Genome Biology

RESEARCH

Open Access

Promoter capture Hi-C identifies promoter-related loops and fountain structures in Arabidopsis

Dingyue Wang[1†], Suxin Xiao[1†], Jiayue Shu[1†], Lingxiao Luo[1], Minqi Yang[1], Myriam Calonje[2], Hang He[1], Baoxing Song[3] and Yue Zhou[1*]

†Dingyue Wang, Suxin Xiao, Jiayue Shu contributed equally to this work.

*Correspondence: yue_zhou@pku.edu.cn

[1] State Key Laboratory of Protein and Plant Gene Research, School of Advanced Agricultural Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China
[2] Institute of Plant Biochemistry and Photosynthesis (IBVF-CSIC), Avenida Américo Vespucio 49, 41092 Seville, Spain
[3] Peking University Institute of Advanced Agricultural Sciences, Weifang 261325, Shandong, China

Abstract

Background: Promoters serve as key elements in the regulation of gene transcription. In mammals, loop interactions between promoters and enhancers increase the complexity of the promoter-based regulatory networks. However, the identification of enhancer-promoter or promoter-related loops in Arabidopsis remains incomplete.

Results: Here, we use promoter capture Hi-C to identify promoter-related loops in Arabidopsis, which shows that gene body, proximal promoter, and intergenic regions can interact with promoters, potentially functioning as distal regulatory elements or enhancers. We find that promoter-related loops mainly repress gene transcription and are associated with ordered chromatin structures, such as topologically associating domains and fountains-chromatin structures not previously identified in Arabidopsis. Cohesin binds to the center of fountains and is involved in their formation. Moreover, fountain strength is positively correlated with the number of promoter-related loops, and the maintenance of these loops is linked to H3K4me3. In *atxr3* mutants, which lack the major H3K4me3 methyltransferases in Arabidopsis, the number of promoter-related loops at fountains is reduced, leading to upregulation of fountain-regulated genes.

Conclusions: We identify promoter-related loops associated with ordered chromatin structures and reveal the molecular mechanisms involved in fountain formation and maintenance.

Background

Chromosomes are organized into specific functional regions, such as promoters and enhancers, which require precise spatial packaging and folding to play important roles in transcription regulation and biological processes [1–3]. Through the development of chromosome conformation capture (3C)-based techniques and microscopic imaging methods [3, 4], multiple scales of three-dimensional (3D) chromatin organization have been identified in eukaryotic genomes, including chromatin territories, A/B

compartments, topologically associating domains (TADs), and chromatin loops [3, 5–8]. Among these, A/B compartments can distinguish transcriptionally active/inactive regions [3]. Chromatin within TADs exhibits significantly higher contact frequencies compared with neighboring regions [5, 9]. In mammals, TAD formation is regulated by CCCTC-binding factor (CTCF) and cohesin complex [10–13]. However, TADs are associated with histone modifications and regarded as compartment domains in *Drosophila* [8, 14]. Recent studies have also identified TADs, or compartment domains, in plants, which can be divided into four types: H3K4me3-TADs, H3K9me2-TADs, H3K27me3-TADs, and null-TADs (without dominant histone modifications) [15–18]. Additionally, chromatin loops, especially promoter-enhancer loops, are involved in the regulation of gene transcription and are confined within TADs in animals [8, 10, 19]. However, this relationship remains unclear in plants. A novel fountain-shaped chromatin feature, appearing perpendicular to the diagonal in contact maps, has been identified in animals. Although referred to by various names, such as "bow-shaped," "hairpin," "flares," and "jet," these structures share the same characteristics: they extend along the diagonal, are regulated by cohesin, and exhibit strong spatial and temporal specificity. The cohesin complex preferentially loads at the target region and begins extruding DNA on both sides. Multiple cohesin extruders can load at the target sites, resulting in the formation of these structures [20–22]. Fountain center regions show H3K27ac signals, suggesting a relationship between fountains and enhancers. A recent preprint reported that enhancer loci are associated with fountain structures located 10–50 kb from those enhancers in *Caenorhabditis elegans* [23].

The Hi-C method captures diverse genomic interactions [3], but promoter-associated interactions are challenging to fully resolve using Hi-C data alone. Although ChIA-PET and HiChIP can generate high-resolution interaction maps for specific loci, they are not suitable for identifying promoter-related loops due to their lack of distinct protein markers [24–26]. In 2015, capture Hi-C was developed and used to map promoter-related loops in human cells [27]. Since then, it has been widely utilized to study promoter-related structures [28, 29]. In plants, the search for enhancers began in pea more than 35 years ago [30]. By analyzing open chromatin and histone acetylation patterns, the enhancers of *GRGP* in maize and *PetE* in pea were identified [31, 32]. However, enhancers in plants do not exhibit distinct histone modifications. As a result, open chromatin distribution combined with *cis*-regulatory element analysis has been used to identify enhancers in rice and Arabidopsis [33, 34]. Although a 2019 study provided insights regarding the functional flexibility of enhancers during plant development in Arabidopsis [35], and more recently, DNase-seq data from different tissues have revealed several super-enhancers in Arabidopsis [36], the identification of enhancers and their corresponding promoter-enhancer loops remains incomplete. Moreover, new methods have been developed to identify chromatin loops in Arabidopsis [37, 38]. Here, we used promoter capture Hi-C (PCHi-C) to gain a genome-wide perspective on promoter-related loops in Arabidopsis. Combined with ATAC-seq data, the PCHi-C results showed that gene body, proximal promoter, and intergenic regions interact with promoters. The non-promoter regions may function as potential enhancers or distal regulatory *cis*-elements. Most promoter-related loops are associated with higher-order chromatin structures. Intriguingly, promoter-related loops were negatively correlated with gene transcription.

Additionally, we detected fountain structures in Arabidopsis. Cohesin binds to the center of these fountains and mediates their formation. The anchor regions surrounding the fountain center were enriched with H3K4me3 and interacted with each other. Finally, we found that ATXR3, the enzyme responsible for regulating H3K4me3, helps maintain loops within fountains to repress fountain-associated genes.

Results

Promoter-related interactions are efficiently captured by PCHi-C in Arabidopsis

Promoters consist of the core promoter region and upstream transcription factor (TF) binding sites. Three promoter elements have been functionally validated in Arabidopsis: the TATA-box, Initiator element, and TC motif or Y patch, respectively. These core promoters, with their specific elements, define the transcription start site (TSS) [39, 40]. To explore the genome-wide regulatory effects of spatial chromatin interactions on the core promoters, we selected the sequence from − 165 to + 5 relative to the TSS (Additional file 1: Figure S1) to design a biotinylated RNA bait library targeting 17,644 core promoter fragments (Fig. 1A). First, we prepared in situ Hi-C libraries using 10-day-old Col-0 plants. Interaction decay exponent (IDE) curves summarize the general features of chromatin compaction. Our new Hi-C data closely matched published data [18], indicating the high quality of the Hi-C libraries (Fig. 1B). Second, we hybridized the Hi-C

