# Enhancer–promoter interactions can form independently of genomic distance and be functional across TAD boundaries

Deevitha Balasubramanian, Pedro Borges Pinto, Alexia Grasso, Séverine Vincent, Hélène Tarayre, Damien Lajoignie, Yad Ghavi-Helm

*Nucleic Acids Research*, 2024, **52**, 1702–1719
https://doi.org/10.1093/nar/gkad1183
Advance access publication date: 12 December 2023
**Gene regulation, Chromatin and Epigenetics**

# Enhancer–promoter interactions can form independently of genomic distance and be functional across TAD boundaries

**Deevitha Balasubramanian[1,2,†], Pedro Borges Pinto[1,†], Alexia Grasso[1,†], Séverine Vincent[1], Hélène Tarayre[1], Damien Lajoignie[1] and Yad Ghavi-Helm[1,*]**

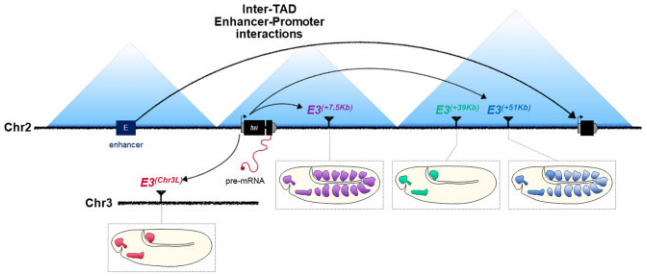
[1]Institut de Génomique Fonctionnelle de Lyon, Ecole Normale Supérieure de Lyon, Centre National de la Recherche Scientifique UMR5242, Université Claude Bernard-Lyon 1; 69364 Lyon, France
[2]Indian Institute of Science Education and Research (IISER) Tirupati; Tirupati 517507 Andhra Pradesh, India

[*]To whom correspondence should be addressed. Tel: +33 426731350; Email: yad.ghavi-helm@ens-lyon.fr
[†]The first three authors should be regarded as Joint First Authors, listed in alphabetic order.

## Abstract

Topologically Associating Domains (TADs) have been suggested to facilitate and constrain enhancer–promoter interactions. However, the role of TAD boundaries in effectively restricting these interactions remains unclear. Here, we show that a significant proportion of enhancer–promoter interactions are established across TAD boundaries in *Drosophila* embryos, but that developmental genes are strikingly enriched in intra- but not inter-TAD interactions. We pursued this observation using the *twist* locus, a master regulator of mesoderm development, and systematically relocated one of its enhancers to various genomic locations. While this developmental gene can establish inter-TAD interactions with its enhancer, the functionality of these interactions remains limited, highlighting the existence of topological constraints. Furthermore, contrary to intra-TAD interactions, the formation of inter-TAD enhancer–promoter interactions is not solely driven by genomic distance, with distal interactions sometimes favored over proximal ones. These observations suggest that other general mechanisms must exist to establish and maintain specific enhancer–promoter interactions across large distances.

## Graphical abstract




## Introduction

Enhancers are short non-coding genomic elements that play a crucial role in regulating precise spatial and temporal gene expression patterns (1). They do so by acting as platforms for transcription factors and activating the transcription of their target gene(s) through the formation of an enhancer–promoter (E–P) chromatin loop. Eukaryotic genomes typically contain over a hundred thousand enhancers interspersed between genes, with enhancers sometimes skipping multiple nearby promoters to regulate the expression of a gene located at a large genomic distance (2,3). It is thus essential for enhancers to specifically target and regulate the expression of the correct gene while avoiding the inappropriate expression of neighboring genes. How E–P interaction specificity is established within such a complex 3D genome is a long-standing question. Several factors, such as the sequence of the core promoter and the presence of insulator or tethering elements have been suggested to contribute to this process (4).

In recent years, there has been growing evidence that genome topology also plays a crucial role in guiding E–P interaction specificity. Indeed, regulatory domains tend to coincide with regions of increased three-dimensional proximity named Topologically Associating Domains (TADs) (5–9). Because E–P interactions are enriched within TADs, it is generally admitted that TAD boundaries are involved in insulating these domains. As a consequence, deletion or rearrangements of TAD boundaries result in the mis-regulation of gene expression due to either loss or hijacking of regulatory elements (10–17). Overall, this suggests that TADs act as functional regulatory units that favor the formation of specific E–P interactions while restricting interactions with regulatory elements located in adjacent TADs (18,19).

**Received:** April 25, 2023. **Revised:** November 21, 2023. **Editorial Decision:** November 24, 2023. **Accepted:** November 29, 2023

However, conflicting observations question the ability of TAD boundaries to properly insulate E–P interactions and their role in gene expression regulation (20). Notably, in some cases, gene expression is resilient to chromosomal rearrangements affecting TAD boundaries (21–24). Moreover, depleting the complexes responsible for TAD boundary formation completely abolishes TAD structures, yet only mildly affects gene expression (25–28). A possible explanation for this apparent discrepancy is that E–P interactions can, to some extent, be established across TAD boundaries, as a result of yet uncharacterized mechanisms. Several observations support this possibility. Up to a third of all long-range promoter interactions identified by Promoter Capture Hi-C (PCHi-C) in mice and humans are established across TAD boundaries (29,30). Although the nature of these interactions and the elements involved have not been studied in detail, promoter-interacting regions were found to be enriched in histone marks associated with active enhancers, active promoters, and transcribed regions, highly suggestive of non-coding regulatory elements (29). Moreover, these interactions are rewired during development, another indicator of their dynamic regulatory nature. Similarly, a Micro-C analysis performed in mouse ES cells reported that over 20% of E–P contacts are established across TAD boundaries (31).

While these results suggest that inter-TAD E–P interactions are not a rare event, the functionality of these interactions (i.e. their ability to drive gene expression) has not yet been established. Indeed, in many cases, E–P interactions are observed in tissues or during stages when the corresponding gene is not transcribed (32–34). Moreover, the existence of long-range inter-TAD E–P interactions contrasts with recent studies quantifying the effect of E–P promoter distance on gene expression which show that increasing the E–P genomic distance within a TAD decreases transcriptional activity (35–38). Overall, these studies clearly show the existence, within the boundaries of a single TAD, of a correlation between transcriptional activity and distance, with the transcriptional output decreasing with increasing E–P distance and dropping to zero when crossing a TAD boundary (38). These observations, therefore, raise important questions about the contribution of inter-TAD E–P interactions to gene expression.

To gain greater insight into the importance of these long-range E–P interactions, we performed a detailed analysis of inter-TAD E–P interactions in *Drosophila* embryos. Micro-C data revealed that a substantial percentage of E–P interactions are established across TAD boundaries. Strikingly, the genes overlapping intra-TAD loop anchors are enriched for developmental functions, while inter-TAD genes failed to show such enrichment. This led us to question the ability of developmental genes to establish functional inter-TAD interactions with their enhancers. To investigate this, we used as a model the *twist* locus in *Drosophila* embryos, a master regulator of mesoderm development, and placed one of its enhancers (hereafter called *E3*) at various ectopic locations. Using circular chromatin conformation capture coupled with next-generation sequencing (4C-seq), Micro-C and DNA-Fluorescent *In Situ* Hybridization (FISH) we determined the ability of the *E3* enhancer to interact with the *twist* promoter when placed at large genomic distances across the boundary of a TAD. The functional outcome of these interactions was determined by analyzing the activity of *twist* and the development of the embryonic somatic musculature, a mesoderm-derived tissue, in the absence of the endogenous *E3* enhancer. Overall, our results revealed that interactions between the *twist* promoter and *E3* can be established across TAD boundaries and even across chromosomes, and that these interactions can be functional. These long-range inter-TAD interactions are established in a distance-independent but context-dependent manner, with distal locations sometimes favored over proximal ones.

## Materials and methods

### Analysis of genome-wide inter-TAD enhancer–promoter loops

We analyzed Micro-C data obtained from nuclear cycle 14 *Drosophila melanogaster* embryos based on two recently published datasets (34,39). The data were processed as described below (section 'Micro-C in *Drosophila* embryos') to generate contact matrices. Given the high concordance between biological replicates within each dataset, we merged them to create high-coverage contact matrices. Significant Interaction Peak caller (SIP, v1.6.2) (40) was used to identify interactions from the contact matrices at various resolutions (500 bp, 1 kb, 2.5 kb, 5 kb, 10 kb, 15 kb and 20 kb), with the following parameters: norm = KR, Gaussian filter = 3, minimum filter = 2, maximum filter = 2, matrix size = 2000, diagonal size = 5, resolution corresponding to the matrix resolution used in bp, % of saturated pixels = 0.01, threshold = 2,000, number of zeros = 6, false discovery rate = 0.05, is Drosophila = false. For the contact matrices from Batut *et al*. (39) of 15 kb and 20 kb resolution, we set Gaussian filter = 1.5 and threshold = 1,500 to enhance the detection of loops in those matrices.

Next, we applied several filters to the identified loops. Loops located in the centromeric regions of the *Drosophila* genome (corresponding to the following dm6 coordinates >22,170,000 for chr2L, <5,650,000 for chr2R, >22,900,000 for chr3L, <4,200,000 for chr3R) were removed due to low read coverage and mapping issues in those regions. The cooltools (version 0.5.4, (41)) coverage function was then used to calculate the coverage of Micro-C data at 2kb resolution. Loops falling within 5 kb of regions with coverage <1000 for the Ing-Simmons *et al*. (34) dataset and <3,500 for the Batut *et al*. (39) dataset were also excluded. After these filtering steps, data from all resolutions were combined, and duplicated loops were removed, to retain only the one from the highest resolution for further analysis.

TAD boundaries were identified from all Micro-C data at a resolution of 20 kb using the hicFindTADs tool of HiCExplorer v2.2.1.1 (42) with default parameters (except delta = 0.05) to avoid considering nested or sub-TADs. Inter-TAD chromatin loops were then identified by examining how many TAD boundaries were crossed by the filtered chromatin interactions. The insulation score was computed by the hicFindTADs tool and used to assess the insulation strength of the TAD boundaries that are (un)crossed by loops. We used publicly available ChIP-Seq data of insulators proteins known to bind to *Drosophila* TAD boundaries and call peaks with MACS2 (version 2.2.6), with a cut-off determined by the –cutoff-analysis argument and visual inspection of the data. The (un)crossed TAD boundaries were then analyzed to check if they fell within 300 bp of an insulator peak.

To identify enhancer–promoter interactions, we obtained enhancer data from the CRM Activity Database, a manually annotated list of *cis*-regulatory modules active during

embryonic stages (43). Promoters were defined as regions –300 bp to +50 bp of the transcription start site (TSS) of a gene. We overlapped the anchors of each chromatin loop with these databases to check if they fell within 300 bp of an enhancer or a promoter, and referred to this set as 'all-possible-E–P' interactions. To obtain a list containing exclusively enhancer–promoter (E–P) interactions (referred to as 'only-E–P'), we excluded from the 'all-possible-E–P' set all chromatin loop anchors that could overlap multiple promoters and enhancers (as these loops could potentially be cases of promoter–promoter and enhancer–enhancer interactions).

The list of enhancers and promoters involved in interactions was then compared to documented expression patterns from FlyBase FB2022_06 (44), modENCODE mRNA-Sequencing (45), Berkeley Drosophila Genome Project (46,47), REDFly (48) and characterization of the Vienna Tile library (49). The 'all-possible-E–P' interactions were categorized into two groups: 'Data on E and P expression,' consisting of interactions involving genes with known expression patterns during embryonic stages, and 'No data,' containing interactions involving genes with no available data on their expression or those that are inactive at embryonic stages. We further analyzed the interacting genes and enhancers in the 'Data on E and P expression' group for temporal and spatial overlap in their expression.

Finally, we conducted a Gene Ontology (GO) enrichment analysis using HOMER v4.11 (50) on the promoters common to the two datasets (34,39) at inter-TAD and intra-TAD interactions, using default parameters. The expression levels of all promoters involved in inter-TAD and intra-TAD interactions were evaluated using RPKM values at 0–24 h after egg-lay (sum of RPKM at 0–2 h, 2–4 h and so on) from modENCODE RNA-Sequencing (45).

## Plasmid construction and transgenic fly generation

All plasmids were constructed using standard cloning methods with New England Biolabs restriction enzymes and T4 DNA ligase (New England Biolabs) or with the NEBuilder HiFi DNA Assembly kit (New England Biolabs). All constructs were verified by sequencing.

Unless specified otherwise, 'wild-type' fly lines used in this study refer to the *yw y[1] w[1118]* line (BDSC_6598). All fly lines were raised on standard food at 25°C.

To create ectopic E3 insertion lines, we used two strategies: First, we took advantage of the popular MiMIC (*Minos* Mediated Integration Cassette) system (51), which consists of a *Minos* transposon carrying a *yellow*$^{+}$ dominant body-color marker and a gene-trap cassette flanked by two inverted ΦC31 integrase *attP* sites. This cassette can be efficiently replaced by another cassette containing the DNA sequence of interest flanked by two inverted ΦC31 integrase *attB* sites using RMCE. This insertion event can be conveniently identified by the loss of body pigmentation in adult flies (corresponding to the replacement of the *yellow*$^{+}$ marker by the sequence of interest). We used this strategy to generate five different fly lines where the 1089 bp-long *E3* enhancer was inserted at different locations. While such MiMIC fly lines are readily available to create insertions at thousands of sites, we had to use a second strategy to specifically insert the *E3* enhancer in the same TAD as *twist* (line *E3(+7.5 kb)*), a region where no MiMIC fly line is available. In this second strategy, we first used CRISPR-Cas9 mediated HDR to create a fly line where a ΦC31 integrase *attP* site is integrated at the desired location. This site was then used to insert the E3 sequence from a donor vector containing a ΦC31 integrase *attB* site.

