

Cooperative and competitive interactions among transcription regulatory elements modulate transcription output

Yiyang Jin^{1,2}, Zhou Zhou¹, Jessica Qin Lin¹, Alden King-Yung Leung², Haiyuan Yu^{2,3*}, John T. Lis^{1*}

¹Department of Molecular Biology and Genetics, Cornell University; Ithaca, NY, USA

²Weill Institute for Cell and Molecular Biology, Cornell University; Ithaca, NY, USA

³Department of Computational Biology, Cornell University; Ithaca, NY, USA

*Correspondence to: Haiyuan Yu (haiyuan.yu@cornell.edu) & John T. Lis (johnlis@cornell.edu)

Abstract

Transcription is modulated by interactions among transcription regulatory elements (TREs), including promoters and enhancers, but their underlying interaction specificities remain opaque. Here, we develop a Chromatin-Integrated, landing-pad-based Enhancer Reporter Assay, CIERA-seq, that interrogates regulatory interactions between target promoters and distal TREs in a controlled genomic and chromatin context. We find both promoter- and enhancer-TREs exhibit enhancer activities on target promoters with a specificity that depends on factors that TREs recruit. Promoters are generally enriched for factors involved in the core transcription process, whereas enhancers are enriched for pioneer factors and chromatin remodelers. Our CIERA-seq assays support a model where TREs activate target promoters constrained at different rate limiting steps in transcription by supplying complementary components of the transcription regulatory machinery. An orthogonal analysis using CRISPRi datasets shows that both promoters and enhancers can positively regulate expression of neighboring genes at their native loci. However, promoters are also more prone than enhancers to exhibit negative regulatory effects, a difference that may arise from competition for limiting transcriptional machinery. Together, these findings support a framework in which promoters and enhancers act as cooperative, or competitive, regulatory elements within local transcription hubs to modulate transcription outputs.

Introduction

Transcription is a dynamic and spatially organized process in which proximity between transcription regulatory elements (TREs), including promoters and enhancers, play a central role in gene regulation^{1,2}. A long-standing question in transcription regulation is whether promoters exhibit selective “compatibility” with specific enhancers. Plasmid-based assays across diverse systems, including drosophila, mouse, and human cells, have produced mixed conclusions, with some studies reporting strong promoter specificity and others supporting broadly permissive promoter–enhancer regulation^{3–8}. However, while plasmid-based assays enable large-scale studies between interactions of enhancers and promoters, they differ fundamentally from the native chromatin environment. In the genome, DNA is packaged into nucleosomes, and chromatin accessibility is tightly regulated by intrinsic DNA features, histone modifications, and the activity of pioneer transcription factors (TF) and chromatin remodelers^{9–12}. Because plasmids exhibit distinct nucleosome organization and lack normal chromatin regulatory constraints, interaction rules inferred from plasmid-based assays may not accurately reflect transcription regulation between TREs in chromatin¹³.

Another unresolved issue in transcriptional regulation arises from insights provided by high-resolution chromatin-conformation assays such as Hi-C and Micro-C, which have revealed multiple classes of TRE–TRE interactions, most prominently promoter–enhancer (P–E) and promoter–promoter (P–P) loops^{14,15}. P–E interactions have been extensively characterized and are widely recognized as key mechanisms for achieving precise spatial and temporal control of gene expression¹⁶. In contrast, the functional significance of P–P interactions remains poorly understood^{15,17}. It is unclear the degree to which P–P loops merely reflect shared engagement with enhancer hubs or whether they play an active regulatory role by enabling promoters to directly influence one another’s transcription output³.

To overcome the limitations of plasmid based assays and to investigate both promoter–enhancer compatibility and the regulatory significance of promoter–promoter interactions, we developed CIERA-seq, a Chromatin-Integrated, massively parallel Enhancer Reporter Assay that here enabled systematic interrogation of regulatory interactions between 16 target promoters and approximately 350 TREs spanning diverse regulatory features^{18,19}. Target promoters differ markedly in their intrinsic chromatin accessibility and basal transcription activity, and these intrinsic properties strongly shape their responsiveness to TREs. Notably, we observed that many target promoters can be activated over a large dynamic range not only by enhancers but also by other promoters³. This widespread promoter–promoter activation suggests that P–P interactions play a more prominent and direct regulatory role at endogenous loci than previously appreciated. Mechanistically, promoters deficient in specific components of the transcription machinery (TFs, remodelers, and Pol II) were selectively activated by TREs that supply the missing factors, indicating that transcription factor sharing is a key mechanism underlying the specificity of TRE-mediated activation of promoters¹³. Consistent with CIERA-seq results, complementary analyses

of CRISPRi datasets show that both promoters and enhancers can activate neighboring genes at their native loci. Interestingly, promoters are also more likely than enhancers to exert negative regulatory effects, perhaps a consequence of competition for limiting local transcription machinery. From both our CIERA-seq assays and CRISPRi data analyses, we propose a model in which promoters and enhancers function as cooperative or competitive elements within local transcriptional hubs.

Results

Design of CIERA-seq

To investigate how target promoters respond to activation by diverse TREs in a native chromatin context, we developed CIERA-seq. Using this landing-pad platform, we quantified the activities of more than 350 TREs paired with 16 target promoters in K562 cells, yielding over 5,000 target promoter–TRE combinations. We first generated a K562 cell line harboring a Bxb1 recombinase landing pad with constitutive BFP expression inserted at the eNMU safe-harbor locus¹⁸. Site-specific integration of a target promoter–EGFP–TRE construct replaces the BFP cassette, such that correctly recombined cells lose BFP and express EGFP under the control of the integrated promoter–TRE pair. For a given target promoter, the expectation is that TREs with stronger regulatory activity drive higher EGFP expression than weaker ones (Fig. 1a). Cells containing different target promoter–TRE combinations were collected into seven EGFP expression bins, and subjected to sequencing of both the target promoter barcodes (PIDs), identifying the target promoters, and the TRE sequences from each bin²⁰. This approach enabled us to derive bin-specific expression distributions for every target promoter–TRE pair and to calculate an activation score based on weighted EGFP bin values (Fig. 1b and Extended Data Fig. 1a). Scores were highly reproducible across two biological replicates, and combined scores were used for all downstream analyses (Extended Data Fig. 1b).

We defined TREs, both enhancers and promoters, using divergent PRO-cap transcription start sites and extended each element by ± 60 bp to capture features required for transcription initiation, pausing, and release (Fig. 1c)^{8,21–24}. Although promoters and enhancers share similar regulatory architectures, promoters are distinguished by their gene-proximal context and the frequent presence of 5' splice sites that stabilize mRNA, whereas enhancers are typically distal or intronic and generate unstable enhancer RNAs (Fig. 1c)^{25,26}. Promoters and enhancers differ in their chromatin states and transcription levels. Promoters exhibited higher PRO-cap, DNase-seq, H3K27ac, and H3K4me3 signals than enhancers, whereas H3K4me1, a frequently used enhancer mark, did not robustly distinguish between the two classes (Fig. 1d). In addition to ORF-derived negative controls, we included a set of “untranscribed” TREs that lacked detectable PRO-cap signal but were annotated as enhancers based on ChromHMM and ENCODE cCRE and enriched for H3K4me1 (Fig. 1d and Extended Data Fig.S4a,b)^{11,27,28}.

Diverse compatibilities of TRE-promoter activation

The 16 target promoters were selected to span a broad range of endogenous transcription states, as defined by PRO-seq at their native loci²¹ (Extended Data Fig. 3a,b). To examine how intrinsic promoter properties influence their responsiveness to TREs, we included both highly expressed promoters, such as *MYC* and *GAPDH*, and promoters that are transcriptionally silent in K562 cells but active in other cell types, including *MNDA*, *BIN2*, *ADGRG5*, and *HNF1B* (hereafter referred to as untranscribed promoters)²⁹. For each target promoter, the distribution and mean score obtained when paired with negative control elements provided a measure of its intrinsic activity at the integration locus (Fig. 2a). Target promoters exhibited wide variation in intrinsic activity, which did not necessarily correspond to their endogenous transcription levels (Fig. 2b, Extended Data Fig. 2a, and Extended Data Fig. 3a,b). Notably, the *MYC* promoter, despite being highly transcribed at its native locus, displayed intrinsically low activity when isolated from its endogenous distal TRE regulatory contacts (Fig. 2a and Extended Data Fig. 3b)¹⁶.

We classified a TRE as an activator of a given target promoter, if it can significantly increase the expression of EGFP from the target promoter compared to negative controls. As expected, the positive control CMV enhancer robustly activated nearly all target promoters, shifting EGFP distributions toward higher expression bins (Extended Data Fig. 2b)³⁰. Promoters differed markedly in their responsiveness to TRE input, revealing two broad classes: promoters that were activated by many TREs and promoters that were largely unresponsive (Fig. 2c). In general, well-expressed promoters were activated by a larger number of TREs than untranscribed promoters. TREs also varied widely in their specificity of enhancer activity, with some activating up to all the target promoters tested, whereas others activated only one or none (Fig. 2d, Extended Data Fig. 4c), consistent with previous studies⁵.

Both enhancer- and promoter-TRE show enhancer activity

In a recent study, we showed that promoters and enhancers display enhancer-like activity in a plasmid-based assay, but whether this property is retained in chromatin remained unclear³. Across all target promoters, both promoter- and enhancer-TREs exhibited substantially stronger enhancer activity than untranscribed elements and negative controls (Fig. 2e). Target promoters displayed highly heterogeneous responses to TRE input: untranscribed target promoters were generally less responsive than more transcribed target promoters (Fig. 2c; Fig. 2e). Among highly-responsive target promoters, promoter-TREs, on average, showed greater enhancer activity and broader generality than enhancer-TREs (Fig. 2e). In contrast, poorly responsive target promoters, particularly the untranscribed group, were activated by only a limited subset of TREs, among which a small number of enhancers exhibited exceptionally strong activity (Fig. 2e). This pattern suggests that certain enhancers possess specific regulatory features that enable them to activate otherwise untranscribed target promoters more effectively than other TREs, highlighting a distinct class of enhancers with elevated activating potential.

Enhancer- and promoter-TREs are enriched for different TFs

In our model, each TRE is comprised of a central region containing transcription factor binding sites flanked by a pair of divergent core promoters that supports RNA polymerase II (Pol II) binding, initiation, pausing, and pause release^{8,25} (Fig. 1c). In the CIERA-seq system, the TRE is positioned in close proximity (< 2 kb) but downstream of the target promoter, leading us to hypothesize that the TREs activate the target promoter by sharing transcription machinery with it^{31,32} (Fig. 1a). To examine how TRE-associated transcription machinery relates to TRE's enhancer activity, we analyzed ChIP-seq profiles for more than 300 TFs across all tested TREs at their native loci and observed pronounced differences among TRE classes^{28,33,34}. K-means clustering partitioned the 350 TREs into seven distinct groups defined by characteristic transcription factor occupancy patterns (Fig. 3a, selected TFs with known activating roles; Extended Data Fig. 6a, all factors showing differential enrichment across clusters)³⁵. Cluster 1 consists of negative controls and untranscribed elements, characterized by uniformly low TF occupancy; clusters 2–4 are composed primarily of promoters, and clusters 5–7 contain mostly enhancers (Fig. 3b). Enhancers (clusters 5–7) were preferentially enriched for pioneer factors (e.g., GATA1), SWI/SNF chromatin remodelers (e.g., ARID1B, SMARCC2), and K562 lineage-specific regulators (e.g., GATA1, STAT5A, LEF1), consistent with roles in chromatin opening and cell-type-specific regulation (Fig. 3c, Extended Data Fig.6-8)^{36–38}. In contrast, promoters (clusters 2–4) were more enriched for factors involved in core transcription processes, including Pol II subunits, general transcription factors required for preinitiation complex (PIC) assembly (TBP, TAF1, TAF9), initiation factors (NRF1, SP1, ELF4), and pausing or early elongation factors (SPT5) (Fig. 3c, Extended Data Fig.6-8)^{39,40}. Differential enrichment of transcription machinery at enhancers and promoters likely explain the distinct activities observed in CIERA-seq when these elements are paired with different target promoters.

Different TF clusters show different target promoter compatibility

We compared the target promoter-specific enhancer activities of TREs from the seven TF-enrichment clusters using violin plots of activity scores across all target promoter–TRE pairs (Fig. 3d). Among those target promoters that are well transcribed at their native loci, consistent activity hierarchies emerged within both promoter- and enhancer-TREs. Both types of TREs displayed a graded increase in activity from promoter cluster 2 to cluster 4 and enhancer cluster 5 to cluster 7. All active clusters exhibited substantially higher activity than cluster 1, which consists primarily of negative controls and untranscribed elements. Based on these trends, we classify the clusters as inactive (cluster 1), weak, moderate and strong promoters (clusters 2–4), and weak, moderate and strong enhancers (clusters 5–7).

Within both promoter- and enhancer-TREs, enhancer activity scaled with the extent and diversity of transcription-factor binding: strong promoters and enhancers showed broader and more robust enrichment of promoter- or enhancer-specific TFs than their moderate or weak counterparts. These functional differences were reflected in distinct transcriptional and chromatin-accessibility

features at their native genomic loci, providing a mechanistic basis for the observed activity trends (Extended Data Fig. 5a-c)²⁸. Among promoter-TREs, PRO-cap and H3K4me3 signals increased progressively from weak to moderate to strong promoters, paralleling the graded enrichment of factors involved in the core transcription process. Enhancer-TREs (clusters 5–7) exhibited progressively increased chromatin accessibility, as measured by DNase-seq, consistent with their enrichment for pioneer factors and chromatin remodelers. H3K27ac signal increased with TRE strength in both promoter and enhancer classes. Together, these patterns suggest that enhancer strength in promoter-TREs is driven primarily by enhanced recruitment of core transcription machinery, whereas enhancer-TREs achieve higher activity through an increased capacity to establish open chromatin via pioneer factors and chromatin remodelers.

Notably, although promoter-TREs were generally more active than enhancer-TREs, strong enhancers in cluster 7 exhibited the highest overall activity and the greatest generality across target promoters (Fig. 3d, Extended Data Fig. 5d). Strikingly, untranscribed promoters were activated almost exclusively by a small subset of these strong cluster 7 enhancers. Together, these observations may reveal fundamental mechanistic differences in how promoters and enhancers activate transcription and suggest that transcription factors enriched at strong enhancers, particularly pioneer factors and SWI/SNF chromatin remodelers, may contribute disproportionately to both their exceptional enhancer activity and broad promoter compatibility.

