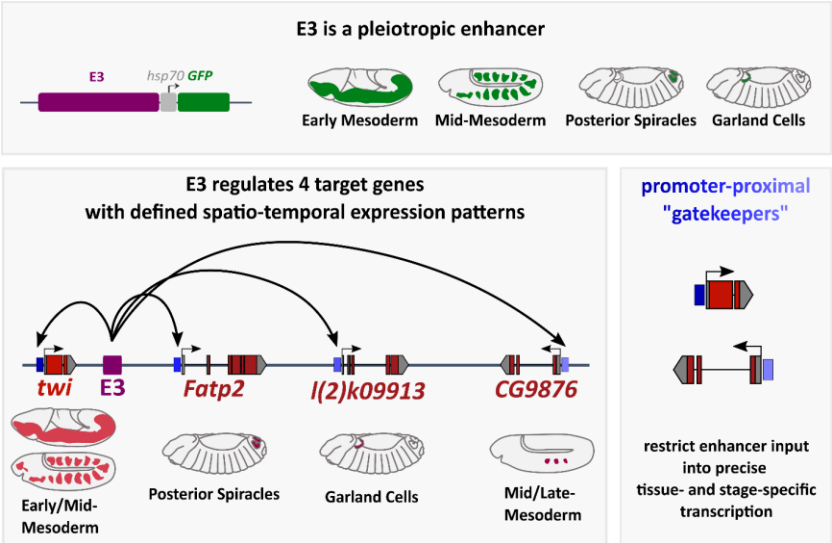
Promoter-proximal gatekeepers restrict pleiotropic enhancer inputs to achieve tissue specificity

Margarita Masoura[1], Deevitha Balasubramanian[1], Charlotte Moretti[1], Matéo Lison[1], Hélène Tarayre[1], Damien Lajoignie[1], Clara Cretet-Rodeschini[1], Laëtitia Cadet[1], Juliette Mendes[1], Séverine Vincent[1], Carlos Flores Aguirre[1], Pedro Borges Pinto[1], and Yad Ghavi-Helm[1,*]

[1]Institut de Génomique Fonctionnelle de Lyon, UMR5242, Ecole Normale Supérieure de Lyon, Centre National de la Recherche Scientifique, Université Claude Bernard-Lyon 1, 46 allée d'Italie F-69364 Lyon, France
*Corresponding author. Email: yad.ghavi-helm@ens-lyon.fr

Developmental enhancers are central regulatory elements that can activate multiple genes, yet how they selectively regulate one gene over its neighbours remains unclear. Using the *Drosophila twist* E3 enhancer as a model, we show that this selectivity is encoded within promoter-proximal regions of target genes. We found that the E3 enhancer is pleiotropic, activating at least four functionally unrelated genes. Despite receiving the same enhancer input each target gene displays distinct and non-overlapping expression patterns. These differences cannot be attributed to context-dependent chromatin looping. Instead, we demonstrate that they arise from the action of each genes' promoter-proximal sequence. These sequences act as "gatekeepers" that restrict enhancer input into precise tissue- and stage-specific transcription, thus functioning as active interpreters rather than passive recipients of enhancer signals. We propose that promoter-proximal gatekeepers provide a critical but under-appreciated layer of regulatory specificity within complex gene expression programs.



INTRODUCTION

Precise spatial and temporal patterns of gene expression are orchestrated by enhancers, *cis*-regulatory elements that integrate diverse regulatory inputs to control transcriptional programs across cell types and developmental stages[1]. Classical studies, such as the dissection of enhancers regulating *even-skipped* expression in *Drosophila*[2–5], led to the prevailing view that enhancers are modular elements, each directing gene expression in a single spatiotemporal context by contacting the promoter of their target gene. However, growing evidence is challenging this paradigm, revealing that enhancer pleiotropy, the ability of an enhancer to drive gene expression across multiple tissues or developmental stages, is far more common than previously appreciated[6–12]. For example, in mice, 52% of putative enhancers are active in multiple tissues[13]. The same trend is observed in humans with at least 40% of the enhancers being pleiotropic[11].

In most cases, enhancer pleiotropy mirrors the multifunctionality of the gene it regulates – a phenomenon often referred to as gene pleiotropy. A well-studied example is the *shavenbaby* (*svb*) locus in *Drosophila*, which contains enhancers driving expression in embryonic, larval, and pupal epidermis, reflecting the gene's diverse roles throughout development[10].

