B: 0.1% formic acid in acetonitrile). The mass spectrometer was operated in positive ionization mode with nanospray voltage set at 2.5 kV and source temperature at 200°C. Ultramark 1621 was used for external calibration of the FT mass analyzer prior the analyses. The background polysiloxane ion signal at m/z 445.1200 was used as lock mass. The instrument was operated in data-dependent acquisition mode, and full MS scans with 1 microscan at resolution of 60,000 were used over a mass range of m/z 350−1,500 with detection in the Orbitrap. Auto gain control (AGC) was set to 106, dynamic exclusion was set at 60 s, and the charge-state filter disqualifying singly charged peptides for fragmentation was activated. Following each survey scan, the 10 most intense ions with multiple charged ions above a threshold ion count of 5000 were selected for fragmentation at normalized collision energy of 35%. Fragment ion spectra produced via collision-induced dissociation were acquired in the linear ion trap, AGC was set to 3·104 and isolation window of 2.0 m/z, activation time of 30 ms, and maximum injection time of 250 ms were used. All data were acquired with Xcalibur software v2.2.

Acquired data were analyzed using the Proteome Discoverer software suite (v1.4, Thermo Fisher Scientific), and the Mascot search engine (v2.5, Matrix Science) was used for peptide identification. Data were searched against the human protein database derived from the SwissProt database plus common contaminants (April 2016; 20,200 sequences). A precursor ion mass tolerance of 7 ppm was used, and up to three missed cleavages were allowed. The fragment ion mass tolerance was set to 0.5 Da, and oxidation (M), and acetylation (Protein N-term) were defined as variable modifications, whereas carbamidomethylation (C) was set as fixed modification. The identified peptides were filtered by FDR < 0.01 (1%).

For the assessment of protein-protein interactors, we used the *S*ignificance *A*nalysis of *INT*eractome **(**SAINT) software (Choi et al., 2011), using rabbit IgG as negative controls. Experiments were performed with samples in triplicate (see Table S1 for results).

Analysis of Breast Cancer Tumor Samples

Kaplan–Meier plots of breast tumor samples were generated at kmplot.com, and analyzed with a Mantel-Cox test (Győrffy et al., 2010). Plots were generated using ''Gene expression-based Outcome for Breast Cancer Online'' (GOBO; co.bmc.lu.se/gobo/) (Ringnér et al., 2011).

DATA AND CODE AVAILABILITY

The accession numbers for the raw sequencing and mass spectrometry data reported in this paper are NCBI GEO: GSE120162 and PRIDE (https://www.ebi.ac.uk/pride/archive/): PXD011250. Original western blots and Coomassie gels were deposited in Mendeley Data and are available at DOI: http://dx.doi.org/10.17632/mzjf96t3gc.5. Custom scripts for data analysis are available upon request, other tools used are indicated in the Key Resources Table and the respective STAR Methods sections. Processed data used for analyses in this manuscript are included as Tables S1, S2, and S3.

Molecular Cell, Volume *77*

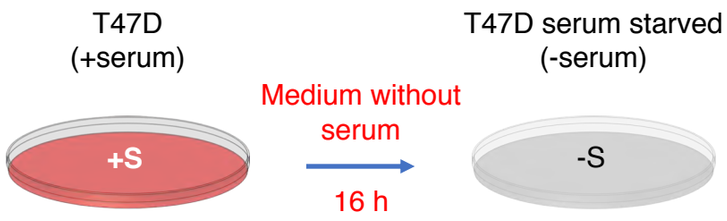
Supplemental Information

TFIIIC Binding to Alu Elements Controls Gene Expression via Chromatin Looping and Histone Acetylation

Roberto Ferrari, Lara Isabel de Llobet Cucalon, Chiara Di Vona, François Le Dilly, Enrique Vidal, Antonios Lioutas, Javier Quilez Oliete, Laura Jochem, Erin Cutts, Giorgio Dieci, Alessandro Vannini, Martin Teichmann, Susana de la Luna, and Miguel Beato