Compatibility of intrinsically-closed and -open target promoters with TREs

Untranscribed target promoters were activated almost exclusively by cluster 7 TREs, the strongest enhancer class most highly enriched for pioneer factors and chromatin remodelers. Because these promoters exhibit low chromatin accessibility at their native loci, we hypothesized that they remain largely inaccessible when integrated at the test locus and therefore require TREs capable of actively opening chromatin for activation. To test this, we generated cell lines containing representative target promoter–TRE combinations from different clusters, each recapitulating the EGFP expression patterns observed in the large-scale CIERA-seq assay (Fig. 4a, Fig.S10a). The *MYC* target promoter is modestly activated by a weak promoter (*UBE2D3* promoter, clusters 2) and a weak enhancer (enhancer_21, cluster 5), whereas a moderate promoter (*RCC1* promoter, cluster 3) or a strong enhancer (enhancer_93, cluster 7) induces substantially higher activation. In contrast, the untranscribed *ADGRG5* promoter is only robustly activated by the strong enhancer (enhancer_93, cluster 7).

We assessed target promoter accessibility using DNase I digestion followed by PCR, where more accessible chromatin regions undergo greater DNase digestion and thus yield weaker PCR signal compared to less accessible ones⁴¹. Endogenous *GAPDH*, CTCF region, and *MNDA* promoters served as controls of different accessibility (Extended Data Fig.9a). DNase I treatment efficiency was assessed in all samples based on signal at CTCF sites and was similar across samples compared within each experimental group (Extended Data Fig.9c-e). For the *MYC* promoter,

DNase treatment nearly abolished PCR signal regardless of the paired TRE, indicating that *MYC* is intrinsically accessible (Fig. 4b). In contrast, the *ADGRG5* promoter exhibited strong PCR signal when paired with an untranscribed TRE but markedly reduced signal when paired with a strong cluster 7 enhancer (enhancer_93), demonstrating that this promoter is intrinsically inaccessible but can be opened by TREs that recruit pioneer factors and chromatin remodelers (Fig. 4c). Notably, the moderate *RCCI* promoter exhibited much weaker ability to increase *ADGRG5* accessibility compared with the strong enhancer, providing a possible explanation for why it failed to activate *ADGRG5* promoter to a similar level as the strong enhancer (Extended Data Fig. 9b).

Role of specific TFs involved in enhancer activity of TREs

To investigate the contribution of transcription factors to promoter-driven enhancer activity in our CIERA-seq assay, we focused on the *RCCI* promoter, which displayed strong enhancer activity when paired with the *MYC* target promoter but showed only weak activity when paired with the *ADGRG5* promoter (Fig. 4a, Extended Data Fig.10a). Because promoters are enriched for transcription factors involved in the core transcription process, we tested whether disrupting the binding motifs of several TFs involved in transcription initiation on the *RCCI* promoter would compromise its enhancer function. This *RCCI* promoter contains high-confidence motifs for SP1, NRF1, and ELF4, which are tightly linked to TSSs and transcription initiation, within the upstream region between the divergent transcription start sites (Fig. 4d)⁴²⁻⁴⁴. We introduced individual and combinatorial mutations by fully replacing each motif with length-matched, non-regulatory sequence, and measured enhancer activity of the mutated *RCCI* promoter-TRE by pairing every mutant with the *MYC* promoter at the same genomic locus and quantifying EGFP expression. (Fig. 4f, Extended Data Fig.10b). Mutating a single SP1, ELF4, or NRF1 motif caused modest reductions in enhancer activity. Double mutations involving ELF4+NRF1 or SP1+ELF4 further reduced activity, whereas disruption of SP1+NRF1 led to a pronounced loss of enhancer function. Simultaneous mutation of all three motifs completely abolished enhancer activity, reducing output to levels comparable to an inactive element (cluster 1). These results demonstrate that recruitment of TFs involved in initiation is required for promoters to function as enhancers and suggest partial functional redundancy or cooperativity among these TFs in supporting enhancer activity⁴⁵.

Because enhancers are enriched for pioneer factors and chromatin remodelers, we next asked whether disrupting pioneer-factor binding sites would compromise enhancer function. We focused on GATA1, a pioneer factor strongly enriched in the strong enhancers (cluster 7) and essential for enhancer activity in K562 cells^{18,46,47}. The strong enhancer (enhancer_93, cluster 7) we tested contains two GATA1 pioneer factor motifs (GATAA) (Fig. 4e)⁴⁸. We mutated either one or both motifs, paired each enhancer variant with *MYC* or *ADGRG5* target promoter, and integrated them into the same landing pad locus. EGFP measurements showed that mutation of either a single GATA1 motif or both motifs similarly reduced enhancer activity, regardless of the target promoter (Fig. 4g, Extended Data Fig.10c). To assess the impact on chromatin accessibility, we performed DNase-PCR at the *ADGRG5* promoter (Extended Data Fig.9b). Compared with the wild-type

enhancer, all GATA1-mutant variants yielded stronger PCR signals after DNase digestion, indicating reduced chromatin accessibility. Together, these results demonstrate that intact GATA1 motifs are required for both chromatin opening and transcription activation by strong enhancers, suggesting that GATA1 contributes not only to chromatin remodeling but also to the recruitment of additional cofactors necessary for enhancer function^{49,50}.

Promoters and enhancers exert overlapping but distinct regulatory roles at their native genomic loci

Given our observation that promoters can exhibit enhancer-like activity in our chromatin-integrated assay, we next asked whether promoters also regulate distal gene expression in their native genomic context. To address this, we analyzed published CRISPRi screening data in K562 cells in which approximately 5,700 enhancers and 350 promoters were individually silenced and the resulting changes in expression on genes within a 1-Mb window were quantified by single-cell RNA sequencing⁵¹. We assessed changes in gene expression as a function of genomic distance (from 20 kb to 1 Mb) following CRISPRi-mediated silencing of enhancers or promoters and visualized these effects using volcano plots and bar plots (Fig. 5a,b).

Strikingly, promoters and enhancers exhibited broadly similar regulatory impacts on the expression of neighboring genes, although clear distinctions between the two classes of TREs were also evident. Both enhancers and promoters frequently acted as activators of distal genes, as indicated by decreased target gene expression upon CRISPRi silencing of the corresponding TRE. For both element types, the fraction of activating interactions declined with increasing genomic distance, consistent with a distance-dependent mode of regulatory influence (Fig. 5b). In addition to their activating roles, a subset of enhancers and promoters also exhibited negative regulatory effects, such that silencing these elements resulted in upregulation of nearby genes. Notably, across all distance thresholds examined, promoters were substantially more likely than enhancers to display such negative regulatory behavior, whereas the overall fraction of activating promoters was comparable to that of enhancers. These findings indicate that although promoters and enhancers can both function as gene activators *in vivo*, promoters are disproportionately enriched for negative regulatory activity. This distinction highlights fundamental differences in the regulatory logic of promoters and enhancers and suggests that promoters may play a broader dual role in modulating gene expression, a property that warrants further mechanistic investigation.

Promoter's regulatory activity is partially determined by transcription features

As promoters were more likely than enhancers to function as negative regulators at their native genomic loci, we then examine promoter features that might predispose a promoter toward positive or negative regulatory roles. Earlier models suggested that when transcription machinery is limited, promoters may compete rather than share transcription machinery, thereby negatively regulating nearby genes⁵². To test this hypothesis, we analyzed PRO-seq derived pausing and elongation metrics for promoters whose CRISPRi-mediated silencing caused up- or down-regulation of target

genes. Promoters that produce transcripts inefficiently, i.e., those in the top quartile of pause index or bottom quartile of gene body signal were significantly less likely to function as negative regulators (Fig. 5c). Collectively, these results indicate that promoters with elevated pausing and low transcription output, properties more similar to enhancers, are less likely to behave as negative regulators. This pattern aligns with a competition model in which highly productive promoters sequester limiting transcription machinery, thereby suppressing the activity of neighboring promoters. Conversely, promoters characterized by strong pausing and weak elongation appear less competitive for transcription machinery and are more likely to exert positive regulatory effects. At the genomic scale, this balance between cooperation and competition enables TREs to work together to determine transcription outcomes.

Discussion

In this study, we developed a chromatin-integrated landing pad-based assay to systematically interrogate regulatory interactions between TREs and target promoters in a native chromatin context^{18,19}. Unlike prior plasmid-based approaches that focused primarily on enhancer-promoter interactions, our screen involves simultaneous analysis of promoter-enhancer and promoter-promoter interactions in a chromatin context, revealing general principles of how TREs cooperate to regulate transcription^{3,5-7}.

Our analyses show that promoters and enhancers differ markedly in their recruitment of transcription machinery. Promoters preferentially engage transcription factors and core transcription components, including Pol II, general transcription factors, pausing, and elongation factors, whereas enhancers are enriched for pioneer factors and chromatin remodelers that promote nucleosome displacement. These distinctions, visible at endogenous loci through chromatin accessibility and nascent transcription profiles, are expected to be at least partially preserved when TREs are relocated to our landing pad, particularly for the transcription factors that recognize and bind specific DNA motifs. Despite these differences, both classes readily act as enhancers in our assay when paired with compatible target promoters. For target promoters that are intrinsically accessible, enhancer potency of enhancer- or promoter-TREs tend to scale with the overall degree of binding of transcription factors. Conversely, target promoters residing in inaccessible chromatin required TREs enriched for pioneer factors and remodelers to achieve activation, suggesting that chromatin opening constitutes a dominant rate-limiting step for this promoter subset.

These observations align with a stepwise model of transcription regulation, in which chromatin opening, PIC assembly, initiation, promoter-proximal pausing, and pause release provide points in the pathway that can be regulated⁵³. The earliest major barrier is the generation of a nucleosome-free region by pioneer factors and remodelers, followed by PIC assembly in that region^{54,55}. Once a PIC forms, initiation proceeds almost immediately, yielding short transcripts before Pol II stalls in a paused state^{56,57}. Efficient pause release, mediated by P-TEFb phosphorylation of Pol II and

pausing factors such as NELF and DSIF, represents the next critical checkpoint leading to productive elongation⁵⁸. Our data suggests that different promoters are constrained at different steps of this pathway. Some promoters are limited at the earliest step due to an inability to recruit chromatin-opening factors, whereas others may be limited at later stages such as PIC assembly or pause release. TREs that recruit complementary transcription machinery can enhance transcription at these target promoters by sharing those missing factors with them. However, the current experimental setup does not allow us to directly identify which rate-limiting step constrains each promoter, and future studies will be required to dissect the contributions of individual transcription processes and factors.

Consistent with the model in which promoter activity relies on cooperation with surrounding TREs, we find that when promoters are removed from their native chromatin environments and placed into a standardized landing pad with minimal flanking TREs, their transcription output frequently decreases from levels inferred from endogenous PRO-seq profiles. Promoters with strong native transcription often show markedly attenuated activity in isolation, underscoring that promoter sequence alone is insufficient to specify transcription potency and that local TRE and chromatin context exerts dominant control. Using our CIERA-seq system, we find that target promoters can be stimulated not only by canonical enhancers but also by neighboring promoters. This extends recent plasmid-based evidence that promoters can adopt enhancer-like roles and suggests that TRE–TRE crosstalk is a general feature of transcription regulation. Complementary CRISPRi perturbations at endogenous loci further support this conclusion: promoter silencing can reduce expression of nearby genes, indicating that promoters can activate distal targets *in vivo* alongside bona fide enhancers⁵⁹.

CRISPRi profiling also uncovers an additional dimension: promoters are substantially more likely than enhancers to function as repressors. This distinction aligns with their biological roles. Promoters sustain engagement of large amounts of transcription machinery to generate long transcripts, and thus are prone to monopolizing limiting transcription machinery, effectively reducing their availability for nearby genes. Enhancers, which produce less stable transcripts and require less sustained Pol II engagement, are less likely to exert such sequestration effects. Notably, we did not observe promoters functioning as negative regulators in our CIERA-seq. The likely explanation is that in our system, promoters are decoupled from their long primary transcripts. Therefore, they do not produce stable transcripts and are not sequestering Pol II and its associated factors as efficiently as native promoters.

Together, these findings support a framework in which transcription output is governed predominantly by interactions among TREs rather than by the intrinsic strength of any single element (Fig. 5d). Promoters and enhancers recruit overlapping but distinct subsets of transcription machinery, enhancers favor pioneer factors and remodelers, while promoters engage initiation and pause-release complexes. When multiple TREs coexist, they may either cooperate to accelerate

rate-limiting steps or compete for limiting factors and decelerate such steps, with outcomes dictated by their relative recruitment capacity and compatibility. Across the genome, this balance between cooperation and competition enables TREs to work together to tune transcription outcomes.

Methods

Genome build

All genomic analyses were performed using the human genome assembly GRCh38/hg38 as the reference.

Construction of Bxb1 recombination vectors

The donor vector for Bxb1-mediated recombination was derived from a modified eSTARR-seq plasmid and the attB1-EF1a-EGFP-attB2 construct described previously^{8,18}. The eSTARR-seq backbone was digested with AflIII (NEB, R0541S) and MfeI (NEB, R3589S) to remove the luciferase reporter and the SV40 polyadenylation signal. In parallel, the EGFP-polyA cassette was PCR-amplified from attB1-EF1a-EGFP-attB2 using primers designed to introduce homology arms compatible with the digested backbone. The forward primer annealed upstream of EGFP and provided homology to the 5' end of the cleaved vector. The reverse primer annealed downstream of the SV40 polyA region and introduced a unique 9-bp randomized barcode, enabling unambiguous identification of each target promoter during sequencing. The PCR product was assembled into the digested backbone using Gibson Assembly® Master Mix (NEB, E2611), and the resulting constructs were transformed into One Shot™ ccdB Survival™ 2 T1R competent cells (Invitrogen, A10460). Transformants were selected on LB agar containing ampicillin and chloramphenicol. Sixteen colonies were picked, plasmid-prepped, and Sanger-sequenced to confirm the presence of distinct 9-bp barcodes lacking transcription-factor motifs or canonical regulatory elements. These barcoded plasmids were subsequently digested with AvrII (NEB, R0174S) and AseI (NEB, R0526S) to allow cloning of sixteen different target promoters.

Target promoters cloning

Target promoters (sequences available in Supplementary Table 2 (target promoter info)) were PCR-amplified from K562 genomic DNA using PrimeSTAR® GXL DNA Polymerase (TaKaRa, R050B) for 32 cycles with primer pairs engineered to introduce AvrII (CCTAAG) and AseI (ATTAAT) restriction sites, with the order of sites determined by whether the primary transcript of the promoter lies in the sense or antisense orientation (primer sequences are provided in Supplementary Table 3 (primers)). Promoter fragments were designed to extend at least 60 bp downstream of the divergent transcription start site to encompass promoter-proximal pausing regions. PCR products were digested with AvrII (NEB, R0174S) and AseI (NEB, R0526S) at 37 °C for 1 h, resolved by electrophoresis on a 1% agarose gel, and purified using a QIAquick Gel Extraction Kit (QIAGEN, 28704). Purified inserts were ligated into barcode-indexed donor vectors

(described above) using T4 DNA Ligase (NEB, M0202S) at 16 °C overnight, transformed into the same competent cells as before, plated on the same selective plates as before, and plasmids were recovered using a QIAprep Spin Miniprep Kit (QIAGEN, 27104).