Enhancer pleiotropy illustrates how biological systems can re-use regulatory elements as a basis for the emergence of evolutionary novelties. In this context, it is conceivable that some pleiotropic enhancers may not only control a single gene across multiple expression patterns but instead regulate multiple genes with distinct and non-overlapping spatial and temporal expression profiles. While this possibility has been proposed[6,14,15], direct experimental evidence remains limited. Moreover, the mechanisms by which a pleiotropic enhancer could drive different transcriptional outputs for its multiple target genes, each active in a different spatio-temporal pattern, remains completely unexplored.

To address these questions, we used as a model the 1 kb-long *Drosophila twist* E3 enhancer, originally characterized as a mesoderm-specific regulatory element controlling *twist* expression during embryogenesis[16]. Our findings reveal that E3 is, in fact, a pleiotropic enhancer that regulates the expression of at least three additional genes in distinct spatiotemporal patterns spanning different germ layers. These target genes perform diverse functions and span a 50 kb window around the E3 enhancer. Intriguingly, despite E3's complex activity, each of its target genes is expressed within a precise and non-overlapping spatiotemporal window. We demonstrate that this specificity is not controlled by the enhancer itself or by its interactions with its target genes, but instead is determined by the differential responsiveness of each promoter to E3's regulatory input. Together, our work shows that promoter-proximal elements provide an additional layer of regulatory specificity, selectively interpreting transcriptional cues to enable gene-specific and context-dependent expression.

RESULTS

E3 is a pleiotropic enhancer whose activity spans different germ layers

The expression of *twist* (*twi*) in the *Drosophila* embryo is regulated by three partially redundant enhancers that together ensure robust transcription in the developing mesoderm (Fig. 1A)[16]. In previous work, we characterized the activity of the most distal *twist* enhancer, referred to here as E3, using a reporter construct that combines the E3 element with an *hsp70* core promoter driving *GFP*. We showed that *GFP* expression, i.e. E3 activity, faithfully recapitulates *twist* expression from stage 5 to stage 11 of embryogenesis (Fig. 1B, Supplementary Fig. 1)[16]. Upon closer inspection, however, we noticed that the reporter remains active beyond this developmental window and, strikingly, in tissues where *twist* itself is not expressed (Fig. 1B). Up to stage 11, its activity largely coincides with endogenous *twist* expression (Fig. 1B, Supplementary Fig. 1). In contrast to earlier stages, from stage 12

onwards, the reporter is expressed in broader mesodermal domains where *twist* is not expressed anymore (Fig. 1B). At stage 14, the reporter becomes active in three distinct tissues: the anterior part of the hindgut (where *twist* is also expressed), the garland cells, and the posterior spiracles (Fig. 1B). While the garland cells are mesodermal derivatives, the hindgut and the posterior spiracles originate from ectoderm cells, indicating that E3 is capable of driving gene expression across different germ layers. We therefore concluded that E3 functions as a pleiotropic enhancer capable of driving gene expression across multiple tissues and developmental stages. While enhancer pleiotropy has been described in various contexts[2,8,17], the ability of a single enhancer to drive expression across different germ layers and developmental stages has not been documented to date. This unique observation raised two major questions which we set out to explore: First, how are such different regulatory activities encoded within the E3 sequence? Second, given the spatial and temporal decoupling between E3 activity and *twist* expression, does E3 actually regulate genes other than *twist*?