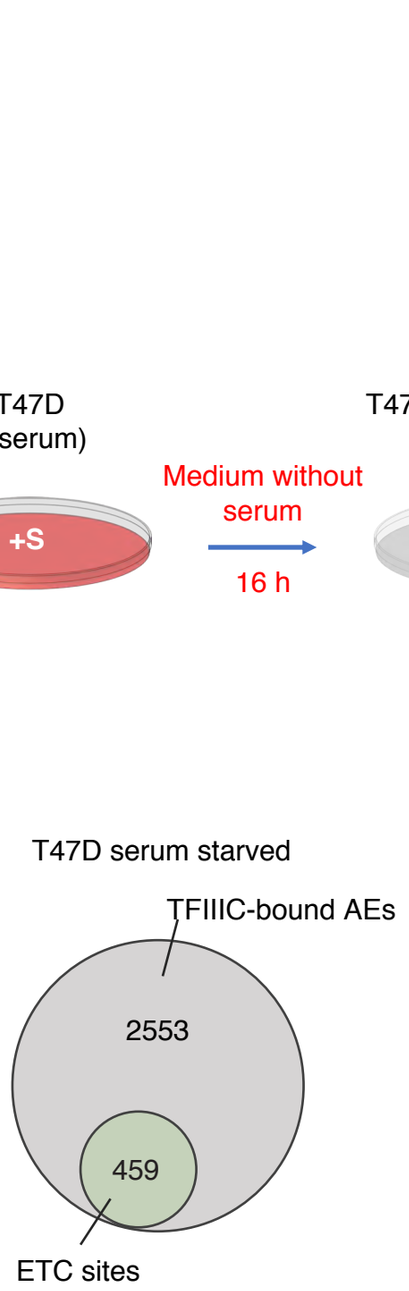
Fig. S1

A

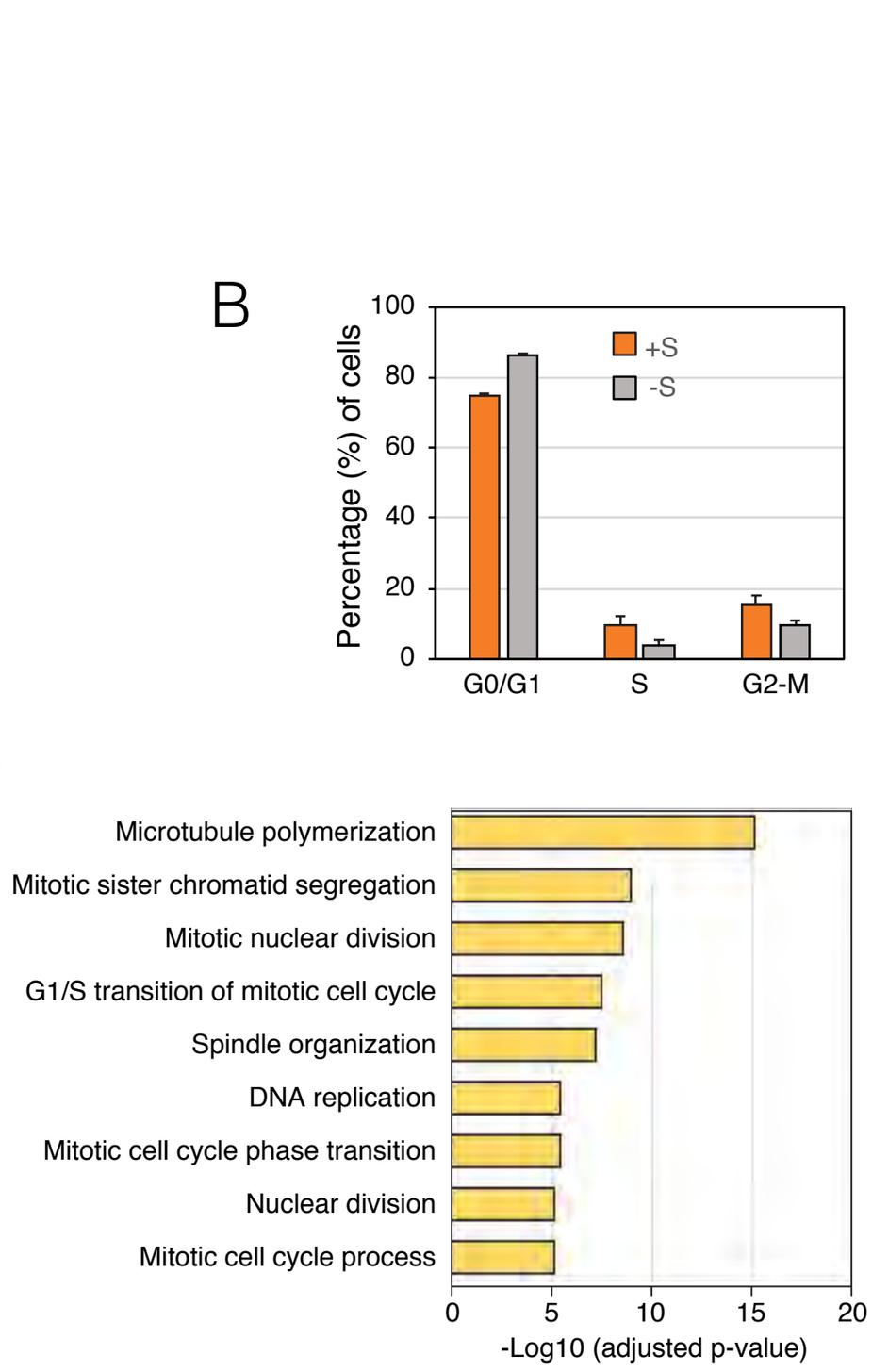


B

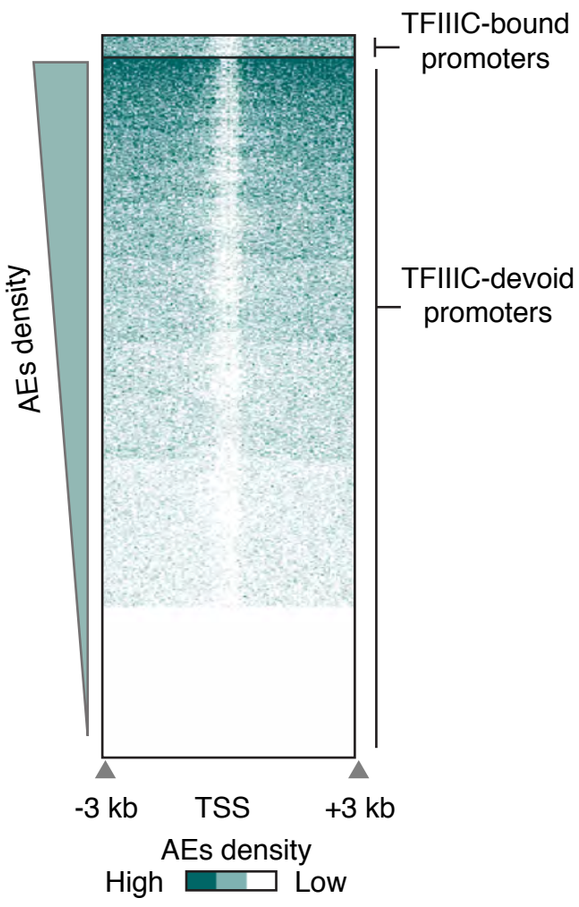
C



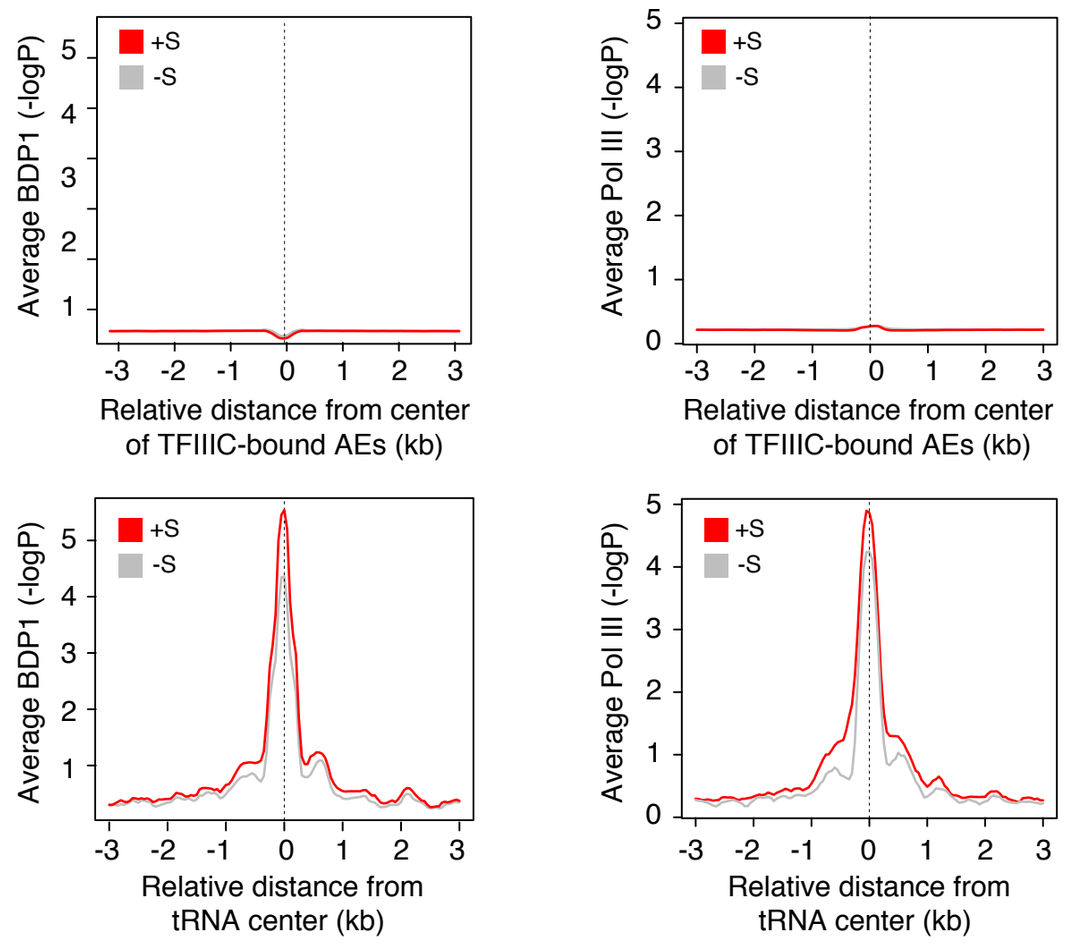
D



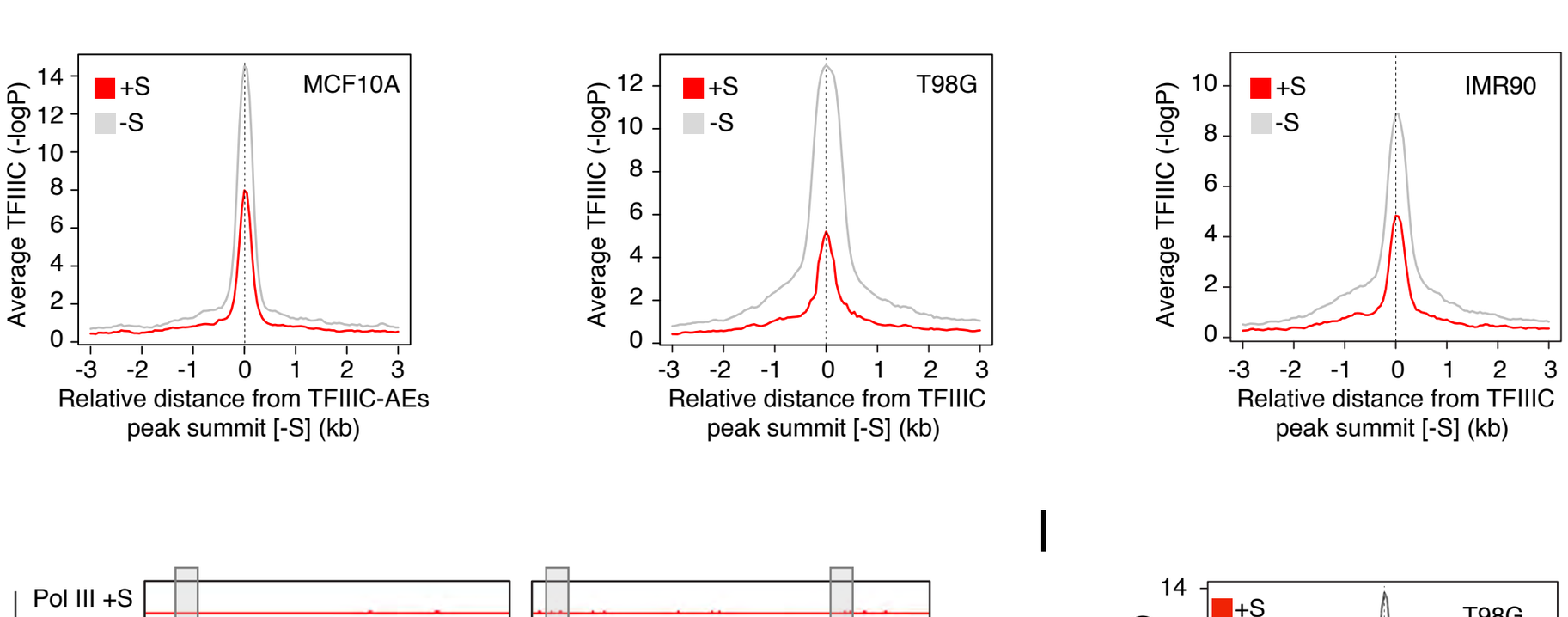
E



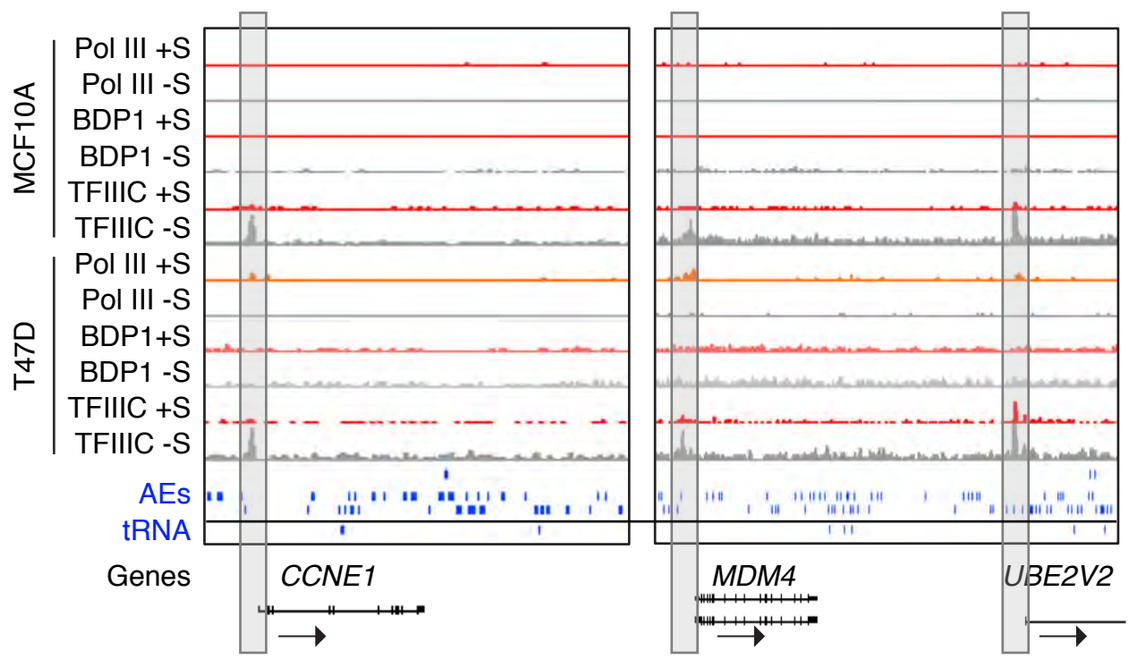
F



G



H



I

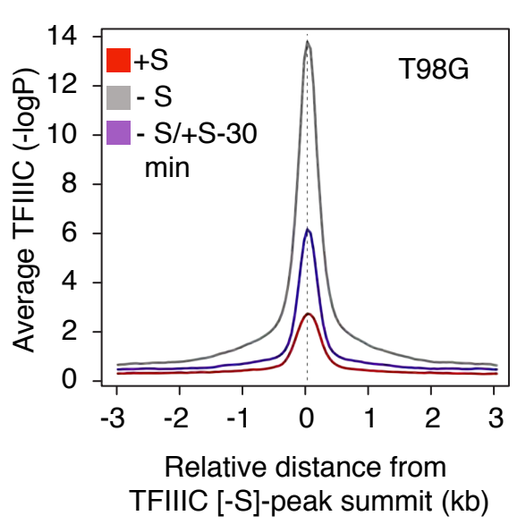


Figure S1. Tumorigenic and normal cell lines show TFIIIC (but not TFIIIB nor Pol III) binding to AEs close to Pol II promoters of cell cycle-related genes upon serum starvation (SS), related to Figure 1.

(**A**) Schematic view of the experimental design.

(**B**) Cell cycle profile of T47D cells grown in normal conditions (+S, orange) or upon 16 h of SS (–S grey). Note that T47D cells show a cell cycle profile highly enriched in the G1 population in the normal growth conditions; although a slight increase in the proportion of cells in G1 phase with a concomitant reduction in the S and G2/M phases was detected following serum depletion, it was not statistically significant (Chi-square test, p-value = 0.09).

(**C**) Proportional Venn diagram of total AEs bound by TFIIIC detected *vs* those containing ETC sites (with only B-box) in the serum-starved cells.

(**D**) Bar plots of gene ontology (GO) enrichment for Pol II genes with TFIIIC-bound within 5 kb of their TSSs) in the serum-starved cells. GO terms were calculated by DAVID (Molecular Function and Biological Processes combined) and ranked from the lowest to the highest p-value of the first nine terms found.

(**E**) Heatmap of AEs density across all human TSSs spanning a 6 kb-region and sorted by high to low AEs density. TFIIIC-associated promoters are shown at the top. Color bar scale with increasing shades of color stands for higher AEs density.

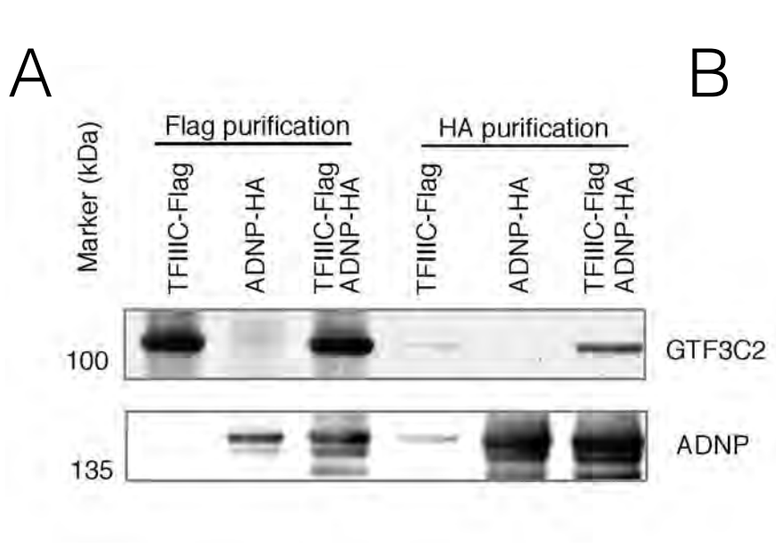
(**F)** CEAS plots of the average binding to AEs bound by TFIIIC (top panels) or to tDNAs bound by TFIIIC (bottom panels) for BDP1 (left) and Pol III/RPC39 (right) in T47D grown in normal condition (+S, red) or upon SS (-S, grey) (plotted is the –log10 of the Poisson p-value).

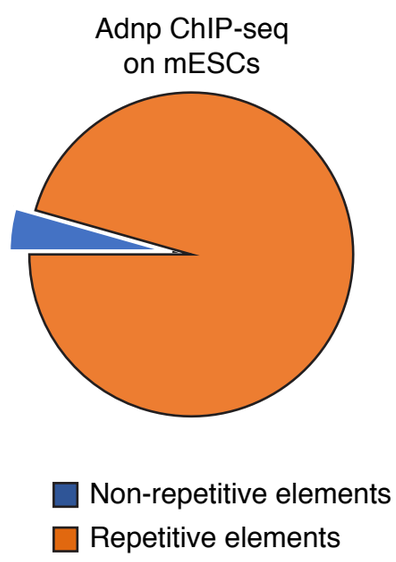
(**G**) CEAS plots of TFIIIC average binding for MCF10A, T98G and IMR90 cells in normal growth conditions (+S, red) and following SS (–S, grey, see Methods for specific conditions for each cell line). The graphs are plotted over the summit of TFIIIC peaks in the –S condition (plotted is the –log10 of the Poisson p-value). The enrichment in peaks corresponds to AEs (not shown).