TRE library cloning

TREs were selected from those characterized in Tippens *et al.*⁸. In total, 354 TREs were chosen, including approximately 142 enhancers, 128 promoters, 50 untranscribed elements, and 34 open reading frame–based negative controls (see Supplementary Table 1 (TRE info) for sequences and genomic coordinates). All TREs were available in a pooled entry plasmid library. For each target promoter, the TRE pool was transferred into the corresponding attB1–promoter–EGFP–barcode–attB2 backbone via LR recombination. Briefly, LR reactions contained 10 µL of the pooled entry plasmids (100 µg µL⁻¹), 10 µL TE buffer, 10 µL of the matched attB1–target-promoter–EGFP–barcode–attB2 plasmid (100 µg µL⁻¹), and 10 µL LR Clonase, and were incubated at 16 °C overnight. Reaction products were transformed into *E. coli* TOP10 cells (5 µL LR reaction per 50 µL cells) and plated on LB–ampicillin agar plates. Colonies were collected by scraping (Avantor, 734-2602), and plasmids were purified using the E.Z.N.A.® Endo-Free Plasmid DNA Maxi Kit (Omega, D6926-00S). The resulting constructs constitute the attB1–target-promoter–EGFP–barcode–TRE–attB2 library.

Genomic integration of target promoter–TRE constructs

For each of the 16 target promoters, 10 µg of the corresponding attB1–target-promoter–EGFP–barcode–TRE–attB2 plasmid was mixed with 10 µg of pFL9_pCAG-NLS-HA-Bxb1 (Addgene 51271). DNA mixtures were concentrated by speed vacuum to ~10 µL to maximize electroporation efficiency. Constructs were delivered into K562 landing-pad cells using the Cell Line Nucleofector® Kit V (Lonza, VCA-1003) and a Lonza Nucleofector 2b device, electroporating 10 million cells per reaction (tenfold above the manufacturer’s recommended cell number). Following electroporation, cells were recovered for 10 min in outgrowth medium composed of 70% fresh RPMI 1640 + GlutaMAX (Gibco, 61870036), 10% FBS (Avantor, CNS89510-194), and penicillin–streptomycin (Corning, 30-001-CI) mixed with 30% conditioned growth medium of identical composition. Cells were plated into 10-cm dishes and cultured for 1 week. Successfully recombined cells were isolated by fluorescence-activated cell sorting (Sony MA900), gating for loss of BFP expression from the landing-pad cassette. Integration efficiencies of ~10% were routinely obtained. To maintain representation of the TRE library for each promoter, at least 5×10^5 BFP-negative cells were collected. Sorted cells were expanded for an additional two weeks to allow degradation of residual plasmid DNA prior to downstream assays.

Sorting cells into EGFP-expression bins

Following three weeks of culture in RPMI supplemented with 10% FBS and penicillin–streptomycin, cells harboring target promoter–TRE integrations were pooled (5×10^6 cells from each promoter pool; 0.8×10^8 cells total), washed twice with 100 mL ice-cold PBS (Gibco,

10010023), resuspended in 10 mL, and filtered through Falcon® round-bottom tubes with 35- μ m cell-strainer caps (Corning, 352235). Cells were sorted on a Sony MA900 cell sorter based on EGFP fluorescence intensity. Seven gates were drawn to correspond to defined percentiles of the EGFP expression distribution, and approximately 5×10^5 cells were collected per bin. Genomic DNA from each sorted population was isolated using the DNeasy Blood & Tissue Kit (QIAGEN, 69504) and stored for downstream library preparation.

Library preparation from genomic DNA of EGFP-expression bins

Genomic DNA from each fluorescence-sorted bin was used as template for PCR amplification of cassette fragments containing both the target-promoter barcode and the integrated TRE sequence. A forward primer binding upstream of the promoter barcode (seq2-EGFP-barcode-2-F) and a reverse primer annealing downstream of the genomic Bxb1 integration junction on the genomic DNA but not on plasmids (seq1-ZZ103-R) were used to amplify the integrated cassette for 10 cycles (25 μ L Q5® High-Fidelity 2X Master Mix, NEB M0492S; 1 ng genomic DNA; 1 μ L 10 mM of each primer; H₂O to 50 μ L total volume). All genomic DNA recovered from each bin was used to ensure sufficient library complexity. Following amplification, reactions were treated with Exonuclease I (NEB M0293S; 50 μ L PCR product + 6 μ L ExoI buffer + 3 μ L H₂O + 1 μ L ExoI) to remove residual primers, then purified with the DNA Clean & Concentrator-5 kit (Zymo, D4014) and eluted in 20 μ L H₂O. Purified products were subsequently re-amplified for 15 cycles to append Illumina P5 and P7 adapters using primers P5-seq1 and P7-barcode-seq2. Final libraries were resolved on a 6% TBE polyacrylamide gel, and the band corresponding to the expected size was excised, crushed, and eluted overnight in TE buffer at 4 °C. Gel eluates were filtered through Spin-X 0.45- μ m centrifuge tubes (Costar, 8163) and purified by phenol–chloroform extraction, then eluted in 20 μ L H₂O. Library concentrations were quantified by Qubit fluorometry, and equal mass from each expression bin was pooled and submitted for 2×150 bp paired-end sequencing (Novogene).

Sequencing alignment and computational analysis

Adaptor trimming was performed using fastp, yielding paired outputs: Read 2 sequences containing the 9-bp barcodes that encode target promoter identity, and Read 1 sequences corresponding to TRE inserts. Each read set was aligned with Bowtie2 to custom reference FASTA files consisting solely of the expected promoter barcodes or TRE sequences, respectively. For every paired-end read, the mapped promoter barcode and TRE identifier were joined and tallied to generate counts of each target promoter–TRE combination within each bin. For each bin, raw counts for all promoter–TRE pairs were summed to obtain the total number of aligned reads. The abundance of each pair was expressed as a fraction of this bin total. Fractions across all seven bins were normalized such that values for each target promoter–TRE pair summed to one, yielding a distribution that reflects its relative representation across the EGFP bins (Extended Data Fig.1a). To quantify regulatory activity, each bin-specific normalized fraction was multiplied by the

median EGFP fluorescence associated with that bin, and weighted values were summed to generate a single activity score per target promoter–TRE pair.

K-means clustering of TREs based on TF ChIP-seq signals

ENCODE ChIP-seq datasets downloaded from the UCSC Genome Browser were used as the foundation for analysis^{28,33,34}. ChIP-seq signals for all available transcription factors (TFs) in K562 cells were used, and TF occupancy within each TRE region was extracted. K-means clustering was then performed on TREs based on their TF enrichment profiles and was optimized to resolve seven clusters with distinct TF-binding signatures. Key TFs contributing to cluster separation were selected, and z-scores were calculated to normalize their enrichment across TREs. Heatmaps of z-scores were generated to visualize the differential enrichment patterns of these TFs among clusters.

Metaplot analysis of chromatin and transcription features

Metaplots of chromatin accessibility and histone modification (DNase-seq (GSM816655), H3K27ac (GSM733656), H3K4me3 (GSM788087), and H3K4me1 (GSM788085)), nascent transcription (PRO-cap and PRO-seq), and selected transcription factor ChIP-seq datasets were generated in R using the GenomicRanges, IRanges, rtracklayer, tibble, dplyr, tidyr, and ggplot2 packages. For all datasets except PRO-seq, signal was extracted in a ± 500 bp window centered on each TRE to capture proximal regulatory activity. For PRO-seq, a ± 5 kb window was used to quantify transcription output across the gene body. For each TRE category, signals were aggregated and normalized to the maximum value observed within the window, enabling comparison across datasets on a standardized 0–1 scale. ChIP-seq tracks for individual transcription factors were downloaded from ENCODE and processed using the same ± 500 bp genomic boundaries.

Gene ontology analysis of transcription factors enriched at promoters and enhancers

Gene ontology analysis of transcription factors differentially enriched between promoters and enhancers was performed using clusterProfiler, org.Hs.eg.db, and enrichplot in R, with a false-discovery rate threshold of $q \leq 0.05$. Representative Gene Ontology terms in the Biological Process and Molecular Function categories were selected for visualization. Network representations were generated using ggraph, ggplot2, and grid, with nodes corresponding to GO terms and transcription factors contributing to each term.

DNase I treatment and DNA purification

DNase I digestion was performed using a protocol modified from He *et al.*⁴¹. Five million K562 cells were collected and washed once in PBS at 900 rpm for 3 min. Cells were resuspended in 3 mL Buffer A⁺ (15 mM Tris-HCl pH 8.0, 15 mM NaCl, 60 mM KCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.15 mM spermine, 0.5 mM spermidine, 1 \times protease inhibitor cocktail, 1 mM PMSF, 0.5 mM DTT), split into two 1.5-mL aliquots—one for digestion and one undigested control—and transferred to 15-mL tubes. To permeabilize nuclei, 0.5 mL of 0.2% NP-40 in Buffer

A (15 mM Tris-HCl pH 8.0, 15 mM NaCl, 60 mM KCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.2% NP-40) was added dropwise, and samples were incubated on ice for 10 min. Cells were pelleted at 2,500 rpm for 3 min at 4 °C, the supernatant removed, and the pellet washed first in 0.5 mL Buffer A⁺, then in 0.5 mL Buffer A, with centrifugation as above after each wash. Pellets were resuspended in 0.5 mL prewarmed DNase I digestion buffer (6 mM MgCl₂, 10 mM NaCl, 1 mM CaCl₂, 40 mM Tris-HCl pH 8.0). Twenty-five units of DNase I (Thermo Scientific, EN0521) were added to the digestion sample, and both digested and control tubes were incubated at 37 °C for 5 min. Digestion was terminated by adding 0.5 mL stop buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1% SDS, 100 mM EDTA pH 8.0, 0.3 mM spermine, 0.1 mM spermidine) and 2 µL RNase A, followed by incubation at 55 °C for 15 min. Proteinase K (2 µL of 20 mg mL⁻¹) was added and samples incubated for an additional 2 h. DNA was purified by sequential extractions with equal volumes of phenol–chloroform and chloroform alone, followed by ethanol precipitation. Final pellets were resuspended in 100 µL nuclease-free water. All results reported in this section were confirmed in at least two biological replicates.

PCR test of the DNase I treated samples

Genomic DNA from DNase I–treated and untreated control samples was diluted to 20 ng µL⁻¹, followed by two sequential 1:4 dilutions to generate input concentrations of 20 ng µL⁻¹, 5 ng µL⁻¹, and 1.25 ng µL⁻¹ for PCR testing. Accessibility was assessed at 3 control genomic loci using primers targeting the *GAPDH* promoter (accessible positive control), CTCF region (less accessible control), *MNDA* promoter (least accessible control), and at promoters of *MYC* and *ADGRG5*. PCR reactions (20 µL) were assembled using Q5® High-Fidelity 2× Master Mix (NEB, M0492L) with the following composition: 10 µL master mix, 1 µL each of 10 µM forward and reverse primers, 1 µL DNA template, and 7 µL H₂O. Thermal cycling conditions were: 1. 98°C for 3 minutes 2. 98 °C for 30 seconds 3. 60 °C for 30 seconds 4. 72 °C for 30 seconds 5. Goto step 2 for 24 more cycles 6. 72 °C for 10 minutes 7. 4°C forever). PCR products were resolved on a 6% TBE polyacrylamide gel at 100 V for 1 h and stained in 1× SYBR Gold (Thermo Fisher, S11494) for 15 min. Bands were visualized under UV illumination, and relative signal intensity between digested and undigested samples was used as a measure of chromatin accessibility: loci showing stronger depletion after DNase I treatment were interpreted as more accessible. ImageJ is used to quantify PCR band size and calculate relative band strength. All results reported in this section were confirmed in at least two biological replicates.

CRISPRi data analysis to identify regulatory effects of TRE silencing

CRISPR interference screening data from Gasperini et al. were analyzed to evaluate the effects of promoter and enhancer perturbation on expression of neighboring genes⁵¹. Fold-change and FDR values were extracted for all guide–gene pairs, and negative-control targeting regions were used as the null reference. Thresholds for both log₂(fold change) and false discovery rate were defined as the 95th percentile of the corresponding negative-control distributions. For each guide–gene pair located within 20 kb, 100 kb, 250 kb, or 1 Mb, log₂(fold change) values were plotted against

$-\log_{10}(\text{FDR})$ using R to generate volcano plots. Pairs exceeding both fold-change and FDR thresholds were classified as significant regulatory interactions. TREs were annotated as promoters or enhancers using the same criteria applied in the primary analyses: promoters were defined by overlap with annotated transcription start sites, whereas enhancers required both DNase I hypersensitivity and divergent PRO-cap transcription. For each genomic distance window, the fraction of promoters and enhancers whose silencing led to significant up- or down-regulation of nearby genes was quantified and visualized using bar plots.

Analysis of pause index and gene body transcription for promoter regulatory activity

A promoter's pause index was calculated using PRO-seq data. Transcription start sites (TSSs) were defined from PRO-cap peaks, and transcription end sites (TESs) for the corresponding genes of the promoter were obtained from GENCODE annotations. The pause region was defined as the first 0–150 bp downstream of the TSS, and PRO-seq read counts within this region were summed and normalized by region length to obtain read density per base pair. The gene body region was defined as +500 bp downstream of the TSS to –500 bp upstream of the TES, and PRO-seq coverage over this interval was likewise summed and normalized by region length. The pause index was calculated as the ratio of pause-region signal to gene-body signal. Promoters whose silencing via CRISPRi produced significant up- or down-regulation of other genes were stratified using these transcription metrics. Promoters in the top and bottom quantiles of pause index or gene-body signal were identified, and the proportion of promoters functioning as activators or repressors upon CRISPRi perturbation was summarized using bar plots.

Data Availability

All raw and processed data of the CIERA-seq can be found at NCB's Gene Expression Omnibus under accession number GSE316493.

Acknowledgements

This work was supported by the National Institutes of Health grant R01 HG012970 to J.T.L. and H.Y.

Author Contributions

Y.J. conceptualized the study with guidance from Z.Z., H.Y, and J.T.L. Y.J. generated the CIERA-seq data with the help from J.L. Z.Z. generated the landing pad containing cell line. A.K-Y.L selected the target promoters tested.

Competing interests

The authors declare no competing interests.

