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

Graphical Abstract

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

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

Roberto Ferrari,1,6,* Lara Isabel de Llobet Cucalon,1,6 Chiara Di Vona,1,6 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 Beato1,6,10,*

1Center for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Dr. Aiguader 88, Barcelona 08003, Spain
2Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain
3The Institute of Cancer Research (ICR), London, UK
4Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy
5Université de Bordeaux, INSERM U1212 CNRS UMR 5320 146, Bordeaux, France
6Universitat Pompeu Fabra (UPF), Barcelona, Spain
7ICREA, Pg. Lluis Companys 23, Barcelona 08010, Spain
8Human Technopole. Via Cristina Belgioioso, 171, 20157 Milano MI, Italy
9These authors contributed equally
10Lead 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.
**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 S1). Thus, TFIIIC is reversibly recruited to AEs close to Pol II promoters of a subset of cell-cycle-related genes in response to SS.
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).

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 suggest 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 in 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 3SA). 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),
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 Hi-C (Rao et al., 2014) (Table S3) in T47D cells growing 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).
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 upstream locus containing the PLIN4/PLIN5 genes, whose

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).
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-range 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 5E). 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 identified 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 S5J). Indeed, almost 70% of the TFIIIC-looped
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 S5 J). 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.
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 rewiring 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 heterogeneous 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 (Policardi et al., 2017), with TFIIIC regulating the production and nuclear localization of AE transcripts, which directly regulates Pol II loci (Crepaldi et al., 2013; Policardi 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 5D).

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 CTCF. 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 transposable-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 RNA-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/CAS9
  - 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 BFU2016-78141-P (to S.d.l.L.); ACER (to M.T.); 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, PRBS; supported by grant PT17/0019 [ISICII and ERDF]). This work was also supported by the Ligue Contre le Cancer (to M.T.).

**AUTHOR CONTRIBUTIONS**


**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**


Chromatin Loops between Promoter and Gene Body Regulate Alternative
Ruiz-Velasco, M., Kumar, M., Lai, M.C., Bhat, P., Solis-Pinson, A.B., Reyes, A.,
s12859-017-1934-z.

Tien, A.L., Senbanerjee, S., Kulkarni, A., Mudbhary, R., Weuts, A.,


Zhao, Y., Wang, J., Liang, F., Liu, Y., Wang, Q., Zhang, H., Jiang, M., Zhang, Z.,