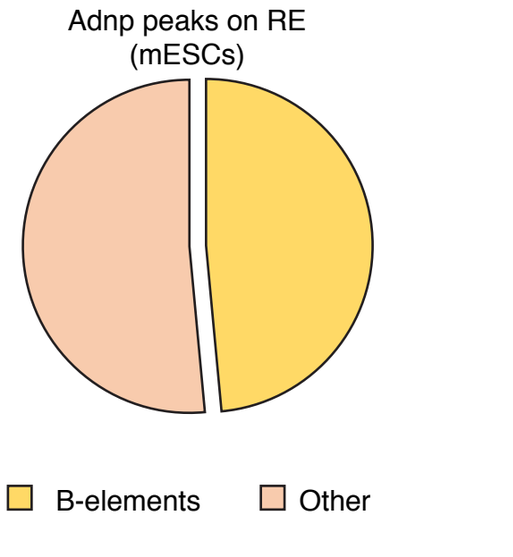
(**H**) Genome browser view of *CCNE1, MDM4 and UBE2V2* loci with ChIP-seq data for Pol III, BDP1 and TFIIIC in MCF10A and T47D breast cell lines. The graph includes the tracks for AEs and tDNAs. Highlighted in grey is the AE bound by TFIIIC close to the TSS of the indicated genes and arrow shows direction of transcription.

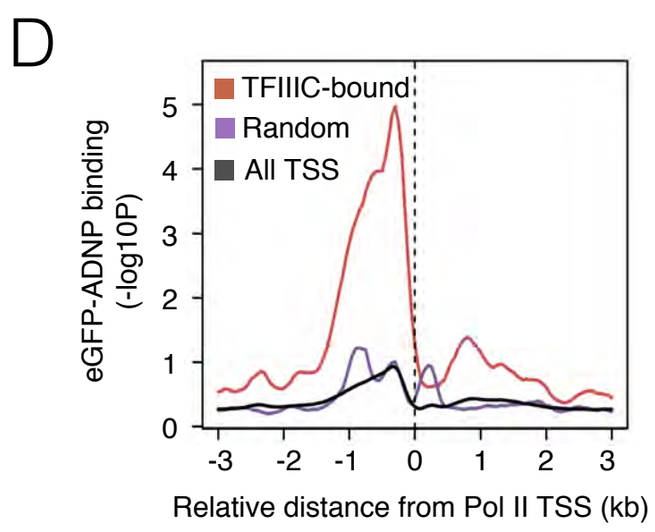
(**I**) CEAS plot of TFIIIC average enrichment for T98G cells grown in normal conditions (+S, red), SS (–S, grey), or SS followed by serum addition for 30 min (–S/+S-30 min, purple). The graphs are plotted over TFIIIC-peaks summit in the –S condition (plotted is the –log10 of the Poisson p-value).

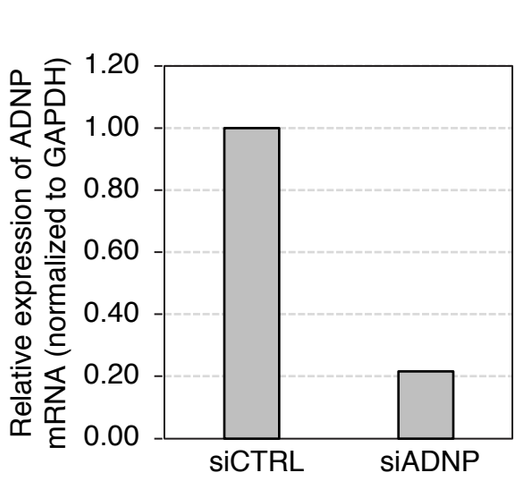
Fig. S2

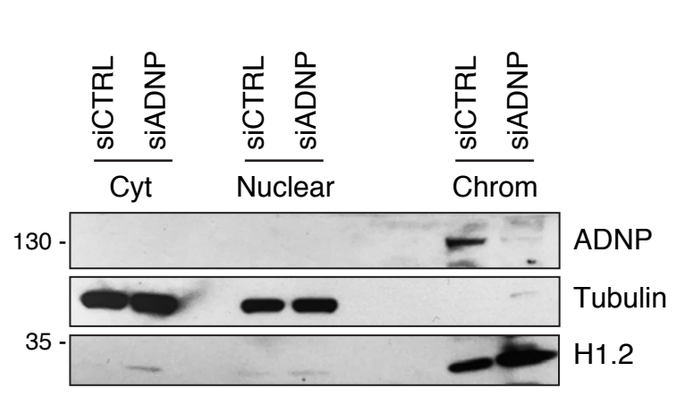

A
Flag purification
HA purification
Marker (kDa)
TFIIIC-Flag
ADNP-HA
TFIIIC-Flag ADNP-HA
TFIIIC-Flag
ADNP-HA
TFIIIC-Flag ADNP-HA
100
GTF3C2
135
ADNP


B
Adnp ChIP-seq on mESCs
Non-repetitive elements
Repetitive elements


C
Adnp peaks on RE (mESCs)
B-elements
Other


D
TFIIIC-bound
Random
All TSS
eGFP-ADNP binding (-log10P)
Relative distance from Pol II TSS (kb)


E
Relative expression of ADNP mRNA (normalized to GAPDH)
siCTRL
siADNP


F
siCTRL
siADNP
Cyt
Nuclear
Chrom
130
ADNP
Tubulin
35
H1.2

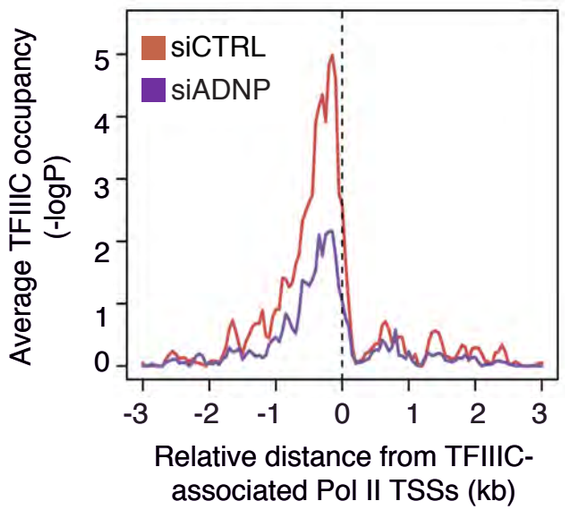

G
siCTRL
siADNP
Average TFIIIC occupancy (-logP)
Relative distance from TFIIIC-associated Pol II TSSs (kb)

Figure S2. ADNP depletion causes TFIIIC loss from TFIIIC-associated promoters upon serum starvation, Related to Figure 2.

(A) Western Blot analysis of tagged-immunopurification of TFIIIC and ADNP complexes as described in Figure 2B. FLAG- and HA-purified complexes were probed for the presence of TFIIIC (GTF3C2 subunit) and ADNP.

(**B**) Pie chart showing the percentage of mouse Adnp peaks found in mESCs belonging to repetitive elements and non-repetitive elements (analysis from GSE97945). Notice that almost all the binding of this factor lays on repetitive elements.

(**C**) Pie chart showing the percentage of Adnp peaks mapping to repetitive elements and belonging to B-elements or to other type of repetitive elements. Notice that more than 50% of Adnp peaks mapping to repetitive elements lay on B-elements.

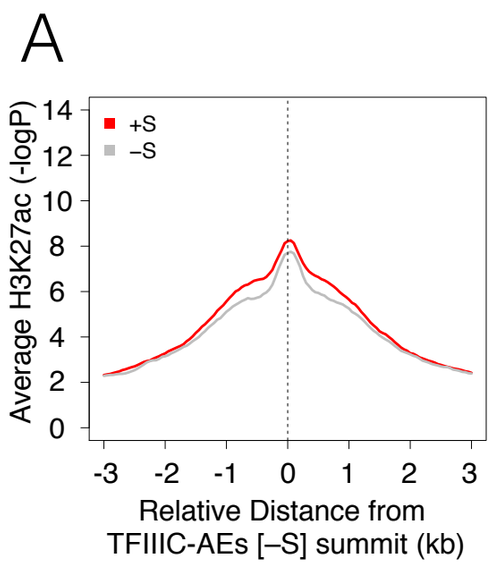
(**D**) Average plot for ADNP-eGFP (GSE105573) enrichment across TFIIIC-associated Pol II promoters (red) spanning a 6 kb-region. The profile of a random set of genes of the same size (purple), as well as the average for all human TSSs (black) is also shown.

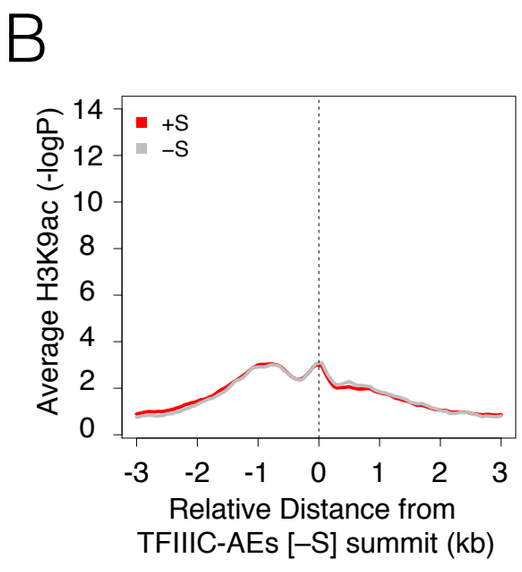
(**E**) qRT-PCR expression analysis of *ADNP* in serum-starved T47D cells (siCTRL and siADNP) corresponding to the experiment in Figure 2D. The value in siCTRL cells was arbitrarily set as 1. Note that the knockdown of *ADNP* reaches values almost 80% of its control.

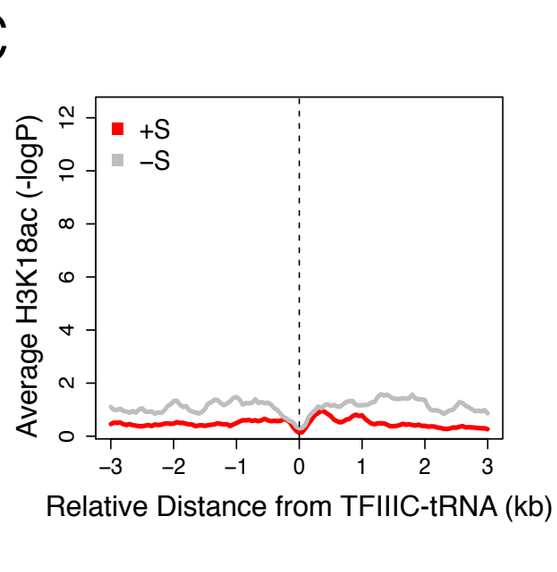
(**F**) Immunoblot probing the levels of ADNP protein across three cellular fractions: cytoplasm (Cyt), nucleoplasm (Nuclear) and chromatin (Chrom) in serum-starved T47D cells transfected with siADNP or control siCTRL. Note that the largest pool of ADNP is found in the chromatin fraction and is strongly reduced upon knock down. α-Tubulin and histone H1.2 are used as controls for the subcellular fractionation.

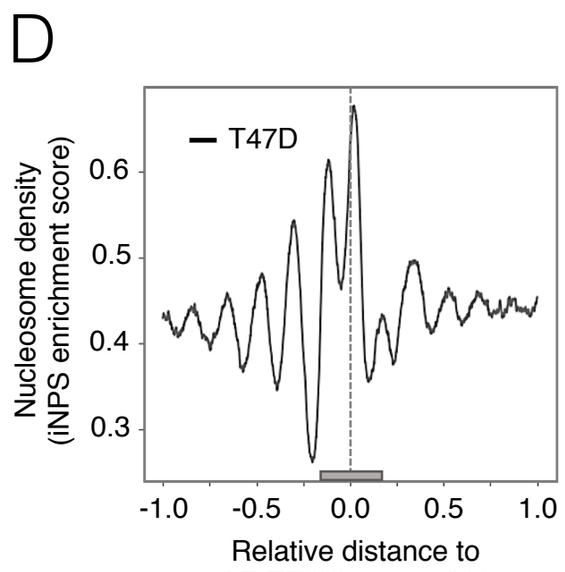
(**G**) CEAS profile of TFIIIC enrichment over TFIIIC-associated Pol II promoters upon depletion of ADNP (siADNP, purple) compared to control cells (siCTRL, red) in serum-starved T47D. Note the strong reduction in TFIIIC occupancy (from -logP ~ 5 to -logP ~ 2.5) in ADNP knocked down cells.

