Denaud *et al. Genome Biology* (2025) 26:116
https://doi.org/10.1186/s13059-025-03587-6

Genome Biology

BRIEF REPORT

Open Access

# Determining the functional relationship between epigenetic and physical chromatin domains in *Drosophila*

Sandrine Denaud[1†], Gonzalo Sabarís[1†], Marco Di Stefano[1†], Giorgio-Lucio Papadopoulos[1], Bernd Schuettengruber[1*] and Giacomo Cavalli[1*]

†Sandrine Denaud, Gonzalo Sabarís and Marco Di Stefano contributed equally to this work.

*Correspondence: bernd.schuttengruber@igh.cnrs.fr; giacomo.cavalli@igh.cnrs.fr

[1] Institute of Human Genetics, UMR9002 CNRS and University of Montpellier, 141 Rue de La Cardonille, Montpellier Cedex 5 34396, France

## Abstract

The tight correlation between topologically associating domains (TADs) and epigenetic domains in *Drosophila* suggests that the epigenome contributes to define TADs. However, it is still unknown whether histone modifications are essential for TAD formation and structure. By either deleting or shifting key regulatory elements needed to establish the epigenetic signature of Polycomb TADs, we show that the epigenome is not a major driving force for the establishment of TADs. On the other hand, physical domains have an important impact on the formation of epigenetic domains, as they can restrict the spreading of repressive histone marks and looping between cis-regulatory elements.

**Keywords:** Chromatin, Genome organization, Epigenome, TADs, *Drosophila*, PcG proteins

## Background

Although the organization of the genome into topologically associating domains (TADs) is well established, the processes that form TADs are still unclear. CTCF/cohesin loop extrusion is emerging as a key mechanism for the formation of TADs in mammals [1]. In contrast, genome organization in *Drosophila* appears to depend primarily on the partitioning of chromatin state domains, whereas insulator proteins are not required for the establishment of TADs in early *Drosophila* embryos [2]. *Drosophila* TAD borders do not correspond to highly localized loop anchor elements, but the interaction between active domains flanking inactive domains may lead to a loose topology based on chromatin state domains rather than on precise loop anchor elements [3, 4]. Indeed, in *Drosophila*, a strong correlation exists between 3D chromatin structure and the epigenome. This has suggested the hypothesis that the partitioning of chromatin states between domains of histone acetylation and methylation, in particular H3 K27 me3, might contribute to TAD

formation [3, 5, 6]. Furthermore, high-resolution Hi-C experiments suggested that TAD organization in *Drosophila* reflects the switch between active and inactive chromatin states [7]. Super-resolution microscopy coupled to DNA labeling (3D-STORM, ORCA) confirmed the strong correlation of TADs with epigenetic domains at a single cell level [8–10]. Finally, H3 K27 me3 is intergenerationally transmitted in flies and the majority of H3 K27 me3 domains can be detected before the major wave of the zygotic genome activation (ZGA) when TADs are established [11, 12].

Altogether, these lines of evidence strongly suggest that histone modifications could be major drivers of TAD formation in *Drosophila*, but this hypothesis has never been directly tested due to an important experimental bottleneck: to directly address the importance of a chromatin state for TAD formation and structure, it is necessary to change this signature without globally affecting gene expression. Here we overcome this challenge by employing two complementary perturbations using CRISPR/Cas9 genome engineering, where we either change the epigenetic signature defining a specific Polycomb-associated TAD or create a new epigenetic domain and analyze its consequences on TAD formation.

## Results and discussion

The key developmental regulator gene *dachshund* (*dac*) is located within a TAD spanning around 180 kb. The chromatin state within this TAD is defined by the presence of PcG proteins and the hallmark of PcG-dependent gene silencing, H3 K27 me3 (Fig. 1a). This epigenetic signature corresponds to the previously identified Polycomb (also defined as "blue") chromatin state [13] and borders of the epigenetic domain correlate with borders of the physical domain. We used CRISPR/Cas9 genome engineering to delete the two Polycomb Response Elements (643 bp for PRE1, and 1090 bp for PRE2)

(See figure on next page.)

**Fig. 1** Erasure of an epigenetic TAD signature. **a** Representation of the *dac* TAD. Hi-C score maps around the *dac* gene locus in *Drosophila* embryos [12]. Dac TAD and neighboring TADs are indicated by a black triangle. Below ChIP-seq tracks and RNA-seq track of late embryos are shown (modENCODE). **b** CUT&RUN profiles for H3 K27 me3 in *Drosophila* embryos and larval imaginal leg discs from WT flies or Double PRE mutant flies. The left panel shows the *dac* domain; the right panel shows a neighboring H3 K27 me3 domain (*esg/sna* genes) which is unaffected by the perturbations. **c** Hi-C score maps of a 1 Mb kb region (upper panel) or 200 kb region (lower panel) around the dac TAD in WT or Double PRE mutant imaginal leg discs. Dac domain is indicated by a black triangle. PRE loop by a black circle. Violet bars indicate the position of PREs. Gray bars indicate the position of the TAD boundary. Black arrows indicate the promoters of the *dac* and the *CG5888* genes. **d** Insulation profile shown at 3 kb resolution along the *dac* 200 kb region in larval WT and Double PRE mutant leg discs is shown as the mean value (line) ± the standard deviation (shaded area) over the insulation scores (IS) computed using 5 different values of the window parameter ($w = 100$, 150, 200, 250, and 300 kb, see the "Methods" section). **e** *P*-values from the comparisons of insulation scores (IS) at the left and right TAD border between the WT and double PRE mutant line (see the "Methods" section). The *p*-values resulted from a two-sided Welch *t*-test between the WT condition and the double PRE mutant at the corresponding locus. **f** Differential Hi-C scores maps (double PRE mutant *vs.* WT) of a 150 kb region of the *dac* gene in third instar imaginal leg disc (see Methods). Red regions indicate higher contacts in Double PRE mutant; blue regions indicate higher contacts in WT. Black circle indicates the position of the PRE loop. Violet bars indicate the position of PREs. Black arrows indicate the promoters of the *dac* and the *CG5888* genes. **g** Quantification of the *dac* PRE loop Hi-C interaction scores in WT ($n = 172$) and double PRE ($n = 92$) mutant flies. Reported *p*-values result from comparing the WT and double PRE mutant distributions using the unpaired two-sided Wilcoxon statistical test. The number of points per boxplot (bottom) is reported (see the "Methods" section). Boxplots show median (central line), Q1 = 25 th and Q3 = 75 th percentiles (box limits), and Q1 + 1.5 × IQR to Q3 + 1.5 × IQR (wiskers), where IQR is the interquartile range