**For the ectopic integration of enhancer *E3* +7.5 kb away from the *twist* promoter (to obtain line *E3 (+7.5 kb)*)**

The *pHD-dsRed-attP* vector (Addgene #51019 (52)) was used to introduce an *attP* docking site at position +7.5 kb. Homology arms (∼1 kb each) surrounding the insertion site were amplified from genomic DNA of the *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* (BDSC_51324 (53)) fly line. The gRNA was designed using the flyCRIPSR target finder (52) (sequence of the gRNA: GTCGAATGTCGGGCATATCTT) and cloned in the *pU6-BbsI-chiRNA* vector (Addgene #45946 (54)) following the flyCRISPR recommendations (https://flycrispr.org/). The vectors were co-injected in-house in embryos of the *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* fly line. The resulting transgenic line was first crossed to a Cre recombinase-expressing line (BDSC_766) to delete the dsRed marker cassette. The E3 sequence was then inserted at position +7.5 kb (*line E3(+7.5 kb)*) by co-injecting the p3xP3-EGFP.vas-int.NLS vector (Addgene #60948 (55)) and the *pattB* vector (DGRC #1420 (56)) containing the E3 sequence (2R:23049440–23050529) amplified from genomic DNA of the *yw y[1] w[1118]* fly line and cloned using the KpnI and XhoI restriction sites.

**For the ectopic integration of enhancer *E3* –1.6 Mb, –181 kb, +39 kb, and +51 kb away from the *twist* promoter and on chromosome 3L (to obtain line *E3(–1.6 Mb)*, *E3(–181 kb)*, *E3(+39 kb)*, *E3(+51 kb)* and *E3(chr3L)*)**

The E3 sequence (2R:23049440–23050529) was amplified from genomic DNA of the *y[1] w[1118]* fly line, except for the *E3(+51kb)* line where the E3 sequence was amplified from the *y[1] w[*]; Mi{y[+mDint2]=MIC}MI01218* (BDSC_55415 (51)) fly line. The PCR product was cloned into the *pBS-KS-attB1-2-PT-SA-SD-0–2xTY1-V5* vector (Addgene #61255 (55)) using HindIII and XbaI restriction sites. The resulting vector and p3xP3-EGFP.vas-int.NLS were injected in-house through ΦC31-mediated recombination(57) in embryos of the following 'MiMIC' fly lines (51):

- *y[1] w[*]; Mi{y[+mDint2]=MIC}MI04814* (BDSC_38170) to obtain line *E3(–1.6 Mb)*
- *y[1] w[*]; Mi{y[+mDint2]=MIC}MI01218* (BDSC_55415) to obtain line *E3(–181 kb)*
- *y[1] w[*]; Mi{y[+mDint2]=MIC}MI11229* (BDSC_55595) to obtain line *E3(+39 kb)*
- *y[1] w[*]; Mi{y[+mDint2]=MIC}MI02100* (BDSC_32829) to obtain line *E3(+51 kb)*
- *y[1] w[*]; Mi{y[+mDint2]=MIC}MI10934* (BDSC_55560) to obtain line *E3(chr3L)*

**For the deletion of the endogenous enhancer E3 in the *w[1118]; PBac{y[±mDint2]=vas-Cas9}VK00027*, *E3(–1.6 Mb)*, *E3(+7.5 kb)*, *E3(+39 kb)* and *E3(+51 kb)* fly lines (to obtain line $twi^{\Delta E3}$;$twi^{\Delta E3}$, *E3(–1.6 Mb)*; $twi^{\Delta E3}$, *E3(+7.5 kb)* and $twi^{\Delta E3}$, *E3(+51 kb)*)**

Two ΦC31 integrase *attP* landing sites were inserted into the *pHD-dsRed* vector (DGRC #1360 (52)) using either a BsiWI

restriction site or an AgeI and a SpeI restriction site. The resulting vector was used to delete the endogenous E3 sequence (2R:23049262–23050567). Homology arms (∼1 kb each) surrounding the deletion site were amplified from genomic DNA of the *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* fly line. gRNAs were designed using the flyCRIPSR target finder (52) and cloned in the *pU6-BbsI-chiRNA* vector (Addgene #45946 (54)) following the flyCRISPR recommendations (https://flycrispr.org/). Sequence of the gRNAs:

- GAAATCAAAGACTTGTATAC and GGGG AAAAATATCTTTGCAG for line *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* and *E3(+7.5 kb)*
- GAAATCAAAGACTTGTATGC and GGGGAAAAAT ATCTTTGAAG for line *E3(–1.6 Mb)* and *E3(+39 kb)*
- GAAATCAAAGACTTGTATAC and GGGGGGAAAT ATCTTTGAAG for line *E3(+51 kb)*

The vectors were co-injected in-house (except for $twi^{\Delta E3}$, *E3(+7.5 kb)* which was generated by the FlyORF Injection Service) in embryos of the following fly lines:

- *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* (to obtain line $twi^{\Delta E3}$)
- *E3(–1.6 Mb)* to obtain line $twi^{\Delta E3}$, *E3(–1.6 Mb)*
- *E3(+7.5 kb)* to obtain line $twi^{\Delta E3}$, *E3(+7.5 kb)*
- *E3(+39 kb)* to obtain line $twi^{\Delta E3}$, *E3(+39 kb)*
- *E3(+51 kb)* to obtain line $twi^{\Delta E3}$, *E3(+51kb)*

Note: The *E3(–181 kb)* line lays very few embryos which made it impossible to obtain the $twi^{\Delta E3}$, *E3(–181 kb)* line. The *E3(–1.6 Mb)*, *E3(+7.5 kb)*, *E3(+39 kb)* and *E3(+51 kb)* fly lines were first crossed with the *w[1118]; PBac{y[+mDint2]=vas-Cas9}VK00027* line to express Cas9 in the germline.

**For the deletion of the endogenous enhancer *E3* in the *E3(chr3L)* fly line (to obtain line $twi^{\Delta E3}$, *E3(chr3L)*)**

The deletion of the endogenous E3 sequence in the line where *E3* was ectopically inserted on chromosome 3L was obtained by crossing the *E3(chr3L)* fly line with the $twi^{\Delta E3}$ fly line.

The final coordinates (dm6) of all insertion sites are as follows:

- line *E3(–1.6 Mb)*: 2R:21,381,841
- line *E3(–181 kb)*: 2R:22,865,023
- line *E3(+7.5 kb)*: 2R:23,053,901
- line *E3(+39 kb)*: 2R:23,084,945
- line *E3(+51 kb)*: 2R:23,097,536
- line *E3(chr3L)*: 3L:6,820,484

**For transgenic reporter assays**

To assess the enhancer activity of the *E1*, *E2* and *E3* enhancers, these enhancers were cloned upstream of the minimal *twist* promoter (chr2R:23046216–23046481) driving a mGFPmut2 reporter gene (58) (codon-optimized for *Drosophila*) in the pBID vector backbone (Addgene #35190). The coordinates of the cloned regions are as follows: enhancer *E1*: chr2R: 23,044,478–23,045,403, enhancer *E2*: chr2R:23,045,827–23,046,215, enhancer *E3*: chr2R: 23,049,440–23,050,529. All constructs were injected in-house through ΦC31-mediated recombination (57) into the *nos-φC31\int.NLS; attP40* line (59). Stably integrated transgenic lines were balanced, and homozygous lines were used for immunostaining and smiFISH to examine *GFP* expression.

## Embryo collections

Freshly hatched adults of the appropriate genotype were placed in embryo collection vials with standard apple cap plates. *Drosophila* embryos were collected on apple juice agar plates at 25°C at the appropriate time-point (after 3 pre-lays of 1 h for stage-specific collections), dechorionated using 50% bleach, and washed alternately with water and PBS + 0.1% Triton X-100. The embryos used for RT-qPCR were directly transferred to RA1 buffer supplemented with 2-Mercaptoethanol provided by the NucleoSpin RNA kit (for RNA purification, Macherey-Nagel) and stored at –80°C. The embryos used for 4C-Seq were covalently crosslinked in 1.8% formaldehyde for 15 min at room temperature and stored at -80°C. The embryos used for Micro-C were covalently crosslinked in 1.8% formaldehyde for 15 min, quenched for 5 min with 2M Tris–HCl pH7.5, then crosslinked again with 3 mM DSG for 45min and stored at –80°C. The embryos used for 3D DNA FISH were covalently crosslinked in 4% formaldehyde for 25 min and stored at –20°C in methanol. The embryos used for immunostaining were covalently crosslinked in 6% formaldehyde for 30 min and stored at –20°C in methanol. The embryos used for smiFISH were covalently crosslinked in 8% formaldehyde for 45 min and stored at –20°C in methanol.

## 4C-seq in *Drosophila* embryos

**Experimental protocol**

Nuclear extraction was carried out as described previously (33). About 100–1000 embryos were used for each 4C template preparation using MboI and NlaIII (New England Biolabs) as the first and second restriction enzymes, respectively. 4C templates were amplified from 320 ng of 4C template using the following primers:

- Twi1_FW: TACGTGCACCAAAAGTTTCTT Twi1_RV: AAAATGGTCGTCAAAGCGC, corresponding to a viewpoint located upstream for the *twist* promoter (chr2R:23,044,043–23,044,500; referred to as viewpoint Twi1)
- Twi2_FW: GGCAACAATCCGAGTGGC Twi2_RV: GTACTCCGAGGGCAGTGG corresponding to a viewpoint located within the *twist* gene (chr2R:23,046,575–23,046,906; referred to as viewpoint Twi2)

An additional 1–8 nucleotides 'shift' sequence was added at the beginning of the primers, to allow optimal base-pair diversity at the beginning of the read after multiplexing.

The PCR product was purified using SPRIselect beads (Beckman Coulter) and 100 ng of each PCR product was used to generate the final libraries using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). The libraries were indexed for multiplexing using NEBNext multiplex oligos kit for Illumina (New England Biolabs). A total of 54 libraries were generated, with two independent biological replicates for each sample. The libraries were multiplexed and sequenced on a NextSeq500 sequencer (Illumina) using 75-bp paired-end reads (at the IGFL sequencing facility), yielding a total of at least 10 million reads per sample.

### Data analysis

The quality of the 4C-seq data was confirmed using the FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapter sequences were trimmed using TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), and the 5′ shift sequence and the primer sequence were trimmed up to the location of the first restriction site (the restriction enzyme cutting site was kept) using Cutadapt version 2.10 (60). As sequencing was performed in paired-end mode, the fastq files corresponding to each pair were merged into a single fastq file. The trimmed reads were then aligned either to the dm6 reference genome or to a custom genome, generated using the reform Python tool (https://github.com/gencorefacility/reform). The percentage of reads mapping to the dm6 genome is provided in Supplementary Table S1. Custom genomes consist of the *Drosophila melanogaster* reference genome (dm6) where the endogenous *E3* enhancer sequence has been deleted and reintroduced at the appropriate location, with the appropriate sequence (Supplementary Figure S1). Six different custom genomes were thus created for each insertion site. Alignment was performed using Bowtie version 1.2.2 (61). As the read length is relatively long, it is possible to obtain reads that result from multiple ligation events between different regions of the genomes. As such reads will not map efficiently to the genome, unmapped reads were retrieved and scanned from the 5′ side for the presence of the first, then of the second restriction site, and trimmed after this location. As a consequence, the reads typically start with the sequence of the first restriction enzyme and have either the sequence of the first or the second restriction site at their 3′ end. These trimmed reads were remapped as previously described and both alignments were merged. The libraries were then normalized by scaling using the read-per-million method and transformed into coverage bedgraph files. 4C-seq data were plotted using pyGenomeTracks version 3.8 (62,63). Reads mapping to the ectopic version of *E3* were differentiated from those mapping to the endogenous *E3* using single nucleotide variants (highlighted by a red asterisk in Supplementary Figure S1). This was used to compute the percentage of reads mapping to the ectopic version of *E3*, out of all reads mapping to some version of *E3* ('ectopic *E3*'). For visualization purposes, the signal displayed over the ectopic E3 regions in 4C interaction maps corresponds to a downsampling of the overall E3 reads to fit the percentage of ectopic reads. The 'strength' of the ectopic *E3* signal was estimated by calculating the ratio of reads overlapping the ∼2 kb ectopic E3 region over the total number of reads in a 10 kb region around the insertion site ('*E3* strength'). Similarly, we computed the 'background' ratio for reads of a central ∼2 kb region over a surrounding 10 kb region, by taking the average ratio over six 10 kb sliding windows to cover a 12kb region upstream and downstream of the ectopic E3 site. Each biological replicate was analyzed independently, and the statistics were averaged across the replicates.

## Micro-C in *Drosophila* embryos

### Experimental protocol

Micro-C libraries were generated based on a previously established protocol (64), with appropriate modifications for *Drosophila* embryos.

Nuclear extraction was carried out as described previously (33) using cold MB#1 buffer (50 mM NaCl, 10 mM Tris–HCl pH 7.5, 5 mM $MgCl_2$, 1 mM $CaCl_2$, freshly added 0.2% NP-40 and 1× PIC) to resuspend the embryos and Dounce homogenization. Nuclei from each replicate were first used for a test Micrococcal Nuclease (MNase, Worthington Biochemical) titration: Nuclei corresponding to 600 ng of chromatin were digested with 45 U of MNase, and then checked for a yield of 300 ng of chromatin with a 90% mononucleosome/10% dinucleosome ratio (64). If required, the original 600 ng of chromatin was adjusted to obtain an appropriate yield.

Following these tests, for each replicate, seven parallel reactions of MNase were set up with the appropriate amount of chromatin and taken forward for end-chewing, end-labeling, and proximity ligation (64). After reverse cross-linking, 150–200 ng of the obtained chromatin was size-selected for ligated mononucleosomes (250–400 bp) on a 3.5% NuSieve agarose gel. The size-selected chromatin was then pulled down with 25 µl of streptavidin beads, and taken forward for library preparation using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). A total of four libraries were generated, with two independent biological replicates for each sample. The libraries were sequenced on a NovaSeq sequencer (Illumina) using 150 bp paired-end reads, yielding at least 100 million reads per sample.

### Data analysis

The quality of the Micro-C data was confirmed using the FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Adaptor sequences were trimmed using TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The paired-end files were then aligned to the dm6 reference genome or a custom genome using the Burrows-Wheeler Alignment (BWA) – maximal exact matches (MEM) tool v7.17–4 (65) (http://bio-bwa.sourceforge.net/bwa.shtml#12). The pairtools v0.3.0 (https://github.com/open2c/pairtools) pipeline was used to detect ligation junctions and quality control the paired sequences: pairtools parse was used to detect the ligation events, pairtools sort was used to block sort the reads, pairtools dedup was used to remove pairs that are PCR duplicates of each other. Pairtools split was then used to generate a Pairs file. Genome-wide contact matrices were generated both using pairix (https://github.com/4dn-dcic/pairix) and cooler (66). HiCExplorer v2.2.1.1 (42) was used to detect and remove genomic regions with low signal or with high noise, and to implicitly address biases in the data by normalizing the matrices using the Knight-Ruiz balancing algorithm. The resultant matrices were merged and used to detect TAD boundaries and compute the insulation score using HiCExplorer. The insulation score and the directionality index were also calculated using FAN-C 0.9.23 (67). A/B compartments were detected using the eigenvector command in Juicer (68).