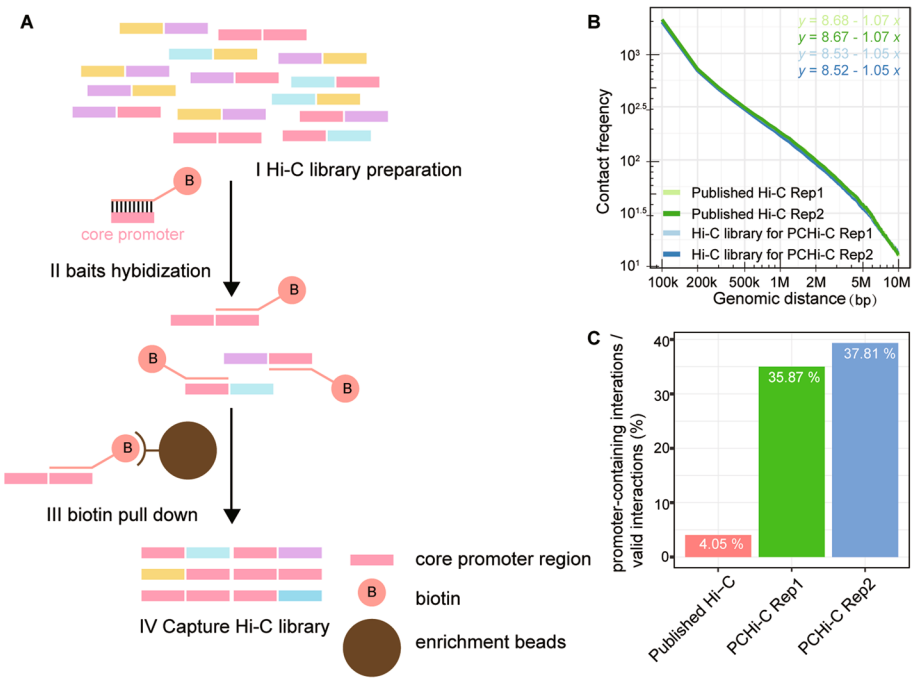

Fig. 1 Promoter capture Hi-C efficiently enriches promoter-related loops. **A** Schematic of the principles of promoter capture Hi-C. Ligated valid interactions prepared by the Hi-C library rarely include promoter-related loops (red). The library of core promoter baits with biotin marks is artificially synthesized and hybridizes with the Hi-C library. By pulling down the biotin-labeled fragments, the promoter capture Hi-C library with highly enriched promoter-related loops can be obtained. **B** Interaction decay exponent (IDE) curves of a published Hi-C library and the Hi-C library used in this study are indicated with green and blue, respectively. **C** Ratios of promoter-related interactions to all valid interactions for published Hi-C data and two replicates of PCHi-C are shown in a bar plot

library DNA to an RNA bait library to capture all promoter-related interactions, then sequenced the resulting PCHi-C libraries. We analyzed the PCHi-C libraries and compared the capture efficiencies of promoter-related interactions between published Hi-C data and our PCHi-C data. The results showed a nearly tenfold increase in the enrichment of promoter-related interactions in the PCHi-C data, compared with conventional Hi-C data, which was sufficient to allow the identification of promoter-related loops (Fig. 1C, Additional file 2: Table S1). This level of enrichment is similar to findings observed with the promoter capture Hi-C method in animals [27]. Overall, our results illustrate that PCHi-C efficiently enriches promoter-related interactions.

Identification of different promoter-related loops

Using a statistical model, we adjusted for distance effects (Additional file 1: Figure S2) and identified 27,946 significant promoter-related loops (Additional file 3: Table S2) from the total interactions. The median length of these loops was approximately 27 kb (Fig. 2A), and there were no noticeable differences among chromosomes. Nearly all promoter-related loops identified by PCHi-C were intra-chromosomal rather than inter-chromosomal, consistent with the previously reported concept of chromosome territories in Arabidopsis, suggesting that promoters preferentially interact with regions on the same chromosome (Fig. 2B). Potential regulatory elements usually exhibit open chromatin signals, representing their ability to bind transcription factors, which mediate interactions between elements. Due to the compact genome of Arabidopsis, it is challenging to identify regulatory elements solely based on genome annotation. To further

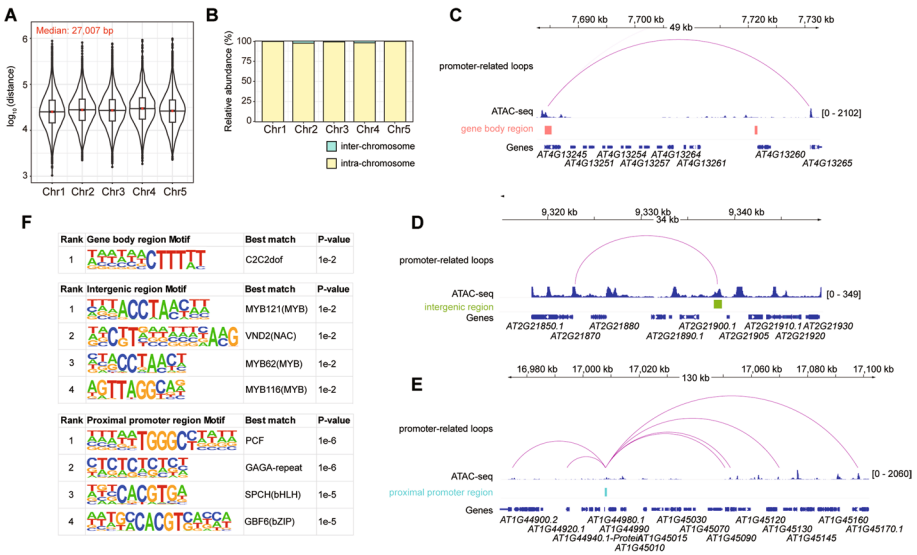

Fig. 2 Promoters interact with gene body, intergenic, and proximal promoter regions. **A** Violin plot and corresponding box plot represent the length distribution of promoter-related loops in individual chromosomes. The median length of promoter-related loops is labeled at the top left of the panel. **B** Bar plot indicates the relative abundances of intra-chromosomal and inter-chromosomal promoter-related loops in individual chromosomes. **C, D, and E** Gene body (**C**), intergenic (**D**), and proximal promoter (**E**) loops are shown. Red, green, and blue boxes represent the gene body, intergenic, and proximal promoter regions, respectively. **F** Motifs enriched in the gene body (top), intergenic (middle), and proximal promoter (bottom) regions are shown. Transcription factors recognizing the top four ranked motifs and their corresponding *P*-values are indicated

characterize the features of regions interacting with promoters, we integrated ATAC-seq data to identify chromatin-accessible regions within promoter-related loop anchors. The promoter-interacting regions were categorized into three types: gene body (located within the gene body region), proximal promoter (located within 200 bp upstream of the TSS), and intergenic (located more than 200 bp upstream of the TSS) (Fig. 2C–E). Of the regions identified by PCHi-C, 104 overlapped with previously published super-enhancers (749) [36] (Additional file 1: Figure S3). The relatively low overlap between our potential enhancers and previously identified super-enhancers may be due to the tissue specificity of enhancers. Super-enhancers were identified by integrating datasets from multiple tissues, whereas the PCHi-C data were derived from 10-day-old seedlings. Considering that promoter-interacting regions could potentially function as regulatory elements in gene expression, we analyzed motif enrichment in these regions (Fig. 2F). *Cis*-elements bound by C2C2dof, MYB121, or PCF were the most enriched in gene body, intergenic, and proximal promoter regions, respectively, suggesting that these transcription factors are involved in mediating promoter-related loops. Taken together, our data indicate that intergenic or gene body regions interacting with promoters may function as potential enhancers or distal regulatory elements.

Promoter-related loops mainly associate with ordered chromatin structures and negatively regulate gene transcription

When we examined the distribution pattern of promoter-related loops, we found that black dots representing the loops were located within shaded triangles, corresponding to three types of TADs (H3K4me3-, H3K9me2-, and H3K27me3-) [18] in the heatmap (Fig. 3A–C). This indicated that most loops were established within larger-scale chromatin structures. We also observed contact structures associated with promoter-related loops outside TADs, which originate from a single genomic locus and expand with increasing distance from the diagonal (Fig. 3D). These structures are known as fountains in animals [20–22]. We identified 1361 fountains based on Hi-C data (Additional file 3: Table S2). Thus, our data revealed the presence of fountain structures in Arabidopsis, which were previously unreported (Fig. 3D). Overall, we found that 73.29% of promoter-related loops were associated with TADs or fountain structures (Fig. 3A–D). Specifically, 31.45, 13.79, and 5.77% of promoter-related loops were associated with H3K27me3-, H3K9me2-, and H3K4me3-TADs, respectively, whereas 22.28% were linked to fountains (Fig. 3E). Because the maintenance of 3D chromatin structures in Arabidopsis is associated with histone modifications [18, 41], we examined the enrichment of histone modifications at promoter-related loop anchors within each type of chromatin structure. We found that whereas anchor regions within TADs were enriched with the histone modification corresponding to the specific TAD, anchors of fountain-associated loops and other loops were enriched in H3K4me3 (Fig. 3F). Altogether, these data indicate that promoter-related loops are closely correlated with TADs and fountain structures.