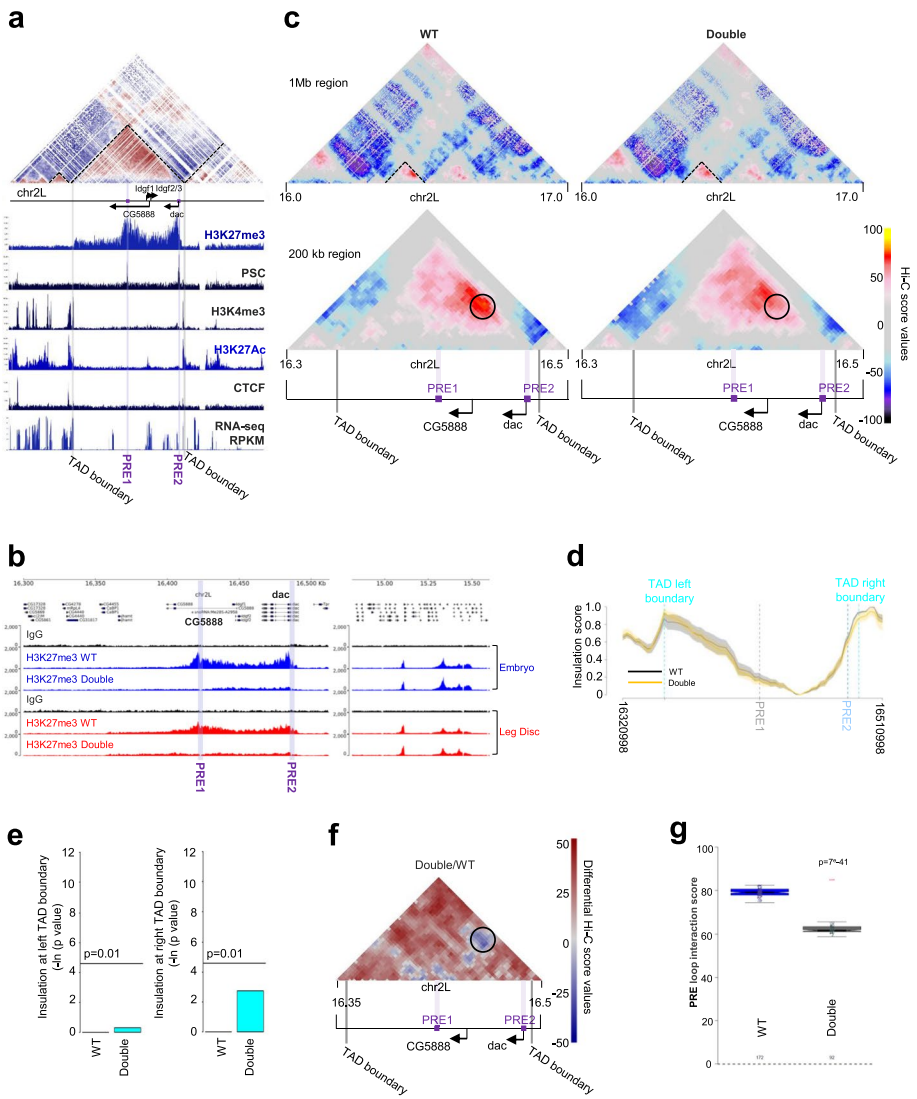


**Fig. 1** (See legend on previous page.)

that act as nucleation sites for PcG protein binding, which is expected to create the epigenetic chromatin domain. The resulting line was defined as "Double PRE mutant" [12]. As previously described, homozygous flies are viable [12]. Male flies show extra sex comb bristles on the first tarsal segment on the first leg, which is due to the specific overexpression of the *dac* gene in the second tarsal segment at the pupal stage, whereas expression is not significantly affected at earlier developmental stages [14]. CUT&RUN experiments in both embryos and third instar larval leg discs showed that, upon deletion of both PREs, only residual levels of H3 K27 me3 can be detected at the *dac* gene locus, whereas other H3 K27 me3 domains are unaffected (Fig. 1b, Additional file 1: Fig. S1a). To analyze the effect of the loss of the PcG-associated epigenetic signature on gene expression, we performed RNA-seq analysis in Double PRE mutant imaginal leg discs. Surprisingly, the substantial reduction of the H3 K27 me3 repressive mark did not significantly impact the expression of genes within or flanking the *dac* TAD, including *dac* itself. The only exception was the CG5888 gene, which showed moderate upregulation