## Two-colour 3D DNA FISH (fluorescent *in situ* hybridization)

3D DNA FISH was performed as previously described (69). Five probe sets were designed, mapping to regions of genomic DNA directly adjacent to the *twist* promoter and the different insertion sites:

- chr2R: 23,040,112–23,051,659 for the *twist* promoter
- chr2R: 21,376,545–21,386,746 for the –1.6 Mb insertion site

- chr2R: 22,860,007–22,870,484 for the –181 kb insertion site
- chr2R: 23,093,753–23,101,622 for the +51 kb insertion site
- chr3L: 6,816,164–6,824,775 for the chr3L insertion site

Each probe set was composed of six 1.2–1.5 kb-long PCR products, which were labeled using the FISH Tag DNA Multicolor kit (Alexa Fluor 488 dye for the *twist* promoter and Alexa Fluor 555 dye for the ectopic insertion sites) (Life Technologies). Mesodermal cells were stained using an anti-Twist antibody (anti-Rabbit polyclonal antibody generated by the Ghavi-Helm lab with assistance from the Protein Sciences Facility of the Lyon SFR Biosciences). Embryos were mounted in ProLong Gold antifade reagent with DAPI (Life Technologies) and imaged on a Leica SP8 confocal microscope using a 40× glycerol objective. For each embryo, several Z-stacks were acquired (section thickness of 0.361 μm) and processed using the Lightning Deconvolution software (Leica). A minimum of 350 nuclei from 3 to 4 independent embryos were analyzed and the relative distances between FISH signals were measured using the Imaris software (Bitplane). A non-parametric two-sample Kolmogorov–Smirnov test was used to verify if the distance distributions were significantly different between samples. Two probes were considered co-localized when the distance between the centers of FISH signal was <0.25 μm.

### Immunostaining

Immunostaining was performed as previously described (70). The following primary antibodies were used: rabbit anti-Twist (1:200, generated by the Ghavi-Helm lab with assistance from the Protein Sciences Facility of the Lyon SFR Biosciences) and rat anti-TM1 (1:200, Developmental Studies Hybridoma Bank). Secondary antibodies were conjugated with Alexa 488 and Alexa 555 (Invitrogen). Confocal images were acquired using a Leica SP8 confocal microscope and processed using the Leica Application Suite X (LAS X) 3D Visualization and Adobe Photoshop CS6 software.

### smiFISH

smiFISH was performed as previously described (71). Briefly, fixed embryos were incubated overnight at 37°C in hybridization buffer containing 320 nM of smiFISH probes. Embryos were washed and immunostained with the appropriate antibody. Probes against *GFP* and *twist* were designed as previously described (71). The X FLAP sequence was 5′ and 3′ end-labeled with Quasar 570. The embryos were mounted in ProLong Gold antifade reagent with DAPI (Life Technologies) and imaged on a Leica SP8 confocal microscope using a 20x objective. The images were processed using the Adobe Photoshop CS6 software.

### Viability tests

Embryo viability was established by scoring the hatching rate of embryos into first instar larvae. Freshly hatched adult flies of the appropriate genotype were placed in embryo collection vials with standard apple cap plates and acclimatized at 25°C for 2 days prior to the experiments. After 3 pre-lays of 1 hour, eggs were collected for 2 h at 25°C and rinsed with water. For each genotype, at least 50 embryos were transferred on a fresh plate and incubated at 25°C. The number of hatched larvae and unhatched embryos was counted 24 h later. Two independent experiments were performed for each genotype. The percentage of viability was calculated by dividing the number of hatched larvae by the total number of embryos.

### Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

RNA extraction was performed by grinding the embryos with a pestle, followed by RNA purification using the Nucleospin RNA kit (for RNA purification, Macherey-Nagel). Reverse transcription of the RNA was performed using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) with random primers. qPCR was performed for three independent biological replicates using the following primers:

- *RpL32*: ATGCTAAGCTGTCGCACAAATG and GTTCGATCCGTAACCGATGT
- *twi*: CCTCAAACTGGCCACAAGATAC and GCCACACCCCGAACAGATAA
- *l(2)k09913*: AAAGGATCCCAGGTGGTGG and GACATCTGCAGCATGATCTCG
- *Fib*: CCAGGTAGAGAACCTTCGAGC and CTCCGTTGAGACCAATGGC
- CG9877: CATCTTCTCGACTTTCTATCCG and TGGCTACGGTGGCTTTAG
- CG9876: GTGGGTGTGGCATATTGG and CCAGAGTTCAGGTCTGGTTTC

$\Delta C_t$ values were computed using $C_t$ (gene of interest) – $C_t$ (*RpL32*). $\Delta\Delta C_t$ were computed using $\Delta C_t$ (*E3(ectopic)*) – $\Delta C_t$ (corresponding MiMIC control). The $2^{-\Delta\Delta Ct}$ values were used as the fold change (FC) values. A two-tailed paired *t*-test was performed to assess the significance of $\Delta C_t$ for each gene in the *E3(ectopic)* line compared to the corresponding MiMIC control.

## Results

### A substantial fraction of enhancer–promoter interactions are established across TAD boundaries

Chromosome conformation capture and high-resolution imaging studies have clearly demonstrated that while chromatin interactions are enriched within TADs, an important fraction of these interactions also occur across TAD boundaries (29,30,72,73). Moreover, recent analyses of high-resolution contact maps obtained in mammalian cells in combination with chromatin state analysis suggest that inter-TAD interactions can be established between putative enhancers and promoters (29,31).

To better characterize the significance of such inter-TAD E–P interactions in gene expression, we extended this analysis to the *Drosophila* genome. We reanalyzed two recently published Micro-C datasets from early *Drosophila* embryos (nuclear cycle 14) (34,39) and overlapped the identified chromatin interactions with the locations of promoters and validated enhancers (Supplementary Table S2). Consistent with previous reports, we found that an average of 17% of the identified interactions occur across TAD boundaries. Among these inter-TAD interactions, 33% are established between a promoter and a validated enhancer. However, as the anchors of chromatin loops vary in size (from 500 bp to 20 kb), each anchor might encompass multiple enhancer and promoter sequences. As a consequence, some of the identified E–P interactions could in fact represent promoter–promoter or

enhancer–enhancer interactions. Having excluded promoter–promoter and enhancer–enhancer loops, our analysis reveals that 17% of inter-TAD interactions occur exclusively between a promoter and a validated enhancer (referred to as 'only E–P,' Figure 1A). This percentage is likely an underestimate as the activity of a vast majority of enhancers remains unknown. A significant proportion of the inter-TAD interactions (61%) and inter-TAD enhancer–promoter interactions (64%) extend across more than one TAD boundary (Figure 1B, Supplementary Figure S2A). Moreover, the median distance for inter-TAD interactions is 720kb (with a maximum of 6.2 Mb) and 830 kb for inter-TAD E–P interactions (with a maximum of 3.9Mb), indicating that these interactions are spanning longer ranges compared to characterized regulatory units within a TAD (Supplementary Figure S2B).

To better characterize these inter-TAD interactions, we first ruled out the possibility that they could be caused by the presence of a weaker boundary. For this purpose, we computed the insulation score from both Micro-C datasets and used it as a proxy for the strength of a TAD boundary. We did not observe any significant difference in the insulation score between the boundaries that are crossed by an interaction and those that are not (Figure 1C). In *Drosophila*, TAD boundaries are associated with a combination of various insulator proteins, including BEAF-32, CP190, CTCF, mod(mdg4) and su(Hw) (74). We characterized the binding of these proteins to boundaries that are crossed by an interaction and to those that are not. We observed no significant difference in binding, and in some cases, even noticed a significantly higher binding at the crossed boundaries (Supplementary Figure S2C). Taken together, these results indicate that the ability for an interaction to cross a TAD boundary is not due to a lack of insulation or the absence of insulator proteins at the boundary.

We next asked whether the genes involved in inter-TAD interactions were functionally different from those involved in intra-TAD interactions. GO term enrichment analysis revealed that intra-TAD interactions are significantly enriched in genes involved in developmental processes and transcription regulation (Figure 1D and Supplementary Table S3). A large proportion of these genes are expressed during embryogenesis (58% have an RPKM value higher than 5; Supplementary Figure S2D). In contrast, genes involved in inter-TAD interactions tend to have a general signaling or metabolic function (Figure 1D and Supplementary Table S3) and are less likely to be expressed in the embryo (36% have an RPKM value higher than 5; Supplementary Figure S2D). To establish the relevance of these inter-TAD E–P interactions for gene expression, we manually examined the spatio-temporal activity (from embryonic *in situ* hybridization and RNA sequencing experiments) of all enhancers and genes overlapping an inter-TAD loop anchor. Unfortunately, embryonic spatio-temporal information about the activity of the gene was available for only 29% of all inter-TAD E–P pairs, as a large proportion of these genes are not expressed in the embryos. Nevertheless, out of these informative E–P pairs, 58% showed a spatio-temporal overlap in their expression pattern (while 37% showed no overlap and 5% were dubious) (Supplementary Table S4).

In summary, we observe that, as in mammals, a significant fraction of *Drosophila* E–P interactions are established across TAD boundaries. Strikingly, there is a high overlap in the spatio-temporal activity of these inter-TAD E–P pairs suggesting that inter-TAD E–P interactions may have a functional role. However, the genes involved in intra- versus inter-TAD interactions are involved in different processes, with a marked enrichment of intra-TAD interactions for developmental genes and transcription factors.

## The *E3* enhancer of *twist* is required during embryonic muscle development

The observation that developmental genes and transcription factors are markedly enriched only in the set of intra-TAD interactions questions their ability to engage in functional long-range inter-TAD interactions. To establish whether developmental genes could function across TAD boundaries, we genetically engineered various fly lines where a given developmental enhancer was placed at increasingly large distances from its endogenous promoter, including relocation to other TADs. We then determined the conditions under which functional E–P interactions were established. To efficiently analyze the effect of distance on gene expression, we searched for a well-characterized developmental gene whose activity is easily tractable during embryogenesis and whose expression is regulated by a tissue-specific enhancer. We, therefore, focused on the *Drosophilatwist* (*twi*) gene, which codes for a highly conserved transcription factor acting as a master regulator of mesoderm development and promoting epithelial-mesenchymal transition in normal and metastatic cells (75). *twist* is strongly expressed from the onset of zygotic transcription in the ventral region of the embryo corresponding to the mesoderm anlage and starts to decline after germ band elongation (76). *twist* mutants are recessive lethal due to abnormal gastrulation characterized by the absence of mesoderm derivatives (77). During early embryogenesis, the expression of the *twist* gene is regulated by three known enhancers: an upstream distal enhancer (*DE*), an upstream proximal enhancer (*PE*) (78), and a downstream distal enhancer (79). For simplicity, we will hereafter refer to these enhancers as *E1*, *E2* and *E3*, respectively (Figure 2A). Previous reports suggested that these regulatory regions might be active in overlapping cell types during the early stages of embryogenesis (49,80,81). To characterize the activity of these regulatory elements in detail, we generated reporter constructs for each of these elements and analyzed reporter gene expression in transgenic flies. Our analysis revealed that all three regulatory elements are active until stage 10. However, by stage 11, *E3* is the only active enhancer controlling *twist* expression in the thoracic and abdominal regions (Figure 2B and Supplementary Figure S3A, B). In line with these results, deletion of the endogenous *E3* enhancer ($twi^{\Delta E3}$) causes a recessive lethal phenotype at embryonic stages. This is associated with a strong reduction in *twist* expression at stage 11 and to severe defects in the embryonic somatic musculature of $twi^{\Delta E3}$ mutant embryos (Figure 2C, D, Supplementary Figure S3C, D).

We next verified that the endogenous *E3* enhancer interacts with the *twist* promoter. The *twist* gene and its three enhancers are located within the same TAD, close to its boundary (Supplementary Figure S4), and in an open chromatin region marked by active histone modifications (Supplementary Figure S4 and S5). The *E3* enhancer is located approximately 3 kb downstream of the *twist* promoter (Supplementary Figure S5). To visualize chromatin interactions at such short distances, we significantly improved our 4C-Seq (circular chromosome conformation capture) protocol and applied it to