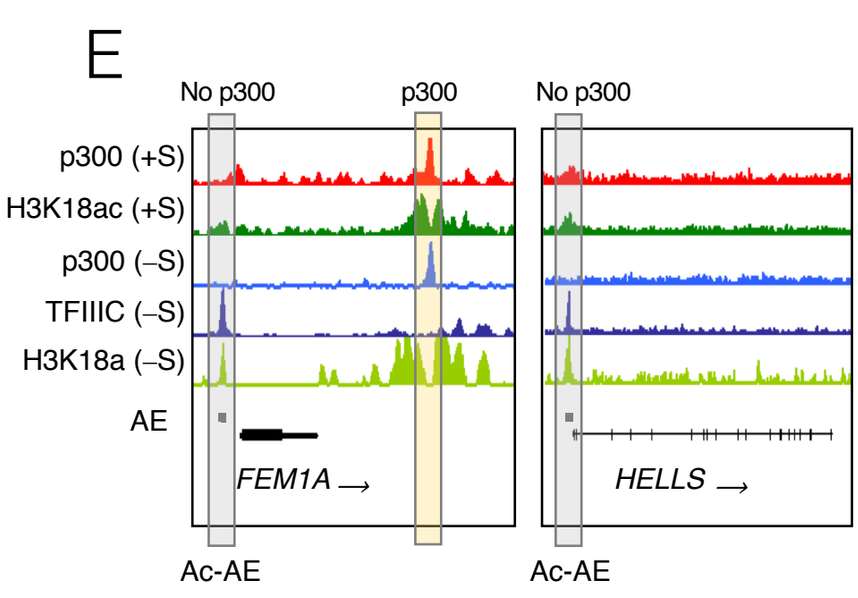
Fig. S3

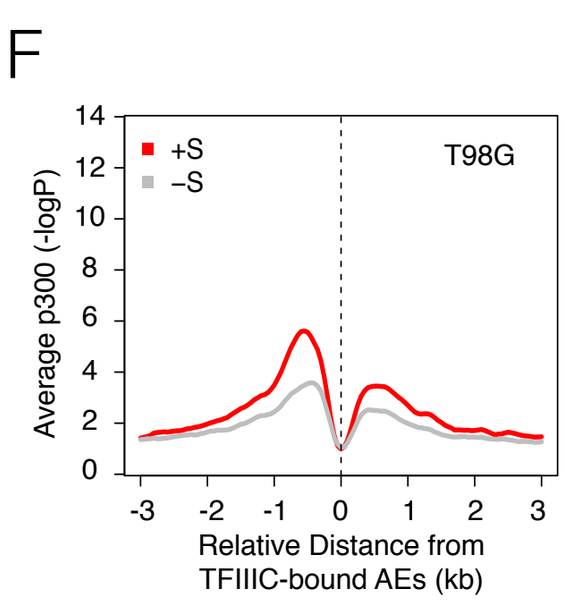

A
Average H3K27ac (-logP)
+S
–S
Relative Distance from
TFIIIC-AEs [–S] summit (kb)

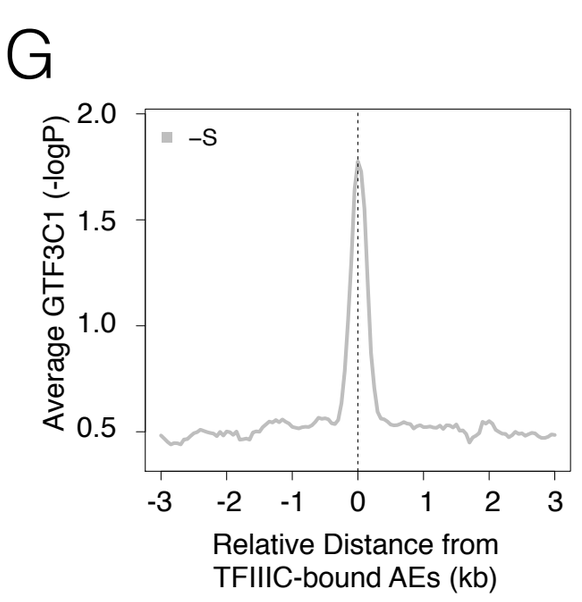

B
Average H3K9ac (-logP)
+S
–S
Relative Distance from
TFIIIC-AEs [–S] summit (kb)

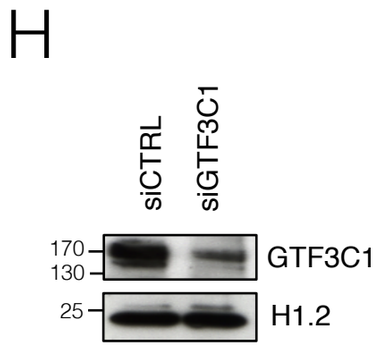

C
Average H3K18ac (-logP)
+S
–S
Relative Distance from TFIIIC-tRNA (kb)

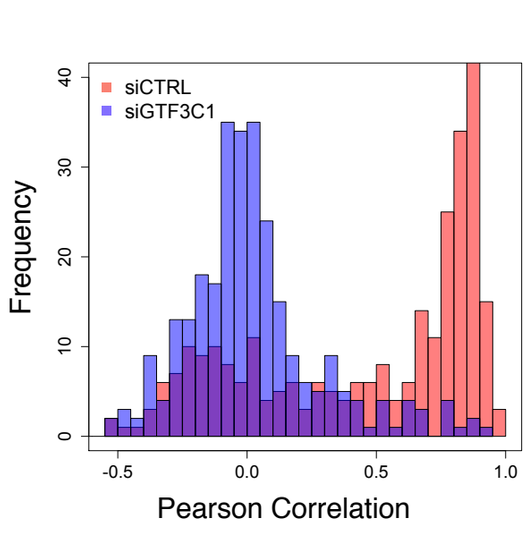

D
T47D
Nucleosome density
(iNPS enrichment score)
Relative distance to
AE-TFIIIC center (kb)

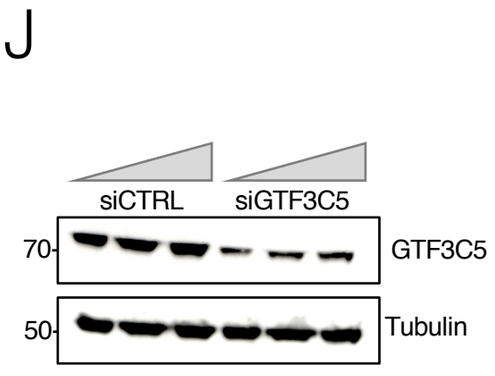

E
No p300
p300
No p300
p300 (+S)
H3K18ac (+S)
p300 (–S)
TFIIIC (–S)
H3K18a (–S)
AE
FEM1A
HELLS
Ac-AE
Ac-AE


F
Average p300 (-logP)
+S
–S
T98G
Relative Distance from
TFIIIC-bound AEs (kb)


G
Average GTF3C1 (-logP)
–S
Relative Distance from
TFIIIC-bound AEs (kb)


H
siCTRL
siGTF3C1
170
130
GTF3C1
25
H1.2


I
siCTRL
siGTF3C1
Frequency
Pearson Correlation


J
siCTRL
siGTF3C5
70
GTF3C5
50
Tubulin

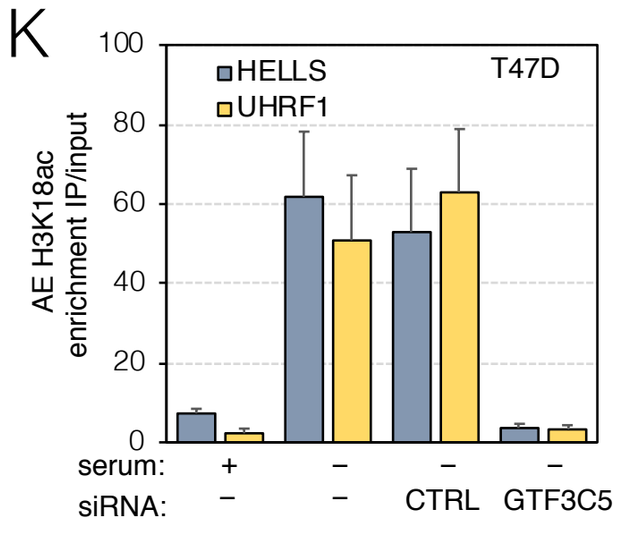

K
HELLS
UHRF1
T47D
AE H3K18ac
enrichment IP/input
serum:
siRNA:
CTRL
GTF3C5

Figure S3. Both H3K18ac and H3K27ac mark AEs occupied by TFIIIC, but only H3K18ac is changed upon serum starvation, Related to Figure 3.

(**A**) Sitepro profile of H3K27ac enrichment in T47D grown in the presence (+S, red) or absence of serum (–S, grey) over TFIIIC-bound AEs (plotted is the –log10 of the Poisson p-value). Note that the H3K27ac levels at these sites are independent of the growth conditions.

(**B**) Sitepro profile of H3K9ac enrichment in T47D grown in the presence (+S, red) or absence of serum (–S, grey) over TFIIIC-bound AEs (plotted is the –log10 of the Poisson p-value). Note the low levels of H3K9ac at these sites and their independence of the growth conditions.

(**C**) Sitepro profile of H3K18ac enrichment in T47D grown in the presence (+S, red) or absence (–S, grey) of serum at tDNAs (plotted is the –log10 of the Poisson p-value).

(**D**) Nucleosome density as mapped by MNase digestion across all TFIIIC-bound AEs in T47D cells using the iNPS score. The plot has been produced using NucMap (Zhao et al., 2019). Note the presence of at least one nucleosome over the region of the AE (represented as a grey rectangle).

(**E**) Genome browser view of representative TFIIIC-associated genes *FEM1A* and *HELLS* with ChIP-seq data for p300 and H3K18ac in T47D in the presence (+S) or absence (-S) of serum. The AE bound by TFIIIC in each locus is shown by a grey rectangle. The gene structure and the direction of transcription (arrow) are shown at the bottom. Note that p300 is not recruited at the AEs bound by TFIIIC (grey boxes) as it is for other adjacent intergenic regions (yellow box).

(**F**) Average plot for p300 occupancy across all TFIIIC-bound AEs spanning a 6 kb-region in T98G grown in the presence (+S, red) or absence (-S, grey) of serum. Data was from GSE21026.

(**G**) Average plot for GTF3C1 occupancy across all TFIIIC-bound AEs spanning a 6 kb-region in T47D grown in the absence of serum.

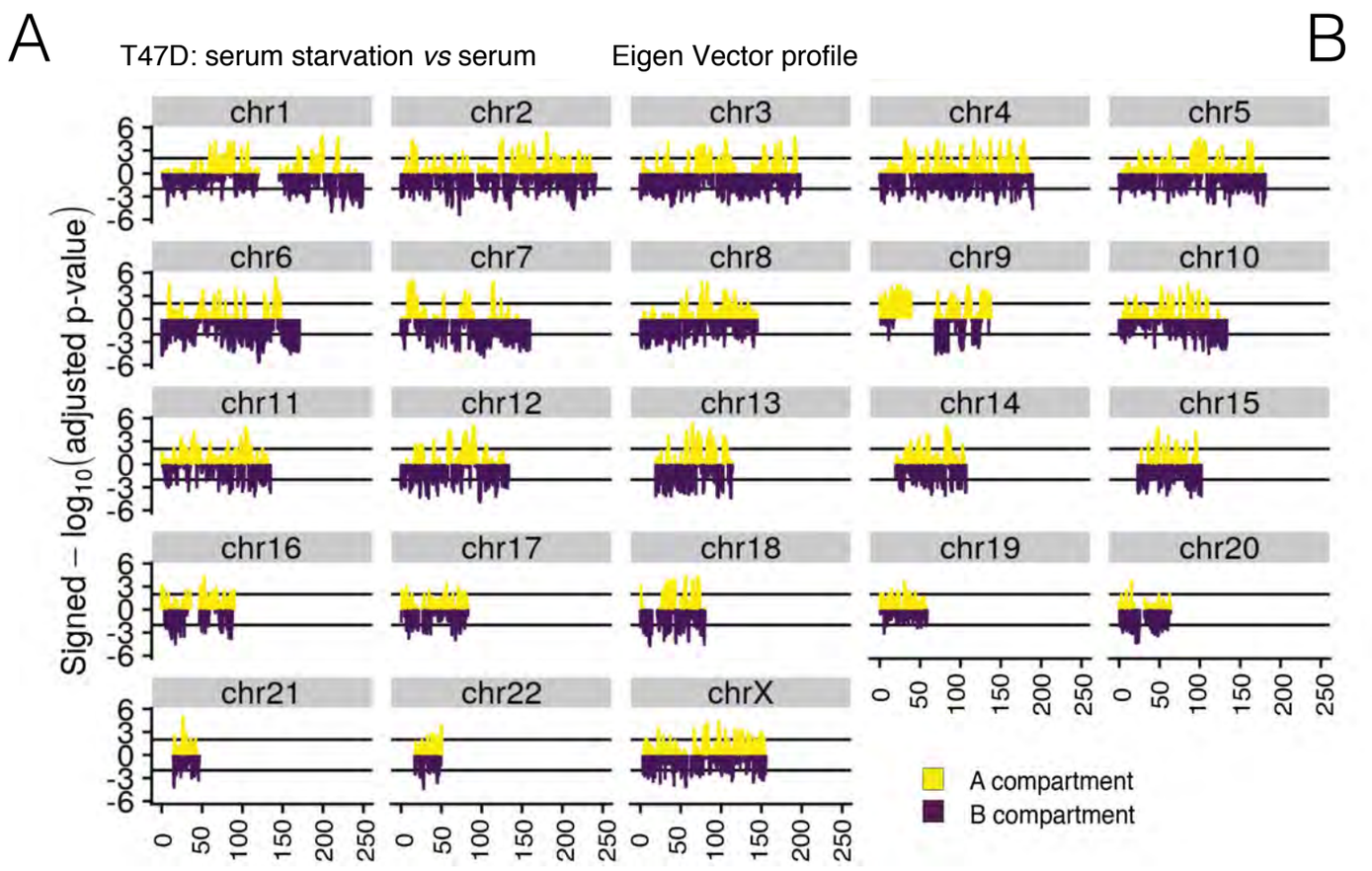
(**H**) Immunoblot probing GTF3C1 protein levels in serum-starved T47D cells transfected with siGTF3C1 or siCTRL as in Figure 3H. Histone H1.2 is shown as loading control.