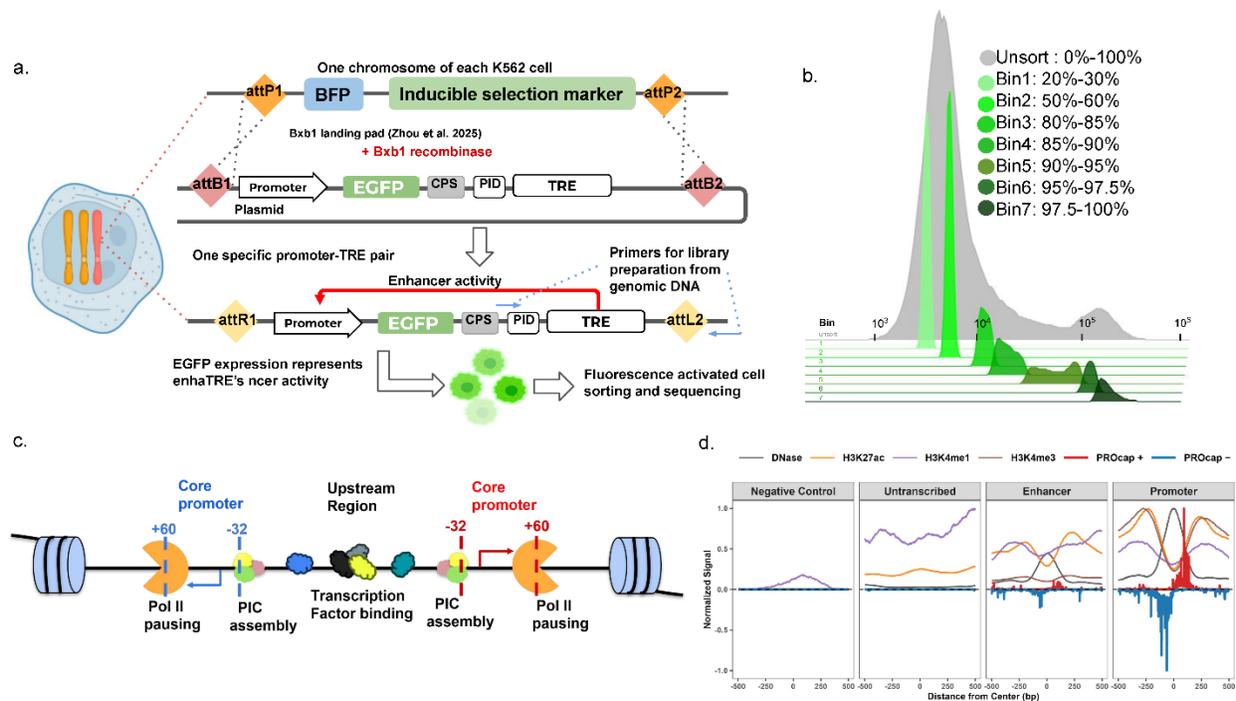


Fig. 1 | Chromatin-integrated assay to quantify promoter-TRE regulatory interactions.

a, Schematic of the chromatin-based landing pad assay used to measure the regulatory activity of >350 transcriptional regulatory elements (TREs paired with 16 promoters). (1) A Bxb1 landing pad is inserted at the enhancer locus of the NMU gene on one chromosome to generate a K562 landing pad cell line. (2) TREs are cloned into vectors containing Bxb1 recombination sites together with distinct promoters and corresponding promoter barcodes (PIDs). (3) Vectors are co-transfected into K562 cells with a Bxb1 recombinase expression plasmid. (4) Cells are cultured for 1 week and sorted for loss of BFP signal, which indicates successful integration (4) Cells are cultured for >3 weeks and sorted into bins based on EGFP expression. (5) Genomic DNA from each bin is extracted and used to generate Illumina sequencing libraries containing PID and TRE sequences with primers show in the illustration. (6) Sequencing data are analyzed as described in Extended Data Fig. 1a.

b, Representative FACS profile of K562 cells. Each cell has a single promoter-TRE pair integrant and gets sorted into seven bins based on EGFP expression. Gray indicates the unsorted population; colored peaks represent individual bins.

c, Definition of promoters and enhancers used in this study. Both promoters and enhancers are defined by divergent PRO-cap signals and representing two divergent core promoters, each defined by where the Pol II PIC begins at -32 bp and extends through the pause region to +60 bp from the TSS. The region between the two core promoters serves as an upstream regulatory region enriched for transcription factor binding. Promoters are defined as TREs overlapping annotated transcription start sites of protein-coding genes, whereas enhancers are located in intergenic regions or within introns.

d, Meta-profiles of PRO-cap, DNase-seq, and ChIP-seq signals for H3K27ac, H3K4me1, and H3K4me3 across TREs of different element types.

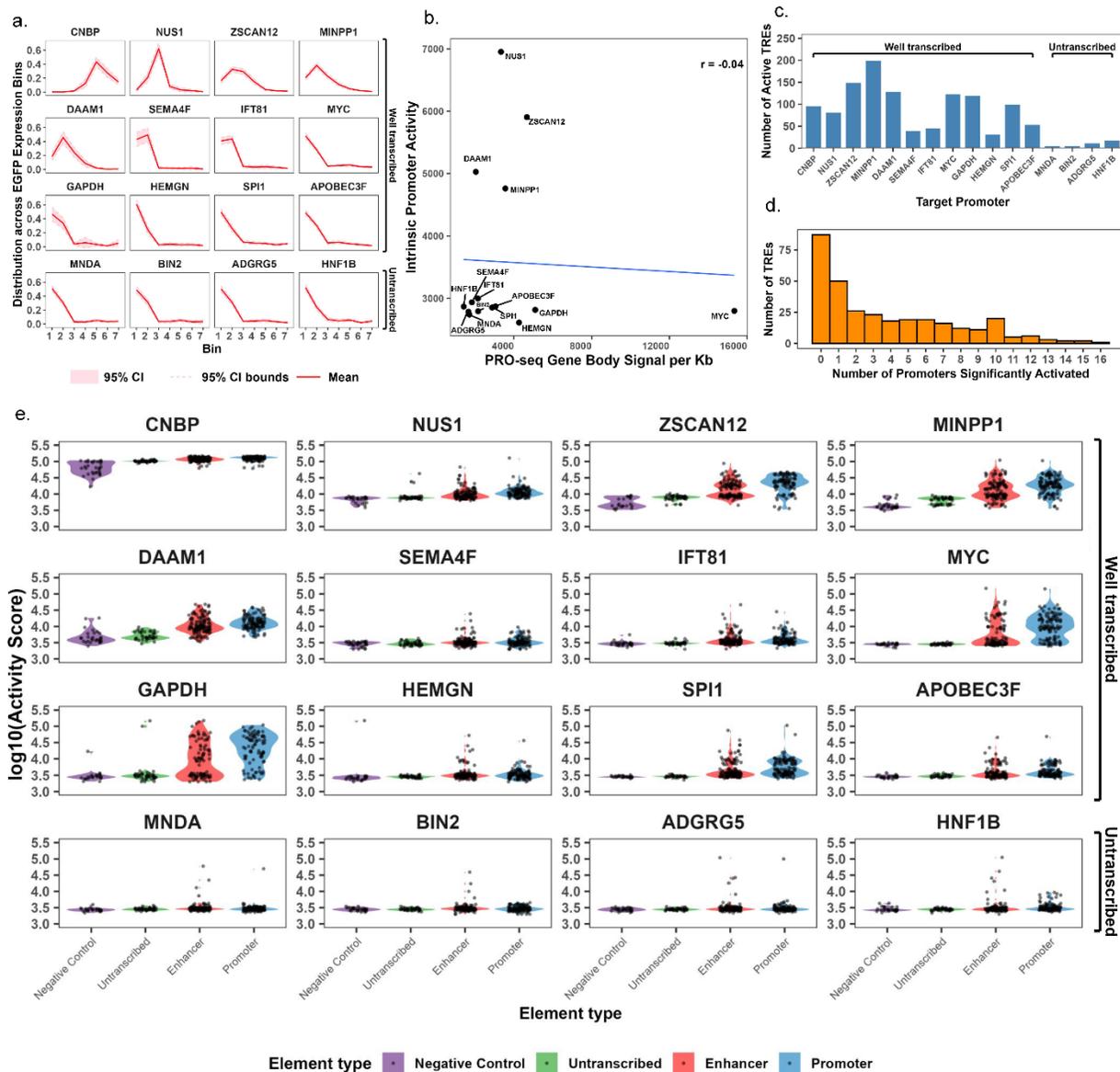


Fig. 2 | Definition and functional properties of promoters and enhancers.

a, Average normalized relative enrichment (red line) and 95% confidence interval (pink shading) across seven expression bins for each of the 16 promoters paired with 34 ORF negative controls, reflecting intrinsic promoter activity at the integration locus.

b, Correlation between intrinsic promoter activity score at the integration locus (from a) calculated based on distribution across the EGFP expression bins and PRO-seq gene body signal per kb at the native genomic locus.

c, Number of TREs capable of activating each of the 16 promoters.

d, Distribution of the number of promoters activated by individual TREs.

e, Violin plot showing the distribution of EGFP expression activity scores for the 16 target promoters when paired with TREs of different element types.

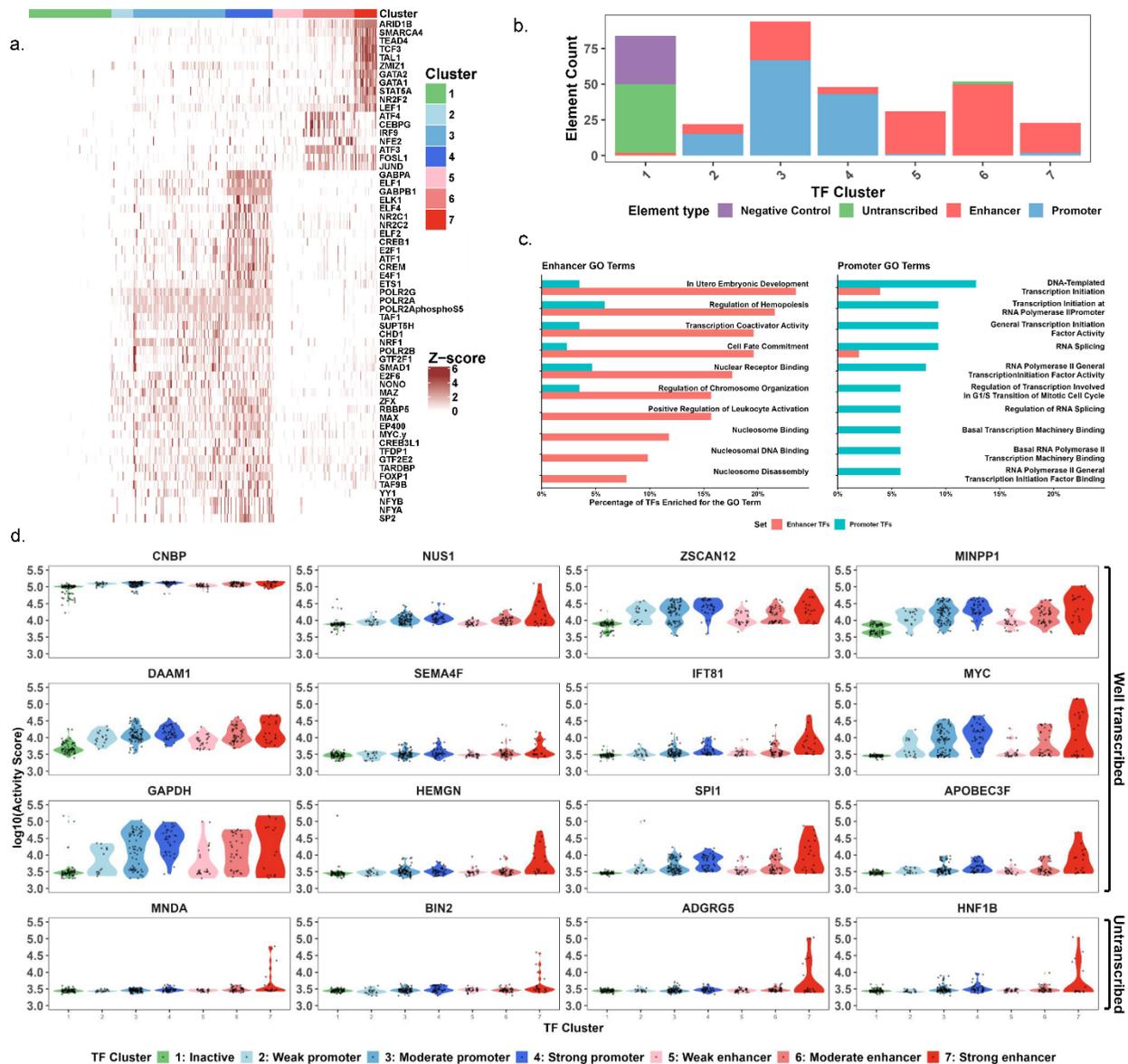


Fig. 3 | Transcription factor landscapes and functional clustering of TREs.

a, Heatmap of ChIP-seq signals for selected transcription factors with known activator roles across TREs. Columns represent individual TREs, ordered by k-means clustering of transcription factor binding profiles, and rows represent transcription factors.

b, Counts of TREs in each transcription factor-based k-means cluster.

c, Top Gene Ontology (GO) biological process (BP) and molecular function (MF) terms enriched among transcription factors preferentially enriched in enhancer or promoter dominated clusters.

d, Violin plot showing the distribution of activity scores for the 16 target promoters when paired with TREs from different transcription factor-based k-means clusters.

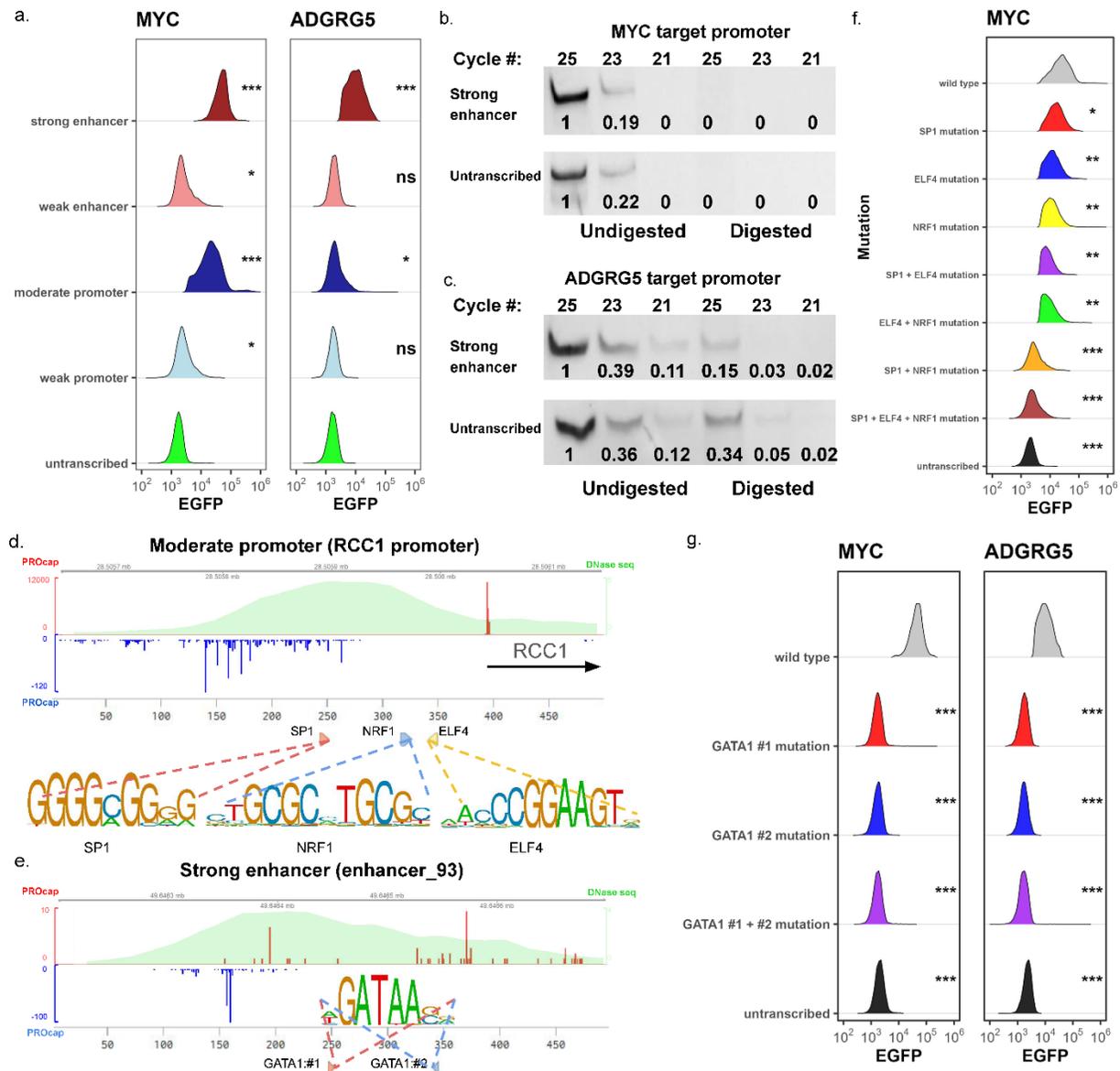


Fig. 4 | Distinct chromatin constraints underlie promoter- and enhancer-mediated activation.