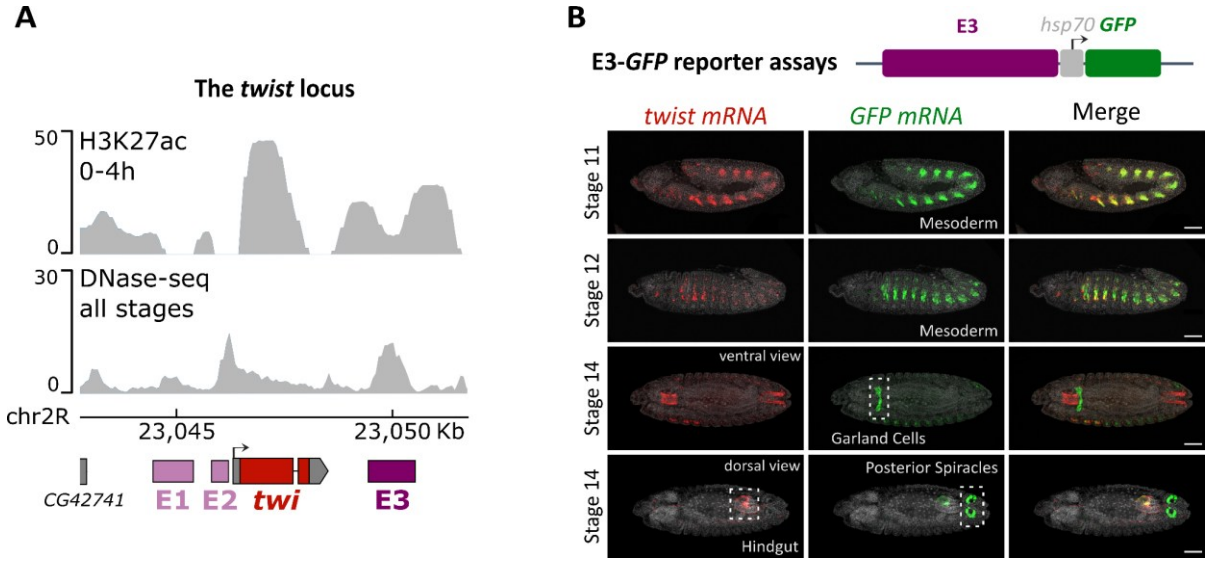


Figure 1. The E3 enhancer lies in an open, active chromatin region and drives expression in various spatio-temporal contexts during embryogenesis.
(A) The *twist* (*twi*) locus, including the E3 enhancer, resides in an open chromatin region, as indicated by DNase-seq signal[18], and is marked by H3K27ac[19], a chromatin modification associated with active enhancers.
(B) Hybridization chain reaction (HCR) RNA *in situ* hybridization of *twist* (red) and *GFP* (green) mRNA in E3-*hsp70-GFP* embryos. Representative images at three developmental stages are shown in lateral orientation unless stated otherwise. Scale bar, 50 µm (shown in merged panels).

E3 consists of three complementary domains driving distinct spatiotemporal outputs

Previous studies have shown that pleiotropic enhancers achieve complex regulatory outcomes by employing different strategies, such as the reuse of transcription factor binding sites (site pleiotropy) or the presence of discrete regulatory modules[6,20,21]. To investigate how the pleiotropic nature of E3 is encoded in its sequence, we generated a series of enhancer fragments of varying lengths and tested their activity in reporter assays (Fig. 2A, Supplementary Fig. S2,3,4,5).

This analysis revealed three main domains within the E3 enhancer, domains A, B, and C, each with distinct spatiotemporal activities, yet with functional interdependency to achieve full enhancer activity (Fig. 2B, Supplementary Fig. S2,3,4,5). Domain B (369-711bp) was active in early mesodermal cells from stage 5 to stage 10 (Fig. 2C, Supplementary Fig. S3), while mesodermal activity at stages 11 and 12 was