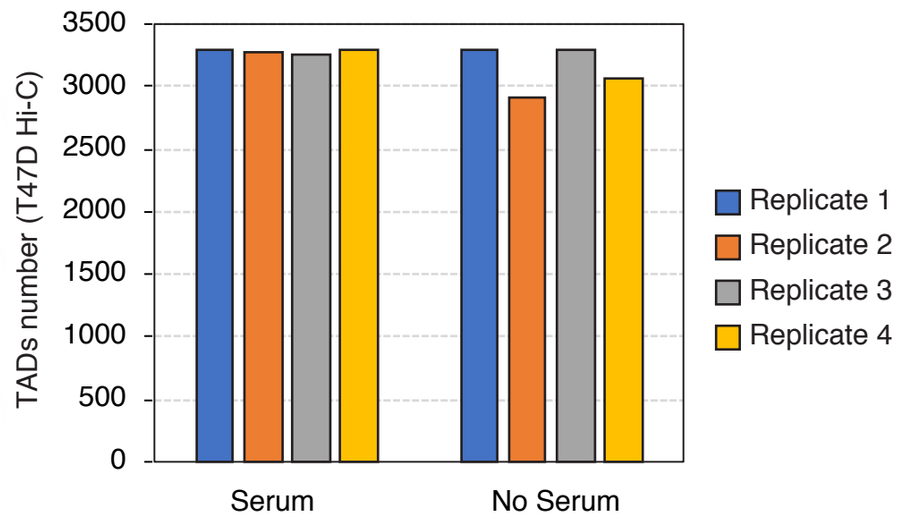
(**I**) Histogram plot of Pearson's correlation frequencies of H3K18ac colocalization with DAPI staining of Figure 3H. The large majority of cells in T47D siCTRL cells had H3K18ac colocalizing with DAPI, whereas cellular ablation of GTF3C1 caused the loss of H3K18ac and consequently its colocalization with DAPI.

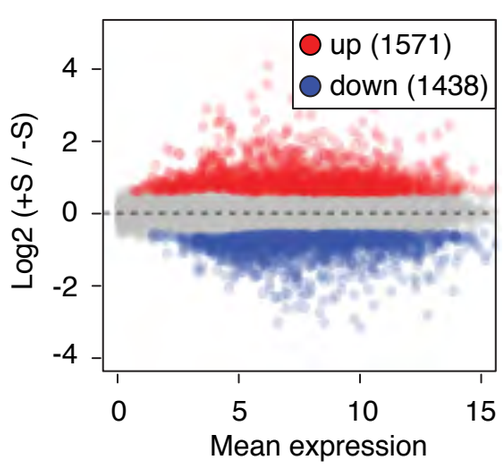
(**J**) Immunoblot probing GTF3C5 protein levels in T47D cells transfected with increasing concentration of siGTF3C5 or control siCTRL.

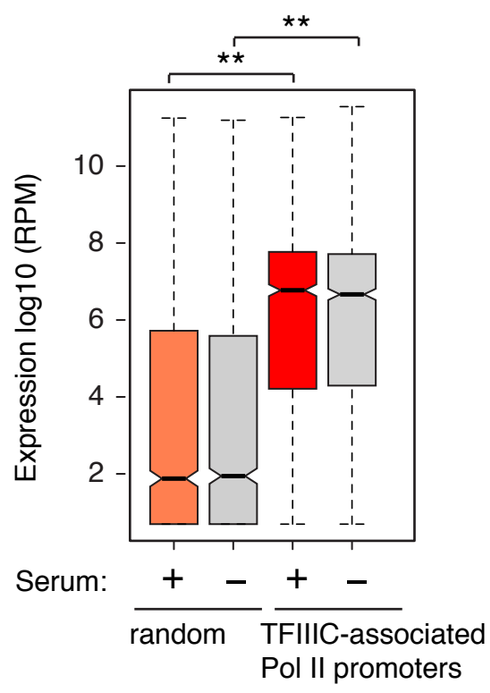
(**K**) ChIP-qPCR showing loss of H3K18ac enrichment at two AEs bound by TFIIIC (*UHRF1* and *HELLS* loci) in serum-starved T47D upon knock down of *GTF3C5* by siRNA. The graph shows mean and SD of 2 independent experiments. Depletion levels are shown in Figure S3J.

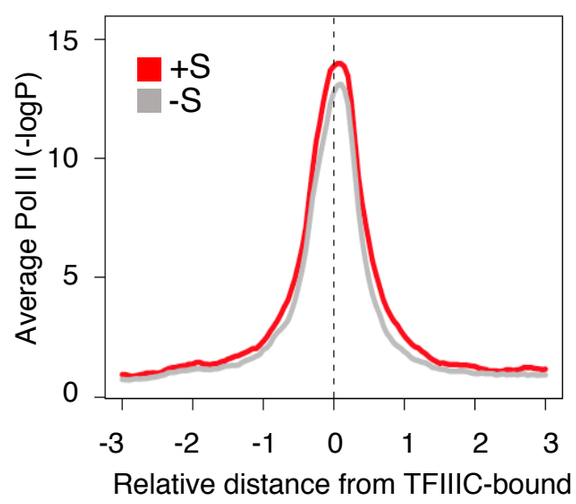
Fig. S4

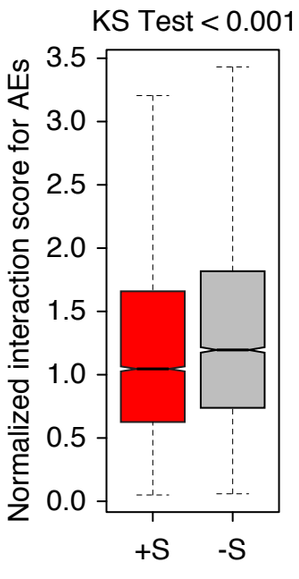

A
T47D: serum starvation vs serum
Eigen Vector profile
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
chrX
Signed − log10(adjusted p-value)
0
50
100
150
200
250
A compartment
B compartment

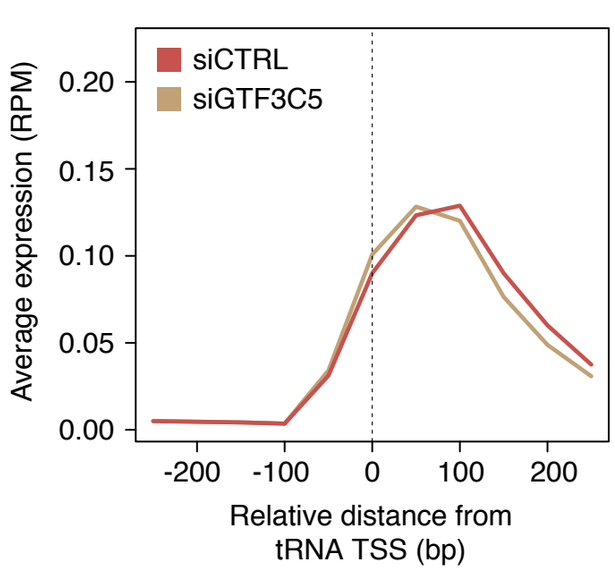

B
TADs number (T47D Hi-C)
3500
3000
2500
2000
1500
1000
500
0
Serum
No Serum
Replicate 1
Replicate 2
Replicate 3
Replicate 4

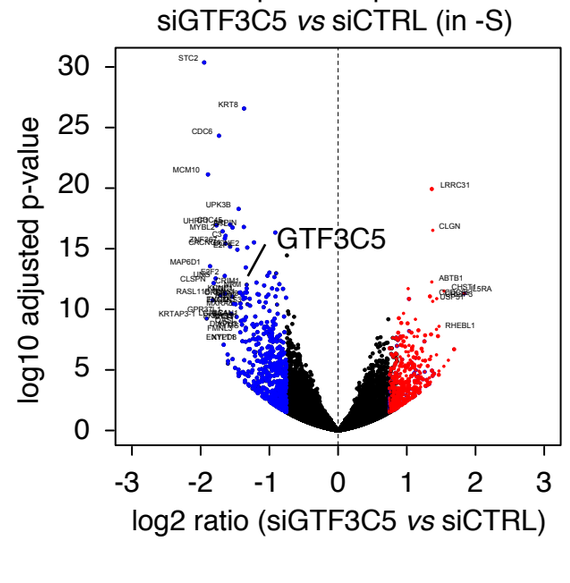

C
up (1571)
down (1438)
Log2 (+S / -S)
Mean expression

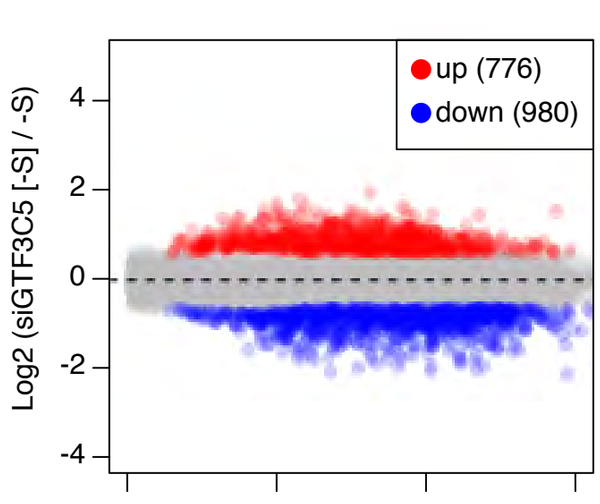

D
**
**
Expression log10 (RPM)
Serum:
+
−
+
−
random
TFIIIC-associated
Pol II promoters

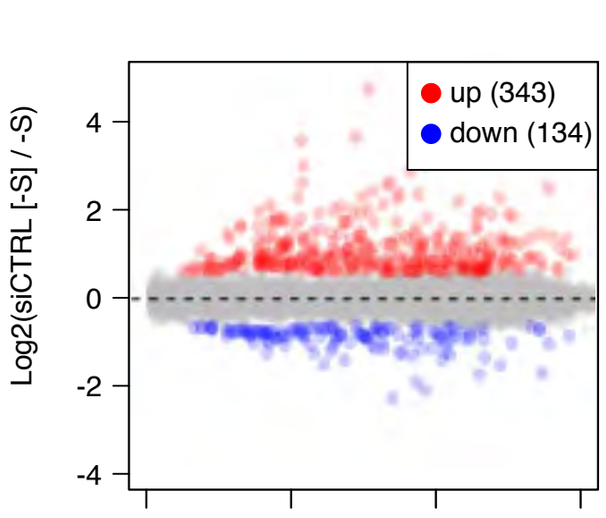

E
+S
-S
Average Pol II (-logP)
Relative distance from TFIIIC-bound
Pol II promoters (kb)

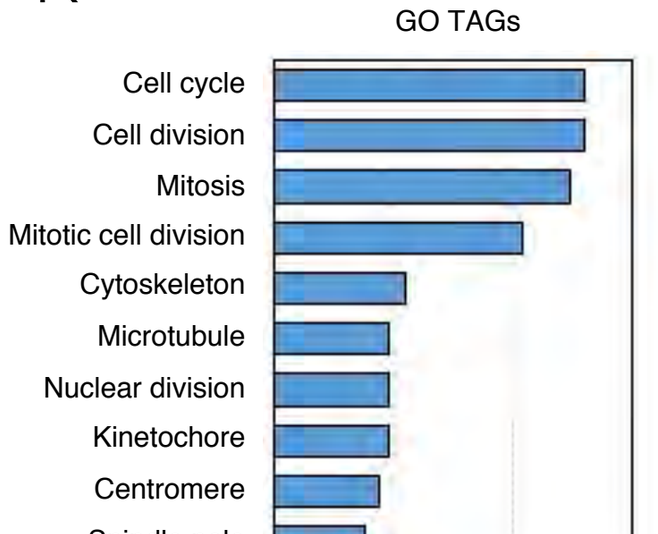

F
KS Test < 0.001
Normalized interaction score for AEs
+S
-S


G
siCTRL
siGTF3C5
Average expression (RPM)
Relative distance from
tRNA TSS (bp)


H
Gene expression profile T47D
siGTF3C5 vs siCTRL (in -S)
GTF3C5
log10 adjusted p-value
log2 ratio (siGTF3C5 vs siCTRL)


I
up (776)
down (980)
Log2 (siGTF3C5 [-S] / -S)
Mean across treatment


J
up (343)
down (134)
Log2(siCTRL [-S] / -S)
Mean across treatment


K
GO TAGs
Cell cycle
Cell division
Mitosis
Mitotic cell division
Cytoskeleton
Microtubule
Nuclear division
Kinetochore
Centromere
Spindle pole
-Log10 (adjusted p-value)

L
G1
S
G2/M
Percentage of cells
−S
siCTRL
siGTF3C5

Figure S4. Global gene expression analysis in T47D grown in normal conditions and after serum deprivation and dependence on TFIIIC, Related to Figure 4.