a, EGFP expression distributions of cells harboring the *MYC* or *ADGRG5* target promoter paired with representative TREs from different transcription factor-based clusters: untranscribed TRE (untranscribed_4, cluster 1), weak promoter (*UBE2D3* promoter, cluster 2), moderate promoter (*RCC1* promoter, cluster 3), weak enhancer (enhancer_21, cluster 5), and strong enhancer (enhancer_93, cluster 7). The x-axis denotes EGFP intensity. Statistical significance was assessed using two one-sided tests (TOST) relative to the untranscribed TRE (*: $-\Delta L > 0.05$; **: $-\Delta L > 0.1$; ***: $-\Delta L > 0.2$).

b, PCR analysis of chromatin accessibility at the *MYC* target promoter paired with either an untranscribed TRE (untranscribed_4, cluster 1) or a strong enhancer (enhancer_93, cluster 7). Genomic DNA from samples that were undigested or digested with DNase I (25 U, 5 min) was amplified using increasing PCR cycle numbers. Relative PCR band size is quantified by ImageJ.

c, As in b, but for the *ADGRG5* target promoter paired with either an untranscribed TRE (untranscribed_4, cluster 1) or a strong enhancer (enhancer_93, cluster 7). Relative PCR band size is quantified by ImageJ.

d, PRO-cap and DNase-seq tracks of the *RCCI* promoter (cluster 3) shown in a, highlighting the locations of SP1, NRF1, and ELF4 binding motifs and their consensus sequences.

e, PRO-cap and DNase-seq tracks of the strong enhancer_93 (cluster 7) shown in a, highlighting the locations of two GATA1 binding motifs and their consensus sequence.

f, EGFP expression distributions of cells with the *MYC* target promoter paired with the *RCCI* promoter (cluster 3) shown in d, carrying different combinations of mutations in the SP1, NRF1, and ELF4 motifs. Statistical significance was assessed using two one-sided tests (TOST) relative to wild type(*: $-\Delta L > 0.05$; **: $-\Delta L > 0.1$; ***: $-\Delta L > 0.2$).

g, EGFP expression distributions of cells with the *MYC* or *ADGRG5* target promoter paired with the strong enhancer_93 (cluster 7) shown in e, carrying different combinations of mutations in the two GATA1 motifs. Statistical significance was assessed using two one-sided tests (TOST) relative to wild type(*: $-\Delta L > 0.05$; **: $-\Delta L > 0.1$; ***: $-\Delta L > 0.2$).

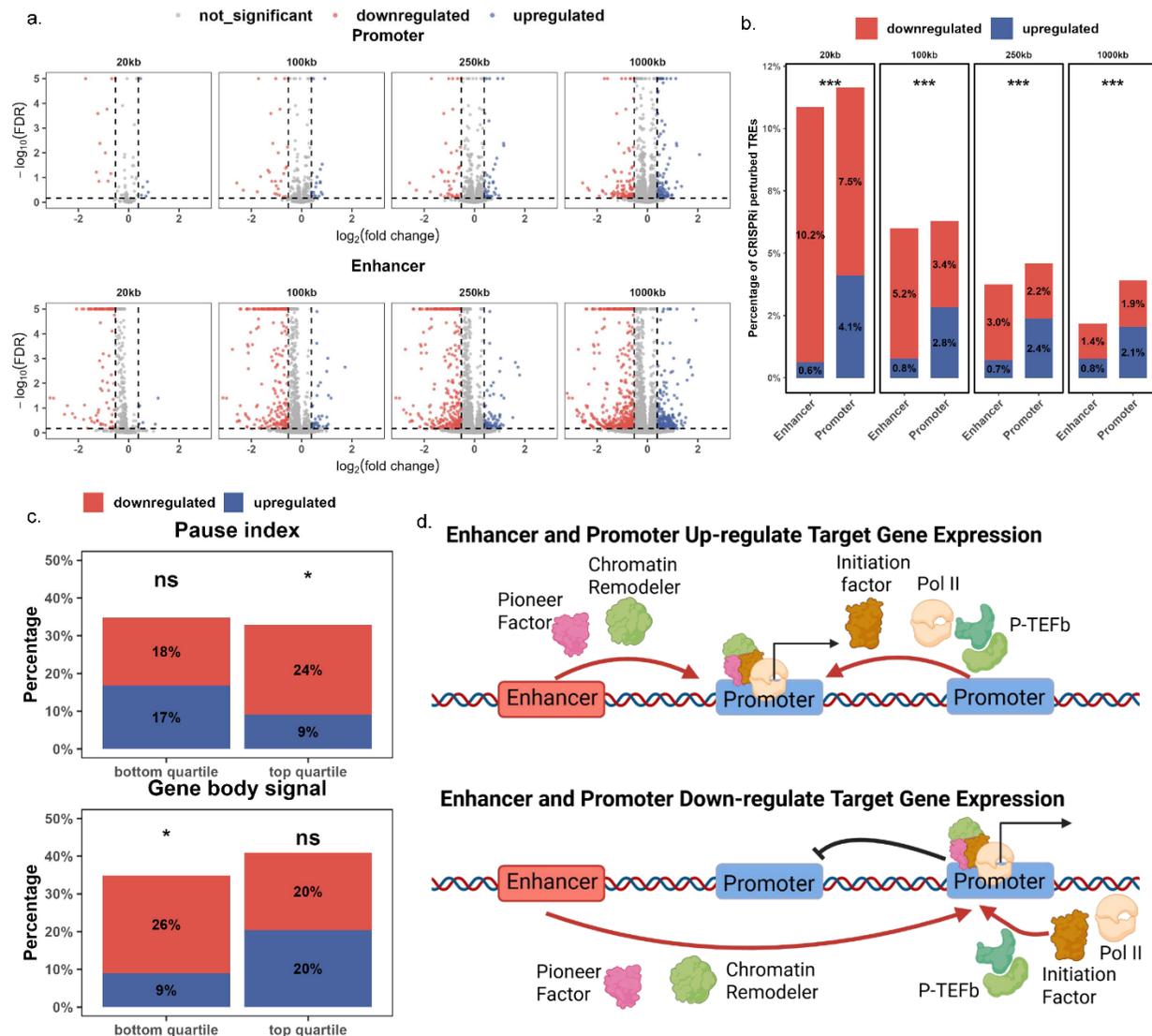


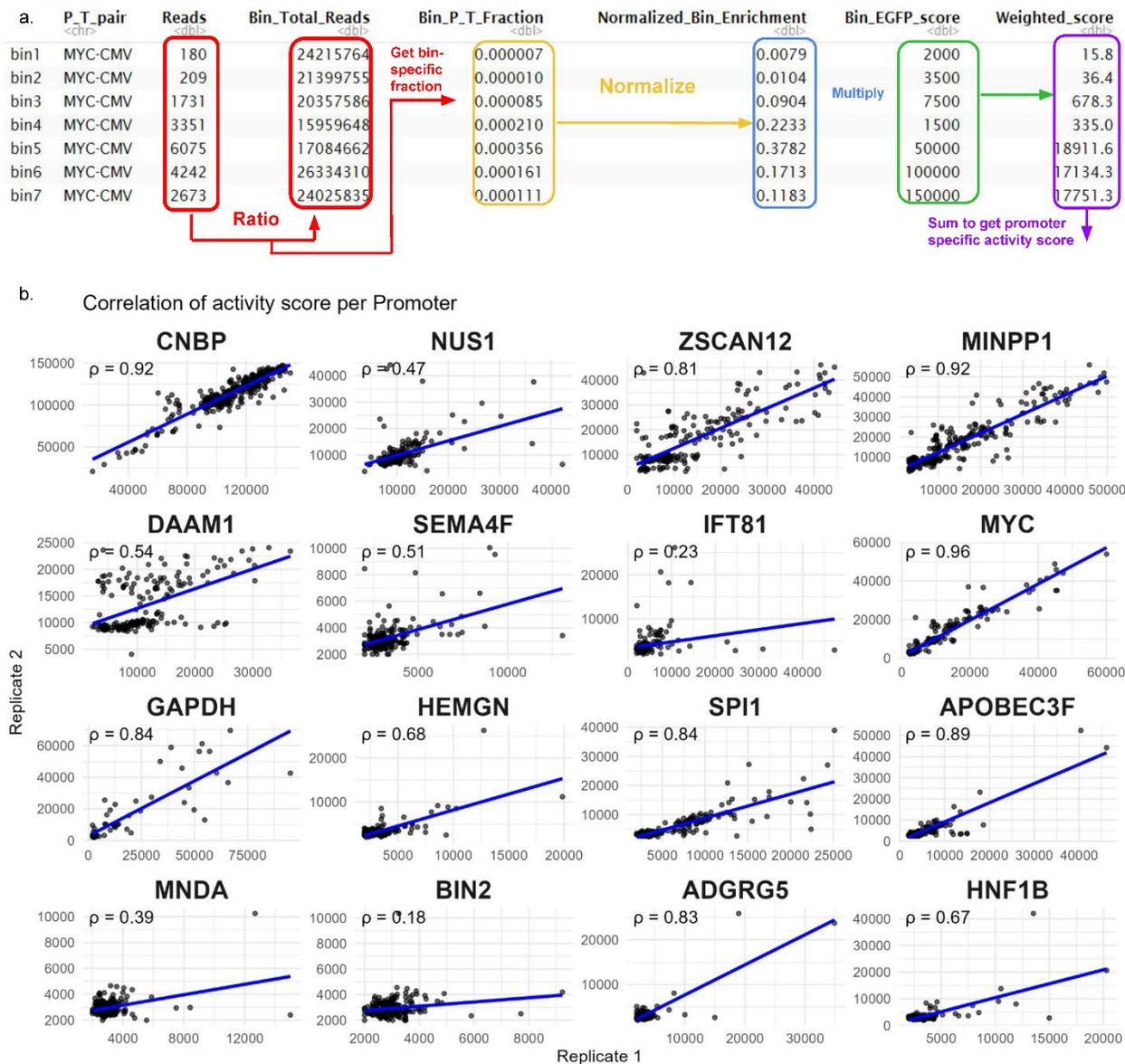
Fig. 5 | CRISPRi reveals that enhancers and promoters both act in a distance-dependent manner but with distinct regulatory outcomes

a, Volcano plots showing changes in target gene expression following CRISPRi-mediated silencing of promoters or enhancers located 20 kb, 100 kb, 250 kb, or 1 Mb from the target gene transcription start site.

b, Percentage of CRISPRi-perturbed promoters or enhancers within each genomic distance window (20 kb, 100 kb, 250 kb, or 1 Mb) that lead to up- or down-regulation of target gene expression. For each distance threshold, the proportions of up- versus down-regulated targets following CRISPRi silencing of enhancers and promoters were compared using a chi-square test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.005$).

c, Bar plots showing the percentage of promoters in the top or bottom quartile of pause index or gene-body (+500 bp from TSS to -500bp from TES) PRO-seq signal that up- or down-regulate target gene expression following CRISPRi silencing. Statistical significance was assessed using chi-square tests of the top and bottom quartiles relative to all promoters. (*, $p < 0.05$)

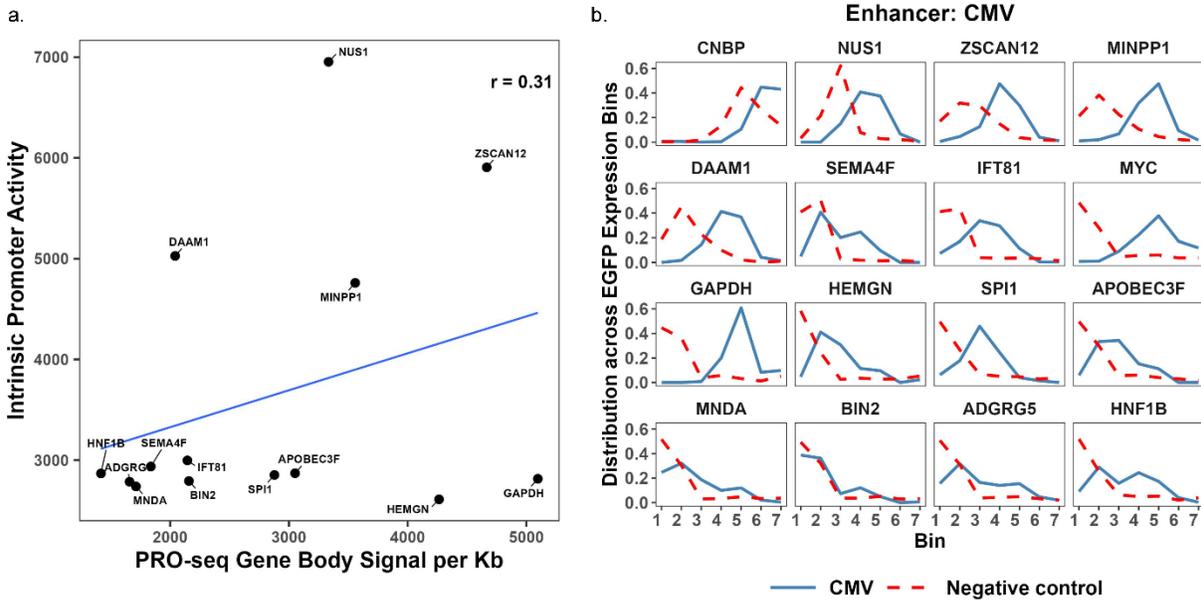
d, Schematic model illustrating how promoters and enhancers can act cooperatively to either increase or decrease transcription of nearby genes.