(Additional file 1: Fig. S1b). The absence of global gene activation upon PRE deletion is consistent with two previous reports deleting PREs at the *inv/en* or *vg* gene locus [15, 16]. Noteworthy, as it is the case of the *dac* gene locus, deletions of the two PREs at the *vg* locus did not result in gene activation, while it left residual H3 K27 me3 levels that were similar to those found in cells where the *vg* gene is active [15]. Therefore, it is unlikely that remaining low levels of repressive histone marks account for maintaining genes repressed. The most likely explanation for these results is that essential activators might be missing in these cells. Finally, RNA-FISH experiments revealed that activation of *CG5888* only occurs in a small subset of cells of the larval leg disc (Additional file 1: Fig. S1c). Together, these results show a clear uncoupling between the gene expression changes and the PcG epigenetic chromatin signature at the *dac* TAD, an important feature which allowed us to analyze the structural consequences of PRE deletion in the absence of transcriptional effects. We next performed Hi-C experiments of wild type (WT) or Double PRE mutant imaginal leg discs (Fig. 1c, Additional file 1: Fig. S1 d) to analyze whether the loss of the epigenetic signature affects the formation and structure of the underlying TAD. Comparison of Hi-C maps and quantification of insulation scores of the TAD borders revealed that global TAD structure and its boundaries are not significantly affected upon changing the underlying chromatin state (Fig. 1d, e). On the other hand, the significant reduction in H3 K27 me3 levels correlates with changes in the intra-TAD organization, specifically the loss of PRE looping interactions, as previously shown [12] (Fig. 1f, g). In summary, these results suggest that the epigenetic signature of a Polycomb TAD is not the major driving force for the formation of a physical domains and TAD boundary delimitation.

As a complementary strategy to dissect the functional link between epigenetic domains and physical domains, we used a previously described fly line where the endogenous PRE1 is deleted (ΔPRE1) [12] and created from this line another one where we inserted the same PRE1 sequence upstream to the right *dac* TAD border, in the flanking black chromatin [13] that has no Polycomb binding despite being transcriptionally inactive (PRE1_UP line). Importantly, the distance of the new PRE insertion to the endogenous PRE2 at the *dac* transcription start site (*dac* TSS) is the same as the distance (65 kb) between the two endogenous PREs (Fig. 2a). This experimental setup allowed us to examine whether the two PRE sequences (endogenous PRE2 and ectopic PRE1) can create a repressive chromatin domain, as it is the case for the endogenous *dac* locus, and whether this affects TAD structure or borders of the endogenous domain. At the same time, this allowed us to assess the importance of TADs for the formation of epigenetic domains or vice versa.

CUT&RUN experiments showed that insertion of the PRE sequence upstream to the *dac* locus results in the local recruitment of the PRC1 component PH, indicating that the PRE is functional (Fig. 2b). Notably, PcG levels at the ectopic PRE1 insertion site are comparable to PcG levels bound to the endogenous PRE1. However, the spreading of H3 K27 me3 from the ectopically inserted PRE1 is strongly limited, resulting in a small H3 K27 me3 domain that does not exceed 10 kb (Fig. 2b). Remarkably, a similar extent of H3 K27 me3 levels is observed when we insert the PRE2 sequence in the upstream domain (Additional file 1: Fig. S2a). Since the PRE2 sequence has stronger PRE activity in terms of PcG binding levels and silencing activity in transgenic reporter assays compared to

PRE1 [12], these data suggest that the reduced H3 K27 me3 spreading from the ectopically inserted PRE1 is not due to the lower affinity of PcG proteins to the PRE1 sequence. Altogether, these results suggest that the endogenous PRE at the *dac* TSS and an ectopically inserted PRE in its upstream TAD do not cooperate to form a large repressive H3 K27 me3 domain, as it is the case when they are located within the same TAD.

The presence of multiple insulator binding sites between the two PREs (Additional file 1: Fig. S2a) might interfere with the ability of the two PREs to efficiently establish a larger epigenetic chromatin domain when located in different physical domains. However, we note that insulator binding sites do not precisely colocalize with borders of the ectopic H3 K27 me3 domain (Fig. S2a), making it unlikely that they directly account for the reduced spreading of H3 K27 me3 from the ectopic PRE. In addition, we previously showed that insertion of a gypsy insulator sequence within the dac domain does not interfere with the deposition of the H3 K27 me3 from the two PREs [14]. We therefore favor the hypothesis that the endogenous chromosomal location of PREs imparts stability to Polycomb domains [16], whereas PREs outside their endogenous context are less efficient in mediating histone methylation of their surrounding chromatin [17].

To examine how TAD structure, their borders, and PRE chromatin contacts are affected upon shifting the position of the PREs, we performed Hi-C experiments in the line where the PRE position is shifted (PRE1_UP). Comparison of Hi-C maps revealed that endogenous TAD structures are largely maintained upon deletion of endogenous PRE1 (ΔPRE1) and insertion of PRE1 in the upstream TAD (PRE1_UP)

(See figure on next page.)