To further explore the biological significance of promoter-related loops, we treated each promoter as a single node and defined the number of interactions with other nodes (loops) as the "interaction degree". A higher interaction degree indicates that more loops are created from a particular node (Fig. 3G). The interaction degrees of promoters were significantly higher in H3K9me2- and H3K27me3-TADs than in H3K4me3-TADs and

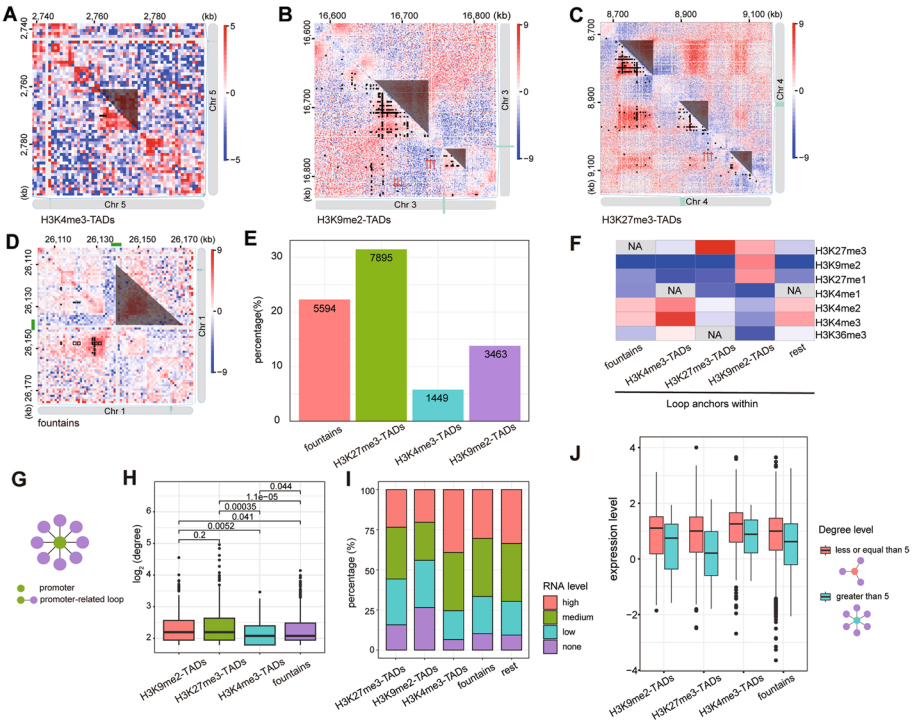

Fig. 3 Promoter-related loops are mainly associated with different TADs and fountain structures; they serve to repress gene transcription. **A–C** Contact heatmaps normalized by O/E for H3K4me3- (**A**), H3K9me2- (**B**), and H3K27me3-TAD (**C**) regions are shown. Promoter-related loops are indicated by black marks in the bottom left corner of each heatmap. Some typical loops not associated with TAD structures are marked by red arrows. Shaded triangles outline TAD patterns in the heatmaps. **D** A typical fountain structure is shown in O/E contact heatmaps. Green boxes represent the base of the fountains. Promoter-related loops are indicated by black marks in the bottom left corner of each heatmap. Shaded triangles outline fountain patterns in the heatmaps. **E** The proportion and number of promoter-related loops associated with each higher-order chromatin structure are shown. **F** The heatmap represents the relative histone modification levels in promoter-related loop anchor genes associated with H3K9me2-TADs, H3K27me3-TADs, H3K4me3-TADs, fountains, or other loops. Pairwise Mann–Whitney *U* tests, corrected by the Bonferroni–Holm method for multiple comparisons, were used to calculate the statistical significance of enrichment or depletion (target regions vs. randomly selected regions) shown in red and blue, respectively. Non-significant modifications are indicated in gray (*P*-value > 0.05). Published data were obtained from the Plant Chromatin State Database and are listed in Additional file 4: Table S3. **G** Schematic of promoter interaction degree**.** Green and purple circles represent the promoter and interacting regions, respectively. Lines between the circles represent promoter-related loops. **H** Box plot representing the interaction degree of each promoter-related loop anchor gene associated with H3K9me2-TADs, H3K27me3-TADs, H3K4me3-TADs, and fountains. Only anchors with an interaction degree greater than five are shown. *P*-values were calculated using the Wilcoxon rank-sum test. **I** Bar plot showing the proportional expression levels of promoter-related loop anchor genes associated with H3K9me2-TADs, H3K27me3-TADs, H3K4me3-TADs, and fountains. **J** Boxplot representing the expression levels of promoter-related loop anchor genes associated with H3K9me2-TADs, H3K27me3-TADs, H3K4me3-TADs, and fountains. Red and blue indicate a loop anchor interaction degree of ≤ 5 or > 5. Expression levels are normalized by $\log_{10}$. The *p*-values are indicated in the Additional file 1: Figure S4

fountains (Fig. 3H). Next, we assessed the transcriptional activities of genes driven by promoters associated with different structures. When we categorized gene transcription levels into four clusters (high, medium, low, and none), we found that, overall, genes driven by promoters in H3K9me2- and H3K27me3-TADs tended to have lower transcription levels than those driven by promoters in H3K4me3-TADs (Fig. 3I). Regarding genes driven by promoters associated with fountains, we observed that the percentage with high expression levels was higher than the percentages in H3K9me2- and

H3K27me3-TADs, but lower than the percentage in H3K4me3-TADs (Fig. 3I). Additionally, we found that most promoters associated with H3K4me3-TADs had low interaction degrees, those associated with fountains had medium interaction degrees, and those associated with H3K9me2- and H3K27me3-TADs had high interaction degrees (Fig. 3H).

To exclude the effect of histone modifications, we compared the expression levels of genes driven by promoters with high interaction degrees (≥ 5 interactions) versus low interaction degrees (< 5 interactions) across TAD- and fountain-associated structures (Fig. 3J). The results showed that genes driven by high interaction degree promoters had significantly lower transcription levels than those driven by low interaction degree promoters. For H3K4me3-TAD-associated promoters, the *P*-value was not statistically significant due to the small number (24) of promoters with a high interaction degree (Additional file 1: Figure S4A). Our results remained consistent under different thresholds (Additional file 1: Figure S4B-C). In summary, our findings suggest that promoter interaction degree is negatively correlated with transcriptional activity.

Fountain structures are associated with promoter-related loops and regulated by cohesin

Fountain structures and their regulatory mechanisms in Arabidopsis are unclear. In animals, the fountain center is bound by RAD21 and NIPBL [21], core components of the cohesin complex. To identify factors that bind to the fountain center in Arabidopsis, we performed motif enrichment analysis and found that the G-box showed the highest enrichment (Fig. 4A,B). This motif reportedly is bound by the cohesin component SISTER CHROMATID COHESION 1 PROTEIN 4 (SYN4) [42]. We then investigated whether SYN4 is involved in the regulation of fountain structures. A metaplot revealed that SYN4 binding is enriched at the center of fountains (Fig. 4A,C). When we compared the strength of fountain structures between Col-0 and the *syn4* mutant, we observed a significant reduction of strength in the *syn4* mutant (Fig. 4D, Additional file 1: Figure S5). Taken together, our results showed that cohesin plays an important role in the formation of fountain structures in Arabidopsis.

In animals, fountains show strong enrichment of enhancer-specific marks (H3K4me1 and H3K27ac) but relatively weak enrichment of the promoter-specific mark H3K4me3. When we examined promoter-related loop anchors on either side of the center, we found that these anchors had high levels of H3K4me3 compared to random selected genes (Figs. 3F and 4A). The H3K4me3-enriched promoter-related loop anchors tended to interact with each other, crossing the fountain center. To further explore the relationship between fountains and promoter-related loops within the fountain, we classified fountains into three groups according to promoter loop number and determined their strength (Fig. 4E). Fountain strength was positively correlated with loop number, suggesting that the regulation of promoter-related loops is involved in the maintenance of fountain structures (Fig. 4F).

ATXR3 maintains loops within fountains to repress fountain-associated gene transcription

Considering that the loop anchors within fountains are enriched in H3K4me3, which significantly differs from observations in animals (Figs. 3F and 4A), we investigated the role of H3K4me3 in regulating fountains and the associated promoter-related loops.