Extended Data Fig. 1 | Quantification of promoter–TRE activity from barcode sequencing.

a, Schematic of the data analysis pipeline using the *MYC–CMV* enhancer pair as an example. (1) For each expression bin, reads corresponding to a given promoter–TRE pair are divided by the total reads in that bin to calculate the fractional representation of that pair. (2) Fractions in bins 1–7 are divided by the fraction in the unsorted population to obtain relative enrichment for each bin. (3) Relative enrichment values across bins 1–7 are normalized to generate a distribution that enables comparison across promoter–TRE pairs. (4) An activity score is calculated as the sum of each bin specific normalized relative enrichment multiplies by the corresponding bin-specific EGFP score.

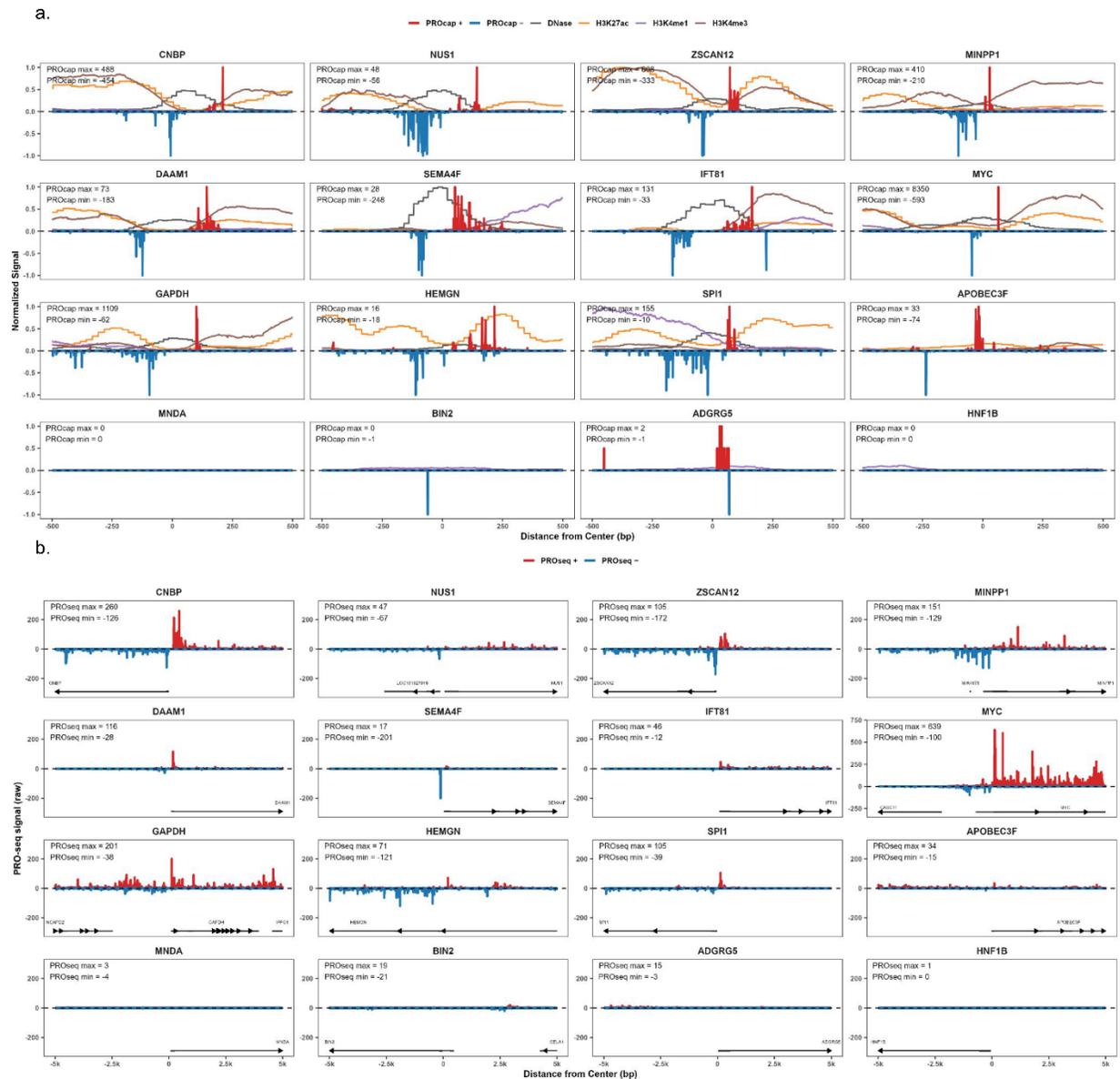
b, Correlation of activity scores for each promoter between biological replicate 1 (rep1) and biological replicate 2 (rep2). Reduced correlation between biological replicates for certain promoters was predominantly attributable to their weak responsiveness to TREs, which yielded low dynamic range in activity scores.



Extended Data Fig. 2 | Reproducibility of promoter activity measurements.

a, Correlation between intrinsic promoter activity at the integration locus (from c) and PRO-seq gene body signal per kb at the native genomic locus without *MYC*.

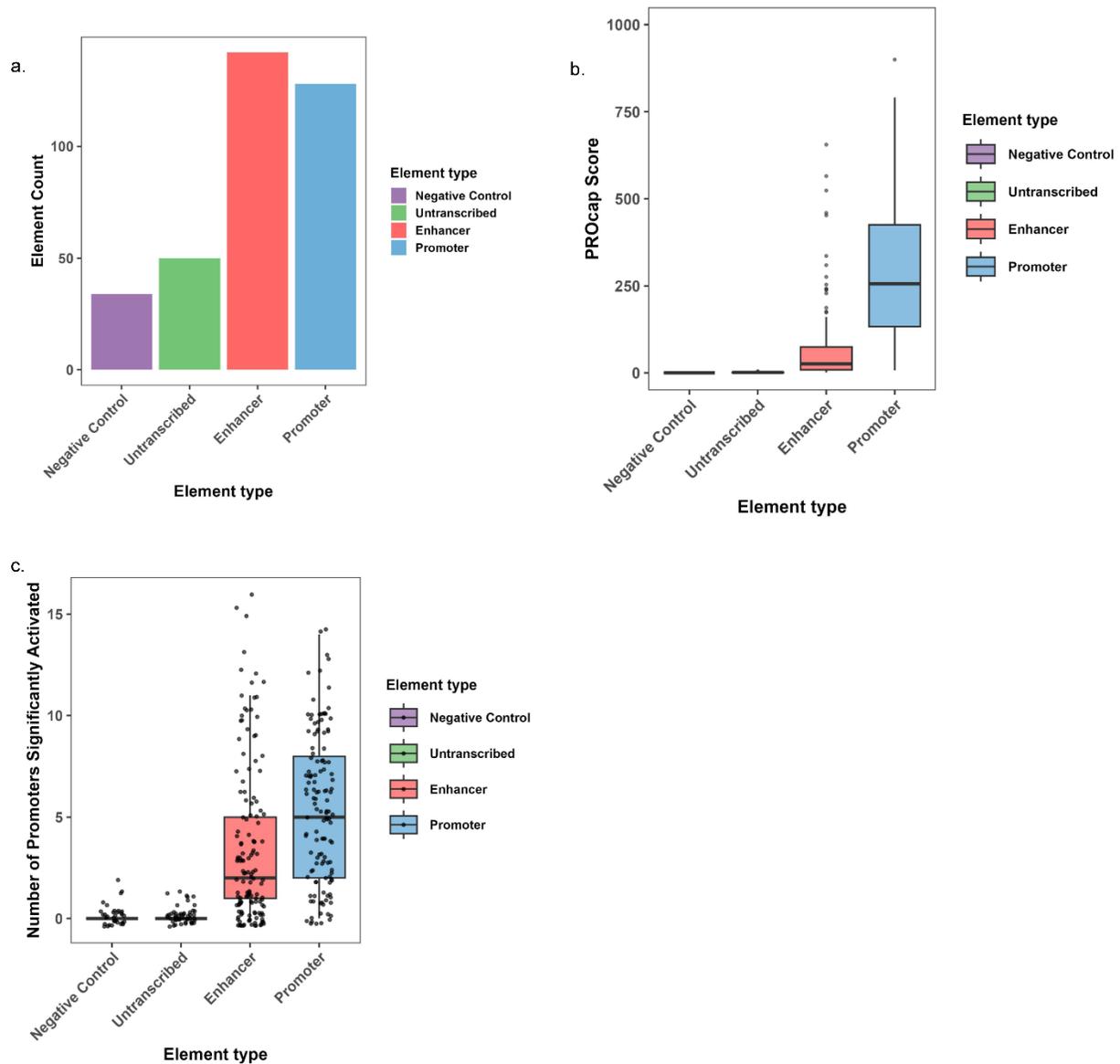
b, Normalized relative enrichment distributions of the 16 promoters when paired with the CMV enhancer. The dotted red line corresponds to the distribution shown in Fig. 2a.



Extended Data Fig. 3 | Chromatin and transcriptional profiles of target promoters.

a, Meta-profiles of PRO-cap, DNase-seq, and ChIP-seq signals for H3K27ac, H3K4me1, and H3K4me3 across the 16 target promoters at their native genomic loci, shown from -500 bp to $+500$ bp relative to the promoter center.

b, PRO-seq meta-profiles across the 16 target promoters, shown from -5 kb to $+5$ kb relative to the promoter center.

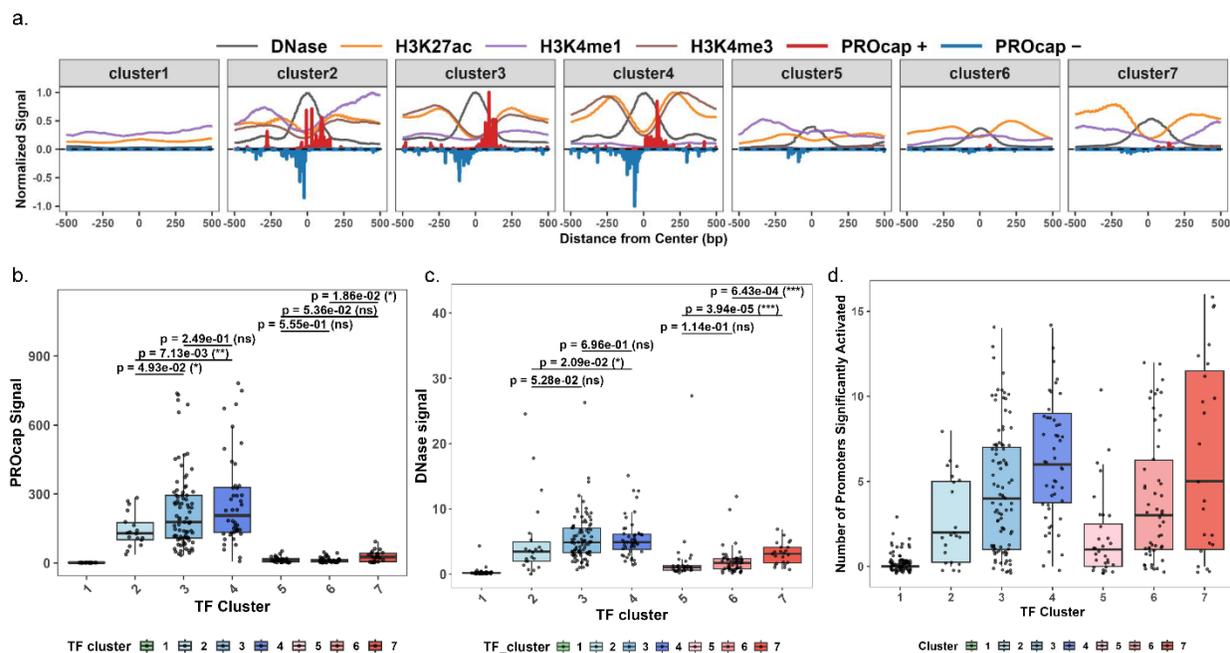


Extended Data Fig. 4 | Distinct regulatory capacities of TREs with different natures.

a, Counts of TREs classified by regulatory nature.

b, Boxplots of maximum PRO-cap signal across TREs of different regulatory natures.

c, Boxplots showing the number of promoters activated by TREs of different regulatory natures.



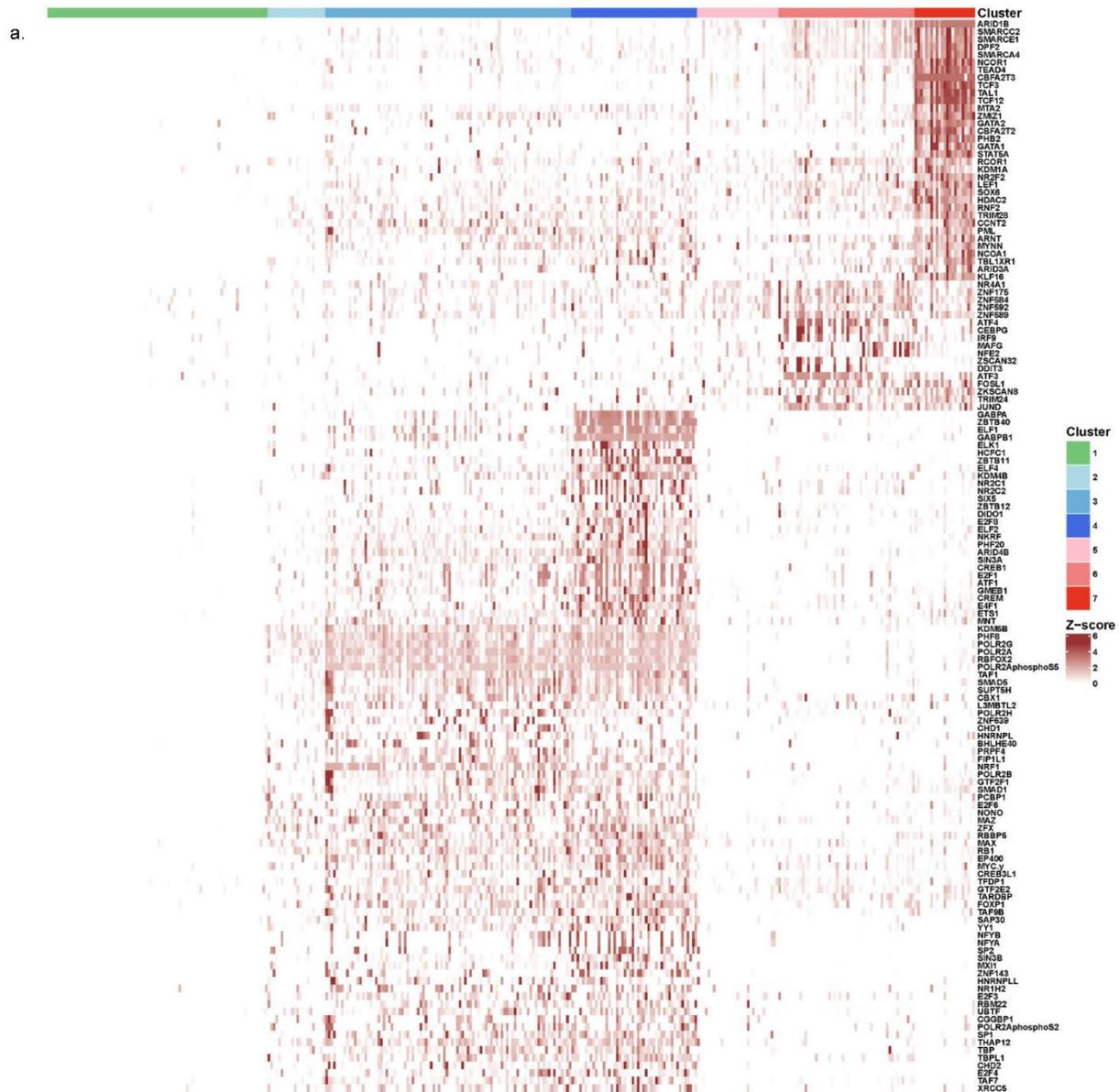
Extended Data Fig. 5 | Functional properties of TREs across transcription factor-based clusters.

a, Meta-profiles of PRO-cap, DNase-seq, and ChIP-seq signals for H3K27ac, H3K4me1, and H3K4me3 across TREs from different transcription factor-based k-means clusters.

b, Boxplots of maximum PRO-cap signal across TREs from each transcription factor-based k-means cluster; P values were calculated using the Wilcoxon rank-sum test (*: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.005$).