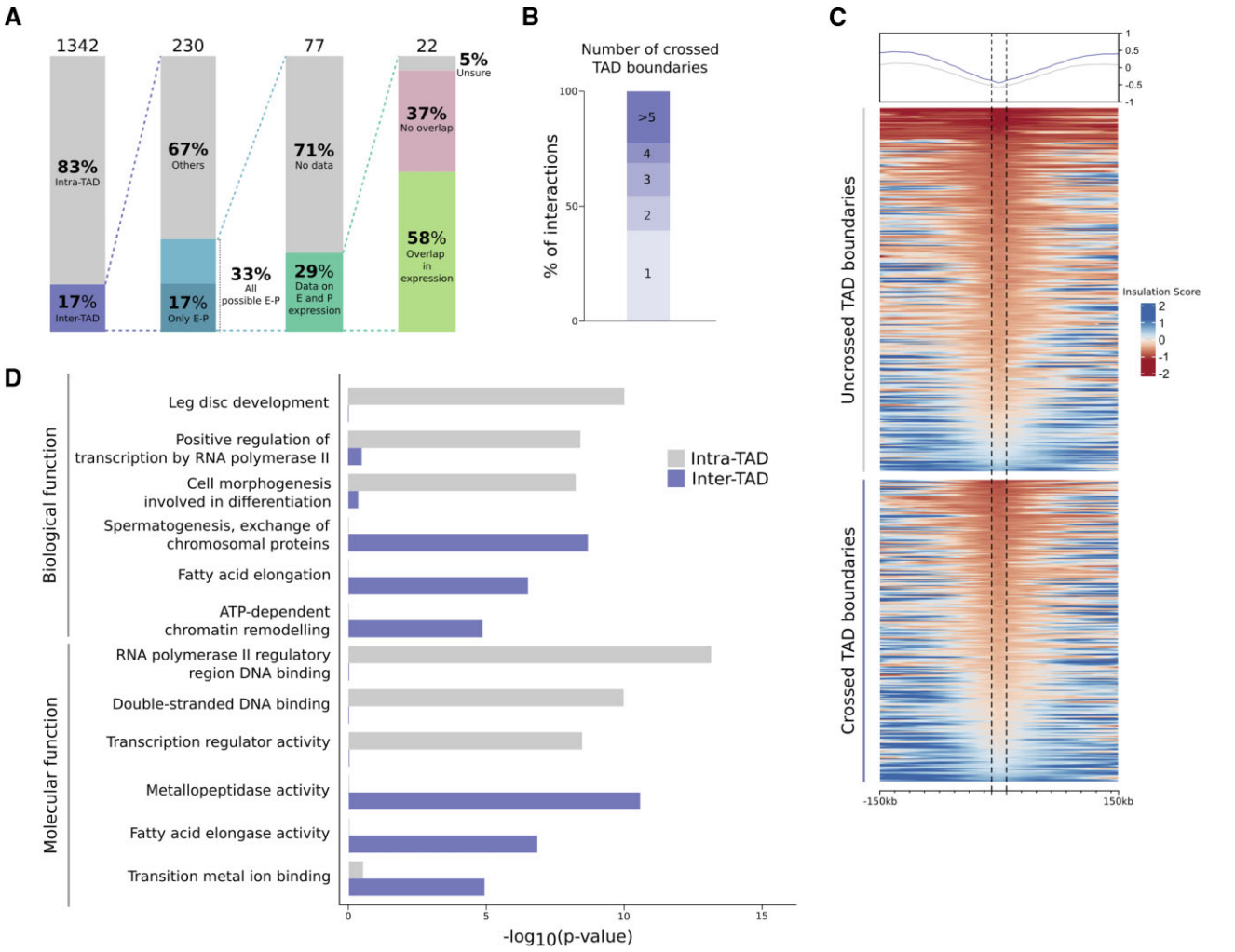



**Figure 1.** A substantial fraction of enhancer–promoter interactions are established across TAD boundaries. (**A**) Barplots showing the percentage of inter-TAD loops (17%, purple) and of inter-TAD enhancer–promoter (E–P) loops (all possible E–P 33%, only E–P 17%, blue). Barplot showing the percentage of inter-TAD E–P interactions with data on the activity of the enhancer and the promoter (29%, green). Barplot with the percentage of inter-TAD E–P interactions showing an overlap in expression pattern (58%, lime green). (**B**) Barplot showing the percentage of interactions crossing one or more TAD boundaries. (**C**) Heatmap and histogram profile of insulation scores spanning a 300kb window for uncrossed (grey) or crossed (purple) TAD boundaries. Color map shows strong insulation in red and weak insulation in blue. (**D**) Gene Ontology (GO) term enrichment analysis for the genes involved in an intra- or inter-TAD interaction based on Micro-C data from (39) and (34). Bar charts showing the − log10 of the p-value of selected GO terms for biological process and molecular function, calculated using a hypergeometric test by HOMER (50). Inter-TAD interactions: purple; intra-TAD interactions: grey. All numbers and percentages represent values after combining loops from the two datasets (34,39) and removing the duplicates.

wild-type *Drosophila* embryos at 5–8 h after egg-lay (stage 10–11) using a viewpoint anchored within the *twist* gene. We observed an interaction between the *twist* promoter and a region overlapping the *E3* enhancer, confirming that the endogenous *E3* enhancer interacts with the *twist* promoter during early embryogenesis (Figure 2E and Supplementary Figure S5).

Taken together, these results show that during stage 11, *twist* activity is specifically regulated by *E3* and that the *E3* enhancer forms a chromatin loop with the *twist* promoter. The deletion of this enhancer causes severe defects in mesoderm development, eventually resulting in embryonic lethality. Therefore, the *twist* locus, and in particular the activation of *twist* by the *E3* enhancer, provides an easily tractable model to determine the outcome of changes in E–P interactions on gene expression and to explore the functionality of inter-TAD E–P interactions involving a developmental gene.

## Relocating the *E3* enhancer to different genomic positions

To determine if *E3* could establish an interaction with the *twi* promoter that spans across TAD boundaries, we performed extensive genomic engineering of the *twist* locus by inserting an ectopic *E3* enhancer in different TADs, at various linear distances from the *twist* promoter (ranging from 7.5 kb to 1.6 Mb and on another chromosome; Figure 3A), generating a total of six different fly lines. To distinguish it from the endogenous *E3* in genomic experiments, the ectopic enhancer contained several naturally-occurring single-nucleotide variants (Supplementary Figure S1). In all cases, the fly lines were homozygous viable. The insertion sites were selected based on their distance to the *twist* promoter and their location with respect to TADs, chromatin domains, and A/B compartments (Supplementary Figure S6 and S7). To minimize deleterious effects, we avoided regions containing annotated genes and regulatory sequences (annotated in the REDfly

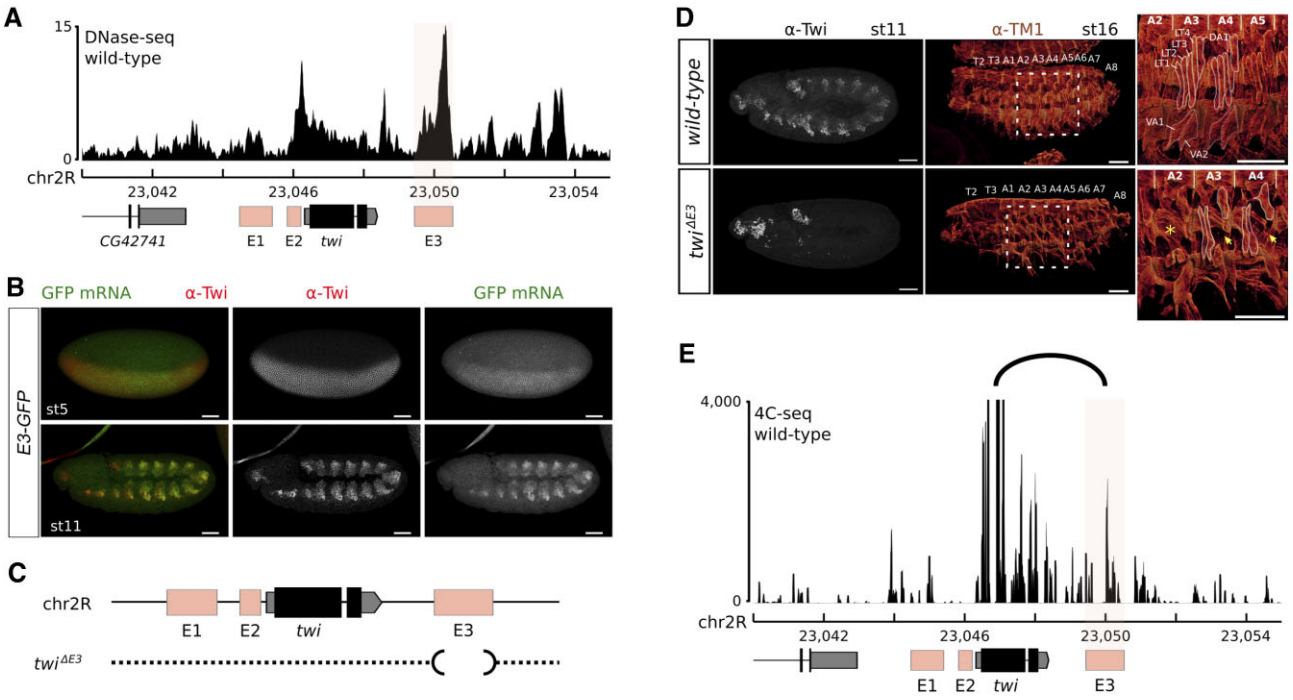



**Figure 2.** The *E3* enhancer activates the expression of *twist* during embryogenesis. (**A**) The *twist E3* enhancer is located in an open chromatin region as defined by DNase-seq signal in wild-type embryos at stage 11 (83). (**B**) Immunostaining with the α-Twist antibody (red) and expression (smiFISH) driven by its *E3* enhancer (*GFP*, green) at stage 5 (top) and 11 (bottom). Scale bars: 50 μm. (**C**) Schematic representation of the *twist* locus and of the $twi^{\Delta E3}$ deletion. (**D**) Immunostaining with the α-Twist antibody at stage 11 (white, left) and the α-TM1 antibody at stage 16 (red, middle) in wild-type (top) and $twi^{\Delta E3}$ embryos (bottom). The location of thoracic segments T2-T3 and abdominal segments A1-A8 is indicated. A blow-up (right) indicates the location of specific embryonic body muscles (LT1–4: lateral transverse muscles, DA1: dorsal acute, VA: ventral acute) and the location of missing muscles in the mutant (yellow asterisk: muscles absent in the whole segment, yellow arrows: absence of specific muscles). Scale bars: 50 μm. (**E**) 4C-seq interaction map at the *twist* locus in wild-type embryos at 5 to 8 hours after egg-lay. The observed interaction between the *twist* promoter and the *E3* enhancer is highlighted by an arc. One representative experiment is shown.

database (82) or overlapping DNase I hypersensitive sites during embryogenesis (83)) (Supplementary Figure S6 and S7). In addition, we ensured that none of the insertion sites interact with the *twist* promoter in wild-type embryos as visualized by 4C-seq (Supplementary Figure S8) and Micro-C (Supplementary Figure S6 and S7). Overall, of the six insertion sites that were generated, the E3(+7.5 kb) is located in the same TAD as twist, while the E3(+39 kb) and E3(+51 kb) insertion sites are both located in the adjacent downstream TAD. Two additional insertion sites, E3(–181 kb) and E3 (–1.6 Mb), are located in more distal upstream regions, and the last insertion site, E3(chr3L), is located on a different chromosome (Figure 3A, Supplementary Figure S6 and S7). Finally, we verified the location of each insertion site with regard to A/B compartments, respectively associated to open and closed chromatin, by calculating the eigenvector of a Hi-C contact matrix obtained from stage 5 to 8 whole embryos (84) (Materials and methods). Except for E3(+51 kb), all insertion sites are located in an A compartment.

## E–P inter-TAD interactions are established in a context-dependent manner

To establish whether ectopically inserted *E3* enhancers could engage in long-range enhancer–promoter interactions with the *twist* promoter, we analyzed the chromatin organization in these fly lines by generating 4C-seq interaction maps from *Drosophila* embryos collected at 2–5 h (stage 5–9) and 5–8 h (stage 10–11) after egg-lay (Figure 3B, C). We used two different viewpoints to have a more comprehensive view of chromatin organization around the *twist* locus: one located at the TAD boundary upstream of the *twist* promoter (viewpoint Twi1) and one within the *twist* gene (viewpoint Twi2).

To characterize the interactions of the *twist* locus with each ectopic *E3* insertion, we measured three parameters in our 4C-seq maps: (i) 'ectopic *E3*', defined as the percentage of interactions exclusively established with the ectopic version of the *E3* enhancer over all versions of *E3*. (ii) '*E3* strength', defined as the percentage of interactions overlapping the ectopic enhancer site as compared to a 10 kb region around the insertion site and used to estimate the strength of the interaction. (iii) 'background', defined as the average '*E3*' strength ratios over six 10 kb sliding windows that cover a 12 kb region upstream and downstream of the ectopic E3 site. We considered that two regions interact if 'ectopic *E3*' was >10% and/or '*E3* strength/background' was >1.7.

We first focused on the fly lines containing the ectopic *E3* positioned downstream of the *twist* locus, corresponding to the insertions at +7.5 kb (line *E3(+7.5 kb)*), +39 kb (line *E3(+39 kb)*) and + 51 kb (line *E3(+51 kb)*) from the *twist* promoter. Placing the ectopic *E3* enhancer in the same TAD as *twist* (line *E3(+7.5 kb)*) did not affect its ability to engage in chromatin interactions with the *twist* promoter. Indeed, at 5–8 h after egg-lay, on average 31.7% of the reads mapping to the *E3* enhancer corresponded to the ectopic *E3* enhancer, indicating that the *twist* promoter interacts with both copies of the *E3* enhancer (Figure 3B, Supplementary Figure S9A, C).