**Fig. 2** Shifting the PRE position to generate a new epigenetic signature. **a** Schematic representation of the *dac* TAD and its upstream TAD in WT, ΔPRE1 flies, and PRE1_UP flies. **b** CUT&RUN profiles for H3 K27 me3 and PH at the *dac* gene locus in WT, ΔPRE1, PRE1_UP flies. PRE1, PRE2, and the upstream PRE insertion site are indicated by gray bars. Note that In the PRE1_UP line, H3 K27 me3 and PH profiles were aligned to a custom genome in which PRE1 was removed from its endogenous location and inserted at the ectopic site (see the "Methods" section). RNA-seq track of late embryos is shown (modENCODE). **c** Hi-C score (see the "Methods" section) maps of a 300 kb region at 3 kb resolution on chromosome 2L, including the *dac* gene locus in WT, ΔPRE1 and PRE1_UP third instar imaginal leg disc. Violet bars indicate the positions of the PREs. Gray bars indicate the position of the TAD boundary. Black arrows indicate gene promoters of the *dac* and the *CG5888*. The black dashed rectangle indicates higher contact regions of chromatin marked by H3 K27 me3. **d** Insulation profile shown at 3 kb resolution along the *dac* 200 kb region in larval WT and ΔPRE1 (top panel, green) and PRE1_UP (bottom panel, light blue) mutant fly lines is shown as the mean value (line) ± the standard deviation (shaded area) over the insulation scores (IS) computed using 5 different values of the window parameter ($w = 100$, 150, 200, 250, and 300 kb, see the "Methods" section). **e** *P*-values from the comparisons of insulation scores (IS) at the left and right TAD borders between the WT and the corresponding fly line (see Methods). The *p-values* resulted from a two-sided Welch t-test between the WT condition and each of the mutants at the corresponding locus. **f** Quantification of *dac* PRE loop (left) Hi-C interaction score in WT, ΔPRE1, and PRE1_UP mutant flies. The left panel shows quantification of the endogenous PRE loop WT, ΔPRE1, and PRE1_UP; the right panel quantifies the PRE loop between endogenous PRE2 and upstream PRE1. Reported *p-values* result from the comparisons between WT, ΔPRE1, and PRE1_UP mutant distributions using the unpaired two-sided Wilcoxon statistical test. The number of points per boxplot is reported. Boxplots show median (central line), Q1 = 25 th and Q3 = 75 th percentiles (box limits), and Q1 + 1.5 × IQR to Q3 + 1.5 × IQR (whiskers), where IQR is the interquartile range. **g** Differential Hi-C interaction score maps (PRE1_UP mutant vs ΔPRE1) of a 300 kb region of the *dac* gene in third instar imaginal leg disc (see the "Methods" section). Red regions indicate higher contacts in PRE1_UP mutant, blue regions indicate higher contacts in ΔPRE1. The black circle indicates the position of the PRE loop. Violet bars indicate the positions of the PREs. Black arrows indicate the promoters of the *dac* and the *CG5888* genes. The black dashed rectangle indicates higher contact regions of chromatin marked by H3 K27 me3

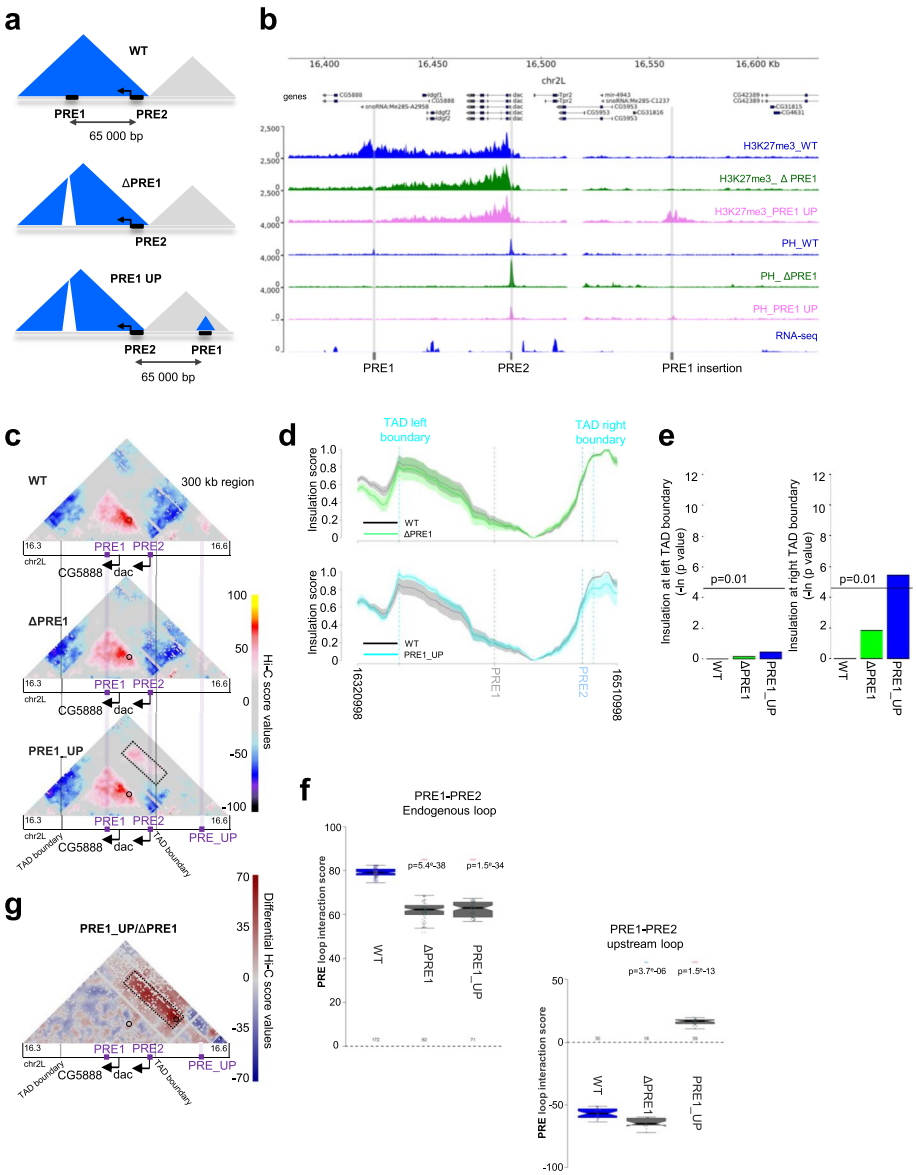


**Fig. 2** (See legend on previous page.)

(Fig. 2c, Additional file 1: Fig. S2b), although a slight but significant decrease in the insulation score between the endogenous *dac* TAD and the upstream TAD can be observed upon insertion of the ectopic PRE (Fig. 2d, e).