(**A**) Eigen Vector profile (-log10 adjusted p-value) for changes in A and B compartments in Hi-C experiments (4 replicates for each condition) across all chromosomes for T47D cells grown in the absence or presence of serum. Yellow represents switches to A compartment and dark purple represents switches to B compartment upon serum deprivation. Black horizontal bars represent threshold of p-val>0.001. Genomic positions are indicated at the bottom.

(**B**) Barplot representing the total number of TADs for Hi-C experiments in T47D grown in the presence or absence of serum (4 independent replicates for each condition). No significant differences were observed comparing replicates and conditions (Fisher exact test = 0.326).

(**C**) Scatter plot of Pol II gene expression in T47D cells grown in the presence (+S) or absence (–S) of serum (red, up-regulated in +S; blue, down-regulated in +S; grey, no changes). The number of genes up- or down-regulated in the +S condition (-1.5 < FC > 1.5; adjusted p-value < 0.05) is indicated.

(**D**) Boxplot of mRNA expression for T47D cells grown in the presence (+S, red) or absence (–S, grey) of serum for TFIIIC-associated Pol II promoters. A random dataset of the same size has been used as a control (**p-value < 0.001; two-tailed paired *t*-test). Notice that TFIIIC-associated promoters show higher level of expression compared to a random control. No significant changes were detected when the two growth conditions were compared (two-tailed paired *t*-test).

(**E**) CEAS profile of average total Pol II enrichment at TFIIIC-associated promoters for T47D cells grown in the presence (+S, red) or absence (-S, grey) of serum. The graphs are plotted over the TSS of the TFIIIC-associated promoters in the –S condition (plotted is the –log10 of the Poisson p-value). No significant differences were detected (two-tailed paired *t*-test).

(**F**) Box plot of oneD-normalized interaction scores calculated for all the AEs bound by TFIIIC for T47D grown in the presence (+S) or absence (-S) of serum. Note the significant increase in the interaction score (Kolmogorov-Smirnov Test).

(**G**) Sitepro profile of strand-specific tRNA expression in T47D in conditions of siCTRL or siGTF3C5 in the absence of serum across all tDNAs spanning a 400 bp region (±200 bp relative to the tRNA TSS; plotted is the average expression in RPM). Note no differences in the two profiles.

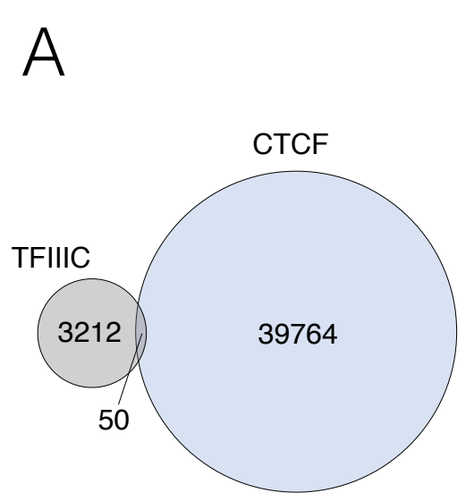
(**H**) Volcano plot comparing mRNA-seq data of siGTF3C5 *vs* siCTRL in T47D grown in the absence of serum (-S) (plotted the –log10 of the adjusted p-value *vs* the -log2 ratio of siGTF3C5 *vs* siCTRL). The genes that scored significant (adjusted p-value < 0.05) are indicated in red (FC > 1.5) and blue (FC < -1.5). *GTF3C5* is found among the most downregulated genes. See Table S2 for more information.

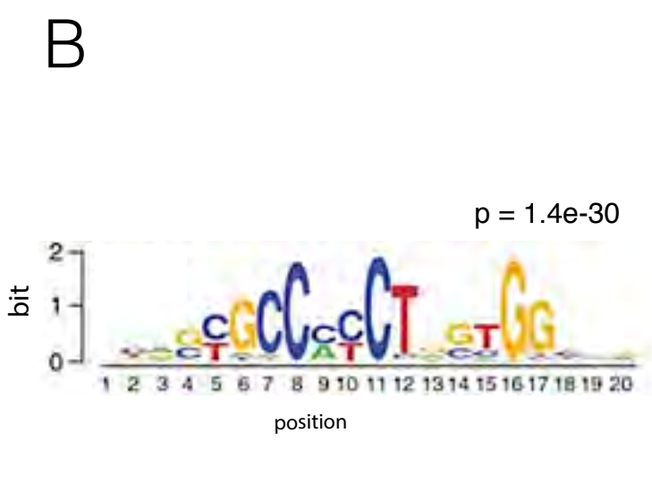
(**I-J**) Scatter plot of gene expression comparing siGTF3C5 (**I**) and siCTRL (**J**) treated cells in the absence of serum *vs* mock-transfected cells in the same growth condition. The number of genes up- or down-regulated (-1.5 < FC > 1.5; adjusted p-value < 0.05) is indicated in red or blue, respectively.

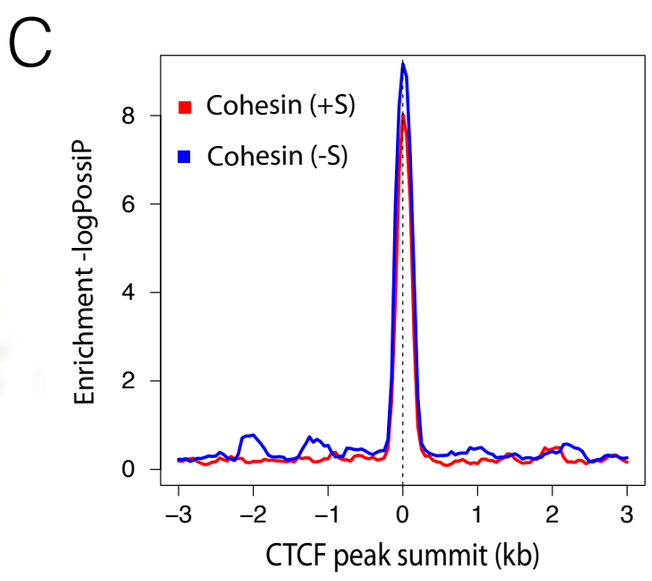
(**K**) Bar plots of GO enrichment (according to DAVID) of TFIIIC-activated genes (genes downregulated in siGTF3C5 cells).

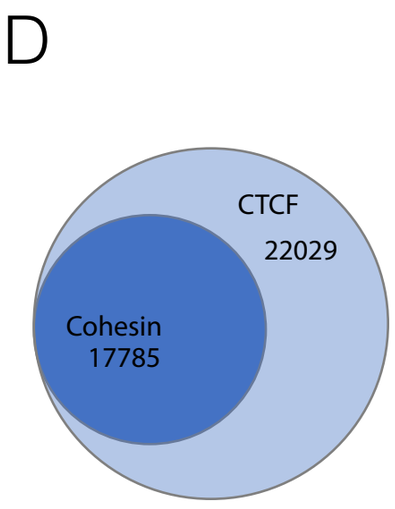
(**L**) FACS cell cycle profile of T47D cells grown in the absence (–S) of serum or transfected with siCTRL or siGT3C5 in the same growth conditions (n = 2 independent experiments). Note that no effect was observed by the siCTRL treatment, but further cell cycle arrest detected in siGTF3C5 cells (*, p-value for squared χ^2 test = 0.00045).