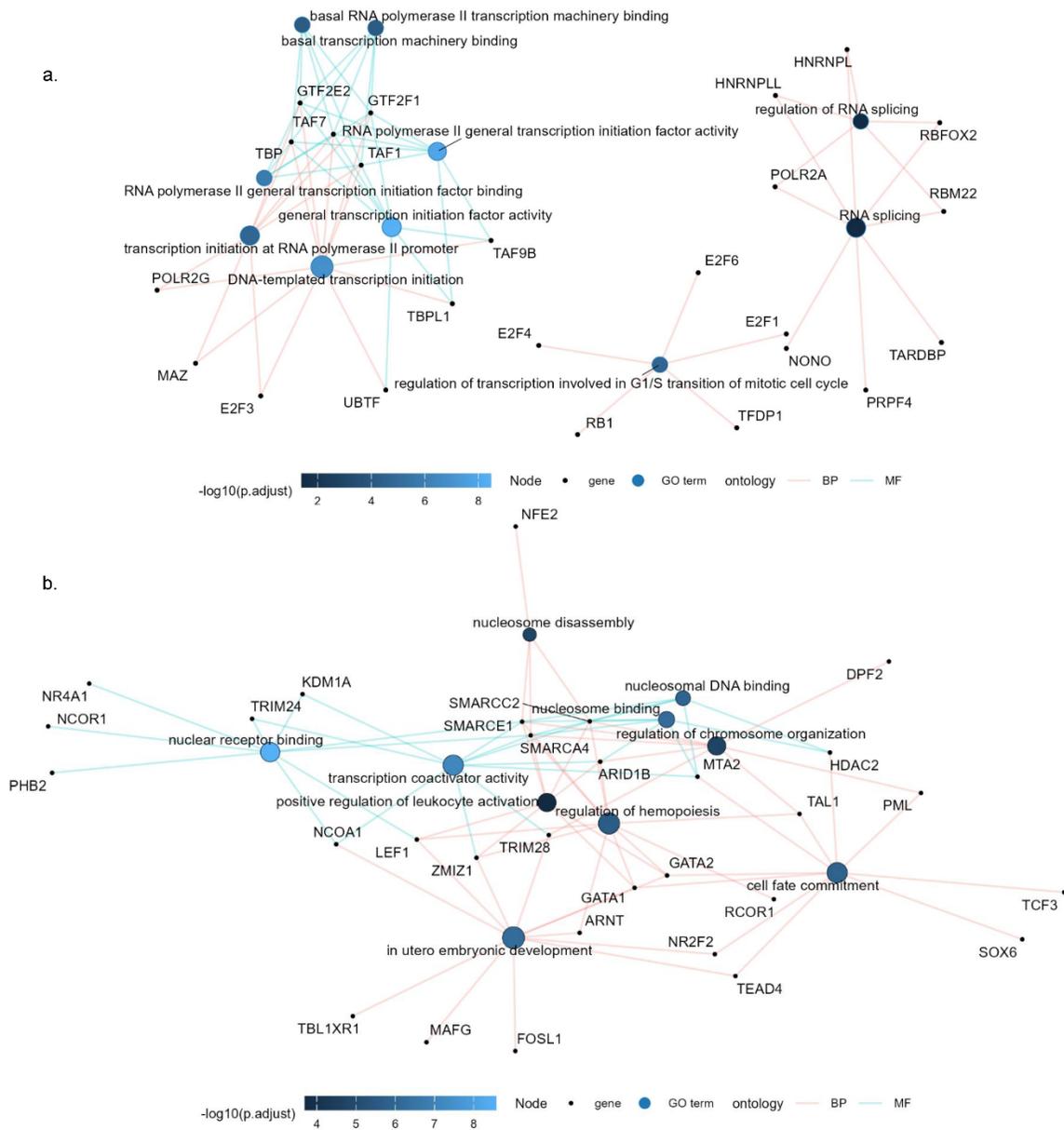
c, Boxplots of maximum DNase-seq signal across TREs from each transcription factor-based k-means cluster; P values were calculated using the Wilcoxon rank-sum test (*: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.005$).

d, Boxplots showing the number of promoters activated by TREs from each transcription factor-based k-means cluster.



Extended Data Fig. 6 | Transcription factor binding landscape across TREs.

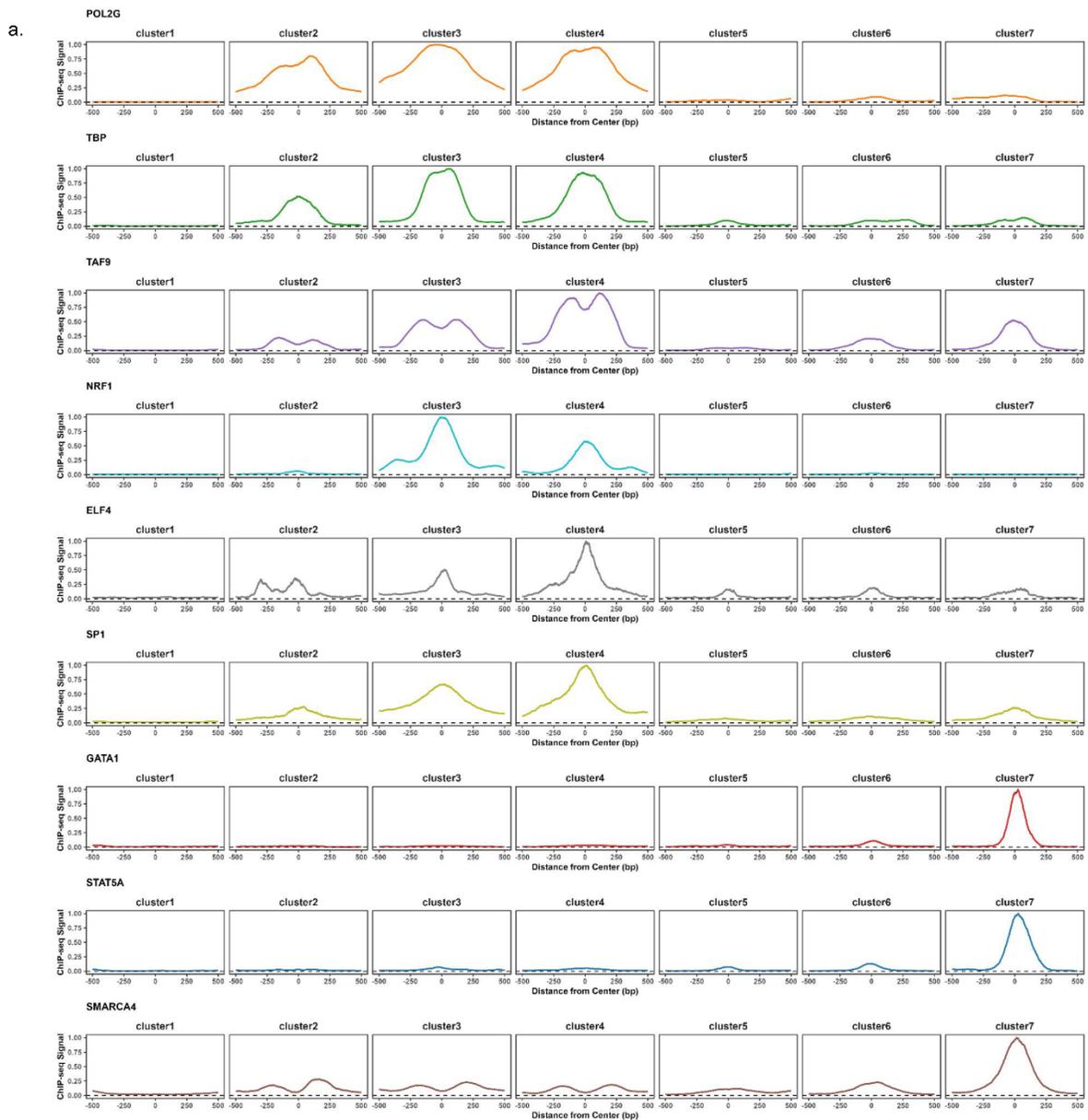
a. Heatmap of ChIP-seq signals for all transcription factors that show difference in enrichment across TREs. Columns represent individual TREs, ordered by k-means clustering based on transcription factor binding profiles, and rows represent transcription factors.



Extended Data Fig. 7 | Gene Ontology enrichment of promoter- and enhancer-biased transcription factors.

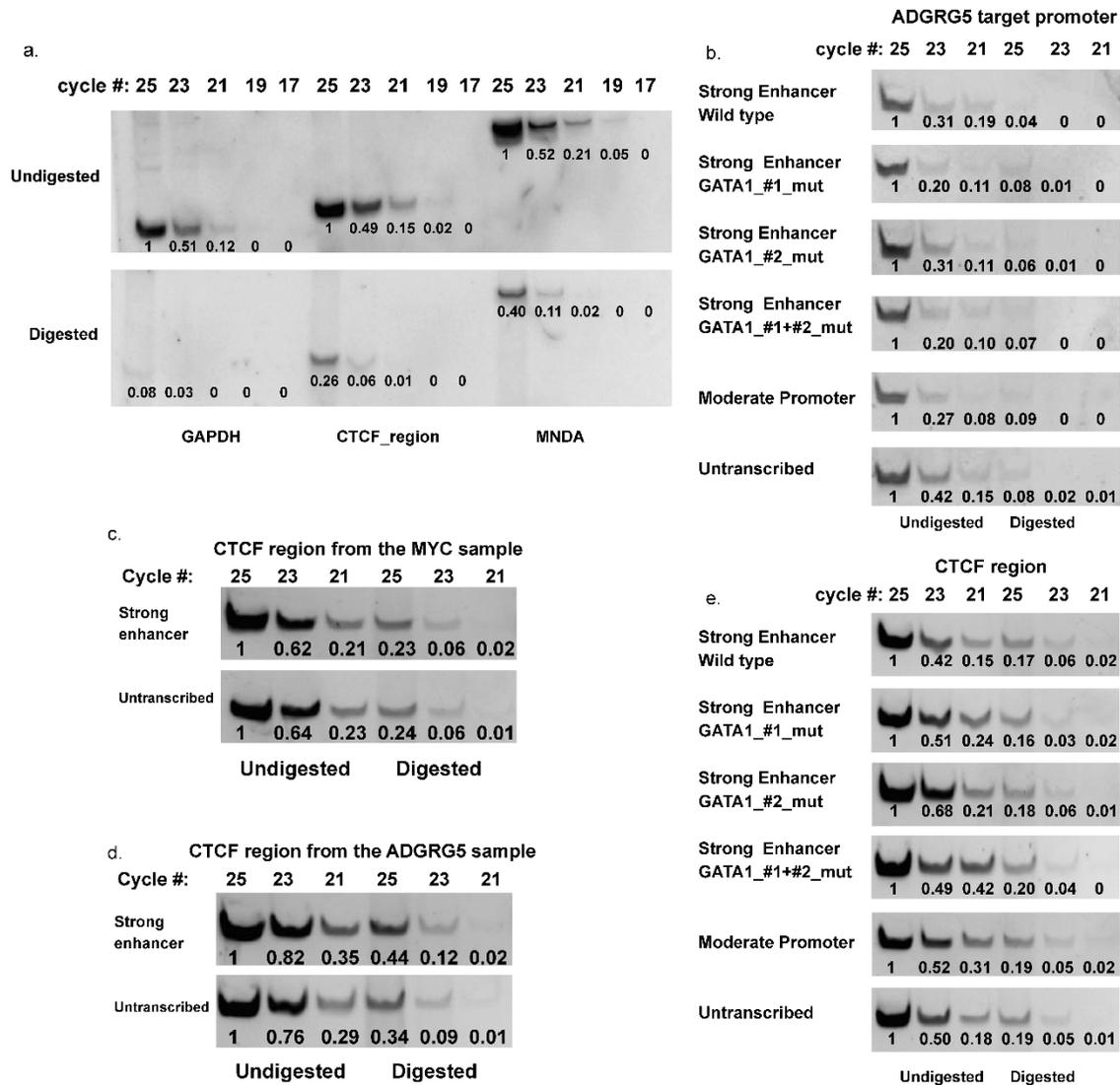
a, Network of gene ontology (GO) molecular function (MF) and biological process (BP) terms enriched among transcription factors showing higher enrichment at promoters than at enhancers. Solid black dots represent transcription factors, and open circles represent GO terms. Blue lines represent molecular function GO terms, and red lines represent biological process GO terms.

b, Network of gene ontology (GO) molecular function (MF) and biological process (BP) terms enriched among transcription factors showing higher enrichment at enhancers than at promoters. Solid black dots represent transcription factors, and open circles represent GO terms. Blue lines represent molecular function GO terms, and red lines represent biological process GO terms.



Extended Data Fig. 8 | Transcription factor ChIP-seq meta-profiles across TRE clusters.

a, Individual meta-profiles of ChIP-seq signals for POLR2G, TBP, TAF9, NRF1, ELF4, SP1, GATA1, STAT5A, and SMARCA4 across transcriptional regulatory elements (TREs) grouped by transcription factor-based k-means clusters.



Extended Data Fig. 9 | DNase I sensitivity assays reveal chromatin accessibility at promoters paired with distinct TREs.