To analyze the effect of a TAD boundary placed between two PRE anchors on their ability to form chromatin loops, we quantified chromatin contacts between the endogenous PRE2 (*dac* TSS) and the PRE1 in its endogenous and ectopic integration site in the upstream TAD in each fly line. As expected, the endogenous PRE loop is disrupted when the PRE 1 is deleted within the *dac* TAD in both ΔPRE1 and PRE1_UP (Fig. 2f, left). On the other hand, a slight increase in chromatin contacts is observed between the endogenous PRE2 and the PRE1 inserted in the upstream TAD in PRE1_UP, although chromatin interactions between these genomic loci stay much

lower than those of the endogenous PRE loop (Fig. 2f, right). This indicates that the two PREs, when positioned in different TADs, have a significantly reduced potential to interact with each other. The inability of the ectopic PRE to loop with the endogenous PRE is unlikely to be due to the absence of an extended H3 K27 me3 domain at the ectopic gene locus. Although, endogenous PRE loops are largely limited to PREs located within the same TAD [12, 18], not all PREs within Polycomb domains form chromatin loops and insertion of a gypsy insulator between the two PREs at the *dac* gene locus decreases PRE looping without affecting the deposition of H3 K27 me3 mark [14]. This favors the hypothesis that PRE looping is restrained by physical constraints like insulator binding or endogenous TAD structures, whereas the simple presence of PREs within H3 K27 me3 marked chromatin is not sufficient to allow PRE looping.

Interestingly, by analyzing the differential score enrichment of interactions in the ΔPRE1 line versus the PRE1_UP line (Fig. 2g), we observed increased interactions of the whole *dac* TAD with a smaller region around the upstream PRE insertion side. These increased contacts might reflect the high affinity for homotypic interactions between H3 K27 me3 marked chromatin domains, which has been previously reported in flies and mammals [19–21], resulting in the compartmentalization of repressive chromatin domains. This, in turn, may explain the slightly reduced insulation activity of the boundary region between the two TADs (Fig. 2d).

## Conclusions

In conclusion, the two experimental approaches show that, although TAD structures are tightly correlated to chromatin states in *Drosophila*, the epigenetic mark H3 K27 me3 is not the major driving force for TAD formation and delimitation but contributes to TAD compartmentalization. Conversely, physical domains impact the formation of epigenetic domains and the interaction of PREs.

## Methods

### Fly work and generation of mutant flies by CRISPR/Cas9 genome engineering

All flies were raised on standard corn meal yeast extract medium at 25 °C. CRISPR/Cas9 mutant fly lines Double and ΔPRE1 are described in [12].

Sequences of gRNAs used to create fly lines PRE1_UP and PRE2_UP are described in Additional file 2: Table S1. Sense and antisense oligonucleotides were annealed and phosphorylated by the T4 polynucleotide kinase (NEB#M0201S) before being inserted inside a pCFD3 plasmid (Addgene #49,410) previously digested by BbsI (NEB#R0539S). To create the pHD-dsRED donor plasmid (Addgene) containing a removable (floxed) 3XP3-dsRED construct flanked by loxP sites and DNA fragments having homology to the target regions (homology arms) serving as a template for homology-directed repair, 1.5 kb genomic DNA fragments were amplified by PCR (Additional file 2: Table S1) and inserted into the pHD-dsRED plasmid using the GIBSON assemble (kit NEBuilder NEB#E2621S).

The PRE1 and PRE2 sequences were PCR amplified and introduced into the donor plasmid cut by SpeI and BglII using GIBSON cloning (Additional file 2: Table S1). To generate mutant fly lines, gRNA-containing pCFD3 and pHD-dsRED donor plasmids

were injected into flies expressing Cas9 in the germline (vas-Cas9(X) RFP-; Bloomington stock #55,821). Injections and dsRED screening were performed by BestGene (https://www.thebestgene.com/). To remove the dsRED reporter construct, mutant flies were crossed with a fly line expressing CRE recombinase (Bloomington stock #34,516).

To generate the PRE1_UP and PRE2_UP mutant line gRNAs targeting the PRE upstream region, the corresponding donor plasmid was injected into ΔPRE1 mutant lines previously generated and expressing Cas9 (vas-Cas9(III)). Genotypes of mutant fly lines were confirmed by PCR genotyping and sequencing analysis of the mutated region.

#### CUT&RUN

CUT&RUN experiments were performed as described by Kami Ahmad in protocilas.io (https://dx.doi.org/https://doi.org/10.17504/protocols.io.umfeu3n) with minor modifications. 50 leg discs were dissected in Schneider medium, centrifuged for 3 min at 700 g, and washed twice with wash + buffer before the addition of Concanavalin A-coated beads. Embryos were collected and homogenized using a glass douncer in Nuclear extraction buffer (20 mM HEPES, 10 mM KCl, 0.1% TritonX, 20% Glycerol), then washed once with wash + buffer before the addition of Concanavalin A-coated beads.

MNase digestion (pAG-MNase Enzyme from Cell Signaling) was performed for 30 min on ice. After ProteinaseK digestion, DNA was recovered using SPRIselect beads and eluted in 50ul TE. DNA libraries for sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina. Sequencing (paired-end sequencing 150 bp, approx. 2 Gb/sample) was performed by Novogene (https://en.novogene.com/).

#### CUT&RUN analysis

Fastq files were aligned using Bowtie 2 (v 2.4.2) [22] with the following parameters: –local –-very-sensitive-local –no-unal –no-mixed –no-discordant –phred33 -I 10 -X 700. The samples for WT and ΔPRE1 were aligned to the *D. melanogaster* reference genome dm6 in https://s3.amazonaws.com/igenomes.illumina.com/ and samples for the PRE1_UP condition were aligned to a modified genome. First, the 1.091 bp region chr2L:16,422,435–16,423,525 including the PRE1 was removed from the dm6 reference genome and replaced with a string of 1091 n's (undetermined nucleotides). Next, the removed 1091 bp-long sequence was inserted at position chr2L:16,560,363 bp to create the modified genome. SAM files were compressed into BAM files using samtools (v 1.16.1) [23] and reads with low mapping quality (Phred score < 30) were discarded. Duplicate reads were removed using sambamba markdup (v 1.0.0) [24] with the following parameters: -r –hash-table-size 500,000 –overflow-list-size 500,000. For visualization, replicates were merged using samtools merge with default parameters and reads per kilo base per million mapped reads (RPKM)-normalized bigWig binary files were generated using the bamCoverage (v 3.5.5) function from deepTools2 [25] with the following parameters: –normalizeUsing RPKM –ignoreDuplicates -e 0 -bs 10. Two biological replicates were used per condition and the IgG condition was used as a control for the differential enrichment analysis in the 131 Drosophila Polycomb domains [19] performing the DESeq2 method from the "DiffBind" R package (v 3.12.0) (Additional file 3: Table S2).