# STAR METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-GTF3C2</td>
<td>This paper</td>
<td>NA</td>
</tr>
<tr>
<td>mouse monoclonal antibody anti-GTF3C2</td>
<td>Abcam</td>
<td>RRID:AB_2043034</td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-BDP1</td>
<td>Wang and Roeder, 1997; Weser et al., 2004</td>
<td>NA</td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-RPC39</td>
<td>Wang and Roeder, 1997; Weser et al., 2004</td>
<td>NA</td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-H3K18ac</td>
<td>Active motif</td>
<td>RRID:AB_2793308</td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-CTCF</td>
<td>Millipore</td>
<td>RRID:AB_441965</td>
</tr>
<tr>
<td>rabbit monoclonal antibody anti-CTCF</td>
<td>Abcam</td>
<td>ab128873</td>
</tr>
<tr>
<td>rabbit polyclonal antibody anti-Pol II</td>
<td>Cell Signaling (NTD) (D8L4Y, Cat#: 14958 RRID:AB_2687876)</td>
<td></td>
</tr>
<tr>
<td>rabbit affinity-purified anti-GTF3C5/TFIIC63</td>
<td>Bethyl</td>
<td>RRID:AB_890669</td>
</tr>
<tr>
<td>rabbit affinity-purified anti-GTF3C1/TFIIC220</td>
<td>Novus Biologicals</td>
<td>RRID:AB_922616</td>
</tr>
<tr>
<td>mouse Monoclonal antibody anti-tubulin</td>
<td>Sigma</td>
<td>RRID:AB_477593</td>
</tr>
<tr>
<td>mouse monoclonal [102C1a] to ADNP</td>
<td>Abcam</td>
<td>RRID:AB_879525</td>
</tr>
<tr>
<td>Rabbit polyclonal HA tag antibody</td>
<td>Abcam</td>
<td>RRID:AB_307019</td>
</tr>
<tr>
<td>rabbit affinity-purified anti-EP300</td>
<td>Santa Cruz</td>
<td>RRID:AB_2231120-Lot: I2815</td>
</tr>
<tr>
<td>rabbit affinity-purified anti-EP300</td>
<td>Santa Cruz</td>
<td>RRID:AB_2293429-Lot: J0915</td>
</tr>
<tr>
<td>Rabbit affinity-purified anti-H3K9ac</td>
<td>Active Motif</td>
<td>RRID:AB_2561017</td>
</tr>
<tr>
<td>Mouse Monoclonal anti-H3K27ac</td>
<td>Millipore</td>
<td>RRID:AB_1977529-Lot.2626032</td>
</tr>
<tr>
<td>Rabbit polyclonal to Histone H1.2</td>
<td>Abcam</td>
<td>RRID:AB_2117983</td>
</tr>
<tr>
<td>Horseradish peroxidase conjugated anti-mouse</td>
<td>Amersham</td>
<td>NA931V</td>
</tr>
<tr>
<td>Horseradish peroxidase conjugated anti-rabbit</td>
<td>Amersham</td>
<td>NA934V</td>
</tr>
<tr>
<td>Anti-DYKDDDDK resin</td>
<td>Genscript</td>
<td>L00432</td>
</tr>
<tr>
<td>Anti-HA Agarose</td>
<td>Pierce</td>
<td>26181</td>
</tr>
<tr>
<td>Bacterial and Virus Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One Shot Stbl3 Chemically Competent E. coli</td>
<td>Invitrogen</td>
<td>C7373-03</td>
</tr>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPMI Red phenol</td>
<td>GIBCO</td>
<td>42401-018</td>
</tr>
<tr>
<td>RPMI No Phenol Red</td>
<td>GIBCO</td>
<td>32404-014</td>
</tr>
<tr>
<td>DMEM Phenol Red</td>
<td>GIBCO</td>
<td>41965-039</td>
</tr>
<tr>
<td>DMEM Phenol Red</td>
<td>GIBCO</td>
<td>21063-029</td>
</tr>
<tr>
<td>DMEM/F12 Phenol Red</td>
<td>GIBCO</td>
<td>11330-032</td>
</tr>
<tr>
<td>DMEM/F12 No Phenol Red</td>
<td>GIBCO</td>
<td>11039-021</td>
</tr>
<tr>
<td>EMEM Phenol Red</td>
<td>GIBCO</td>
<td>31095-029</td>
</tr>
<tr>
<td>EMEM No Phenol Red</td>
<td>GIBCO</td>
<td>51200-038</td>
</tr>
<tr>
<td>Fetal Bovine Serum</td>
<td>GIBCO</td>
<td>10270-106</td>
</tr>
<tr>
<td>Fetal Bovine Serum, charcoal stripped</td>
<td>GIBCO</td>
<td>12676029</td>
</tr>
<tr>
<td>0.5% Trypsin-EDTA 1x</td>
<td>GIBCO</td>
<td>25300-054</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE SOURCE IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Glutamine 200 mM 100x GIBCO 25030-024</td>
</tr>
<tr>
<td>Penicillin-Streptomycin GIBCO 15140-122</td>
</tr>
<tr>
<td>Human Insulin (Humulin regular) Lilly U100</td>
</tr>
<tr>
<td>EGF SIGMA E-9644</td>
</tr>
<tr>
<td>Hydrocortisone SIGMA H-0888</td>
</tr>
<tr>
<td>Horse Serum Life technologies 16050122</td>
</tr>
<tr>
<td>Cholera Toxin SIGMA C8052</td>
</tr>
<tr>
<td>Trizol Reagent Ambion 15596018</td>
</tr>
<tr>
<td>Lipofectamine 3000 Invitrogen 11668-019</td>
</tr>
<tr>
<td>Proteinase K ThermoFisher Scientific AM2546</td>
</tr>
<tr>
<td>Protein G Plus / Protein A Agarose Millipore IP05</td>
</tr>
<tr>
<td>protease inhibitor cocktail, cComplete EDTA-free Roche 05 892 791 001</td>
</tr>
<tr>
<td>Na3VO4</td>
</tr>
<tr>
<td>MboI New England BioLabs r0147-mboi</td>
</tr>
<tr>
<td>Herculase II Fusion DNA Polymerase Agilent 600675</td>
</tr>
<tr>
<td>AMPure XP beads Beckman Coulter A63881</td>
</tr>
<tr>
<td>Endopeptidase LysC, Wako 125-05061</td>
</tr>
<tr>
<td>Sequencing grade Trypsin Promega VS111</td>
</tr>
</tbody>
</table>