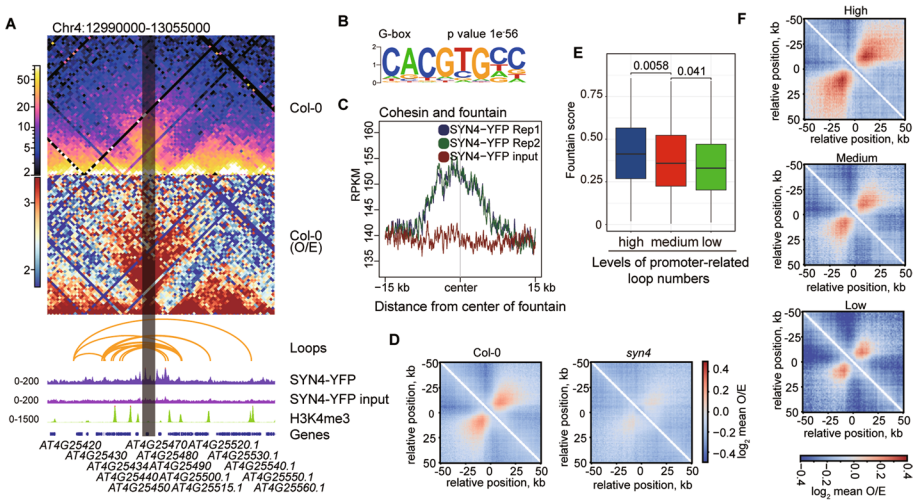

Fig. 4 Fountain structures are correlated with promoter-related loops and regulated by cohesin. **A** The contact matrix, O/E contact matrix, promoter-related loops, SYN binding sites, H3K4me3 modified regions, and gene-occupied regions are presented from top to bottom. Typical fountain structures are represented by raw (top) and O/E (bottom) contact heatmaps. Yellow arcs indicate promoter-related loops, and the shaded area marks the location of fountain center. **B** The G-box motif feature at the center of the fountain. **C** The meta plot shows SYN4-YFP binding to the center of the fountain structures. **D** Heatmaps representing the average contact strength of fountains and their 50-kb flanked regions in Col-0 and the *syn4* mutant. Values are shown as $\log_2$ observed/expected. **E** Box plot representing the fountain strength of three clusters. *P*-values are calculated using the Wilcoxon rank-sum test. **F** Heatmaps representing the average contact strength of fountains and their 50-kb flanked regions. From top to bottom, fountains are categorized as high, medium, or low based on the number of promoter-related loops. Values are shown as $\log_2$ observed/expected

ARABIDOPSIS TRITHORAX-RELATED3 (ATXR3) is a major H3K4 tri-methyltransferase in Arabidopsis; accordingly, the level and distribution of H3K4me3 are significantly altered in the *atxr3* mutant [43, 44]. To investigate changes in 3D chromatin structure, we first performed Hi-C in the *atxr3* mutant (Additional file 2: Table S1). The results showed that contact frequency was significantly reduced in the *atxr3* mutant (Additional file 1: Figure S6). Aggregate TAD analysis (ATA), which calculates interaction intensity within TADs, showed that whereas the strengths of H3K27me3- and H3K9me2-TADs were unchanged (Additional file 1: Figure S7), the strength of H3K4me3-TADs was reduced in the *atxr3* mutant (Additional file 1: Figure S7). Consistent with this observation, we found that H3K4me3 levels tended to decrease in the *atxr3* mutant at the anchors of fountain-associated promoter-related loops, as well as genome-wide (Fig. 5A, Additional file 1: Figure S8). Thus, we hypothesized that the number of promoter-related loops would be reduced. We performed PCHi-C in the *atxr3* mutant and identified 23,477 loops, fewer than the number identified in Col-0. Compared with Col-0, PCHi-C in *atxr3* showed a significant reduction in the interaction degree of fountain-associated promoter-related loops (Fig. 5B). Because our results suggested that promoter-related loops play a repressive role in gene expression (Fig. 3), we analyzed the transcription levels of genes driven by fountain-associated promoter-related loops in the *atxr3* mutant using RNA-Seq (Additional file 1: Figure S9). We also examined the correlation between gene expression and H3K4me3 levels in Col-0 and the *atrx3* mutant. We found that the correlation between fountain-associated gene transcription and

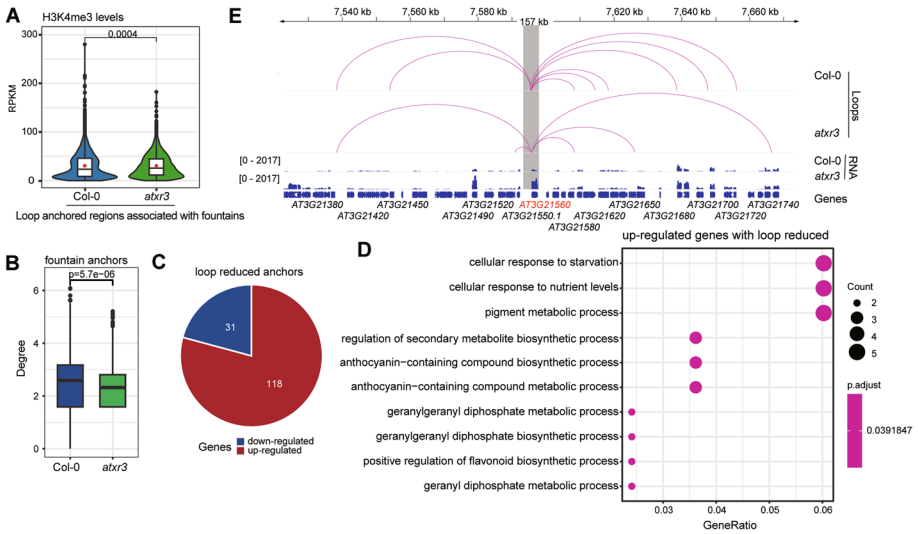

Fig. 5 ATXR3 functions to maintain promoter-related loops in fountains, thereby repressing gene transcription. **A** Box plot representing H3K4me3 levels at fountain-associated loop anchors in Col-0 and the *atxr3* mutant. Red diamonds indicate mean H3K4me3 levels. The *p*-value is calculated by Wilcoxon rank-sum test. **B** Box plot representing the interaction degree of loop anchors within fountains in Col-0 and the *atxr3* mutant. *P*-values were calculated using the Wilcoxon rank-sum test. **C** Differential expression analysis was conducted on anchored genes with reduced interaction degrees as shown in **B**. Differentially expressed genes were defined by the following criteria: $|\log_2$ (fold change)$|> 1.5$ and p value < 0.05. **D** GO analysis was performed on the upregulated genes (red) identified in **C**. **E** Promoter-related loops, gene expression levels in Col-0 and the *atxr3* mutant, and gene regions are presented from top to bottom. Only *BRT1* (*At3G21560*)-associated loops are highlighted in red. RNA-Seq data are presented in RPKM. The shaded area indicates the location of the *BRT1* core promoter region

H3K4me3 levels was decreased in the *atxr3* mutant relative to Col-0 (Additional file 1: Figure S10). The *atxr3* mutant exhibits severe growth and developmental phenotypes compared with the wild type (Additional file 1: Figure S11A). Approximately, 20% of differentially expressed genes (DEGs) are associated with promoter-associated loops (Additional file 1: Figure S11B). For instance, loop numbers are reduced at the *AT5G50800* (*AtSWEET13*) and *AT5G46700* (*TORNADO 2*) loci. Correspondingly, their expression levels are upregulated in the *atxr3* mutant (Additional file 1: Figure S11C). AtSWEET13 and TORNADO 2 are reportedly involved in seed and seedling development [45] and root differentiation regulation [46], respectively. Their aberrant expression may contribute to the phenotypic abnormalities observed in the *atxr3* mutant. Additionally, the reduced number of promoter-related loops in the *atxr3* mutant led to the upregulation of 79% (118/(118 + 31)) of all DEGs within fountains (Fig. 5C). Gene Ontology (GO) analysis showed that upregulated genes were enriched for metabolic processes involved in plant development and adaption [47] (Fig. 5D). Misregulation of these genes may be responsible for the developmental phenotype of the *atxr3* mutant (Additional file 1: Figure S11A). For example, the promoter of *Bright Trichomes 1* (*BRT1*), a UDP-glucosyltransferase and putative cytoplasmic enzyme involved in phenylpropanoid metabolism [48], is located at a fountain anchor. In the *atxr3* mutant, the interaction degree of this promoter was reduced, leading to the upregulation of *BRT1* (Fig. 5E) and neighboring anchor genes. Overall, the results demonstrate that fountain-associated promoter-related loops serve to repress gene transcription.