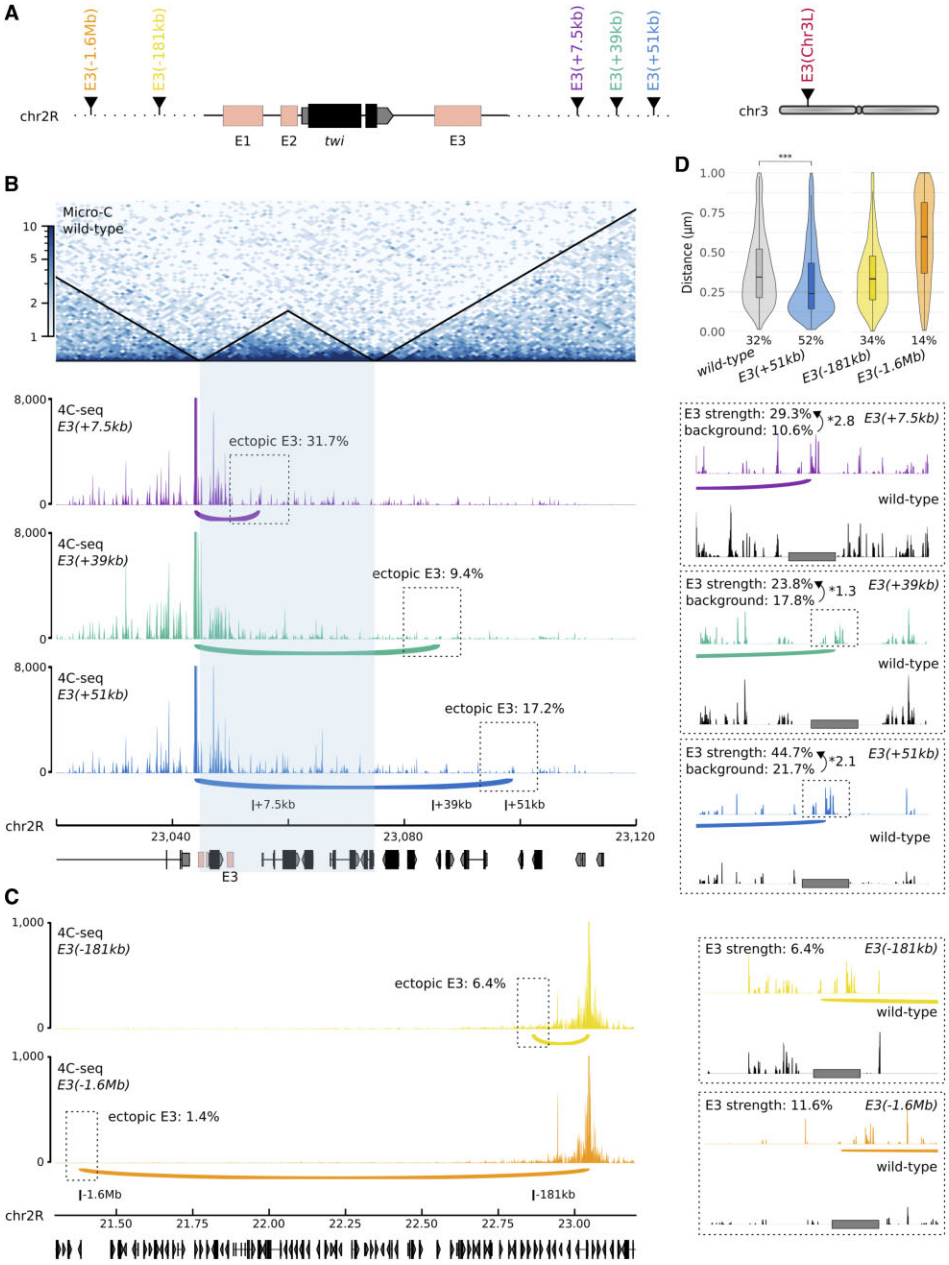



**Figure 3.** The *twist* promoter interacts with the *E3* enhancer across large genomic distances. (**A**) Schematic representation of the different ectopic E3 insertion sites on chr2R and chr3L. (**B**) High-resolution chromatin organization around the *twist* locus. Top to bottom: normalized Micro-C contact map at 1000 bp resolution in wild-type embryos at 5–8 h after egg-lay (two biological replicates merged), 4C-seq interaction maps in *E3(+7.5 kb)* (purple), *E3(+39 kb)* (green) and *E3(+51 kb)* (blue) embryos at 5–8 h after egg-lay (one representative experiment is shown). The TAD containing *twist* and the *E3* enhancer is highlighted in light blue. A 10 kb region surrounding the ectopic E3 sites is highlighted by a dotted box and is shown as an inset (right). Potential interactions between the *twist* promoter and the ectopic *E3* enhancer are highlighted by an arc. The percentage of ectopic *E3* reads (ectopic *E3*), the percentage of reads mapping on the 2-kb ectopic *E3* over a 10-kb window (*E3* strength), and the percentage of reads mapping on an adjacent control region (background) are indicated. Insets: 4C-seq interaction maps in a 10 kb region around the ectopic E3 sites in *E3(+7.5 kb)* (purple), *E3(+39 kb)* (green) and *E3(+51 kb)* (blue) embryos compared to wild-type (black) embryos at the same stage. The location of the ectopically-inserted sequence is highlighted by a dotted box. A grey rectangle indicates the position of the *E3* ectopic insertion in the wild-type line. (**C**) 4C-seq interaction maps in *E3(–181 kb)* (yellow) and *E3(–1.6 Mb)* (orange) embryos at 5–8 h after egg-lay (one representative experiment is shown). A 10 kb region surrounding the ectopic E3 sites is highlighted by a dotted box and shown as an inset (right). Potential interactions between the *twist* promoter and the ectopic *E3* enhancer are highlighted by an arc. The percentage of ectopic *E3* reads (ectopic *E3*) and the percentage of reads mapping on the 2-kb ectopic *E3* over a 10-kb window (*E3* strength) are indicated. Insets: 4C-seq interaction maps in a 10 kb region around the ectopic E3 sites in *E3(–181 kb)* (yellow) and *E3(–1.6 Mb)* (orange) embryos compared to wild-type (black) embryos at the same stage. A grey rectangle indicates the position of the *E3* ectopic insertion in the wild-type line. (**D**) Left: violin plots representing 3D DNA FISH distances measured in mesodermal nuclei between a probe located next to the *twist* promoter and a probe located next to the +51 kb insert site in wild-type ($n = 2346$; grey) and *E3(+51 kb)* ($n = 3034$; blue) embryos at stage 11. A non-parametric two-sample Kolmogorov–Smirnov test was used to assess the significant difference between DNA FISH distance distributions ($p = 2.2e^{-16}$). Right: Violin plots representing 3D DNA FISH distances between a probe located next to the *twist* promoter and a probe located next to the –181 kb insert site in *E3(–181 kb)* ($n = 642$; yellow) embryos and to the –1.6 Mb insert site in *E3(–1.6 Mb)* ($n = 570$; orange) embryos at stage 5. The percentage of colocalization (defined as the percentage of probe pairs with a distance $< 0.25$ µm) is indicated for each condition. Boxplots within the violin plots show median, edges are 25th, 75th percentiles, whiskers extend to non-outlier data points. $^{***}P \leq 0.001$.

When the ectopic *E3* enhancer was inserted in a different TAD, however, we observed two opposite situations. At position +39 kb (line *E3(+39 kb)*), no significant interaction was observed at 5–8 h after egg-lay when compared to a wild-type control as only 9.36% of the reads mapped to the ectopic *E3* enhancer, and the interaction strength ('*E3* strength') was similar to the background control (23.8% versus 17.8%, respectively) (Figure 3B, Supplementary Figure S9B–D). In contrast, when the ectopic *E3* enhancer was inserted at position +51 kb (line *E3(+51 kb)*), we detected a specific interaction between the enhancer and the *twist* promoter with 17.2% of the reads mapping the *E3* enhancer corresponding to the ectopic version of *E3* (Figure 3B, Supplementary Figure S9C, Supplementary Figure S10A). In addition, the interaction strength ('*E3* strength') was nearly doubled in the *E3(+51 kb)* line (44.7%) compared to the background control (21.7%) (Figure 3B, Supplementary Figure S9D). We also observed an increase in the ectopic *E3* interaction frequency between 2–5 and 5–8 h after egg-lay, from 26% to 31.7% in line *E3(+7.5 kb)* and from 13.7% to 17.2% in line *E3(+51 kb)*. This was not the case for the non-interacting *E3(+39 kb)* line (from 9.8% to 9.4%; Supplementary Figure S9C). The increase in ectopic *E3* interaction correlates with the specific activity of *E3* during stage 11. We further validated the interaction between the *twist* promoter and the ectopic *E3(+51 kb)* enhancer by performing 3D DNA fluorescence *in situ* hybridization (FISH) in *E3(+51 kb)* embryos at stage 5 and 11 and measuring the 3D distance between the *twist* promoter and the ectopic *E3* insertion in mesodermal cells (Figure 3D, Supplementary Figure S11). As a control, we measured the same distance in a wild-type line. The distance distribution was significantly different between the two lines ($p = 2.2e^{-16}$). Overall, we observed a decrease in the distance between the region containing the ectopic *E3* and the *twist* promoter in the *E3(+51 kb)* line as compared to the wild-type control (0.36 μm in the wild-type control, 0.26 μm in the *E3(+51 kb)* line at stage 11) with a corresponding increase in percentage of colocalization – 32% in the wild-type control versus 52% in the *E3(+51 kb)* line (Figure 3D, Supplementary Figure S11).

In lines *E3(–181 kb)* and *E3(–1.6 Mb)*, the ectopic *E3* enhancer is located at a much larger distance from the *twist* locus, with several TADs in between. In these lines, we did not detect any significant interaction between the ectopic *E3* enhancer and the *twist* promoter (Figure 3C, D, Supplementary Figure S10B). Compared to a wild-type control, the ectopic E3 site was not enriched for 4C-seq interaction in those fly lines (Figure 3C, inset). The absence of interaction was also observed by 3D DNA FISH in the *E3(–1.6 Mb)* and *E3(–181 kb)* fly lines (Figure 3D; 34% and 14% colocalization, respectively), confirming that the *E3* enhancer is not able to establish a long-range interaction with the *twist* promoter when located at the –1.6 Mb and –181 kb sites.

Taken together, our results show that the *E3* enhancer is able to establish an interaction with the *twist* promoter when placed on a different TAD. However, the analysis of the different lines suggests that these inter-TAD E–P interactions are established in a context-dependent manner.

## The *E3* enhancer establishes functional inter-TAD E–P interactions

Close 3D proximity between an enhancer and a promoter does not always result in gene expression. In fact, E–P interactions can be detected in tissues or during stages when the corresponding gene is not actively transcribed (32–34). Furthermore, transcriptional activation can sometimes lead to an increase in the distance between the enhancer and its target gene (85,86). To demonstrate the biological relevance of the observed interactions between the *twist* promoter and various ectopic *E3* insertions, we determined if they could rescue the deletion of the endogenous *E3* enhancer ($twi^{\Delta E3}$) (Figure 4A, B and Supplementary Figure S12). The insertion of the ectopic *E3* enhancer 7.5 kb downstream of the *twist* promoter, fully rescued viability (Figure 4B, Supplementary Figure S12A; 81% versus 0%), *twist* expression (Supplementary Figure S12B), and muscle formation (Supplementary Figure S12C) of $twi^{\Delta E3}$ embryos. There was however a more modest effect with the ectopic *E3* enhancer positioned 51 kb downstream of the *twist* promoter, with $twi^{\Delta E3}$, *E3(+51 kb)* embryos displaying a partial rescue in embryo viability (Figure 4B and Supplementary Figure S12A; 14% versus 0%), *twist* expression (Figure 4C and Supplementary Figure S12B), and muscle formation (Figure 4D and Supplementary Figure S12C) relative to $twi^{\Delta E3}$ embryos. Finally, as expected, the *E3* insertion lines that failed to interact with the *twist* promoter were unable to rescue the embryonic lethality of $twi^{\Delta E3}$ embryos (Supplementary Figure S12A), the expression of *twist* (Supplementary Figure S12B), and muscle formation (Supplementary Figure S12C).

Taken together, these results demonstrate that the *E3* enhancer is able to interact and activate the *twist* promoter when placed in different genomic locations, including when bypassing TAD boundaries.

## Ectopic *twist-E3* interactions are not dependent on the endogenous *E3* enhancer

To confirm that the activation of *twist* by the ectopic *E3* enhancers is dependent on enhancer–promoter looping, we repeated 4C-seq experiments in $twi^{\Delta E3}$, *E3(+7.5 kb)* and $twi^{\Delta E3}$, *E3(+51 kb)* embryos, where the endogenous *E3* was deleted. In both cases, we were able to detect an interaction between the *twist* promoter and the +7.5 kb or +51 kb ectopic *E3* enhancers (Figure 4E, F and Supplementary Figure S13). Mesoderm-specific 3D DNA FISH experiments performed at stage 5 and 11 further validated this observation (Figure 4G, Supplementary Figure S11A), with a significant difference in the distance distribution ($p = 2.2e^{-16}$) and an increased colocalization from 32% in the wild-type control to 51% in the $twi^{\Delta E3}$, *E3(+51kb)* line. Deleting the endogenous *E3* enhancer in the *E3(+51 kb)* line however resulted in a slight decrease in 4C-seq interaction frequency (Figure 4F; *E3* strength from 44.7% to 33.3%) as well as a slight increase in the distance between the *twist* promoter and the +51kb insertion site (Supplementary Figure S11; colocalization from 56% to 43%). This effect appeared more pronounced at stage 5 than at stage 11 (Supplementary Figure S11A), suggesting that the presence of the endogenous *E3* enhancer might favor a more compact conformation at earlier stages.

Overall, these data show that the *twist* promoter can engage in functional enhancer–promoter interactions across large distances—greater than about 10 kb, which would be the typical average distance between known enhancer–promoter pairs in *Drosophila*—and that these interactions do not depend on the presence of the endogenous *E3* enhancer.