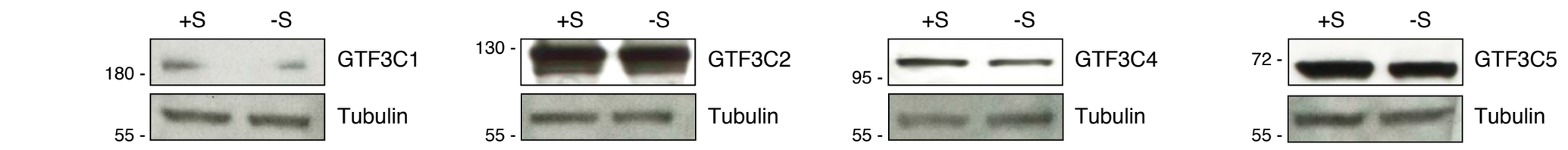
Fig. S5

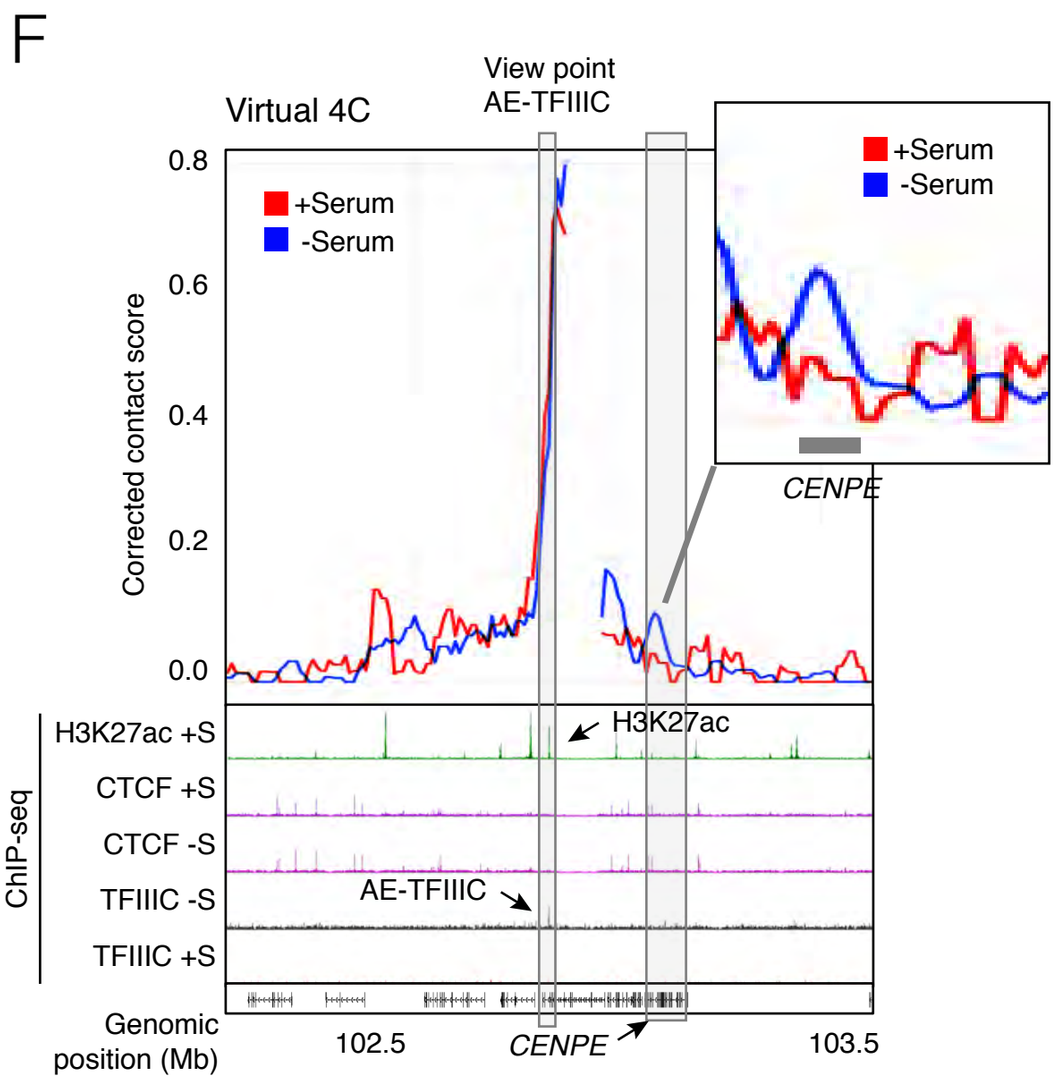

A
CTCF
TFIIIC
3212
39764
50

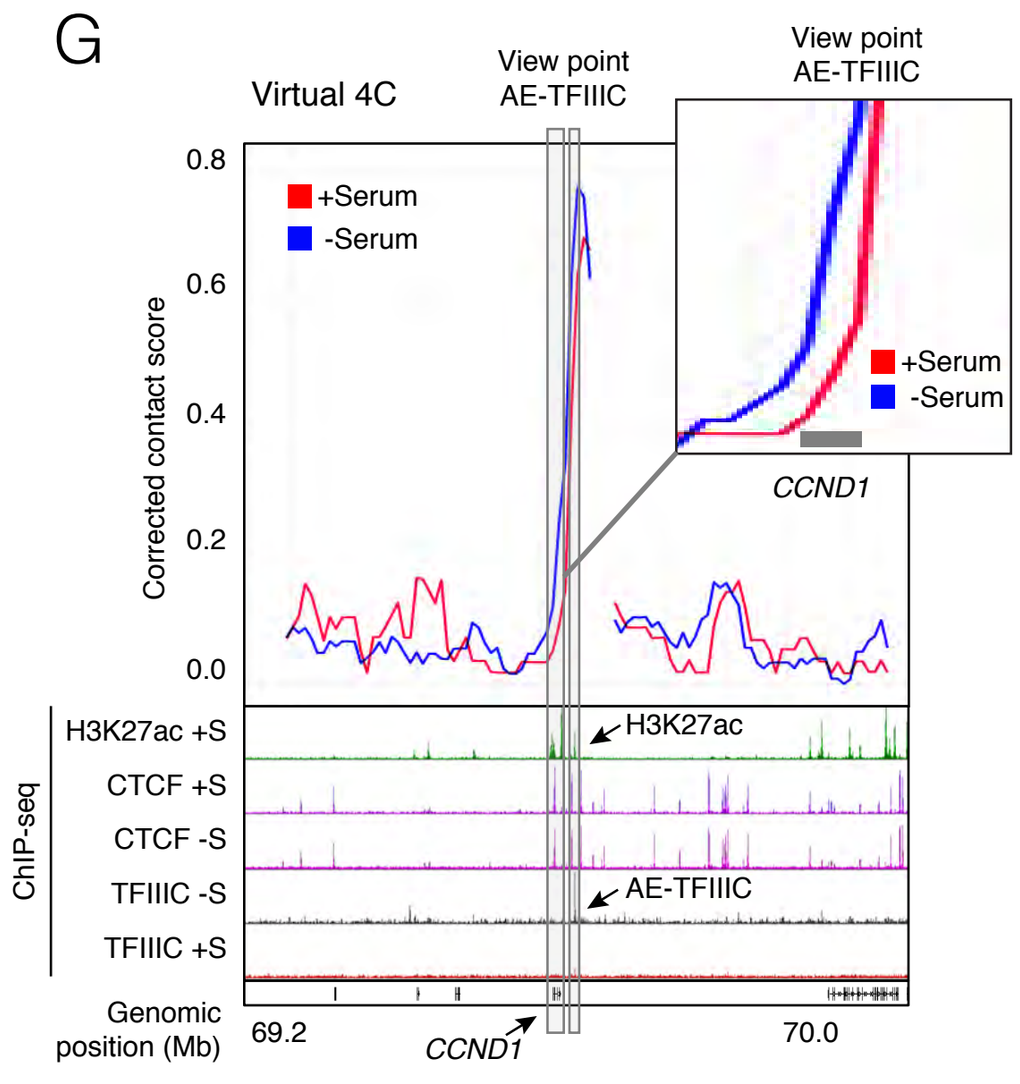

B
p = 1.4e-30
bit
position

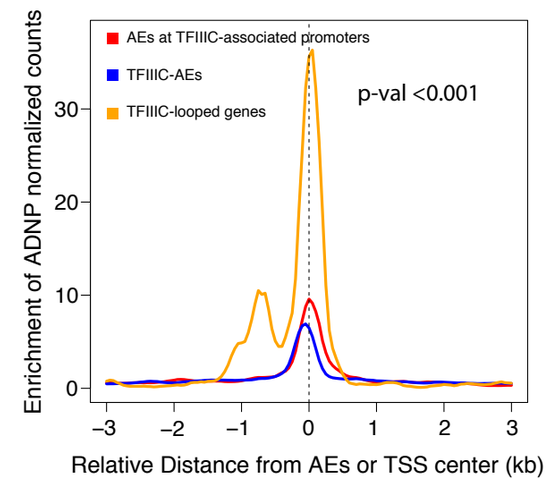

C
Cohesin (+S)
Cohesin (-S)
Enrichment -logPossiP
CTCF peak summit (kb)

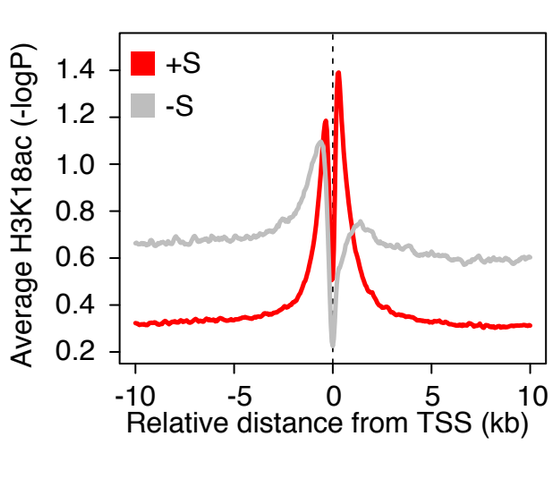

D
CTCF
22029
Cohesin
17785


E
+S
-S
180 -
GTF3C1
Tubulin
55 -
130 -
GTF3C2
95 -
GTF3C4
72 -
GTF3C5


F
Virtual 4C
View point
AE-TFIIIC
+Serum
-Serum
Corrected contact score
CENPE
H3K27ac +S
CTCF +S
CTCF -S
TFIIIC -S
TFIIIC +S
ChIP-seq
H3K27ac
AE-TFIIIC
Genomic
position (Mb)
102.5
103.5


G
Virtual 4C
View point
AE-TFIIIC
+Serum
-Serum
Corrected contact score
CCND1
H3K27ac +S
CTCF +S
CTCF -S
TFIIIC -S
TFIIIC +S
ChIP-seq
H3K27ac
AE-TFIIIC
Genomic
position (Mb)
69.2
70.0


H
AEs at TFIIIC-associated promoters
TFIIIC-AEs
TFIIIC-looped genes
p-val <0.001
Enrichment of ADNP normalized counts
Relative Distance from AEs or TSS center (kb)


I
+S
-S
Average H3K18ac (-logP)
Relative distance from TSS (kb)

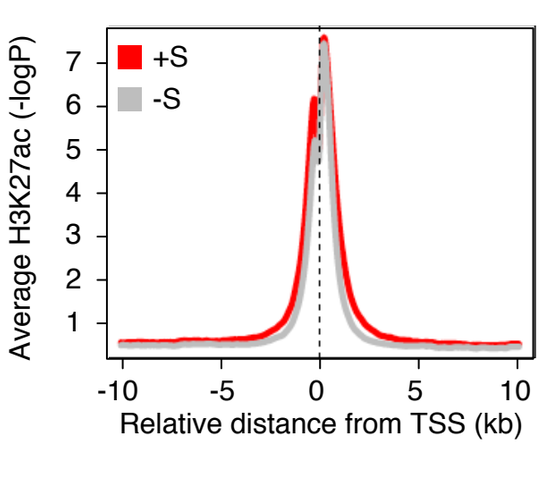

J
+S
-S
Average H3K27ac (-logP)
Relative distance from TSS (kb)

Figure S5. TFIIIC and CTCF increased interaction upon SS does not depend on changes in total levels of TFIIIC, Related to Figure 5.

(A) Venn diagram of overlapping peaks for CTCF and TFIIIC in T47D grown in the absence of serum.

(B) Motif enrichment analysis of CTCF-bound regions in TFIIIC-looped genes. The *p*-value of the motif's enrichment is shown.

(C) Sitepro analysis of RAD21 enrichment at CTCF-bound regions in the TFIIIC-looped genes (red) compared to a random set of regions of the same size (blue). Note that there is no difference in the RAD21 enrichment between the two conditions.

(D) Venn diagram of overlapping peaks for the Cohesin component RAD21 and the CTCF peaks in condition of serum starvation.

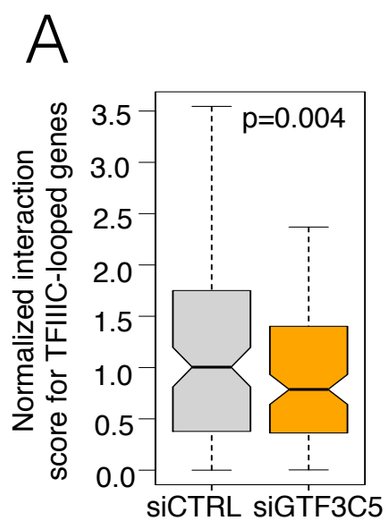
(E) Western blot for different TFIIIC subunits (GTF3C1, GTF3C2, GTF3C4 and GTF3C5) in T47D cells grown in the presence (+S) or absence (–S) of serum. For each panel, a loading control with α-tubulin is also shown.

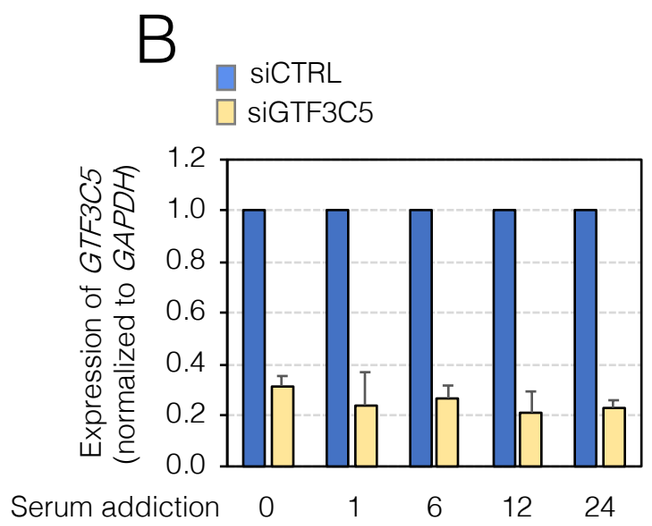
(**F-G**) Virtual 4C (from Hi-C data) representation of the *CENPE* (**F**) and *CCND1* (**G**) loci, two of the genes repressed upon siGTF3C5 (Figure 4C). The data is represented for the conditions of presence (+S red) and absence (–S, blue) of serum. The viewpoint of the virtual 4C is the AE bound by TFIIIC. Arrows point to the location of the gene, the AE bound by TFIIIC (AE-TFIIIC) and the peak of H3K27ac. The y-axis corresponds to the corrected contact score. Regions with changes in the frequency of interaction of the AE-TFIIIC and the target genes have been zoomed out for a clearer view. Grey boxes represent the location of the gene compared to the 4C signal.