Discussion

Promoters are essential sequences that can be bound by RNA polymerase II (RNAPII) and transcription factors to initiate gene transcription [39]. In mammals, promoter-enhancer loops are regulated by RNAPII, mediator complexes, and transcription factors [49–51]. H3K27ac has been identified as a marker of enhancers in animals [52–54]. Therefore, methods based on H3K27ac immunoprecipitation are widely used to identify enhancers and promoter-enhancer loops in these organisms. However, enhancers enriched with specific histone modifications have not been detected in plants; thus, identification of enhancer- and promoter-related loops has remained challenging. Techniques based on chromatin accessibility, such as DNase I-Seq and ATAC-Seq, have been used to detect enhancer candidates [34, 55, 56]. However, few enhancers have been validated. Using the PCHi-C method, we found that gene body and intergenic regions can interact with promoters and may function as potential enhancers. Additionally, the Arabidopsis genome is very dense, with more than 30,000 genes in 120 Mb of DNA. Thus, it has been suggested that only proximal promoter regulation elements are present or that enhancer-based gene regulation is not the main mechanism of gene regulation in Arabidopsis [57]. Our results generally support these hypotheses. Among all promoter-related loops, we found that a portion of loop anchors were located in the proximal promoter region. Notably, another set of loops identified by CAP-C also contained a high number of promoter-promoter interactions (PPIs) in Arabidopsis [37]. Both findings suggest that PPIs are widespread in Arabidopsis. Furthermore, TFs such as bHLH were enriched at the loop anchor regions in both promoter-related loops and PPIs, indicating that TFs play an important role in regulating these interactions.

By combining PCHi-C and Hi-C data, we found that most promoter-related loops are associated with higher-order chromatin structures (TADs or fountains) (Fig. 6A). We also found that the interaction degree of a promoter is negatively correlated with the transcription levels of the corresponding gene (Fig. 6C); accordingly, promoters within H3K9me2- or H3K27me3-TADs have a higher interaction degree relative to those within H3K4me3-TADs. Promoters associated with high-interaction-degree loops form condensed regions where genes exhibit relatively lower transcription levels. This finding aligns with previous studies that combined Hi-C and ChIP-seq of Pol II to explore the characteristics of higher-order chromatin structures containing both active and repressive genes [57]. These studies also suggested that the regulatory mechanism in Arabidopsis favors co-repression over co-activation. Additionally, we identified fountain structures in plants; these structures have been identified in other organisms, such as mammals and zebrafish [20–22]. Extensive functional evidence now supports a role for fountain structures in mediating and controlling promoter-enhancer interactions [58–60]. Moreover, fountain structures are highly dynamic, consistent with the stage-specific characteristics of enhancers. Similar to fountains in animals, we found that the centers of fountains in Arabidopsis are bound by cohesin. In cohesin component mutants, fountain structures are significantly diminished, indicating that fountain formation is driven by extrusion of the cohesin complex. Regions enriched with H3K4me3 tend to interact with each other across the fountain center. Cohesin regulates fountain structure, whereas H3K4me3 helps to maintain loops and interactions within the fountain (Fig. 6B). However, we observed notable differences between animals and Arabidopsis. First, fountain structures in Arabidopsis are mainly

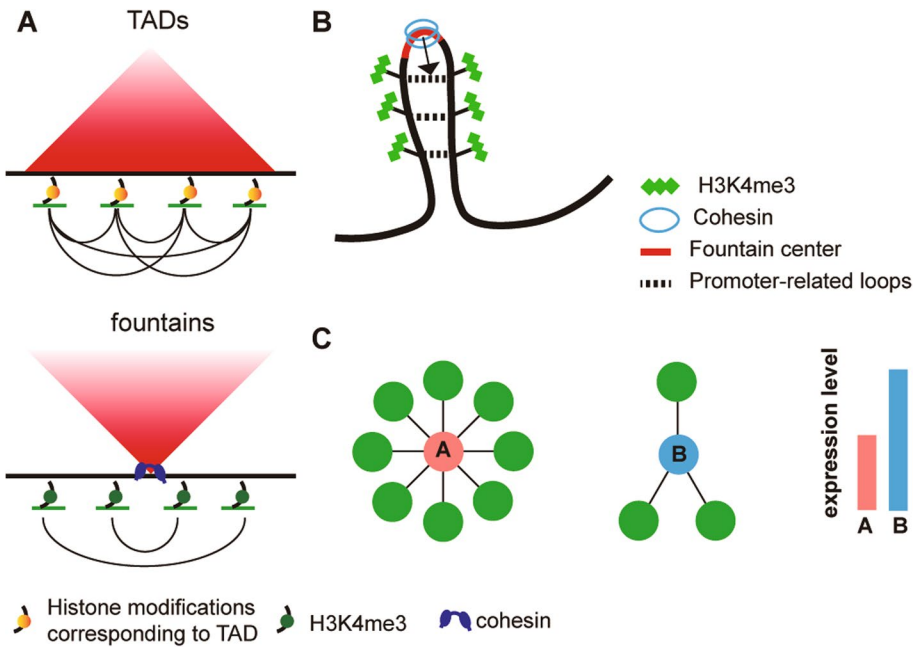

Fig. 6 Working models for promoter-related loops and their functions. **A** Both TADs and fountains are associated with promoter-related loops in Arabidopsis. Heatmap models depict different chromosome structures; black arcs indicate anchoring promoter-related loops, and green lines represent gene regions. TAD-associated promoter-related loops are enriched with histone modifications corresponding to the TAD (top). At the center of the fountain, cohesin binds to mediate loop formation. The anchors of fountain-associated promoter-related loops are enriched with H3K4me3 (bottom). **B** Working model of promoter-related loops in fountains. Cohesin binds at the center to mediate loop extrusion within the fountain; H3K4me3 enrichment at the anchors contributes to the maintenance of promoter-related loops. **C** Promoter-related loop number and gene expression level are negatively correlated. Green circles represent regions that interact with promoters (red or blue). Genes driven by higher interaction degree promoters (red) exhibit lower expression levels relative to genes driven by lower interaction degree promoters (blue)

associated with promoter-related loops. Second, the H3K4me3 levels at anchor regions play a role in maintaining fountain strength. Through this maintenance, chromatin interactions crossing the fountain center repress the associated genes, ensuring that their expression levels remain low. In support of this notion, we found that fountain strength and the numbers of promoter-related loops are reduced in the *atxr3* mutant, leading to the upregulation of fountain-associated genes.

Conclusions

This study constitutes the first application of PCHi-C technology to Arabidopsis. We identified novel promoter-related regulatory loops, potential enhancers, and fountain structures in Arabidopsis. Moreover, we showed that cohesin and H3K4me3 are involved in fountain formation and maintenance, respectively. The establishment of PCHi-C technology offers a new approach for studying distal regulatory elements or enhancers in crops with large genomes, such as wheat and maize.

Methods

Plant material and growth conditions

The mutant and WT were established on the Columbia-0 (Col-0) background. The T-DNA insertional mutant used in this study, *atxr3* mutant (*sdg2, SALK_021008*), was previously described [44]. The *syn4* mutant (*SALK_130085*) was obtained from the European Arabidopsis Stock Centre (NASC, https://arabidopsis.info). For all experiments, seeds were sown on Murashige and Skoog (MS) plates containing 1% sucrose and 0.7% agar, incubated at 4 °C for 2 days, then grown for 10 days under long-day conditions (8 h dark and 16 h light) at 22 °C. Whole seedlings were used for experiments.

RNA-Seq assay

First, total RNA was extracted from 10-day-old seedlings using the E.Z.N.A. Plant RNA Kit (Omega, R6827-01). RNA-Seq libraries were prepared with VAHTS® mRNA Capture Beads (Vazyme, N401) and VAHTS® Universal V8 RNA-Seq Library Prep Kit for MGI (Vazyme, NRM605). Briefly, messenger RNA capture, fragment, double-stranded cDNA synthesis, adapter ligation, and polymerase chain reaction (PCR) were performed in accordance with manufacturer protocols. Libraries were sequenced using NovaSeq 6000 guidelines (Illumina). Three biological replicates were performed for RNA-Seq.