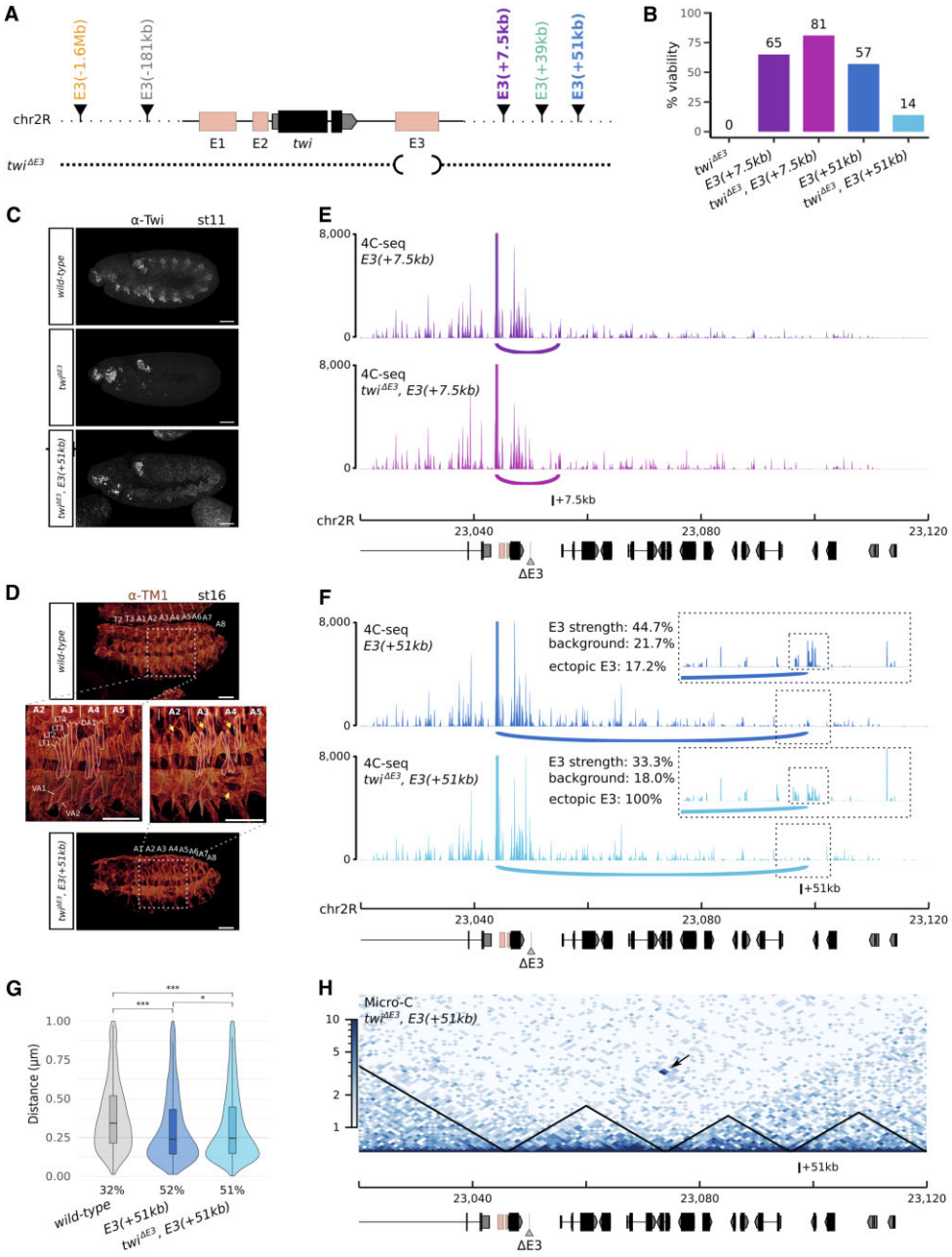



**Figure 4.** Long-range interactions between the ectopic *E3* and the *twist* promoter can rescue $twi^{\Delta E3}$mutants. (**A**) Schematic representation of the different ectopic E3 insertion sites on chr2R and of the $twi^{\Delta E3}$ deletion. (**B**) Bar plot representing the percentage of viable embryos on the $twi^{\Delta E3}$ (black), *E3(+7.5 kb)* (dark purple), $twi^{\Delta E3}$*, E3(+7.5 kb)* (light purple), *E3(+51 kb)* (dark blue) and $twi^{\Delta E3}$*, E3(+51 kb)* (light blue) lines. For each condition, at least two independent experiments were performed, with at least 50 embryos each. (**C**) Immunostaining with the α-Twist antibody in wild-type (top), $twi^{\Delta E3}$ (middle) and $twi^{\Delta E3}$*, E3(+51 kb)* embryos (bottom) at stage 11. Scale bars: 50 µm. (**D**) Immunostaining with the α-TM1 antibody in wild-type (top) and $twi^{\Delta E3}$*, E3(+51 kb)* embryos (bottom) at stage 16. The location of thoracic segments T2–T3 and abdominal segments A1–A8 is indicated. A blow-up (dotted scare) indicates the location of specific embryonic body muscles (LT1–4: lateral transverse muscles, DA1: dorsal acute, VA: ventral acute) and the location of missing muscles in the mutant (yellow arrows). Scale bars: 50 µm. (**E**) 4C-seq interaction maps in *E3(+7.5 kb)* (dark purple) and $twi^{\Delta E3}$*, E3(+7.5 kb)* (light purple) embryos at 5–8 h after egg-lay (one representative experiment is shown). (**F**) 4C-seq interaction maps in *E3(+51 kb)* (dark blue), and $twi^{\Delta E3}$*, E3(+51kb)* (light blue) embryos at 5–8 h after egg-lay (one representative experiment is shown). A 10 kb region surrounding the ectopic E3 sites is highlighted by a dotted box and shown as an inset. Potential interactions between the *twist* promoter and the ectopic *E3* enhancer are highlighted by an arc. The percentage of reads mapping on the 2-kb ectopic *E3* over a 10-kb window (E3/bkgd), the percentage of reads mapping on an adjacent control region (control), and the percentage of ectopic *E3* reads (ectopic *E3*) are indicated. (**G**) Violin plots representing 3D DNA FISH distances measured in mesodermal nuclei between a probe located next to the *twist* promoter and a probe located next to the +51 kb insert site in wild-type ($n = 2346$; grey), *E3(+51 kb)* ($n = 3034$; dark blue), and $twi^{\Delta E3}$*, E3(+51 kb)* ($n = 5460$; light blue) embryos at stage 11. A non-parametric two-sample Kolmogorov–Smirnov test was used to assess the significant difference between DNA FISH distance distributions (wild-type versus *E3(+51 kb)*: $p = 2.2e^{-16}$, wild-type versus $twi^{\Delta E3}$*, E3(+51 kb)*: $p = 2.2e^{-16}$, *E3(+51 kb)* versus $twi^{\Delta E3}$*, E3(+51 kb)*: $p = 0.04$). The percentage of colocalization (defined as the percentage of probe pairs with a distance $< 0.25$ µm) is indicated for each condition. Boxplots within the violin plots show median, edges are 25th, 75th percentiles, whiskers extend to non-outlier data points. $^{*}P \leq 0.05$; $^{***}P \leq 0.001$. (**H**) Normalized Micro-C contact map at 1000 bp resolution in $twi^{\Delta E3}$*, E3(+51 kb)* embryos at 5–8 h after egg-lay (two biological replicates merged). The interaction between the *twist* locus and the +51 kb insertion site is indicated by a black arrow.

### The ectopic *E3(+51kb)* insertion creates a new boundary but does not affect the insulation properties of existing boundaries

TAD boundaries have been suggested to play an important role in blocking E–P interactions, by insulating genes from regulatory elements present within neighboring TADs. To exclude the possibility that the interaction between the ectopic *E3* enhancer and the *twist* promoter observed in line *E3(+51 kb)* results from a weakening of the TAD boundary, we analyzed global changes in chromatin organization as a consequence of repositioning *E3*. For this purpose, we generated Micro-C contact maps at 5–8 h after egg-lay in the wild-type and $twi^{\Delta E3}$, *E3(+51 kb)* lines. The Micro-C maps provided us with an additional opportunity to confirm the presence of long-range interactions between the *twist* locus and the ectopic *E3* enhancer at the +51 kb site (Figure 4H, black arrow) that were absent in the wild-type line (Figure 3B, Supplementary Figure S14). However, the ectopic *E3* insertion did not affect the presence of the boundary located between the *twist* locus and the +51 kb site. Instead, we observed the formation of an additional boundary at the +51 kb site as confirmed by a strong shift in the directionality index (Supplementary Figure S14, blue arrow), likely driven by the strong interaction between *twist* and the *E3(+51 kb)* enhancer. This insertion also resulted in a shift of the region between the *twist* locus and the +51 kb site from an A compartment to a B compartment, while the position of the +51 kb site itself shifted from a B compartment to an A compartment (Supplementary Figure S14, red arrow).

To further validate the robustness of the boundary between the *twist* gene and the +51 kb insertion site even after the insertion of the ectopic *E3* enhancer, we conducted RT-qPCR analysis to examine the expression of several genes on both sides of the boundary in *E3(+51 kb)* embryos. We observed a striking contrast between two genes located in the same TAD as *twist* (*l(2)k009913* and *Fib*) which remained unaffected, and two genes located in the same TAD as the +51 kb insertion (*CG9877* and *CG9876*) showing a significant upregulation compared to the wild-type control (Supplementary Figure S15). Of note, a similar pattern was observed in the *E3(+39 kb)* line (Supplementary Figure S15), indicating that the lack of interaction between the *twist* promoter and the *E3(+39 kb)* enhancer is not due to the inactivity of this enhancer at this specific location.

Overall, our results confirm that an inter-TAD interaction is established between the *twist* promoter and the ectopic *E3* enhancer at position +51 kb and that this interaction is not the result of a weakening of the TAD boundary between these two loci. However, the *E3(+51 kb)* insertion did result in small changes in chromatin organization, with the E3(+51 kb) site shifting from a B to an A compartment.

### The *E3* enhancer establishes *trans*-interactions with the *twist* promoter

The last E3 insertion site included in our analysis is located on a different chromosome than *twist*, allowing us to investigate the possibility of inter-chromosomal enhancer–promoter interactions. Surprisingly, 4C-seq experiments in *E3(chr3L)* embryos revealed that at 5–8 h after egg-lay, an average of 46.9% of the reads mapping to the *E3* enhancer corresponded to the ectopic *E3* enhancer (Figure 5A, B, Supplementary Figure S9C, and Supplementary Figure S16). The interaction between the *twist* promoter and the ectopic *E3(chr3L)* enhancer was also validated by performing 3D DNA FISH in stage 5 *E3(chr3L)* embryos, and measuring the distance between the *twist* promoter and the ectopic *E3* insertion in mesodermal and non-mesodermal cells (Figure 5C and Supplementary Figure S16C). As a control, we measured the same distance in a wild-type line. The distance distribution was significantly different between the two lines ($p = 1.39e^{-3}$). We observed a decrease in the distance between the region containing the ectopic *E3* and the *twist* promoter in the *E3(chr3L)* line as compared to the wild-type control, with a corresponding increase in the percentage of colocalization, from 17% in the wild-type control to 24% in the *E3(chr3L)* line (Figure 5C). A similar trend was observed in non-mesodermal nuclei (Supplementary Figure S16C). We observed two populations of nuclei in *E3(chr3L)* embryos: a population where the two loci are very distant (as in the wild-type condition), and a population where the two loci are highly colocalized. This might indicate that, while this interaction can be very strong in some cells, it is also highly unstable. This observation is further supported by the inability of the *E3(chr3L)* insertion to rescue the viability (Supplementary Figure S12A) and the expression of *twist* in $twi^{\Delta E3}$ mutant embryos (Figure 5D).

In sum, these results reaffirm that the ectopic *E3* enhancer is able to establish a specific interaction with the *twist* promoter across large distances, even when positioned in a different chromosome. In this particular situation however, while there is a strong association between the two regulatory elements, the inability of the *E3(chr3L)* enhancer to rescue the expression of Twist in $twi^{\Delta E3}$ mutants suggests that this *trans*-interaction is likely to be unstable.

## Discussion

Since their discovery, TADs have been proposed to play a central role in promoting and/or constraining E–P interactions within their boundaries (20). However, Hi-C contact maps and single-cell imaging studies suggest that TAD boundaries only result in a two-fold depletion in interaction frequency (87,88), questioning the nature of inter-TAD interactions and their importance for gene expression. Our genome-wide analysis of Micro-C data from *Drosophila* embryos revealed that, as in mammals (31), a substantial percentage of E–P interactions are established across TAD boundaries. Analysis of these interactions revealed that intra-TAD interactions are enriched in genes involved in developmental processes and transcription regulation whereas genes involved in inter-TAD interactions are more related to signaling and metabolic functions. This suggests that TADs are of particular importance for developmental genes, potentially ensuring their faithful regulation by constraining E–P interactions. To challenge the functional requirement of this constraint, we used *twist*, a master regulator of mesoderm development, and analyzed its ability to establish inter-TAD E–P interactions. Repositioning the *E3* enhancer within the TAD where *twist* is located caused no changes in *twist* expression. The ectopic *E3(+7.5 kb)* enhancer was able to fully rescue both the expression of *twist* as well as the lethality in $twi^{\Delta E3}$ mutant embryos. However, when positioned in a different TAD, the *E3* enhancer showed context-dependent effects. Of the five inter-TAD insertions, only *E3(+51 kb)*—located in the adjacent TAD of *twist*, and *E3(chr3L)* - located in chromosome 3, were able to interact with the *twist* promoter. Even among these two lines, only

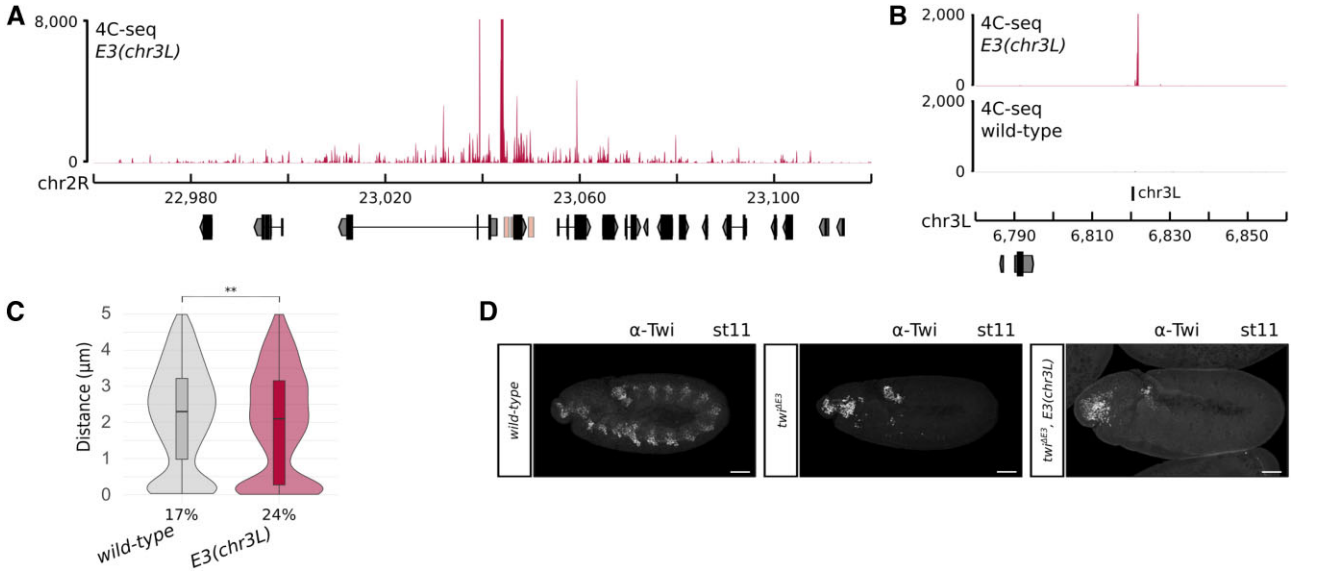



**Figure 5.** The *twist* promoter interacts with the *E3* enhancer placed on a different chromosome. (**A**) 4C-seq interaction map in *E3(chr3L)* (red) embryos at 5–8 h after egg-lay around the *twist* locus (one representative experiment is shown). (**B**) 4C-seq interaction maps in *E3(chr3L)* (red) and wild-type (black) embryos at 5–8 h after egg-lay around the chr3L insert site (one representative experiment is shown). (**C**) Violin plots representing 3D DNA FISH distances measured in mesodermal nuclei between a probe located next to the *twist* promoter and a probe located next to the chr3L insert site in wild-type ($n = 1299$; grey) and *E3(chr3L)* ($n = 1375$; red) embryos at stage 5. A non-parametric two-sample Kolmogorov–Smirnov test was used to assess the significant difference between DNA FISH distance distributions ($p = 1.39e^{-3}$). The percentage of colocalization (defined as the percentage of probe pairs with a distance $< 0.25$ µm) is indicated for each condition. Boxplots within the violin plots show median, edges are 25th, 75th percentiles, whiskers extend to non-outlier data points. $^{**}P \leq 0.01$. (**D**) Immunostaining with the α-Twist antibody at stage 11 in wild-type (left), $twi^{\Delta E3}$ (middle) and $twi^{\Delta E3}$, *E3(chr3L)* (right) embryos. Scale bars: 50 µm.