### RNA-seq experiments

Third instar larval leg discs were dissected in Schneider medium on ice. Total RNA was extracted using TRIzol reagent. RNA purification was performed using the RNA Clean & Concentrator kit (Zymo Research, #R1015). Finally, poly-A RNA selection, library preparation, and Illumina sequencing (20 M paired-end reads, 150nt) were performed by Novogene (https://en.novogene.com/). All experiments were performed in triplicate.

### RNA-seq processing and differential analysis

RNA-seq data quality was assessed using FastQC (v0.12.1). Stranded RNA-seq data were mapped to the Drosophila melanogaster dm6 genome (FlyBase dmel_r6.34) using Subread-align (Subread v2.0.6) with default parameters. Samtools merge was used to merge bam files from the same sample sequenced in different lanes. (Samtools v1.16.1). Aligned sequencing reads mapped to gene transcripts were counted using featureCounts (Subread v 2.0.6) with -s 2 (reverse stranded) and default parameters. The gene transcript annotation file was obtained from FlyBase (release 6.34). Prior to statistical analysis, genes with fewer than 10 reads (cumulating all samples analyzed) were removed. Differentially expressed genes (DEGs) were identified using the DESeq2 R package (Additional file 4: Table S3). Genes with adjusted $p$-value < 0.01 (using the Benjamini–Hochberg FDR method) and |log2 FC|> 1 were considered differentially expressed. Volcano plots were generated using the "EnhancedVolcano" R package (https://doi.org/10.18129/B9.bioc.EnhancedVolcano). modENCODE RNA-seq data from whole embryo 14–16 h post synchronization at egg laying stage (ENCSR246 JXB) was used for visualization purposes.

### Hi-C

Hi-C experiments were performed using the EpiTect Hi-C Kit (Quiagene#59,971). All Hi-C experiments were performed in two or three independent experiments using 50 3rd instar imaginal leg discs. Briefly, discs were homogenized and fixed in activated Buffer T and 2% Formaldehyde using Tissue Masher tubes (Biomasher II (EOG-sterilized) 320,103 Funakoshi). Tissue was digested by adding 25 µl Collagenase I and II (40 mg/ml) for 1 h at 37 °C. Samples were centrifuged and supernatant was carefully aspirated, leaving ~250 µl of solution in the tube. Then 250 µl QIAseq Beads equilibrated to room temperature were added to bind nuclei to the beads, and all subsequent reactions were performed on the beads according to the manufacturer's protocol. Libraries were sequenced at BGI (https://www.bgi.com/) PE 150 (approx. 400 million reads per replicate).

### Hi-C analysis

Raw data from Hi-C sequencing were processed by using the "shHiC2" pipeline. Sequencing statistics are summarized in Additional file 5: Table S4. Valid interactions were stored in a database using the "misha" R package (https://github.com/msauria/misha-package). Extracting the valid interactions from the *misha* database, the "shaman" R package [https://bitbucket.org/tanaylab/shaman] has been used for computing the Hi-C expected models, Hi-C scores with parameters $k = 250$ and $k\_exp = 500$ (Figs. 1c and 2c), and differential Hi-C interaction scores with parameters $k = 250$ and $k\_exp = 250$ and for each comparison down-sampling the compared datasets to have the

same number of valid pairs in chr2L (Figs. 1f and 2g). Specifically, Hi-C scores quantify the contact enrichment (positive values) or depletion (negative values) of each bin of the map with respect to a statistical model used to evaluate the expected number of counts. To generate this expected model, we (randomized) shuffled the observed Hi-C contacts using a Markov Chain Monte Carlo-like approach per chromosome [26]. Shuffling is done such that the marginal coverage and decay of the number of observed contacts with the genomic distance are preserved, but any features of genome organization (e.g., TADs or loops) are not. These expected maps were generated for each biological replicate separately and contain twice the number of observed cis-contacts. Next, the score for each contact in the observed contact matrix was calculated using the *k* nearest neighbors (kNN) strategy [26]. In brief, the distributions of two-dimensional Euclidean distances between the observed contact and its nearest *k_exp* neighbors in the pooled observed and pooled expected (per cell type) data are compared, using Kolmogorov–Smirnov *D* statistics to visualize positive (higher density in observed data) and negative (lower density in observed data) enrichments. These *D*-scores are then used for visualization ($-100$ to $+100$ scale) and are referred to as Hi-C scores in the text. Accordingly, the color scale of the Hi-C scores comprises both positive and negative values. When computing the differential Hi-C scores maps, the reference dataset (WT in Fig. 1f and ΔPRE1 in Fig. 2g) was used as the expected model.

For each condition, the Hi-C interaction quantifications at the dac PRE loop (Fig. 1g, 2f) were performed by considering the Hi-C scores between two regions of 6 kb: in Figs. 1g and 2f (*left*), we considered the locations of the two endogenous PREs (PRE1 in chr2L:16,419,514–16,425,515 bp and PRE2 in chr2L:16,482,929–16,488,930 bp) and in Fig. 2f (*right*) we considered the locations of the endogenous PRE2 and the insertion site of the relocated PRE1 (PRE2 in chr2L:16,482,929–16,488,930 bp and PRE1_UP in chr2L:16,557,363–16,563,364 bp). Each of the comparisons of the Hi-C interaction quantifications at the *dac* PRE loop was performed between the WT (control line) and the indicated mutant line. An unpaired two-sided Wilcoxon statistical test (H0: true median shift is equal to 0. The two variables are not normally distributed) was used to estimate the reported *p*-values.