In situ Hi-C assay

In situ Hi-C was performed as previously described [18]. Seedlings (~3 g) were fixed with 1% formaldehyde solution in MC buffer (10 mM potassium phosphate, pH 7.0; 50 mM NaCl; 0.1 M sucrose) at room temperature for 2 × 10 min. After fixation, 0.1 M glycine was added for 5 min to stop the reaction. The fixed tissue was homogenized with liquid nitrogen and resuspended in nuclear isolation buffer (20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid [HEPES], pH 8.0, 250 mM sucrose, 1 mM $MgCl_2$, 5 mM KCl, 40% [v/v] glycerol, 0.25% Triton X-100, 0.1 mM phenylmethylsulfonyl fluoride, 0.1% [v/v] 2-mercaptoethanol), then filtered through two layers of Miracloth (Merck Millipore). Nuclei were resuspended in 0.5% sodium dodecyl sulfate (SDS) and denatured at 62 °C for 5 min. Next, Triton X-100 was added to quench the SDS; the mixture was digested with 50 units of DpnII at 37 °C overnight. On the following day, the digested DNA was blunt-end digested with Klenow enzyme (Thermo Scientific); biotin-14-dCTP (Invitrogen) was also incorporated. After ligation with T4 DNA ligase, proteinase K was added and the tubes were incubated at 55 °C for 30 min, then held at 65 °C overnight. On the third day, DNA isolation was performed via phenol–chloroform extraction, followed by ethanol precipitation. DNA was then sheared by sonication with a Bioruptor Pico (Diagenode) to 200–600 bp; size selection was performed with VAHTS DNA Clean Beads (N411-01). The DNA was then pulled down with DynaBeads MyOne Streptavidin C1 beads (Invitrogen); on-bead end repair and adapter ligation were performed. After washing, the beads were resuspended in 15 μL of 10 mM Tris–HCl buffer (pH 8.0) and incubated

at 98 °C for 10 min to detach DNA. Library molecules were amplified by 11 cycles of PCR. The resulting PCR products were purified with VAHTS® DNA Clean Beads (Vazyme, N411-01), and libraries were sequenced on an MGI DNBSEQ T7 PE150 platform, which generated 2 × 150 bp paired-end reads.

Capture Hi-C assay

For Capture Hi-C, in situ Hi-C libraries of Col-0 and *atxr3* were generated. The hybridization of 1 μg Hi-C libraries and 18,329 biotinylated RNA baits (Dynegene Technologies) was performed using QuarHyb One Reagent Kit Box 2 (Dynegene Technologies). Capture libraries were enriched and amplified with QuarAcces Hyper Enrichment beads and QuarHyb One Reagent Kit Box 1 (Dynegene Technologies). PCR products were subsequently purified with VAHTS® DNA Clean Beads (Vazyme, N411-01), and libraries were sequenced on an MGI DNBSEQ T7 PE150 platform, which generated 2 × 150 bp paired-end reads.

Hi-C data analysis

Raw sequencing data were filtered using fastp, and reads were mapped to the reference genome (TAIR10, https://www.arabidopsis.org) using Bowtie [61]. Valid contacts were filtered and identified using HiC-Pro [62] with default parameters. TAD identification was performed as described in our previous study [18]. Aggregate analyses of TADs were performed with coolpup.py 0.9.5 using the following parameters: -n_proc 5, -rescale, -local, -rescale_size 45, and -rescale_pad 1; they were visualized using plotpup.py with -scale linear [63]. The center of the fountain structures was defined and identified using Fontanka (https://github.com/agalitsyna/fontanka) with a snip resolution of 5 kb and snip size of 100-kb flanked regions. The correlations between each potential fountain snip and manually selected fountain structure snips were used to calculate fountain strength (fountain score). For loop classification, we annotated the anchors associated with loops that cross the fountain center as fountain-associated; those located within TADs were considered TAD-associated. Next, the genome-wide identified fountain structures were aggregated and visualized using cooltools (https://github.com/open2c/cooltools). Chromatin interactions at specific positions were visualized using Juicebox (https://github.com/aidenlab/Juicebox) and pyGenomeTracks (https://github.com/deeptools/pyGenomeTracks) [64]. The WashU Epigenome Browser and Integrative Genomic Viewer were used to visualize ChIP-seq results [65].

Capture Hi-C data analysis

After raw data acquisition, the filtration and mapping steps were identical to those used for Hi-C. Subsequently, valid pairs were identified and deduplicated using HiCUP-0.9. [66]. Based on the valid pairs, we utilized CHiCAGO, a method designed for capture Hi-C, to detect promoter-associated interactions and identify significant promoter-related loops [67]. CHiCAGO can detect significant chromatin loops weighted by genomic distance; however, the default settings were not optimized for Arabidopsis. Based on diagnostic plots and loop visualization on Hi-C heatmaps, we adjusted the weightAlpha, weightBeta, weightGamma, and weightDelta parameters to 13.5319239, -1.3100426, -10.3516115, and 0.1635212, respectively, to eliminate distance-induced

bias. At the anchors of promoter-related loops, potential regulatory regions were annotated using ATAC-seq. Motif enrichment analysis was performed by the findMotifsGenome.pl script in HOMER software with the parameters -length 8 and -size 200 [68]. The interaction degree was calculated using the igraph package in R [69].

Motif enrichment analysis

To identify potential protein regulatory sites, we first extracted accessible regions from the target areas for input into the motif enrichment analysis. Bias can often exist and substantially influence enrichment outcomes. HOMER, a widely used software, provides auto-normalization to address this bias by assigning weights to background sequences. For the enrichment of each motif, HOMER considers the number of target sequences that are "bound" to the background sequences and applies hypergeometric or binomial calculations to determine significance. The findMotifsGenome.pl was used with the parameters -length 8 and -size 200 for motif discovery.

RNA-Seq data analysis

Fastp was initially used to filter raw reads, which were then mapped with gaps to the reference genome (TAIR10, https://www.arabidopsis.org) via HISAT2 [70]. After filtering, indexing, and format conversion using SAMtools [71], we utilized StringTie [72] to annotate reads with gene data from Araport11 (https://www.arabidopsis.org). DEGs were identified using the R package DESeq2 [73]. GO analysis was performed and visualized by clusterProfiler [74]. Correlations between RNA-seq and ChIP-seq data were assessed by Spearman correlation. Transcript levels were normalized by FPKM and H3K4me3 levels were normalized by RPKM.

Public genomics data analysis

Enrichment of histone modifications were performed using published ChIP-Seq data [17, 75–78] from the Plant Chromatin State Database [79] (Additional file 4: Table S3). Bedmap [80] with the -wmean parameter was used to determine the mean occupancy of each modification in targeted and random regions. Enrichment was calculated as $\log_2$ (target/random). Pairwise Mann–Whitney *U* tests were used to calculate *p*-values, which were then corrected for multiple comparisons via the Bonferroni–Holm method [81]. The H3K4me3 and the SYN4 ChIP-seq data are obtained from published paper [44, 82] and listed in Additional file 4: Table S3.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-024-03465-7.

Additional file 1. Supplementary figures S1-S11.

Additional file 2: Table S1. Summary for PCHi-C and Hi-C data.

Additional file 3: Table S2. Locations of promoter-related loops and fountains.

Additional file 4: Table S3. Summary of public datasets used in this study.

Acknowledgements

We thank Dr. Soichi Inagaki (The University of Tokyo) for providing atxr3 mutant seeds. This study was conducted at the Peking University High Performance Computing Platform, and calculations were performed on CLS-HPC.

Peer review information
Qingxin Song and Wenjing She were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer review history is available in the online version of this article.

Authors' contributions
DW and YZ conceived the study; SX, JS, and LL conducted the experiments; DW, MY, HH, BS, and YZ analyzed the data; DW, SX, MC, and YZ wrote the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by the Biological Breeding-National Science and Technology Major Project (2023ZD04073), grant 32370612 (Yue Zhou) from the National Natural Science Foundation of China, grant JCTD-2022–06 (Yue Zhou) supported by the CAS Youth Interdisciplinary Team, as well as startup funds to Yue Zhou from the State Key Laboratory for Protein and Plant Gene Research, the School of Advanced Agricultural Sciences, and the Peking-Tsinghua Center for Life Sciences at Peking University.