Critical Commercial Assays

<table>
<thead>
<tr>
<th>Critical Commercial Assays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovation Ultralow DR Multiplex System 9-16 kit NUGEN 0535-32</td>
</tr>
<tr>
<td>Micro BCA Protein Assay Kit ThermoFisher Scientific 23235</td>
</tr>
<tr>
<td>qScript cDNA Synthesis kit Quanta Biosciences 95047-025</td>
</tr>
<tr>
<td>LightCycler FastStart DNA Master SYBR Green I kit Roche 03 515 885 001</td>
</tr>
<tr>
<td>Illumina TruSeq Kit Illumina 20020594</td>
</tr>
<tr>
<td>TruSeq Stranded Total RNA Library Prep Human/Mouse/Rat Illumina 20020596</td>
</tr>
<tr>
<td>Pierce Coomassie (Bradford) kit Thermo Fisher 23200</td>
</tr>
<tr>
<td>NEBNext DNA Library Prep Kit New England Biolabs NEB E7645</td>
</tr>
<tr>
<td>Qubit HS kit Thermo Fisher Q32854</td>
</tr>
<tr>
<td>FISH Tag RNA Multicolor Kit Thermo Fisher mp32956</td>
</tr>
<tr>
<td>Gibson Assembly Master New England Biolabs E2611S</td>
</tr>
<tr>
<td>Bsmbl Thermo Fisher ER0451</td>
</tr>
<tr>
<td>Anti-DYKDDDDK G1 Affinity Resin Genscript L00432</td>
</tr>
<tr>
<td>Anti-HA Agarose Thermo Fisher Scientific 26181</td>
</tr>
<tr>
<td>NuPAGE MOPS SDS Running Buffer Novex by Life Technology NP0001</td>
</tr>
<tr>
<td>NuPAGE 4–12% Bis Tris Gels Invitrogen NP0335box</td>
</tr>
<tr>
<td>Skim Milk Powder Sigma 70166-500G</td>
</tr>
<tr>
<td>XCell SureLock Mini-Cell Electrophoresis System Thermo Fisher EI0001</td>
</tr>
</tbody>
</table>

Deposited Data

<table>
<thead>
<tr>
<th>Deposited Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChIP-seq, RNA-seq, Hi-C data This study GSE120162</td>
</tr>
<tr>
<td>CTCF ChIP-seq of T47D in -S Le Dily et al., 2014 GSE53463</td>
</tr>
<tr>
<td>ChIP-seq of eGFP-ADNP in K562 cells Consortium, 2012 GSE105573</td>
</tr>
<tr>
<td>ChIP-seq of mouse Adnp Ostapcuk et al., 2018 GSE97945</td>
</tr>
<tr>
<td>Gene expression data of breast cancer samples Ringnér et al., 2011 <a href="http://co.bmc.lu.se/gobo/">http://co.bmc.lu.se/gobo/</a></td>
</tr>
<tr>
<td>EP300 ChIP-seq in T98G Ramos et al., 2010 GSE21026</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteomic data: PRIDE</td>
<td>This study</td>
<td>PXD011250</td>
</tr>
<tr>
<td>Mendelely</td>
<td>This study</td>
<td><a href="https://doi.org/10.17632/mzjf96t3gc.3">https://doi.org/10.17632/mzjf96t3gc.3</a></td>
</tr>
</tbody>
</table>

Experimental Models: Cell Lines

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Source</th>
<th>Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>T47D</td>
<td>ATCC</td>
<td>CRL-2865</td>
</tr>
<tr>
<td>T98G</td>
<td>ATCC</td>
<td>CRL-1690</td>
</tr>
<tr>
<td>IMR90</td>
<td>ATCC</td>
<td>CCL-186</td>
</tr>
<tr>
<td>MCF10A</td>
<td>ATCC</td>
<td>CRL-10317</td>
</tr>
</tbody>
</table>

Oligonucleotides

- **UHRF1**-associated AE H3K18ac ChIP Forward: ATTGTAAT
  - This study
  - NA

- **UHRF1**-associated AE H3K18ac ChIP Reverse: CGGGTTC
  - This study
  - NA

- **UHRF1**-expression Forward: GCCATACCCCTCTTTTCGACTACG
  - This study
  - NA

- **UHRF1**-expression Reverse: GCCCCTAATCCCCCTCTCATTCC
  - This study
  - NA

- **HELLS**-associated AE H3K18ac ChIP: Forward: TAGCTGGAAATGGGCTAAT
  - This study
  - NA

- **HELLS**-associated AE H3K18ac ChIP: Reverse: TCAGTTGATCTCCACCTTC
  - This study
  - NA

- PIPIA primers: For – GCCGAGGAAAACCGTGTACT
  - This study
  - NA

- PIPIA 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: GCCGGAACGATACGCGTCAT
  - 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 forward: TGGGATTCCTGGATCCGGTGTTTCTTACCAAGAT
  - This study
  - NA

- Oligo2 forward: CACCGGATCCAGGAATCCCCTCTGTTTTCAGAGTAGGAGAC
  - This study
  - NA