The insulation scores [27] were computed on the observed Hi-C datasets binned at 2 kb resolution with windows of 100, 150, 200, 250, and 300 kb, resulting in five values per bin, and were stored in the *misha* database using an *in-house* R script. The mean and standard deviation for each of the 2 kb-bins were computed and were used for the plots in Figs. 1d and 2d. The quantification of the insulation scores (IS) at the TAD borders was performed by applying the pair-wise statistical comparison of the five IS quantifications per 2 kb-bins. The *p*-values in Figs. 1e and 2e resulted from a Welch *t*-test (H0: true difference in means is equal to 0. The variances of the samples are thought not to be equal) between the WT condition and each of the mutant conditions at the corresponding locus. Fig. S1 d and Additional file 1: Fig. S2b show the maps obtained by applying the SCALE normalization of the command juicer tool (version 1.22.01) (Command: *java -Xmx300G -jar scripts/juicer_tools_1.22.01.jar addNorm -d -F -k SCALE*) on the.*hic* files. All plots of Hi-C maps, Hi-C interaction scores comparisons, insulation score (IS) profiles with window 100 kb, and *p*-values of IS comparisons were obtained with *in-house* R and gnuplot scripts provided in Github (see Availability of data and materials).

### RNA-FISH

RNA FISH was performed as described in [14].

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03587-6.

Additional file 1. Fig S1. Data set related to Fig. 1. Fig S1a: MA plot showing the log2 fold change of H3 K27 me3 in Double PRE mutant versus WT. Fig S1b: Volcano plot showing the number of differentially expressed genes in Double PRE mutant *vs* WT. Fig S1c: RNA FISH images of WT and Double PRE mutant third instar imaginal leg discs. Fig S1 d: Normalized Hi-C contact maps of a 1Mb kb regionor 200kb regionaround the *dac* TAD in WT or Double PRE mutant imaginal leg discs. Fig S2. Data set related to Fig. 2. Fig S2a: Snapshots of the H3 K27 me3 CUT&RUN profiles. Fig S2b: Normalized Hi-C contact maps of a 300 kb region at 3kb resolution on chromosome 2L including the *dac* gene locus in WT, ΔPRE1 and PRE1_UP third instar imaginal leg disc.

Additional file 2. Table S1. List of Primers used in the study. Table S2. Differential enrichment analysis of H3 K27 me3 in embryo and third instar larval leg imaginal disc of Double PRE mutant versus WT flies in the 131 Drosophila Polycomb domains. Table S3. Differential expression analysis results and DESeq2 normalized counts of Double PRE mutant versus WT flies. Table S4. Hi-C alignment statistics

**Author information**
Sandrine Denaud, Gonzalo Sabarís, and Marco Di Stefano contributed equally to this work. Correspondence to Bernd Schuettengruber and Giacomo Cavalli.

**Acknowledgements**
This work was supported by grants from the European Research Council (Advanced Grant 3DEpi and WaddingtonMemory), the European CHROMDESIGN ITN project (Marie Skłodowska-Curie grant agreement No 813327), the European E-RARE NEURO DISEASES grant "IMPACT", by the Agence Nationale de la Recherche (PLASMADIFF3D, grant N. ANR-18-CE15-0010, LIVCHROM, grant N. ANR-21-CE45-0011), by the Fondation pour la Recherche Médicale (EQU202303016280), by the MSD Avenir Foundation (Project GENE-IGH), and by the French National Cancer Institute (INCa, PIT-MM grant N. INCA-PLBIO18-362).

**Peer review information**
Andrew Cosgrove was the primary editor 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**
B.S. and G.C conceived and designed the study. B.S. generated mutant fly lines and performed RNA-seq and CUT&RUN experiments. S.D. performed Hi-C and RNA-FISH experiments. G.S. established CUT&RUN protocol and performed computational analysis of CUT&RUN experiments and RNA-seq experiments. M.D.S. and G.P performed computational analysis of Hi-C experiments. B.S. and G.C wrote the manuscript with input from all the authors. All authors read and approved the final manuscript.

**Funding**
This work was supported by grants from the European Research Council (Advanced Grant 3DEpi and WaddingtonMemory), the European CHROMDESIGN ITN project (Marie Skłodowska-Curie grant agreement No 813327), the European E-RARE NEURO DISEASES grant "IMPACT" by the Agence Nationale de la Recherche (PLASMADIFF3D, grant N. ANR-18-CE15-0010, LIVCHROM, grant N. ANR-21-CE45-0011), by the Fondation pour la Recherche Médicale (EQU202303016280), by the MSD Avenir Foundation (Project GENE-IGH), and by the French National Cancer Institute (INCa, PIT-MM grant N. INCA-PLBIO18-362).

**Data availability**
All genomic data have been deposited on GEO: GSE274157, GSE274158 and GSE274159 [28–30]. Hi-C data for WT and gypsy 2 are available on GEO: GSE247377. All original codes were released under a GPL 3.0 license and have been deposited on GitHub (https://github.com/cavallifly/Denaud_et_al_GenomeBiol_2025/tree/main) [31] and Zenodo (DOI 10.5281/zenodo.15236064) [32]. Fly lines are available on request.