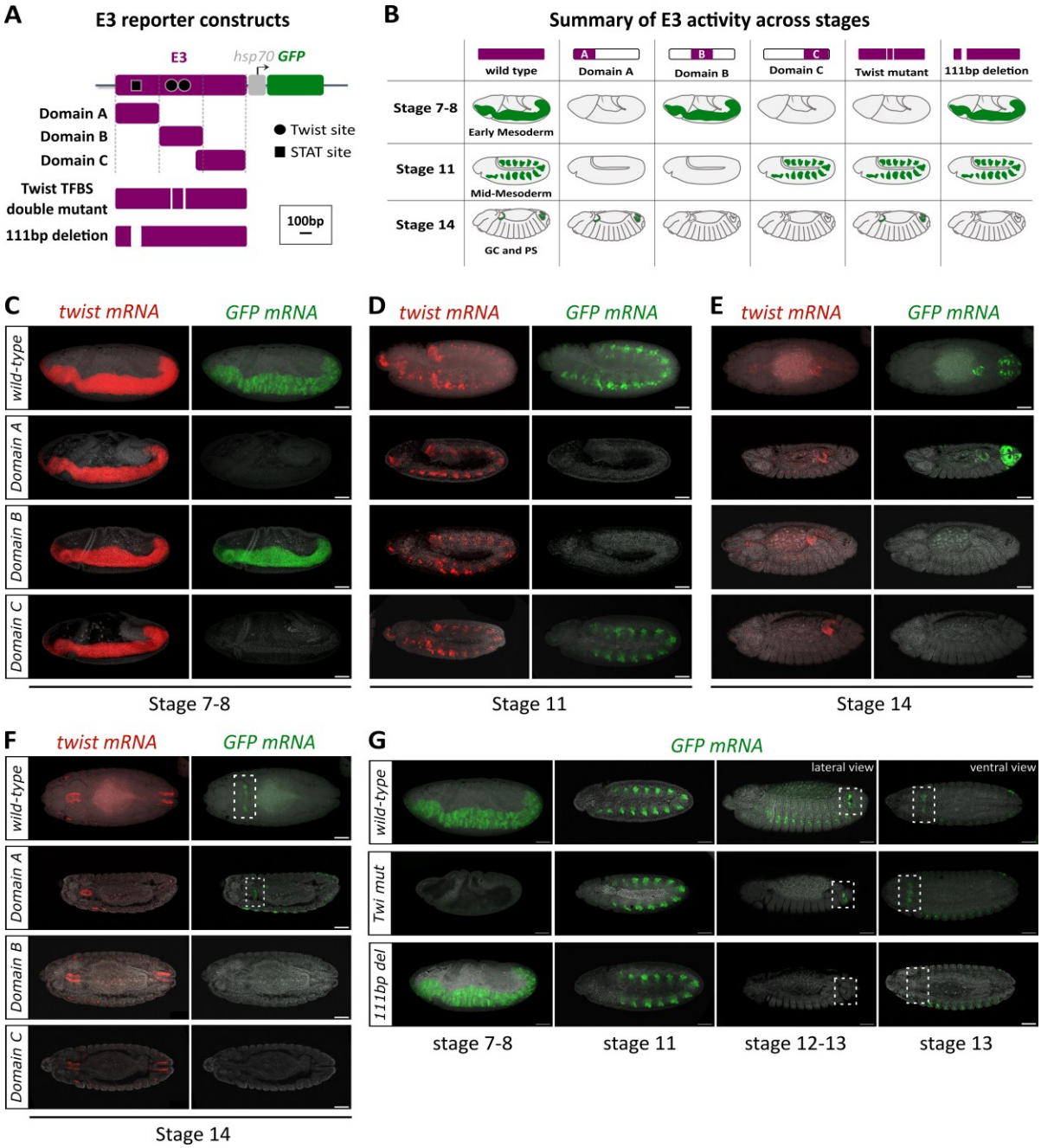


Figure 2. Dissection of the E3 enhancer reveals three complementary domains driving distinct spatiotemporal outputs.
(A) Schematic of the *E3-hsp70-GFP* reporter construct, indicating the positions of Domains A, B, and C, and the location of the Twist Transcription Factor Binding Sites (TFBS) and 111bp deletions. Binding sites for Twist and STAT transcription factors are marked in black. Scale bar: 100bp.
(B) Schematic representation of the activity of the different domains and mutants of E3. GC: Garland cells, PS: Posterior Spiracles.
(C-F) HCR RNA *in situ* hybridization of *twist* (red) and *GFP* (green) mRNA in *hsp70-GFP* reporter embryos expressing *GFP* under the control of full-length E3 or individual Domains A, B, or C. Representative embryos are shown at stages 7–8 (C), stage 11 (D), and stage 14 (E). (C) and (D) are maximum intensity projections. (E) and (F) represent two different focal planes. A dotted rectangle highlights the posterior spiracles and the garland cells. Scale bars, 50 µm.
(G) HCR RNA *in situ* hybridization of *GFP* (green) mRNA in *hsp70-GFP* reporter embryos carrying the full-length E3 or mutant enhancers with the deletion of two Twist Transcription Factor Binding Sites (TFBS) in Domain B or of a 111 bp region in Domain A (corresponding to the STAT site). Representative embryos are shown at stages 7–8 and 12–13. A dotted rectangle highlights the posterior spiracles in the lateral view and the garland cells in the ventral view. Scale bars, 50 µm.

driven exclusively by Domain C (633-1084bp) (Fig. 2D, Supplementary Fig. S4). In contrast, Domain A (1-368bp) was active in cell types of different embryonic origin: garland cells (mesoderm), posterior spiracles cells (ectoderm), and the anterior part of the hindgut (ectoderm) (Fig. 2E-F, Supplementary Fig. S5). Despite this apparent modularity, our results also indicate that the activity of E3 consists of more than the additive action of these modules. Indeed, we observed both activating and repressive interactions between different enhancer regions. For example, the posterior spiracle expression was stronger in Domain A than in the full-length enhancer, suggesting that sequences within Domain B or C may down-regulate spiracle-specific activity (Fig. 2E). Similarly, early mesodermal activity was enhanced in Domain B relative to the full-length enhancer, implying that Domain A or C may exert a repressive influence on early mesodermal expression (Fig. 2C). These results suggest that the E3 fragment is an integrated regulatory element, rather than a collection of separate enhancers.