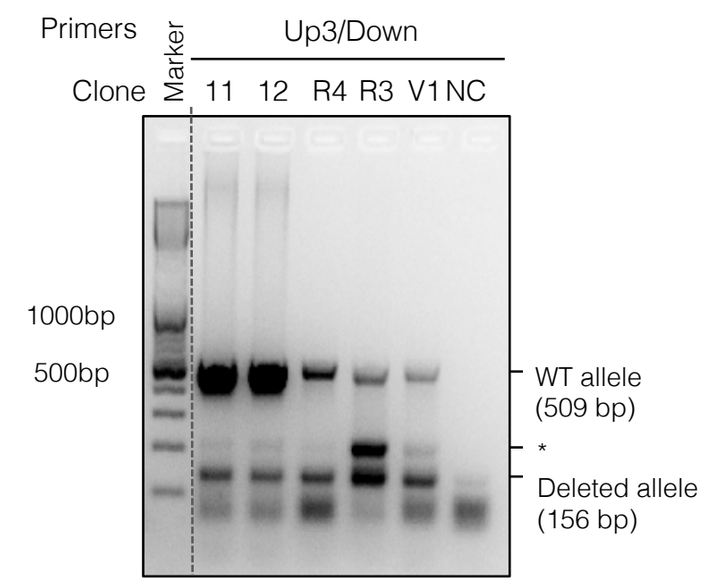
(**H**) Sitepro analysis of ADNP-eGFP binding in K562 cells (GSE105573) at AEs of TFIIIC-associated promoters (red), all AEs-bound by TFIIIC (TFIIIC-AEs, blue) and TFIIIC-looped genes (yellow). Significant higher levels of ADNP-eGFP are found at TFIIIC-looped genes (p-values for squared Chi test comparing TFIIIC-looped genes vs AEs at TFIIIC-associated promoters and TFIIIC-AEs is reported). X-axis is either the center of the AE (for the AEs at TFIIIC-associated promoters or TFIIIC-AEs, or the TSS of the TFIIIC looped genes).

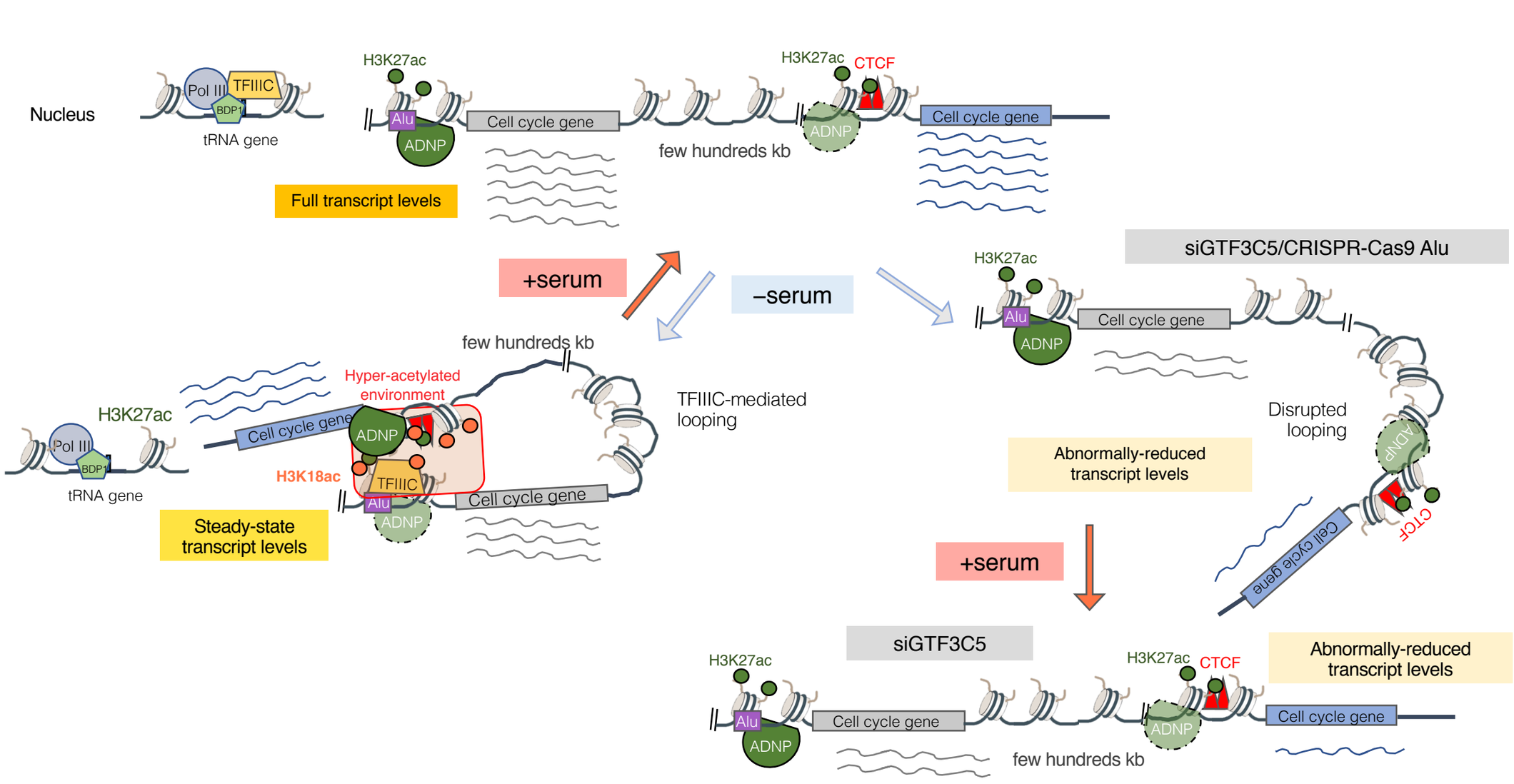
(**I-J)** CEAS plot of H3K18ac (**I**) and H3K27ac (**J**) average at the TSS of all human genes in T47D grown in the presence (+S, red) or absence (–S, grey) of serum (plotted is the –log10 of the Poisson p-value). Note how H3K18ac drastically changed upon serum starvation, whereas H3K27ac remained unaffected.

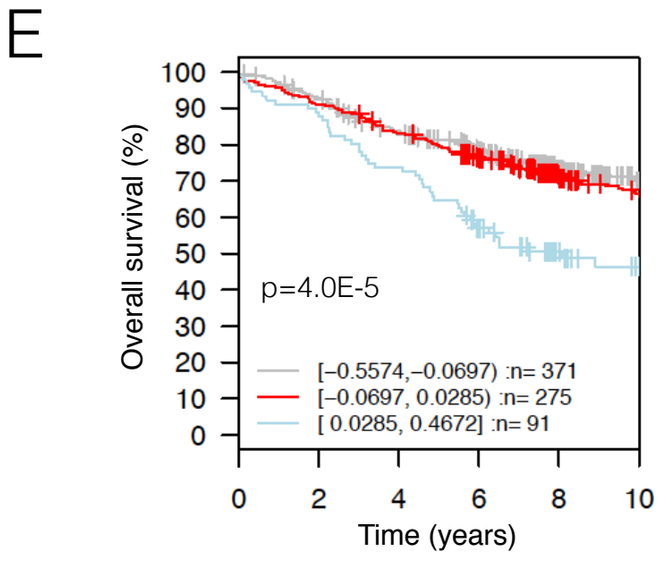
Fig. S6

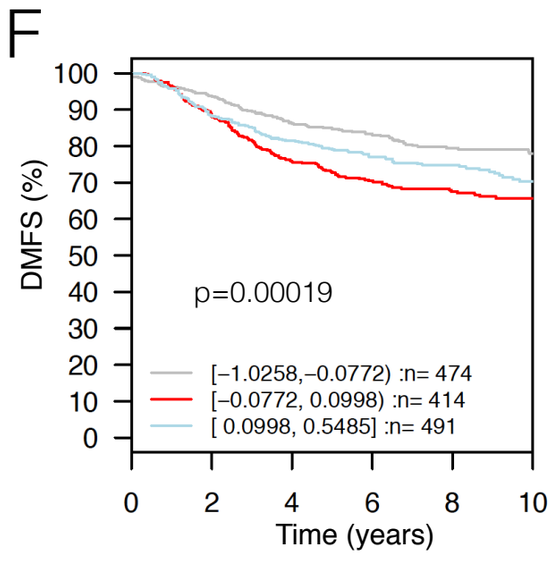

A
Normalized interaction score for TFIIIC-looped genes
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
p=0.004
siCTRL
siGTF3C5


B
siCTRL
siGTF3C5
Expression of GTF3C5 (normalized to GAPDH)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Serum addiction (h)
0
1
6
12
24
Time


C
Primers
Up3/Down
Clone
Marker
11
12
R4
R3
V1
NC
1000bp
500bp
WT allele (509 bp)
*
Deleted allele (156 bp)


D
Nucleus
Pol III
TFIIIC
BDP1
tRNA gene
H3K27ac
Alu
ADNP
Cell cycle gene
few hundreds kb
CTCF
Full transcript levels
+serum
–serum
siGTF3C5/CRISPR-Cas9 Alu
Hyper-acetylated environment
H3K18ac
TFIIIC
TFIIIC-mediated looping
Steady-state transcript levels
Disrupted looping
Abnormally-reduced transcript levels
siGTF3C5


E
Overall survival (%)
p=4.0E-5
[–0.5574,–0.0697) :n= 371
[–0.0697, 0.0285) :n= 275
[0.0285, 0.4672] :n= 91
Time (years)


F
DMFS (%)
p=0.00019
[–1.0258,–0.0772) :n= 474
[–0.0772, 0.0998) :n= 414
[0.0998, 0.5485] :n= 491
Time (years)

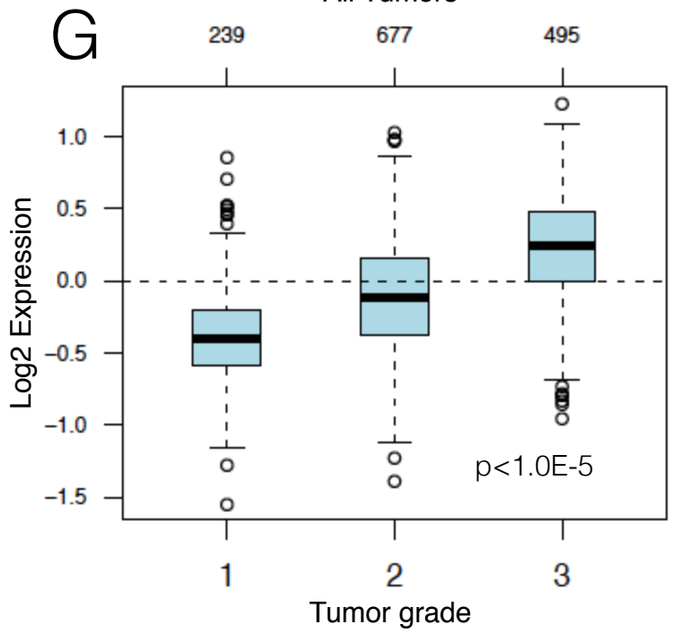

G
All Tumors
239
677
495
Log2 Expression
p<1.0E-5
Tumor grade

Figure S6. The deletion of a TFIIIC-bound AE affects DNA looping and expression of the distal *UHRF1* locus, Related to Figure 6.

(**A**) Boxplot of the normalized interaction score of Hi-C data for siCTRL and siGTF3C5 between promoters and TFIIIC-bound AEs for TFIIIC-looped genes. Two Hi-C biological replicates were used (p-value from Friedman X^2 test is indicated).

(**B**) qRT-PCR expression analysis of *GTF3C5* in T47D cells treated with siCTRL and siGT3C5 in the absence of serum and released from serum starvation by serum addition for the indicated times. Samples correspond to Figure 6D. The graph represents the mean ± SEM from two biological experiments, in which the value in siCTRL cells was arbitrarily set as 1 at each time point. Note that the knockdown of TFIIIC always reaches values of more than 70% at each time point analyzed.

(**C**) PCR result for the screen of CRISPR-Cas9 T47D clones with primers Up3 and Down (see Supplementary Materials and Methods for details and schematic representation in Figure 6E): the upper band corresponds to the WT allele, whereas the lower band correspond to the deleted allele. Representative clones are shown, but almost all clones analyzed were heterozygous for the deletion. For further analysis, clone 11 was selected. The DNA marker size is shown. * indicates a non-specific band. NC corresponds to no DNA sample. Dotted line represents a cut of the gel. However, all the remaining lanes are from the same gel.

(D) Cartoon model for the proposed TFIIIC mechanism of action.

(**E-F**) Kaplan–Meier plots of breast tumor samples for TFIIIC-associated promoters or TFIIIC-activated genes expression, respectively. TFIIIC-associated promoters were divided in three main groups according to their expression levels within brackets (with blue being the highest, red the intermediate and grey the lowest). P-values from a Mantel-Cox test are indicated. Higher expression of TFIIIC-associated promoters is associated with poor prognosis for overall survival and distance metastasis free survival (DMFS), respectively. Plots are generated using (GOBO) (Ringner et al., 2011).

(**G**) Boxplots of expression of TFIIIC-associated promoters from all tumor samples across the three breast cancer grades. Box plots are generated by using (GOBO). TFIIIC-associated promoters show higher expression in most aggressive tumors (3rd grade), p-value is also indicated.