*E3(+51kb)* was able to partially rescue the expression of *twist*, the abnormal development of the somatic musculature, and the embryonic lethality of $twi^{\Delta E3}$ mutants. Overall, by using the *twist* locus as a model, we demonstrate that developmental genes can engage in E–P interactions across TAD boundaries and even between different chromosomes. However, the inability of these interactions to fully rescue the deletion of the endogenous *E3* enhancer suggests that inter-TAD E–P interactions result in lower levels of transcriptional activation. This is in line with recent attempts of artificially inserting a TAD boundary between an E–P pair at the *Sox2* (89) or *HoxD* (90) loci, which caused a significant but not complete loss of their ability to interact and only partially affected the expression of the gene. Therefore, we propose that although developmental genes can establish inter-TAD E–P interactions, higher transcriptional rates and/or activation efficiencies can only be achieved within the context of the same TAD.

Our finding that in *Drosophila* a large proportion of E–P interactions are established across TAD boundaries is in agreement with a recent similar report in mammals (31). While this suggests that the extent of inter-TAD regulation might have been under-estimated, the function of these interactions remains to be systematically established. Indeed, only few examples of naturally-occurring inter-TAD regulation have been reported so far (91,92), and these involve lincRNAs rather than E–P interactions. Alternatively, inter-TAD interactions might often have a structural rather than a regulatory function. In fact, for 40% of the inter-TAD E–P interactions, we don't observe a transcriptional output in the embryo. This is exemplified by the *E3(chr3L)* line, where the *E3* enhancer interacts with the *twist* promoter but doesn't activate its expression. Such stable interactions have also been observed within TAD boundaries (32–34), but their function has remained elusive. Whether these inter-TAD interactions can be considered as background noise, or if they have a structural importance in the three-dimensional organization of the genome remains to be established.

While TAD-mediated enhancer–promoter proximity certainly favors rapid gene activation (38,93), especially in cells characterized by very fast cell cycles such as *Drosophila* embryonic cells, other mechanisms must exist to promote interactions across large distances. Overall, neither the genomic distance, the location of TAD boundaries, nor the enhancer sequence can solely dictate enhancer–promoter interaction specificity. Instead, we propose that long-range enhancer–promoter interactions would be favored between specific genomic sites, a feature which is reminiscent of architectural proteins such as insulators (94) and tethering elements (39,95). However, the presence of such proteins is not sufficient to drive long-range interactions. Indeed, the +51 kb insertion site does not interact with the *twist* promoter in wild-type conditions, but only upon the insertion of the *E3* ectopic enhancer. Therefore, we propose a model where enhancer–promoter interaction specificity across large distances is governed by an interplay between the sequence of the enhancer itself and the sequence of the surrounding genomic locus.

## Data availability

All raw data were submitted to ArrayExpress (https://www.ebi.ac.uk/arrayexpress/browse.html) under accession numbers: E-MTAB-12153 (4C-seq), and E-MTAB-12146 (Micro-C). The following publicly available databases and datasets were used: FlyBase r6.40 (https://flybase.org/) using the dm6 reference genome; Micro-C (E-MTAB-9784, GEO:GSE171396); DNase-seq (SRA:SRX020697,

SRA:SRX020698); Histone ChIP-seq (SRA:ERP000560); Hi-C (GEO:GSE103625); ChIP-seq of insulator proteins (CP190 GEO:GSE55257, BEAF-32 ENCODE:ENCSR533CXT, su(Hw) ENCODE:ENCSR761TCG, Mod(mdg4) ENCODE:ENCSR861XSB, CTCF GEO:GSE136407); RNA-seq (2–4 h AEL SRA:SRX246401, 6–8 h AEL SRA:SRX246422).

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We are very grateful to Mounia Lagha, Samir Merabet, Julie Soutourina, Michalis Averof and François Leulier for critically reading the manuscript and to Frank Schnorrer for valuable input during the course of the study. We thank all members of the Ghavi-Helm lab for discussions and comments on the manuscript. This work was technically supported by the IGFL sequencing facility (PSI), the IGFL microscopy facility, the PLATIM imaging facility, the Arthrotools facility and Protein Sciences Facility of the Lyon SFR Biosciences (UAR3444/US8).

*Author contributions*: Y.G.-H. conceived and supervised the study. A.G., P.B.P., S.V. and Y.G.-H. designed experiments. A.G. and S.V. generated all the transgenic lines with help from D.L. and H.T to perform microinjections. A.G., S.V. and Y.G-H. performed 4C-seq experiments. Y.G.-H. analyzed 4C-seq experiments. D.B. performed and analyzed Micro-C and RT-qPCR experiments. P.B.P. performed and analyzed all imaging experiments. P.B.P. and S.V. performed viability tests. All of the authors discussed the results and implications and commented on the manuscript at all stages. Y.G.-H., D.B. and A.G. acquired funding.

## Funding

Fondation pour la Recherche Médicale [AJE20161236686 to Y.G-H, FDT202012010607 to A.G., ECO202206015555 to D.B.]; European Research Council [ERC starting grant Enhancer3D 759708]; an internship fellowship from the UFR Biosciences Université Claude Bernard Lyon 1 (to D.B.). Funding for open access charge: [ERC starting grant Enhancer3D 759708].

## Conflict of interest statement

None declared.

## References

1. Furlong,E.E.M. and Levine,M. (2018) Developmental enhancers and chromosome topology. *Science*, **361**, 1341–1345.
2. van Arensbergen,J., van Steensel,B. and Bussemaker,H.J. (2014) In search of the determinants of enhancer–promoter interaction specificity. *Trends Cell Biol.*, **24**, 695–702.
3. Sanyal,A., Lajoie,B.R., Jain,G. and Dekker,J. (2012) The long-range interaction landscape of gene promoters. *Nature*, **489**, 109–113.
4. Pachano,T., Haro,E. and Rada-Iglesias,A. (2022) Enhancer-gene specificity in development and disease. *Development*, **149**, dev186536.
5. Le Dily,F., Baù,D., Pohl,A., Vicent,G.P., Serra,F., Soronellas,D., Castellano,G., Wright,R.H.G., Ballare,C., Filion,G., *et al.* (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.*, **28**, 2151–2162.
6. Shen,Y., Yue,F., McCleary,D.F., Ye,Z., Edsall,L., Kuan,S., Wagner,U., Dixon,J., Lee,L., Lobanenkov,V.V., *et al.* (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature*, **488**, 116–120.
7. Symmons,O., Uslu,V.V., Tsujimura,T., Ruf,S., Nassari,S., Schwarzer,W., Ettwiller,L. and Spitz,F. (2014) Functional and topological characteristics of mammalian regulatory domains. *Genome Res.*, **24**, 390–400.
8. Zhan,Y., Mariani,L., Barozzi,I., Schulz,E.G., Blüthgen,N., Stadler,M., Tiana,G. and Giorgetti,L. (2017) Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res.*, **27**, 479–490.
9. Rao,S.S.P., Huntley,M.H., Durand,N.C., Stamenova,E.K., Bochkov,I.D., Robinson,J.T., Sanborn,A.L., Machol,I., Omer,A.D., Lander,E.S., *et al.* (2014) A 3D map of the Human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
10. Lupiáñez,D.G., Kraft,K., Heinrich,V., Krawitz,P., Brancati,F., Klopocki,E., Horn,D., Kayserili,H., Opitz,J.M., Laxova,R., *et al.* (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**, 1012–1025.
11. Franke,M., Ibrahim,D.M., Andrey,G., Schwarzer,W., Heinrich,V., Schöpflin,R., Kraft,K., Kempfer,R., Jerković,I., Chan,W.-L., *et al.* (2016) Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, **538**, 265–269.
12. Tsujimura,T., Klein,F.A., Langenfeld,K., Glaser,J., Huber,W. and Spitz,F. (2015) A discrete transition zone organizes the topological and regulatory autonomy of the adjacent Tfap2c and Bmp7 genes. *PLoS Genet.*, **11**, e1004897.
13. Narendra,V., Rocha,P.P., An,D., Raviram,R., Skok,J.A., Mazzoni,E.O. and Reinberg,D. (2015) CTCF establishes discrete functional chromatin domains at the hox clusters during differentiation. *Science*, **347**, 1017–1021.
14. Guo,Y., Xu,Q., Canzio,D., Shou,J., Li,J., Gorkin,D.U., Jung,I., Wu,H., Zhai,Y., Tang,Y., *et al.* (2015) CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell*, **162**, 900–910.
15. Flavahan,W.A., Drier,Y., Liau,B.B., Gillespie,S.M., Venteicher,A.S., Stemmer-Rachamimov,A.O., Suvà,M.L. and Bernstein,B.E. (2016) Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature*, **529**, 110–114.
16. Rodríguez-Carballo,E., Lopez-Delisle,L., Zhan,Y., Fabre,P.J., Beccari,L., El-Idrissi,I., Huynh,T.H.N., Ozadam,H., Dekker,J. and Duboule,D. (2017) The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. *Genes Dev.*, **31**, 2264–2281.
17. de Wit,E., Vos,E.S.M., Holwerda,S.J.B., Valdes-Quezada,C., Verstegen,M.J.A.M., Teunissen,H., Splinter,E., Wijchers,P.J., Krijger,P.H.L. and de Laat,W. (2015) CTCF binding polarity determines chromatin looping. *Mol. Cell*, **60**, 676–684.
18. Fudenberg,G., Imakaev,M., Lu,C., Goloborodko,A., Abdennur,N. and Mirny,L.A. (2016) Formation of chromosomal domains by loop extrusion. *Cell Rep.*, **15**, 2038–2049.
19. Sun,F., Chronis,C., Kronenberg,M., Chen,X.-F., Su,T., Lay,F.D., Plath,K., Kurdistani,S.K. and Carey,M.F. (2019) Promoter-enhancer communication occurs primarily within insulated neighborhoods. *Mol. Cell*, **73**, 250–263.
20. Ghavi-Helm,Y. (2020) Functional consequences of chromosomal rearrangements on gene expression: not so deleterious after all?*J. Mol. Biol.*, **432**, 665–675.
21. Despang,A., Schöpflin,R., Franke,M., Ali,S., Jerković,I., Paliou,C., Chan,W.-L., Timmermann,B., Wittler,L., Vingron,M., *et al.* (2019) Functional dissection of the Sox9 – Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.*, **51**, 1263–1271.