To further elucidate the regulatory architecture of E3, we deleted two Twist transcription factor binding sites (TFBS) within Domain B. This led to a marked reduction of E3 activity in the early mesoderm, without affecting expression in the late mesoderm, posterior spiracles, or garland cells (Fig. 2G). Similarly, deletion of a 111 bp fragment in Domain A abolished reporter expression in the garland cells and in most cells of the posterior spiracles (Fig. 2G). This deletion includes a putative binding motif for STAT (Fig. 2A), a transcription factor previously implicated in posterior spiracle formation[22], suggesting that E3's activity in this tissue is at least partly mediated by JAK/STAT signaling.

Taken together, these findings indicate that E3 contains three regulatory domains, each capable of driving distinct spatial and temporal expression patterns. While trunk mesoderm activity could be clearly partitioned into two modules (Domains B and C), a 111 bp fragment within Domain A drives pleiotropic activity across germ layers, suggesting a more intricate integration of regulatory information. Despite the present of these domains, their activity is clearly fine-tuned through activating and repressive interactions between the domains, suggesting E3 is active as a single pleiotropic enhancer.

E3 regulates the expression of multiple functionally unrelated genes

Having observed that E3 is active in tissues where *twist* is not expressed, we set out to identify other potential E3 target genes. For this purpose, we surveyed the genomic region flanking E3, spanning 100 kb upstream and downstream, for genes expressed in the mesoderm, garland cells, and/or posterior spiracles. We identified three candidate genes: *Fatp2*, *l(2)k09913*, and *CG9876*, located approximately 5kb, 25kb, and 50kb downstream of E3, respectively (Fig. 3A). *Fatp2* encodes a fatty acid transport protein and *CG9876* is predicted to encode a paired-like homeobox transcription factor[23]. Both genes were reported to be expressed in the posterior spiracles from stage 13 to 16[24–26]. *l(2)k09913*, which encodes a metabolic enzyme enabling serine hydrolase activity, was reported to be expressed in the trunk mesoderm during stages 7-10, and in the garland cells from stages 11 to 16[24–26].

We verified the expression of these genes by *in situ* hybridization experiments, confirming the expression of *Fatp2* in the posterior spiracles and of *l(2)k09913* in the garland cells. Both mRNAs showed a clear colocalization with the E3 reporter (Fig. 3B-D, Supplementary Fig. S6). While *Fatp2* is expressed in the entire ring of the spiracle, E3 is active in slightly fewer cells, suggesting that *Fatp2* expression in the posterior spiracle might also be controlled by additional enhancers. Contrary to previous reports, we were unable to detect *l(2)k09913* expression in the trunk mesoderm, but instead observed a weak ubiquitous expression (Fig. 3B). We also found that *CG9876* is expressed in the mesoderm from stages 11 to 15 (Fig. 3D, Supplementary Fig. S6), consistent with its potential role as a muscle identity gene (personal communication, Guillaume Junion). Within the