Data availability
The datasets generated during the current study are available in the Genome Sequence Archive (GSA) database under accession numbers CRA014427 [83]. Previously generated Hi-C and RNA-Seq data for Col-0 have been deposited in the European Nucleotide Archive (ENA) under accession code PRJEB52473 [84] and the Sequence Read Archive (SRA) under accession code PRJNA780072 [85], respectively. The published ChIP-Seq data for histone modifications were obtained from the Plant Chromatin State Database [86]. The data of ChIP-seq for H3K4me3 in Col-0 and syn4 were downloaded from SRA under the accession number of PRJNA681034 [87]. The SYN4 ChIP-seq data was downloaded from SRA under the accession number of PRJNA681034 [88]. No custom scripts or software were utilized other than those mentioned in the "Methods" section.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 May 2024 Accepted: 19 December 2024
Published online: 31 December 2024

References

1. Tiang CL, He Y, Pawlowski WP. Chromosome organization and dynamics during interphase, mitosis, and meiosis in plants. Plant Physiol. 2012;158:26–34.
2. Mueller F, Stark H, van Heel M, Rinke-Appel J, Brimacombe R. A new model for the three-dimensional folding of Escherichia coli 16 S ribosomal RNA. III. The topography of the functional centre. J Mol Biol. 1997;271:566–87.
3. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326:289–93.
4. Rust MJ, Bates M, Zhuang XW. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods. 2006;3:793–5.
5. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80.
6. Feng S, Cokus SJ, Schubert V, Zhai J, Pellegrini M, Jacobsen SE. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. Mol Cell. 2014;55:694–707.
7. Grob S, Schmid MW, Grossniklaus U. Hi-C analysis in Arabidopsis identifies the KNOT, a structure with similarities to the flamenco locus of Drosophila. Mol Cell. 2014;55:678–93.
8. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012;148:458–72.
9. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485:381–5.
10. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159:1665–80.
11. Sanborn AL, Rao SSP, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A. 2015;112:E6456–65.
12. Nichols MH, Corces VG. A CTCF code for 3D genome architecture. Cell. 2015;162:702–5.
13. Rao SSP, Huang SC, St Hilaire BG, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. Cohesin loss eliminates all loop domains. Cell. 2017;171:305–20.

14. Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun. 2018;9:189.
15. Baile F, Gomez-Zambrano A, Calonje M. Roles of Polycomb complexes in regulating gene expression and chromatin structure in plants. Plant Commun. 2022;3:100267.
16. Huang Y, Sicar S, Ramirez-Prado JS, Manza-Mianza D, Antunez-Sanchez J, Brik-Chaouche R, Rodriguez-Granados NY, An J, Bergounioux C, Mahfouz MM, et al. Polycomb-dependent differential chromatin compartmentalization determines gene coregulation in Arabidopsis. Genome Res. 2021;31:1230–44.
17. Yang T, Wang D, Tian G, Sun L, Yang M, Yin X, Xiao J, Sheng Y, Zhu D, He H, Zhou Y. Chromatin remodeling complexes regulate genome architecture in Arabidopsis. Plant Cell. 2022;34:2638–51.
18. Yin X, Romero-Campero FJ, Yang M, Baile F, Cao Y, Shu J, Luo L, Wang D, Sun S, Yan P, et al. Binding by the Polycomb complex component BMI1 and H2A monoubiquitination shape local and long-range interactions in the Arabidopsis genome. Plant Cell. 2023;35:2484–503.
19. Bonev B, Cohen NM, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu XL, Lv XD, Hugnot JP, Tanay A, Cavalli G. Multiscale 3D genome rewiring during mouse neural development. Cell. 2017;171:557–72.
20. Matthey-Doret C, Baudry L, Breuer A, Montagne R, Guiglielmoni N, Scolari V, Jean E, Campeas A, Chanut PH, Oriol E, et al. Computer vision for pattern detection in chromosome contact maps. Nat Commun. 2020;11:5759.
21. Guo Y, Al-Jibury E, Garcia-Millan R, Ntagiantas K, King JWD, Nash AJ, Galjart N, Lenhard B, Rueckert D, Fisher AG, et al. Chromatin jets define the properties of cohesin-driven in vivo loop extrusion. Mol Cell. 2022;82:3769-3780 e3765.
22. Wike CL, Guo YX, Tan MY, Nakamura R, Shaw DK, Díaz N, Whittaker-Tademy AF, Durand NC, Aiden EL, Vaquerizas JM, et al. Chromatin architecture transitions from zebrafish sperm through early embryogenesis. Genome Res. 2021;31:981–94.
23. Isiaka BN, Semple JI, Haemmerli A, Thapliyal S, Stojanovski K, Das M, Gilbert N, Glauser DA, Towbin B, Jost D, Meister P. Cohesin forms fountains at active enhancers in *C. elegans*. bioRxiv preprint 2023.
24. Li X, Luo OJ, Wang P, Zheng M, Wang D, Piecuch E, Zhu JJ, Tian SZ, Tang Z, Li G, Ruan Y. Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions. Nat Protoc. 2017;12:899–915.
25. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods. 2016;13:919–22.
26. Fang R, Yu M, Li G, Chee S, Liu T, Schmitt AD, Ren B. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. Cell Res. 2016;26:1345–8.
27. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet. 2015;47:598–606.
28. Martin P, McGovern A, Orozco G, Duffus K, Yarwood A, Schoenfelder S, Cooper NJ, Barton A, Wallace C, Fraser P, et al. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. Nat Commun. 2015;6:10069.
29. Orlando G, Law PJ, Cornish AJ, Dobbins SE, Chubb D, Broderick P, Litchfield K, Hariri F, Pastinen T, Osborne CS, et al. Promoter capture Hi-C-based identification of recurrent noncoding mutations in colorectal cancer. Nat Genet. 2018;50:1375–80.
30. Simpson J, Timko MP, Cashmore AR, Schell J, Montagu MV, Herrera-Estrella L. Light-inducible and tissue-specific expression of a chimaeric gene under control of the 5'-flanking sequence of a pea chlorophyll a/b-binding protein gene. EMBO J. 1985;4:2723–9.
31. Valles MP, Bernues J, Azorin F, Puigdomenech P. Nuclease sensitivity of a maize Hrgp gene in chromatin and in naked DNA. Plant Sci. 1991;78:225–30.
32. Chua YL, Watson LA, Gray JC. The transcriptional enhancer of the pea plastocyanin gene associates with the nuclear matrix and regulates gene expression through histone acetylation. Plant Cell. 2003;15:1468–79.
33. Zhang WL, Wu YF, Schnable JC, Zeng ZX, Freeling M, Crawford GE, Jiang JM. High-resolution mapping of open chromatin in the rice genome. Genome Res. 2012;22:151–62.
34. Zhu B, Zhang WL, Zhang T, Liu B, Jiang JM. Genome-Wide Prediction and validation of intergenic enhancers in arabidopsis using open chromatin signatures. Plant Cell. 2015;27:2415–26.
35. Yan WH, Chen DJ, Schumacher J, Durantini D, Engelhorn J, Chen M, Carles CC, Kaufmann K. Dynamic control of enhancer activity drives stage-specific gene expression during flower morphogenesis. Nat Commun. 2019;10:1705.
36. Zhao HN, Yang MY, Bishop J, Teng YH, Cao YX, Beall BD, Li SL, Liu TX, Fang QX, Fang C, et al. Identification and functional validation of super-enhancers in *Arabidopsis thaliana*. P Natl Acad Sci USA. 2022;119:e2215328119.
37. Zhang YY, Dong QL, Wang Z, Liu QZ, Yu HP, Sun WQ, Cheema J, You QC, Ding L, Cao XF, et al: A fine-scale chromatin landscape reveals chromatin conformation-associated transcriptional dynamics. Nat Commun 2024;15:3235.
38. Sun LH, Zhou JR, Xu X, Liu Y, Ma N, Liu YT, Nie WC, Zou L, Deng XW, He H: Mapping nucleosome-resolution chromatin organization and enhancer-promoter loops in plants using Micro-C-XL. Nat Commun 2024;15:35.
39. Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annu Rev Biochem. 2003;72:449–79.
40. Wasylyk B, Derbyshire R, Guy A, Molko D, Roget A, Teoule R, Chambon P. Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence. Proc Natl Acad Sci U S A. 1980;77:7024–8.
41. Wang C, Liu C, Roqueiro D, Grimm D, Schwab R, Becker C, Lanz C, Weigel D. Genome-wide analysis of local chromatin packing in *Arabidopsis thaliana*. Genome Res. 2015;25:246–56.
42. Zhang Y, Ma M, Liu M, Sun AQ, Zheng XY, Liu KP, Yin CM, Li CS, Jiang CZ, Tu XY, Fang YD: Histone H2A monoubiquitination marks are targeted to specific sites by cohesin subunits in Arabidopsis. Nat Commun 2023;14:1209.
43. Guo L, Yu Y, Law JA, Zhang X. SET DOMAIN GROUP2 is the major histone H3 lysine [corrected] 4 trimethyltransferase in Arabidopsis. Proc Natl Acad Sci U S A. 2010;107:18557–62.
44. Oya S, Takahashi M, Takashima K, Kakutani T, Inagaki S. Transcription-coupled and epigenome-encoded mechanisms direct H3K4 methylation. Nat Commun. 2022;13:4521.
45. Kanno Y, Oikawa T, Chiba Y, Ishimaru Y, Shimizu T, Sano N, Koshiba T, Kamiya Y, Ueda M, Seo M. AtSWEET13 and AtSWEET14 regulate gibberellin-mediated physiological processes. Nat Commun. 2016;7:13245.