22. Dong,Z., Wang,H., Chen,H., Jiang,H., Yuan,J., Yang,Z., Wang,W.-J., Xu,F., Guo,X., Cao,Y., *et al.* (2018) Identification of balanced chromosomal rearrangements previously unknown among participants in the 1000 Genomes Project: implications for interpretation of structural variation in genomes and the future of clinical cytogenetics. *Genet. Med.*, **20**, 697.
23. Ghavi-Helm,Y., Jankowski,A., Meiers,S., Viales,R.R., Korbel,J.O. and Furlong,E.E.M. (2019) Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.*, **51**, 1272–1282.
24. Laugsch,M., Bartusel,M., Rehimi,R., Alirzayeva,H., Karaolidou,A., Crispatzu,G., Zentis,P., Nikolic,M., Bleckwehl,T., Kolovos,P., *et al.* (2019) Modeling the pathological long-range regulatory effects of human structural variation with patient-specific hiPSCs. *Cell Stem Cell*, **24**, 736–752.
25. Nora,E.P., Goloborodko,A., Valton,A.-L., Gibcus,J.H., Uebersohn,A., Abdennur,N., Dekker,J., Mirny,L.A. and Bruneau,B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.
26. Rao,S.S.P., Huang,S.-C., Glenn St Hilaire,B., Engreitz,J.M., Perez,E.M., Kieffer-Kwon,K.-R., Sanborn,A.L., Johnstone,S.E., Bascom,G.D., Bochkov,I.D., *et al.* (2017) Cohesin loss eliminates all loop domains. *Cell*, **171**, 305–320.
27. Schwarzer,W., Abdennur,N., Goloborodko,A., Pekowska,A., Fudenberg,G., Loe-Mie,Y., Fonseca,N.A., Huber,W., Haering,C.H., Mirny,L., *et al.* (2017) Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, **551**, 51–56.
28. Soshnikova,N., Montavon,T., Leleu,M., Galjart,N. and Duboule,D. (2010) Functional analysis of CTCF during mammalian limb development. *Dev. Cell*, **19**, 819–830.
29. Freire-Pritchett,P., Schoenfelder,S., Várnai,C., Wingett,S.W., Cairns,J., Collier,A.J., García-Vílchez,R., Furlan-Magaril,M., Osborne,C.S., Fraser,P., *et al.* (2017) Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. *eLife*, **6**, e21926.
30. Javierre,B.M., Burren,O.S., Wilder,S.P., Kreuzhuber,R., Hill,S.M., Sewitz,S., Cairns,J., Wingett,S.W., Várnai,C., Thiecke,M.J., *et al.* (2016) Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell*, **167**, 1369–1384.
31. Hsieh,T.-H.S., Cattoglio,C., Slobodyanyuk,E., Hansen,A.S., Darzacq,X. and Tjian,R. (2022) Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nat. Genet.*, **54**, 1919–1932.
32. Espinola,S.M., Götz,M., Bellec,M., Messina,O., Fiche,J.-B., Houbron,C., Dejean,M., Reim,I., Cardozo Gizzi,A.M., Lagha,M., *et al.* (2021) Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. *Nat. Genet.*, **53**, 477–486.
33. Ghavi-Helm,Y., Klein,F.A., Pakozdi,T., Ciglar,L., Noordermeer,D., Huber,W. and Furlong,E.E.M. (2014) Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*, **512**, 96–100.
34. Ing-Simmons,E., Vaid,R., Bing,X.Y., Levine,M., Mannervik,M. and Vaquerizas,J.M. (2021) Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning. *Nat. Genet.*, **53**, 487–499.
35. Rinzema,N.J., Sofiadis,K., Tjalsma,S.J.D., Verstegen,M.J.A.M., Oz,Y., Valdes-Quezada,C., Felder,A.-K., Filipovska,T., Elst,S.v., Ramos,Z.d.A.d., *et al.* (2021) Building regulatory landscapes: enhancer recruits cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nat. Struct. Mol. Biol.*, **29**, 563–574.
36. Yokoshi,M., Segawa,K. and Fukaya,T. (2020) Visualizing the role of boundary elements in enhancer–promoter communication. *Mol. Cell*, **78**, 224–235.
37. Barinov,L., Ryabichko,S., Bialek,W. and Gregor,T. (2020) Transcription-dependent spatial organization of a gene locus. arXiv doi: https://arxiv.org/abs/2012.15819, 31 December 2020, preprint: not peer reviewed.
38. Zuin,J., Roth,G., Zhan,Y., Cramard,J., Redolfi,J., Piskadlo,E., Mach,P., Kryzhanovska,M., Tihanyi,G., Kohler,H., *et al.* (2022) Nonlinear control of transcription through enhancer–promoter interactions. *Nature*, **604**, 571–577.
39. Batut,P.J., Bing,X.Y., Sisco,Z., Raimundo,J., Levo,M. and Levine,M.S. (2022) Genome organization controls transcriptional dynamics during development. *Science*, **375**, 566–570.
40. Rowley,M.J., Poulet,A., Nichols,M.H., Bixler,B.J., Sanborn,A.L., Brouhard,E.A., Hermetz,K., Linsenbaum,H., Csankovszki,G., Aiden,E.L., *et al.* (2020) Analysis of hi-C data using SIP effectively identifies loops in organisms from C. elegans to mammals. *Genome Res.*, **30**, 447–458.
41. Venev,S., Abdennur,N., Goloborodko,A., Flyamer,I., Fudenberg,G., Nuebler,J., Galitsyna,A., Akgol,B., Abraham,S., Kerpedjiev,P., *et al.* (2021) open2c/cooltools: v0.4.1. 10.5281/zenodo.5214125.
42. Ramírez,F., Bhardwaj,V., Arrigoni,L., Lam,K.C., Grüning,B.A., Villaveces,J., Habermann,B., Akhtar,A. and Manke,T. (2018) High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.*, **9**, 189.
43. Khoueiry,P., Girardot,C., Ciglar,L., Peng,P.-C., Gustafson,E.H., Sinha,S. and Furlong,E.E. (2017) Uncoupling evolutionary changes in DNA sequence, transcription factor occupancy and enhancer activity. *eLife*, **6**, e28440.
44. Gramates,L.S., Agapite,J., Attrill,H., Calvi,B.R., Crosby,M.A., dos Santos,G., Goodman,J.L., Goutte-Gattat,D., Jenkins,V.K., Kaufman,T., *et al.* (2022) FlyBase: a guided tour of highlighted features. *Genetics*, **220**, iyac035.
45. Graveley,B.R., Brooks,A.N., Carlson,J.W., Duff,M.O., Landolin,J.M., Yang,L., Artieri,C.G., van Baren,M.J., Boley,N., Booth,B.W., *et al.* (2011) The developmental transcriptome of *Drosophila melanogaster*. *Nature*, **471**, 473–479.
46. Hammonds,A.S., Bristow,C.A., Fisher,W.W., Weiszmann,R., Wu,S., Hartenstein,V., Kellis,M., Yu,B., Frise,E. and Celniker,S.E. (2013) Spatial expression of transcription factors in Drosophila embryonic organ development. *Genome Biol.*, **14**, R140.
47. Tomancak,P., Beaton,A., Weiszmann,R., Kwan,E., Shu,S., Lewis,S.E., Richards,S., Ashburner,M., Hartenstein,V., Celniker,S.E., *et al.* (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis. *Genome Biol.*, **3**, research0088.1-88.14.
48. Gallo,S.M., Gerrard,D.T., Miner,D., Simich,M., Des Soye,B., Bergman,C.M. and Halfon,M.S. (2011) REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Res.*, **39**, D118–D123.
49. Kvon,E.Z., Kazmar,T., Stampfel,G., Yáñez-Cuna,J.O., Pagani,M., Schernhuber,K., Dickson,B.J. and Stark,A. (2014) Genome-scale functional characterization of *drosophila* developmental enhancers *in vivo*. *Nature*, **512**, 91–95.
50. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
51. Venken,K.J.T., Schulze,K.L., Haelterman,N.A., Pan,H., He,Y., Evans-Holm,M., Carlson,J.W., Levis,R.W., Spradling,A.C., Hoskins,R.A., *et al.* (2011) MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. *Nat. Methods*, **8**, 737–743.
52. Gratz,S.J., Ukken,F.P., Rubinstein,C.D., Thiede,G., Donohue,L.K., Cummings,A.M. and O'Connor-Giles,K.M. (2014) Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics*, **196**, 961–971.
53. Gratz,S.J., Wildonger,J., Harrison,M.M. and O'Connor-Giles,K.M. (2013) CRISPR/Cas9-mediated genome engineering and the promise of designer flies on demand. *fly.*, **7**, 249–255.
54. Gratz,S.J., Cummings,A.M., Nguyen,J.N., Hamm,D.C., Donohue,L.K., Harrison,M.M., Wildonger,J. and

O'Connor-Giles,K.M. (2013) Genome engineering of drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, **194**, 1029–1035.
55. Zhang,X., Koolhaas,W.H. and Schnorrer,F. (2014) A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in Drosophila. *G3 (Bethesda)*, **4**, 2409–2418.
56. Bischof,J., Björklund,M., Furger,E., Schertel,C., Taipale,J. and Basler,K. (2013) A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. *Development*, **140**, 2434–2442.
57. Bischof,J., Maeda,R.K., Hediger,M., Karch,F. and Basler,K. (2007) An optimized transgenesis system for Drosophila using germ-line-specific φC31 integrases. *Proc. Natl Acad. Sci.*, **104**, 3312–3317.
58. Balleza,E., Kim,J.M. and Cluzel,P. (2018) Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods*, **15**, 47–51.
59. Markstein,M., Pitsouli,C., Villalta,C., Celniker,S.E. and Perrimon,N. (2008) Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.*, **40**, 476–483.
60. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, 10–12.
61. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
62. Ramírez,F., Bhardwaj,V., Arrigoni,L., Lam,K.C., Grüning,B.A., Villaveces,J., Habermann,B., Akhtar,A. and Manke,T. (2018) High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.*, **9**, 189.
63. Lopez-Delisle,L., Rabbani,L., Wolff,J., Bhardwaj,V., Backofen,R., Grüning,B., Ramírez,F. and Manke,T. (2021) pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics*, **37**, 422–423.
64. Hsieh,T.-H.S., Cattoglio,C., Slobodyanyuk,E., Hansen,A.S., Rando,O.J., Tjian,R. and Darzacq,X. (2020) Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell*, **78**, 539–553.
65. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
66. Abdennur,N. and Mirny,L.A. (2020) Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics*, **36**, 311–316.
67. Kruse,K., Hug,C.B. and Vaquerizas,J.M. (2020) FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biol.*, **21**, 303.
68. Durand,N.C., Shamim,M.S., Machol,I., Rao,S.S.P., Huntley,M.H., Lander,E.S. and Aiden,E.L. (2016) Juicer provides a one-click system for analyzing loop-resolution hi-C experiments. *Cels*, **3**, 95–98.
69. Zenk,F., Loeser,E., Schiavo,R., Kilpert,F., Bogdanović,O. and Iovino,N. (2017) Germ line-inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition. *Science*, **357**, 212–216.
70. Pinto,P.B., Domsch,K., Gao,X., Wölk,M., Carnesecchi,J. and Lohmann,I. (2022) Specificity of the Hox member deformed is determined by transcription factor levels and binding site affinities. *Nat. Commun.*, **13**, 5037.
71. Calvo,L., Ronshaugen,M. and Pettini,T. (2021) smiFISH and embryo segmentation for single-cell multi-gene RNA quantification in arthropods. *Commun. Biol.*, **4**, 352.
72. Bintu,B., Mateo,L.J., Su,J.-H., Sinnott-Armstrong,N.A., Parker,M., Kinrot,S., Yamaya,K., Boettiger,A.N. and Zhuang,X. (2018) Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, **362**, eaau1783.
73. Nagano,T., Lubling,Y., Stevens,T.J., Schoenfelder,S., Yaffe,E., Dean,W., Laue,E.D., Tanay,A. and Fraser,P. (2013) Single-cell hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, **502**, 59–64.
74. Van Bortle,K., Nichols,M.H., Li,L., Ong,C.-T., Takenaka,N., Qin,Z.S. and Corces,V.G. (2014) Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.*, **15**, R82.
75. Yang,J., Mani,S.A., Donaher,J.L., Ramaswamy,S., Itzykson,R.A., Come,C., Savagner,P., Gitelman,I., Richardson,A. and Weinberg,R.A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*, **117**, 927–939.
76. Thisse,B., Stoetzel,C., Gorostiza-Thisse,C. and Perrin-Schmitt,F. (1988) Sequence of the twist gene and nuclear localization of its protein in endomesodermal cells of early Drosophila embryos. *EMBO J.*, **7**, 2175–2183.
77. Simpson,P. (1983) Maternal-zygotic gene interactions during formation of the dorsoventral pattern in drosophila embryos. *Genetics*, **105**, 615–632.
78. Jiang,J., Kosman,D., Ip,Y.T. and Levine,M. (1991) The dorsal morphogen gradient regulates the mesoderm determinant twist in early Drosophila embryos. *Genes Dev.*, **5**, 1881–1891.
79. Zinzen,R.P., Girardot,C., Gagneur,J., Braun,M. and Furlong,E.E.M. (2009) Combinatorial binding predicts spatio-temporal cis -regulatory activity. *Nature*, **462**, 65–70.
80. Pan,D.J., Huang,J.D. and Courey,A.J. (1991) Functional analysis of the Drosophila twist promoter reveals a dorsal-binding ventral activator region. *Genes Dev.*, **5**, 1892–1901.
81. Thisse,C., Perrin-Schmitt,F., Stoetzel,C. and Thisse,B. (1991) Sequence-specific transactivation of the Drosophila twist gene by the dorsal gene product. *Cell*, **65**, 1191–1201.
82. Rivera,J., Keränen,S.V.E., Gallo,S.M. and Halfon,M.S. (2019) REDfly: the transcriptional regulatory element database for Drosophila. *Nucleic Acids Res.*, **47**, D828–D834.
83. Thomas,S., Li,X.-Y., Sabo,P.J., Sandstrom,R., Thurman,R.E., Canfield,T.K., Giste,E., Fisher,W., Hammonds,A., Celniker,S.E., *et al.* (2011) Dynamic reprogramming of chromatin accessibility during Drosophila embryo development. *Genome Biol.*, **12**, R43.
84. Ogiyama,Y., Schuettengruber,B., Papadopoulos,G.L., Chang,J.-M. and Cavalli,G. (2018) Polycomb-dependent chromatin looping contributes to gene silencing during Drosophila development. *Mol. Cell*, **71**, 73–88.
85. Benabdallah,N.S., Williamson,I., Illingworth,R.S., Kane,L., Boyle,S., Sengupta,D., Grimes,G.R., Therizols,P. and Bickmore,W.A. (2019) Decreased enhancer–promoter proximity accompanying enhancer activation. *Mol. Cell*, **76**, 473–484.
86. Alexander,J.M., Guan,J., Li,B., Maliskova,L., Song,M., Shen,Y., Huang,B., Lomvardas,S. and Weiner,O.D. (2019) Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *eLife*, **8**, e41769.
87. Dekker,J. and Mirny,L. (2016) The 3D genome as moderator of chromosomal communication. *Cell*, **164**, 1110–1121.
88. Finn,E.H., Pegoraro,G., Brandão,H.B., Valton,A.-L., Oomen,M.E., Dekker,J., Mirny,L. and Misteli,T. (2019) Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell*, **176**, 1502–1515.
89. Chakraborty,S., Kopitchinski,N., Zuo,Z., Eraso,A., Awasthi,P., Chari,R., Mitra,A., Tobias,I.C., Moorthy,S.D., Dale,R.K., *et al.* (2023) Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. *Nat. Genet.*, **55**, 280–290.
90. Rodríguez-Carballo,E., Lopez-Delisle,L., Willemin,A., Beccari,L., Gitto,S., Mascrez,B. and Duboule,D. (2020) Chromatin topology and the timing of enhancer function at the HoxD locus. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 31231–31241.
91. Groff,A.F., Barutcu,A.R., Lewandowski,J.P. and Rinn,J.L. (2018) Enhancers in the Peril lincRNA locus regulate distant but not local genes. *Genome Biol.*, **19**, 219.
92. Galupa,R., Nora,E.P., Worsley-Hunt,R., Picard,C., Gard,C., van Bemmel,J.G., Servant,N., Zhan,Y., El Marjou,F., Johanneau,C., *et al.* (2020) A conserved noncoding locus regulates random monoallelic xist expression across a topological boundary. *Mol. Cell*, **77**, 352–367.

93. Chen,H., Levo,M., Barinov,L., Fujioka,M., Jaynes,J.B. and Gregor,T. (2018) Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.*, **50**, 1296–1303.
94. Cubeñas-Potts,C. and Corces,V.G. (2015) Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett.*, **589**, 2923–2930.
95. Pachano,T., Sánchez-Gaya,V., Ealo,T., Mariner-Faulí,M., Bleckwehl,T., Asenjo,H.G., Respuela,P., Cruz-Molina,S., Muñoz-San Martín,M., Haro,E., *et al.* (2021) Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nat. Genet.*, **53**, 1036–1049.

**Received:** April 25, 2023. **Revised:** November 21, 2023. **Editorial Decision:** November 24, 2023. **Accepted:** November 29, 2023