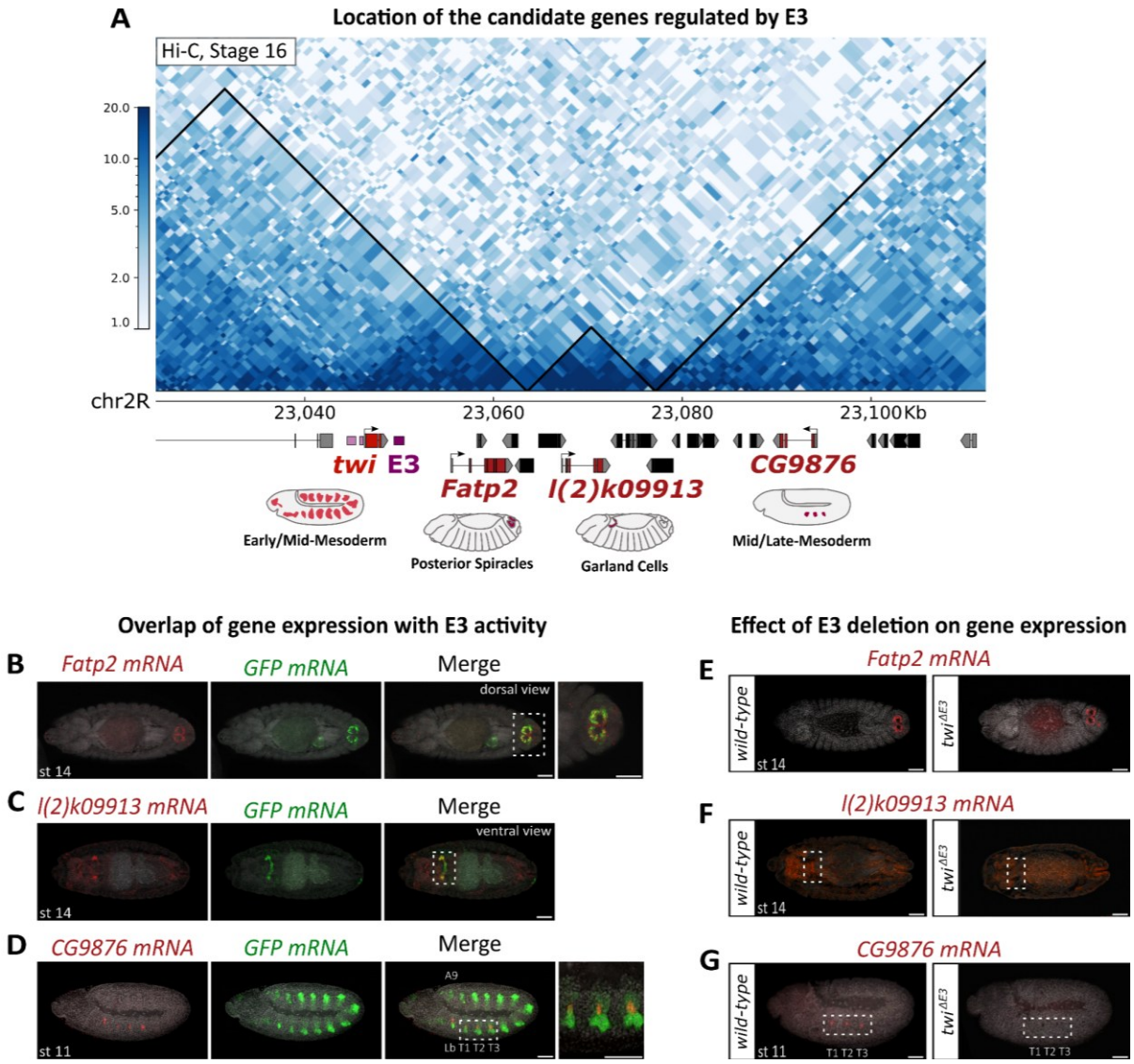


Figure 3. E3 regulates the expression of multiple functionally unrelated genes during embryogenesis.
(A) Chromatin organization around E3 and the genes it regulates – *twist*, *Fatp2*, *l(2)k09913* and *CG9876*, as visualized by Hi-C contact map at stage 16 (two biological replicates merged)[28]. The location of TADs is highlighted in black.
(B-D) HCR RNA *in situ* hybridization of *Fatp2* (B), *l(2)k09913* (C), *CG9876* (D) (red) and *GFP* (green) mRNA in E3-*hsp70-GFP* embryos. Representative embryos are shown at stage 14 (B-C) and stage 11 (D). Scale bar, 50 µm (shown in merged panels). Dotted rectangles indicated the region of interest that is shown in the zoom in for (B) and (D). The location of the Labial (LB), first three thoracic segments (T1, T2, T3) and ninth Abdominal segment (A9) is indicated.
(E-G) HCR RNA *in situ* hybridization of *Fatp2* (E), *l(2)k09913* (F), and *CG9876* (G) mRNA in wild type and $twi^{\Delta E3}$ embryos. Representative embryos are shown at stage 14 (E-F) and stage 11 (G). Scale bar, 50 µm. Dotted rectangles indicated the region of interest. The location of the first three thoracic segments (T1, T2, T3) is indicated in (G).

mesoderm, *CG9876* expression partially overlapped with the E3 reporter expression in five embryonic segments during stages 11–12 (Fig. 3D). During later stages, *CG9876* expression persisted in the mesoderm, but no longer overlapped with the E3 reporter expression, suggesting the existence of additional enhancers driving its late expression (Supplementary Fig. S6). Detailed analysis of *CG9876* expression also revealed that this gene is active in a tissue adjacent to, but not overlapping, the posterior spiracles (Supplementary Fig. S6).

To determine whether these genes are regulated by E3, we analyzed their expression in embryos

lacking the endogenous E3 enhancer (Fig. 3E-G). We previously demonstrated that E3 deletion abolishes *twist* expression specifically in the trunk mesoderm at stage 11[16]. Strikingly, in E3-deficient embryos, *CG9876* expression was completely lost in the mesoderm at stage 11 and *l(2)k09913* expression was absent from the garland cells at stage 14 (Fig. 3F-G). In the case of *l(2)k09913*, the loss of expression is likely a direct consequence of the E3 deletion, as *twist* is not normally expressed in garland cells. In contrast to the strong effects observed for *CG9876* and *l(2)k09913*, *Fatp2* expression in the posterior spiracles remained unchanged in the absence of E3 (Fig. 3E). This observation suggests either that E3 does not regulate *Fatp2*, or that other enhancers that act redundantly with E3 can sustain *Fatp2* expression in this tissue. Based on the genomic and 3D proximity between E3 and the *Fatp2* promoter that we present below, we assumed that E3 is more likely contributing to *Fatp2* expression redundantly with other enhancers, which can compensate for the absence of E3.