46. Chiu WH, Chandler J, Cnops G, Van Lijsebettens M, Werr W. Mutations in the TORNADO2 gene affect cellular decisions in the peripheral zone of the shoot apical meristem of *Arabidopsis thaliana*. Plant Mol Biol. 2007;63:731–44.
47. Huang ACC, Jiang T, Liu YX, Bai YC, Reed J, Qu BY, Goossens A, Nützmann HW, Bai Y, Osbourn A: A specialized metabolic network selectively modulates root microbiota. Science 2019;364:eaau6389.
48. Langenbach C, Campe R, Schaffrath U, Goellner K, Conrath U. UDP-glucosyltransferase UGT84A2/BRT1 is required for Arabidopsis nonhost resistance to the Asian soybean rust pathogen Phakopsora pachyrhizi. New Phytol. 2013;198:536–45.
49. Deng WL, Lee J, Wang HX, Miller J, Reik A, Gregory PD, Dean A, Blobel GA. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell. 2012;149:1233–44.
50. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, et al. YY1 is a structural regulator of enhancer-promoter loops. Cell. 2017;171:1573–88.
51. Richter WF, Nayak S, Iwasa J, Taatjes DJ. The mediator complex as a master regulator of transcription by RNA polymerase II. Nat Rev Mol Cell Biol. 2022;23:732–49.
52. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39:311–8.
53. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009;459:108–12.
54. Wang ZB, Zang CZ, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui KR, Roh TY, Peng WQ, Zhang MQ, Zhao KJ. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet. 2008;40:897–903.
55. Deng L, Zhou Q, Zhou J, Zhang Q, Jia Z, Zhu G, Cheng S, Cheng L, Yin C, Yang C, et al. 3D organization of regulatory elements for transcriptional regulation in Arabidopsis. Genome Biol. 2023;24:181.
56. Zhao HN, Zhang WL, Zhang T, Lin Y, Hu YD, Fang C, Jiang JM. Genome-wide MNase hypersensitivity assay unveils distinct classes of open chromatin associated with H3K27me3 and DNA methylation in *Arabidopsis thaliana*. Genome Biol. 2020;21:24.
57. Liu C, Wang C, Wang G, Becker C, Zaidem M, Weigel D. Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. Genome Res. 2016;26:1057–68.
58. Miao L, Tang Y, Bonneau AR, Chan SH, Kojima ML, Pownall ME, Vejnar CE, Gao F, Krishnaswamy S, Hendry CE, Giraldez AJ. The landscape of pioneer factor activity reveals the mechanisms of chromatin reprogramming and genome activation. Mol Cell. 2022;82:986–1002.
59. Veil M, Yampolsky LY, Grüning B, Onichtchouk D. Pou5f3, SoxB1, and Nanog remodel chromatin on high nucleosome affinity regions at zygotic genome activation. Genome Res. 2019;29:383–95.
60. Pálfy M, Schulze G, Valen E, Vastenhouw NL. Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. PLoS Genet. 2020;16:e1008546.
61. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
62. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 2015;16:259.
63 Flyamer IM, Illingworth RS, Bickmore WA. Coolpup.py: versatile pile-up analysis of Hi-C data. Bioinformatics. 2020;36:2980–5.
64. Lopez-Delisle L, Rabbani L, Wolff J, Bhardwaj V, Backofen R, Gruning B, Ramirez F, Manke T. pyGenomeTracks: reproducible plots for multivariate genomic datasets. Bioinformatics. 2021;37:422–3.
65. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29:24–6.
66 Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, Andrews S. HiCUP: pipeline for mapping and processing Hi-C data. F1000Res. 2015;4:1310.
67. Cairns J, Freire-Pritchett P, Wingett SW, Varnai C, Dimond A, Plagnol V, Zerbino D, Schoenfelder S, Javierre BM, Osborne C, et al. CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. Genome Biol. 2016;17:127.
68. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38:576–89.
69. Csardi G, Nepusz T: The igraph software package for complex network research. Interjournal Complex Syst 2006;1695.
70. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37:907–15.
71. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.
72. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33:290–5.
73. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
74. Yu GC, Wang LG, Han YY, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.
75. Cui X, Lu FL, Qiu Q, Zhou B, Gu LF, Zhang SB, Kang YY, Cui XK, Ma X, Yao QQ, et al. REF6 recognizes a specific DNA sequence to demethylate H3K27me3 and regulate organ boundary formation in Arabidopsis. Nat Genet. 2016;48:694–9.
76. Moreno-Romero J, Jiang H, Santos-González J, Köhler C. Parental epigenetic asymmetry of PRC2-mediated histone modifications in the endosperm. EMBO J. 2016;35:1298–311.
77. Inagaki S, Takahashi M, Hosaka A, Ito T, Toyoda A, Fujiyama A, Tarutani Y, Kakutani T. Gene-body chromatin modification dynamics mediate epigenome differentiation in Arabidopsis. Embo J. 2017;36:970–80.

78. Bewick AJ, Ji LX, Niederhuth CE, Willing EM, Hofmeister BT, Shi XL, Wang L, Lu ZF, Rohr NA, Hartwig B, et al. On the origin and evolutionary consequences of gene body DNA methylation. P Natl Acad Sci USA. 2016;113:9111–6.
79. Liu Y, Tian T, Zhang K, You Q, Yan H, Zhao N, Yi X, Xu W, Su Z. PCSD: a plant chromatin state database. Nucleic Acids Res. 2018;46:D1157–67.
80. Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, et al. BEDOPS: high-performance genomic feature operations. Bioinformatics. 2012;28:1919–20.
81. Holm S. A simple sequentially rejective multiple test procedure. SJS. 1979;6:65–70.
82. Marbouty M, Le Gall A, Cattoni DI, Cournac A, Koh A, Fiche JB, Mozziconacci J, Murray H, Koszul R, Nollmann M. Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. Mol Cell. 2015;59:588–602.
83. Wang D, Xiao S, Shu J, Luo L, Yang M, Calonje M, He H, Song B, Zhou Y. Promoter capture Hi-C identifies promoter-related loops and fountain structures in Arabidopsis. CRA014427. Genome Sequence Archive. 2024. https://ngdc.cncb.ac.cn/gsa/browse/CRA014427.
84. Yin X, Romero-Campero, F J, Yang M, Baile F, Cao Y, Shu J, Luo L, Wang D, Sun S, Yan P, Gong Z, Mo X, Qin G, Calonje M, Zhou Y. Binding by the Polycomb complex component BMI1 and H2A monoubiquitination shape local and long-range interactions in the Arabidopsis genome. PRJEB52473. European Nucleotide Archive. https://www.ebi.ac.uk/ena/browser/view/PRJEB34105 (2023).
85. Yang T, Wang D, Tian G, Sun L, Yang M, Yin X, Xiao J, Sheng Y, Zhu D, He H, Zhou Y. Chromatin remodeling complexes regulate genome architecture in Arabidopsis. PRJNA780072. Sequence Read Archive. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA780072 (2022).
86. Liu Y, Tian T, Zhang K, You Q, Yan H, Zhao N, Yi X, Xu W, Su Z. PCSD: a plant chromatin state database. https://systemsbiology.cau.edu.cn/chromstates (2018).
87. Oya S, Takahashi M, Takashima K, Kakutani T, Inagaki S. Transcription-coupled and epigenome- encoded mechanisms direct H3K4 methylation. PRJNA681034. Sequence Read Archive. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA681034 (2022).
88. Zhang Y, Ma M, Liu M, Sun A, Zheng X, Liu K, Yin C, Li C, Jiang C, Tu X, Fang Y. Histone H2A monoubiquitination marks are targeted to specific sites by cohesin subunits in Arabidopsis. PRJNA681034. Sequence Read Archive. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA681034 (2024).

Publisher's Note