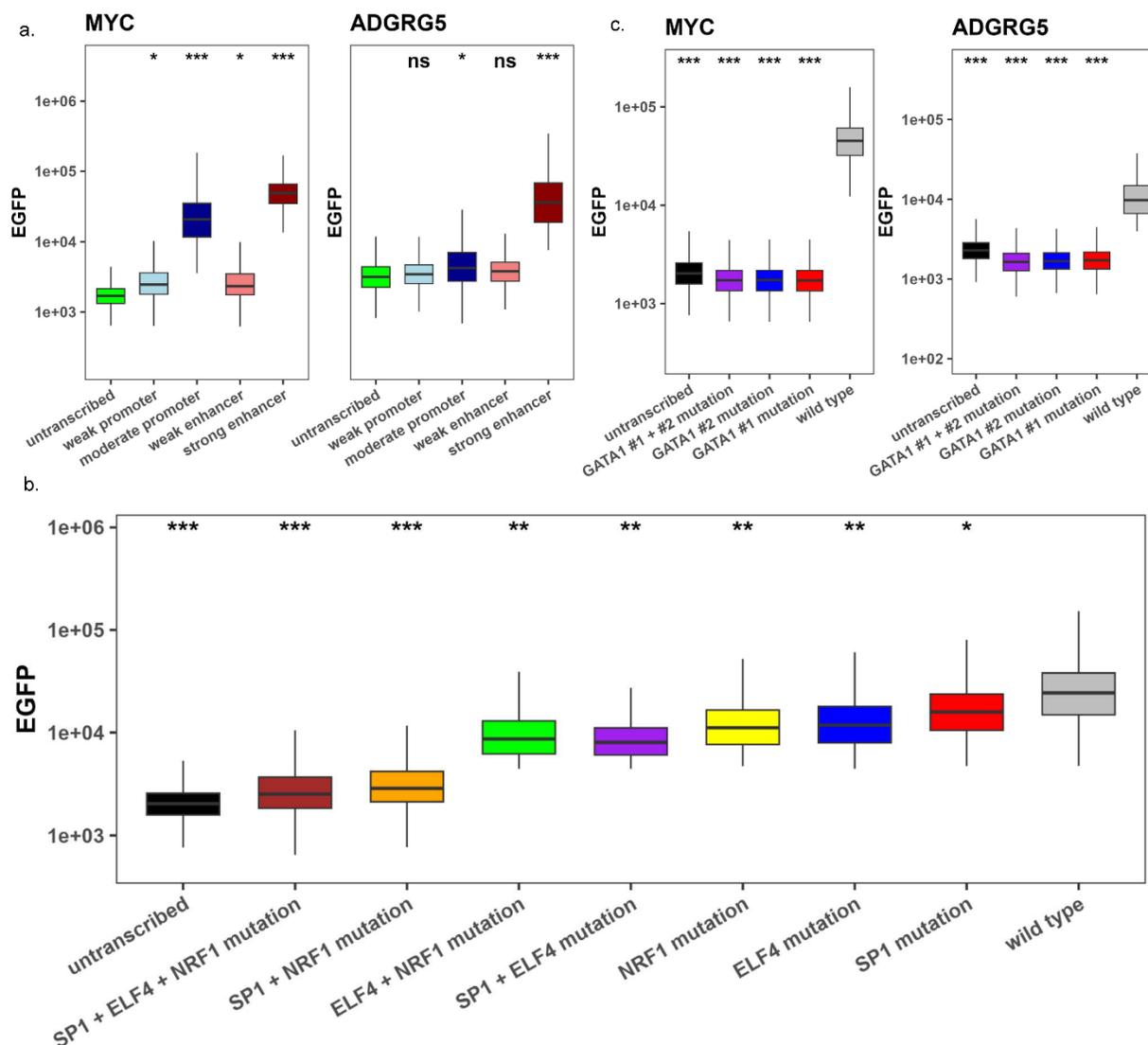
a, PCR analysis of chromatin accessibility at the *GAPDH* promoter (highly accessible), a CTCF peak region (less accessible), and the *MNDA* promoter (least accessible). Genomic DNA from samples that were either undigested or digested with DNase I (25 U, 5 min) was amplified using increasing PCR cycle numbers. Relative PCR band size is quantified by ImageJ.

b, PCR analysis of chromatin accessibility at the *ADGRG5* target promoter paired with an untranscribed TRE (cluster 1), *RCCI* promoter (cluster 3), strong enhancer_93 (cluster 7), strong enhancer_93 carrying a GATA1 motif #1 mutation, strong enhancer_93 carrying a GATA1 motif #2 mutation, or strong enhancer_93 carrying mutations in both GATA1 motifs (#1 and #2). Genomic DNA from samples that were either undigested or digested with DNase I (25 U, 5 min) was amplified using increasing PCR cycle numbers. Relative PCR band size is quantified by ImageJ.

c, PCR analysis of chromatin accessibility at CTCF peak region (Extended Data Fig. 9a) in samples in Fig. 4b. Relative PCR band size is quantified by ImageJ.

d, PCR analysis of chromatin accessibility at CTCF peak region (Extended Data Fig.9a) in samples in Fig. 4c. Relative PCR band size is quantified by ImageJ.

e, PCR analysis of chromatin accessibility at CTCF peak region (Extended Data Fig.9a) in samples in Extended Fig. 9b. Relative PCR band size is quantified by ImageJ.



Extended Data Fig. 10 | Quantitative effects of TRE class and motif mutations on promoter activity.

a, Boxplots of EGFP expression for cells harboring the *MYC* or *ADGRG5* target promoter paired with representative TREs from different transcription factor–based clusters: untranscribed₄ TRE (cluster 1), *UBE2D3* promoter (cluster 2), *RCC1* promoter (cluster 3), weak enhancer₂₁ (cluster 5), and strong enhancer₉₃ (cluster 7). EGFP intensity is shown on the x-axis. Statistical significance was assessed using two one-sided tests (TOST) relative to the untranscribed TRE (*, $-\Delta L > 0.05$; **, $-\Delta L > 0.1$; ***, $-\Delta L > 0.2$).

b, Boxplots of EGFP expression for cells harboring the *MYC* target promoter paired with the *RCC1* promoter (cluster 3) carrying different combinations of mutations in the SP1, NRF1, and ELF4 motifs. EGFP intensity is shown on the x-axis. Statistical significance was assessed using TOST relative to the wild-type promoter (*, $-\Delta L > 0.05$; **, $-\Delta L > 0.1$; ***, $-\Delta L > 0.2$).

c, Boxplots of EGFP expression for cells harboring the *MYC* or *ADGRG5* target promoter paired with the strong enhancer₉₃ (cluster 7) carrying different combinations of mutations in the two GATA1 motifs. EGFP intensity is shown on the x-axis. Statistical significance was assessed using TOST relative to the wild-type enhancer (*, $-\Delta L > 0.05$; **, $-\Delta L > 0.1$; ***, $-\Delta L > 0.2$).

Reference

1. Benoist, C. & Chambon, P. In vivo sequence requirements of the SV40 early promoter region. *Nature* **290**, 304–310 (1981).
2. Banerji, J., Rusconi, S. & Schaffner, W. Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a β -globin gene is enhanced by remote. (1981).
3. Paramo, M. I. *et al.* Dual promoter–enhancer activities reflect a unified regulatory logic. *Nat. Commun.* <https://doi.org/10.1038/s41467-026-68780-y> (2026) doi:10.1038/s41467-026-68780-y.
4. Zabidi, M. A. *et al.* Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556–559 (2015).
5. Martinez-Ara, M., Comoglio, F., Van Arensbergen, J. & Van Steensel, B. Systematic analysis of intrinsic enhancer-promoter compatibility in the mouse genome. *Mol. Cell* **82**, 2519–2531.e6 (2022).
6. Bergman, D. T. *et al.* Compatibility rules of human enhancer and promoter sequences. *Nature* **607**, 176–184 (2022).
7. Arnold, C. D. *et al.* Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. *Science* **339**, 1074–1077 (2013).
8. Tippens, N. D. *et al.* Transcription imparts architecture, function and logic to enhancer units. *Nat. Genet.* **52**, 1067–1075 (2020).
9. Fuda, N. J. *et al.* GAGA Factor Maintains Nucleosome-Free Regions and Has a Role in RNA Polymerase II Recruitment to Promoters. *PLoS Genet.* **11**, e1005108 (2015).
10. Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172 (2009).
11. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).
12. Barral, A. & Zaret, K. S. Pioneer factors: roles and their regulation in development. *Trends Genet.* **40**, 134–148 (2024).
13. Mladenova, V. & Mladenov, E. Organization of Plasmid DNA into Nucleosome-Like Structures after Transfection in Eukaryotic Cells. *Biotechnol. Bioengineering Equip.* **23**, 1044–1047 (2014).
14. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).
15. Hsieh, T.-H. S. *et al.* Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol. Cell* **78**, 539–553.e8 (2020).
16. Fulco, C. P. *et al.* Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).
17. Uyehara, C. M. & Apostolou, E. 3D enhancer-promoter interactions and multi-connected hubs: Organizational principles and functional roles. *Cell Rep.* **42**, 112068 (2023).
18. Zhou, Z. Robust regulatory interplay of enhancers, facilitators, and promoters in a native chromatin context.
19. Brosh, R. *et al.* A versatile platform for locus-scale genome rewriting and verification. *Proc. Natl. Acad. Sci.* **118**, e2023952118 (2021).
20. Koberstein, J. N., Stewart, M. L., Mighell, T. L., Smith, C. B. & Cohen, M. S. A Sort-Seq Approach to the Development of Single Fluorescent Protein Biosensors. *ACS Chem. Biol.* **16**, 1709–1720 (2021).

21. Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing. *Science* **339**, 950–953 (2013).
22. Yao, L. *et al.* A comparison of experimental assays and analytical methods for genome-wide identification of active enhancers. *Nat. Biotechnol.* **40**, 1056–1065 (2022).
23. Zhang, J. *et al.* Comprehensive Evaluation of Diverse Massively Parallel Reporter Assays to Functionally Characterize Human Enhancers Genome-wide. Preprint at <https://doi.org/10.1101/2025.03.25.645321> (2025).
24. Shah, S. R. *et al.* The regulatory landscape of nascent transcription in human health and disease.
25. Core, L. J. *et al.* Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* **46**, 1311–1320 (2014).
26. Kim, T.-K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
27. Rada-Iglesias, A. Is H3K4me1 at enhancers correlative or causative? *Nat. Genet.* **50**, 4–5 (2018).
28. The ENCODE Project Consortium *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).
29. Yang, X. & Vingron, M. Classifying human promoters by occupancy patterns identifies recurring sequence elements, combinatorial binding, and spatial interactions. *BMC Biol.* **16**, 138 (2018).
30. Boshart, M. *et al.* A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* **41**, 521–530 (1985).
31. Zabidi, M. A. & Stark, A. Regulatory Enhancer–Core–Promoter Communication via Transcription Factors and Cofactors. *Trends Genet.* **32**, 801–814 (2016).
32. Karr, J. P., Ferrie, J. J., Tjian, R. & Darzacq, X. The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer–promoter communication. *Genes Dev.* **36**, 7–16 (2022).
33. Perez, G. *et al.* The UCSC Genome Browser database: 2025 update. *Nucleic Acids Res.* **53**, D1243–D1249 (2025).
34. Karolchik, D. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, 493D–496 (2004).
35. Likas, A., Vlassis, N. & J. Verbeek, J. The global k-means clustering algorithm. *Pattern Recognit.* **36**, 451–461 (2003).
36. Tóthová, Z., Tomc, J., Debeljak, N. & Solár, P. STAT5 as a Key Protein of Erythropoietin Signaling. *Int. J. Mol. Sci.* **22**, 7109 (2021).
37. Jain, S. U. *et al.* A SWI/SNF-specific Ig-like domain, SWIFT, is a transcription factor binding platform. *Science* eab3627 (2026) doi:10.1126/science.eab3627.
38. Metzeler, K. H. *et al.* High expression of lymphoid enhancer-binding factor-1 (LEF1) is a novel favorable prognostic factor in cytogenetically normal acute myeloid leukemia. *Blood* **120**, 2118–2126 (2012).
39. Kwak, Y. T. *et al.* Methylation of SPT5 Regulates Its Interaction with RNA Polymerase II and Transcriptional Elongation Properties. *Mol. Cell* **11**, 1055–1066 (2003).
40. He, A. Y. & Danko, C. G. Dissection of core promoter syntax through single nucleotide resolution modeling of transcription initiation.
41. He, H. H. *et al.* Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. *Nat. Methods* **11**, 73–78 (2014).

42. Beishline, K. & Azizkhan-Clifford, J. Sp1 and the ‘hallmarks of cancer’. *FEBS J.* **282**, 224–258 (2015).
43. Ruvkun, G. & Lehrbach, N. Regulation and Functions of the ER-Associated Nrf1 Transcription Factor. *Cold Spring Harb. Perspect. Biol.* **15**, a041266 (2023).
44. Suico, M. A., Shuto, T. & Kai, H. Roles and regulations of the ETS transcription factor ELF4/MEF. *J. Mol. Cell Biol.* jmc;bjw051v3 (2016) doi:10.1093/jmc;bjw051.
45. An, Y. *et al.* TRIM59 expression is regulated by Sp1 and Nrf1 in LPS-activated macrophages through JNK signaling pathway. *Cell. Signal.* **67**, 109522 (2020).
46. Boyes, J., Byfield, P., Nakatani, Y. & Ogryzko, V. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* **396**, 594–598 (1998).
47. Ferreira, R., Ohneda, K., Yamamoto, M. & Philipsen, S. GATA1 Function, a Paradigm for Transcription Factors in Hematopoiesis. *Mol. Cell. Biol.* **25**, 1215–1227 (2005).
48. Im, H. *et al.* Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc. Natl. Acad. Sci.* **102**, 17065–17070 (2005).
49. Cheng, Y. *et al.* Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res.* **19**, 2172–2184 (2009).
50. Guyot, B. *et al.* Characterization of a Megakaryocyte-specific Enhancer of the Key Hemopoietic Transcription Factor GATA1. *J. Biol. Chem.* **281**, 13733–13742 (2006).
51. Gasperini, M. *et al.* A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. *Cell* **176**, 1516 (2019).
52. Eder, M., Moene, C. J. I., Dauban, L., Leemans, C. & Van Steensel, B. Functional maps of a genomic locus reveal confinement of an enhancer by its target gene. Preprint at <https://doi.org/10.1101/2024.08.26.609360> (2024).
53. Lis, J. T. A 50 year history of technologies that drove discovery in eukaryotic transcription regulation. *Nat. Struct. Mol. Biol.* **26**, 777–782 (2019).
54. Farnung, L. & Vos, S. M. Assembly of RNA polymerase II transcription initiation complexes. *Curr. Opin. Struct. Biol.* **73**, 102335 (2022).
55. Adams, C. C. & Workman, J. L. Nucleosome displacement in transcription. *Cell* **72**, 305–308 (1993).
56. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
57. Pugh, B. F. & Venters, B. J. Genomic Organization of Human Transcription Initiation Complexes. *PLOS ONE* **11**, e0149339 (2016).
58. Zhou, Q., Li, T. & Price, D. H. RNA Polymerase II Elongation Control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
59. Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).