- Oligo3 reverse: GAGCGGGATCCATGCCCTGATCGGCCTGACCTCCAC
  - This study
  - NA

- Oligo4 forward: TCCCTAGAGATCCTCACCTCTGCTGTAGGAC
  - This study
  - NA

- Oligo5 reverse: AACAAACCAAAAAAACCACCAAGGGAGAGATGGAGCTCAGAT
  - This study
  - NA

- Oligo6 forward: GCCGGAATCTGGACGTTTTGTTGTTGTTGTTTAGGCTAGAA
  - This study
  - NA

- C557F: GTTTTGAAGCTAGAATTAGCAAG
  - This study
  - NA

- C557R: GTGGCACTCATACAAGAATCTATAAG
  - This study
  - NA

- C542F: GTCCCCCAATGGCCGTTCCGC
  - This study
  - NA

- C542R: ATGTGCTACTATTTCCC
  - This study
  - NA

- gRNA1: GATCAGGAAATCCACTCCACTTCT
  - This study
  - NA

- gRNA2: TTTGAGGGGTTTTTGTGTGGTTGTTTAGAGCTAGAA
  - This study
  - NA

- Alu_up: CGCAAGCTAAAGGAGCAGT
  - This study
  - NA

- Alu_down: ACCTGGCAAGGATTTGAAG
  - This study
  - NA

- **GTF3C1** expression_for: GGGAAGCTGCACTATCACAGA
  - This study
  - NA

- **GTF3C1** expression_rev: GGTCAATGGGATCAGTCATGTGGGACT
  - This study
  - NA

(Continued on next page)
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 for 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).
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 900xg 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 MgCl₂) 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 (10 mM Tris-HCl pH 7.4, 0.1% Igepal CA-630) followed by 50 µl of RBS buffer to resuspend the cell nuclei.

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 Genes and Antibodies for Key Resources Table). The ChIP for p300 was carried out mixing the two Santa Cruz antibodies (ab9110) and mouse monoclonal anti-GTF3C2 Abcam (ab89113).

**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; Ferrai 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); β-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 C. Cells were harvested by centrifugation, resuspended in lysis buffer (500 mM NaCl, 20 mM HEPES pH 8, 1 mM MgCl₂, 10% glycerol, 5 mM β-mercaptoethanol) and sonicated for 10 s. The lysates were centrifuged at 16,000xg 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 1,200xg 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 (Ventin 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 a coverage of 4 reads/bp of the human genome for each condition. The T47D ChIPs for TFIIIC, BDP1, RPC39, and DR Multiplex System 1-8 and 9-16 kit (NuGen). To minimize false positives calling, several input libraries were sequenced to reach a coverage of 4 reads/bp of the human genome for each condition. The T47D ChIPs for TFIIIC, BDP1, RPC39, and DR Multiplex System 1-8 and 9-16 kit (NuGen).

**ATAC-seq**

ATAC-seq reaction and library preparation

For nuclei preparation, 5x10⁶ 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 900xg 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 MgCl₂) 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 assembled as follows:
- 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 follows:
(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 follows:
(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 close 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 (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).

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 (Wang et al., 2012) (see Oligo-nucleotides 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://ggo.io/VzooQg) (trimmomatic PE raw_fasta trimmed_fasta 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--runMode genomeGenerate--genomeDir GENOME_DIR--genomeLoad NoShared Memory--runThreadN slots--outFilterType “BySJout”--outFilterMultimapNmax 20--alignSJoverhangMin 8–alignSJoverhangMin 1–13.e1–e11, February 6, 2020

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://docs.google.com/document/d/1yRZevDdjxkEdna9WF5-qaARjROZmcmdPl3xtFrYe/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 using a DESeq.R script for RNA analysis. The script is provided as a downloadable

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 MgCl2, 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 Na2VO4 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-001266-14 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 x 10^7 – 2 x 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 was 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 cComplete 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 MgCl2, 0.34 M Sucrose, 10% Glycerol, 1 mM DTT and protease inhibitor cocktail cComplete 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 cComplete 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 supernatant (after discarding the supernatant) was resuspended 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 0.45μm 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: 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 x 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 CiaI 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., 2019). 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
Identification of Intra-TAD contacts

Using Hi-C matrices at 5 kb resolution, we focused on TADs containing TFIIC. Each bin was labeled according to TFIIC 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/CAS9

Generation of Alu-deleted T47D cells by CRISPR/Cas9

The generation of the targeting vector was carried out by DECKO 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 Polystyrene, 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: cComplete EDTA-free as recommended by manufacturer (1 tablet for 50mL), 10 µM phenylmethylsulfonyl fluoride and 10 µM Na3VO4. 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 107 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 iodoacetamide 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°C; 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 Significance Analysis of INTeractome (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 (Györffy 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.