## Declarations

**Ethics approval and consent to participate**
Not applicable.

**Competing interests**
The authors declare no competing interests.

Received: 12 August 2024 Accepted: 25 April 2025
Published online: 08 May 2025

**References**


1. da Costa-Nunes JA, Noordermeer D. TADs: dynamic structures to create stable regulatory functions. Curr Opin Struct Biol. 2023;81:102622.
2. Cavalheiro GR, Girardot C, Viales RR, Pollex T, Cao TBN, Lacour P, Feng S, Rabinowitz A, Furlong EEM. CTCF, BEAF-32, and CP190 are not required for the establishment of TADs in early Drosophila embryos but have locus-specific roles. Sci Adv. 2023;9:eade1085.
3. Matthews NE, White R. Chromatin architecture in the fly: living without CTCF/Cohesin Loop Extrusion?: alternating chromatin states provide a basis for domain architecture in Drosophila. BioEssays. 2019;41: e1900048.
4. Szabo Q, Jost D, Chang JM, Cattoni DI, Papadopoulos GL, Bonev B, Sexton T, Gurgo J, Jacquier C, Nollmann M, et al. TADs are 3D structural units of higher-order chromosome organization in Drosophila. Sci Adv. 2018;4: eaar8082.
5. El-Sharnouby S, Fischer B, Magbanua JP, Umans B, Flower R, Choo SW, Russell S, White R. Regions of very low H3K27me3 partition the Drosophila genome into topological domains. PLoS ONE. 2017;12: e0172725.
6. Ulianov SV, Khrameeva EE, Gavrilov AA, Flyamer IM, Kos P, Mikhaleva EA, Penin AA, Logacheva MD, Imakaev MV, Chertovich A, et al. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. Genome Res. 2016;26:70–84.
7. Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y, Corces VG. Evolutionarily conserved principles predict 3D chromatin organization. Mol Cell. 2017;67(837–852): e837.
8. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu CT, Zhuang X. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature. 2016;529:418–22.
9. Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN. Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature. 2019;568:49–54.
10. Szabo Q, Donjon A, Jerkovic I, Papadopoulos GL, Cheutin T, Bonev B, Nora E, Bruneau BG, Bantignies F, Cavalli G. Regulation of single-cell genome organization into TADs and chromatin nanodomains. Nat Genet. 2020;52:1151–7.
11. Zenk F, Loeser E, Schiavo R, Kilpert F, Bogdanovic O, Iovino N. Germ line-inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition. Science. 2017;357:212–6.
12. Ogiyama Y, Schuettengruber B, Papadopoulos GL, Chang JM, Cavalli G. Polycomb-dependent chromatin looping contributes to gene silencing during Drosophila development. Mol Cell. 2018;71(73–88): e75.
13. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell. 2010;143:212–24.
14. Denaud S, Bardou M, Papadopoulos GL, Grob S, Di Stefano M, Sabaris G, Nollmann M, Schuettengruber B, Cavalli G. A PRE loop at the dac locus acts as a topological chromatin structure that restricts and specifies enhancer-promoter communication. Nat Struct Mol Biol. 2024;31(12):1942–54.
15. Ahmad K, Spens AE. Separate polycomb response elements control chromatin state and activation of the vestigial gene. PLoS Genet. 2019;15: e1007877.
16. De S, Cheng Y, Sun MA, Gehred ND, Kassis JA. Structure and function of an ectopic Polycomb chromatin domain. Sci Adv. 2019;5:eaau9739.
17. Schuettengruber B, Cavalli G. Polycomb domain formation depends on short and long distance regulatory cues. PLoS ONE. 2013;8: e56531.
18. Eagen KP, Aiden EL, Kornberg RD. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. Proc Natl Acad Sci U S A. 2017;114:8764–9.
19. 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.
20. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, Cavalli G. Multiscale 3D genome rewiring during mouse neural development. Cell. 2017;171(557–572): e524.
21. Kraft K, Yost KE, Murphy SE, Magg A, Long Y, Corces MR, Granja JM, Wittler L, Mundlos S, Cech TR, et al. Polycomb-mediated genome architecture enables long-range spreading of H3K27 methylation. Proc Natl Acad Sci U S A. 2022;119: e2201883119.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
23. 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.
24. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics. 2015;31:2032–4.
25. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44:W160-165.
26. Olivares-Chauvet P, Mukamel Z, Lifshitz A, Schwartzman O, Elkayam NO, Lubling Y, Deikus G, Sebra RP, Tanay A. Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. Nature. 2016;540:296–300.
27. Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J, Meyer BJ. Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature. 2015;523:240–4.
28. Denaud S, Sabarís G, Di Stefano M, Papadopoulos GL, Schuettengruber B, Cavalli G. Determining the functional relationship between epigenetic and physical chromatin domains in Drosophila. Datasets. Gene Expression Omnibus. 2024. Available from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274157.
29. Denaud S, Sabarís G, Di Stefano M, Papadopoulos GL, Schuettengruber B, Cavalli G. Determining the functional relationship between epigenetic and physical chromatin domains in Drosophila. Datasets (2024) Gene Expression Omnibus. Available from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274158/.
30. Denaud S, Sabarís G, Di Stefano M, Papadopoulos GL, Schuettengruber B, Cavalli G. Determining the functional relationship between epigenetic and physical chromatin domains in Drosophila. Datasets (2024) Gene Expression Omnibus. Available from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274159.
31. Denaud S, Sabarís G, Di Stefano M, Papadopoulos GL, Schuettengruber B, Cavalli G. Determining the functional relationship between epigenetic and physical chromatin domains in Drosophila. Github; 2025. Available from https://github.com/cavallifly/Denaud_et_al_GenomeBiol_2025/tree/main.

32. Denaud S, Sabarís G, Di Stefano M, Papadopoulos GL, Schuettengruber B, Cavalli G. Determining the functional relationship between epigenetic and physical chromatin domains in Drosophila. Zenodo. 2025. https://doi.org/10.5281/zenodo.15236064. Available from https://zenodo.org/records/15236065.

## Publisher's Note