All together, these findings support a model in which the E3 enhancer functions as a shared pleiotropic enhancer, capable of regulating the expression of multiple genes with distinct spatio-temporal expression profiles across different germ layers. Enhancer sharing is quite prevalent for co-expressed genes[27]. However, remarkably, E3 target genes encompass diverse spatio-temporal expression patterns and molecular functions, ranging from transcription factors to metabolic enzymes. Although all targets are located within 50 kb of E3, *l(2)k09913* and *CG9876* are located in a different Topologically Associating Domain (TAD) (Fig. 3A), an observation that is consistent with our previous findings that E3 forms functional enhancer-promoter interactions across TAD borders[16].

Tissue-specific expression of E3 target genes cannot be solely attributed to chromatin accessibility

Having established that E3 activates several functionally unrelated genes, we next wondered how these target genes could display such different expression patterns despite being regulated by the same enhancer. We first investigated whether stage and/or tissue specific accessibility of E3 and of its target promoters could explain their different expression patterns. To explore this, we examined tissue-specific chromatin accessibility at the E3 enhancer and at the promoters of *twist*, *Fatp2*, *l(2)k09913*, and *CG9876*, using previously published single-cell ATAC-seq (scATACseq) data across embryonic development[29]. We particularly focused on mesodermal cell clusters, which should encompass *twist*, *l(2)k09913*, and *CG9876* expressing cells, and ectodermal, epidermis and tracheal cell clusters which encompass *Fatp2* expressing cells.

As expected from its complex activity in reporter assays, the E3 enhancer is accessible both in mesodermal clusters, and in ectodermal/epidermis/tracheal clusters (Fig. 4). We observed a strong accessibility peak in mesodermal clusters at 6-8h after egg lay (AEL; corresponding to stages 11-12), which becomes slightly weaker over time. In the epidermis and tracheal clusters, E3 is weakly but consistently accessible across all stages. As E3's four target genes display distinct expression patterns, we expected that their promoter regions might display distinct, cluster-specific accessibility profiles. This was indeed the case for the *Fatp2* promoter, which becomes accessible only in the epidermis and tracheal clusters at 10-12h AEL (corresponding to stages 13-15) but is inaccessible in mesodermal clusters (Fig. 4). In contrast, the *twist* promoter is broadly accessible across all clusters throughout embryogenesis, but accessibility along the gene body is restricted to mesodermal clusters (Fig. 4). This phenomenon is in line with promoter-proximal pausing of RNA polymerase II, a common feature of developmental genes[30,31]. Similarly, the *l(2)k09913* promoter remains accessible across all stages and clusters (Fig. 4). Finally, the *CG9876* promoter displays an intermediate behavior, being accessible exclusively in mesodermal clusters across all stages, including prior to the gene's activation (Fig. 4).

Together, these observations indicate that, except for *Fatp2*, enhancer or promoter accessibility does not precisely match gene expression, suggesting that chromatin accessibility alone cannot explain the tissue-specific activation of E3 target genes.

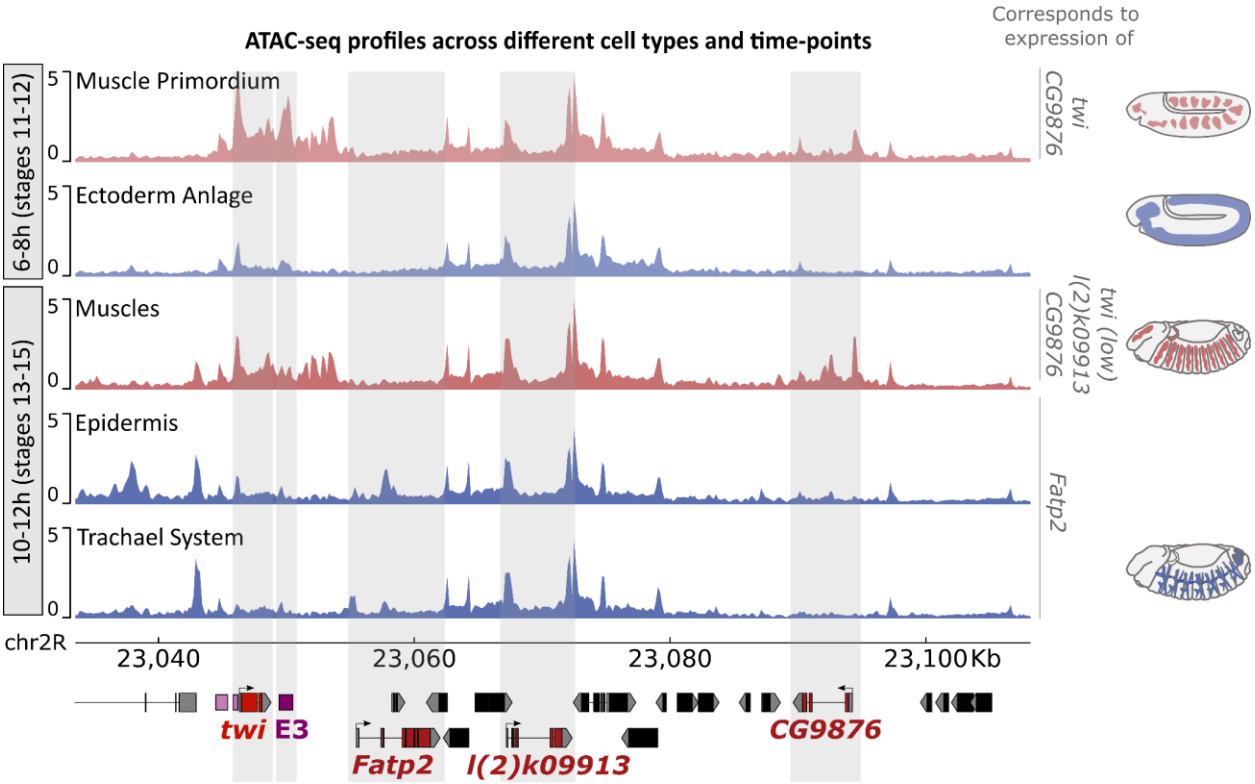


Figure 4. Stage- and tissue-specific chromatin accessibility at the E3 enhancer and its target genes. Pseudo-bulk single-cell ATAC-seq data from the muscle primordium (light red), ectoderm anlage (light blue), somatic and visceral muscle (pooled into muscles; red), epidermis (blue), and tracheal system (blue) at 6-8h and 10-12h after egg lay (AEL)[30]. Vertical grey bars indicate regions of interest. The genes expressed in each cell type at each time point are indicated on the side together with an embryo sketch highlighting the cell type of interest.

E3 forms stable permissive interactions with most of its target genes

Given the broad activity of E3 and the restricted spatio-temporal expression of its target genes, we expected that stage- and tissue-specific gene expression might be mediated by the formation of tissue- and/or stage-specific chromatin loops between the E3 enhancer and its target genes. Enhancer-mediated transcription often involves some degree of physical proximity between enhancers and their target promoters. However, such proximity does not always confer transcriptional activity[32–34]. During early stages of development, many enhancer–promoter interactions appear to be "permissive", i.e., are present across different cell types regardless of gene activity. By contrast, "instructive" interactions, which more directly correlate with gene expression, tend to emerge during terminal differentiation[35].

To determine whether E3-mediated gene expression depends on such instructive looping interactions, we performed 4C-seq (chromosome conformation capture sequencing) at different stages of *Drosophila* embryogenesis encompassing the expression of *twist* (2-5 hours AEL), *CG9876* (5-8 hours AEL), and *Fatp2*/*l(2)k09913* (8-11 hours AEL) (Fig. 5A-B, Supplementary Fig. 7). To comprehensively assess the 3D organization around the E3 locus,

we used three different viewpoints: one upstream of the *twist* promoter, one just upstream of the E3 enhancer, and one within the *Fatp2* promoter.

Our 4C-seq data revealed strong interactions between E3 enhancer and the *twist*, *Fatp2*, and *l(2)k09913* loci at all time points (Fig. 5A-B light grey bars, Supplementary Fig. 7), consistent with a permissive chromatin architecture. The interaction between E3 and *Fatp2* further supports that E3 is indeed regulating the expression of *Fatp2* despite its continued expression upon the deletion of E3. In contrast, E3 shows a weaker interaction with *CG9876* at 2–5 hours compared to 8-11 hours (Fig. 5B light blue bar), suggesting a more dynamic and context-specific interaction. This finding was supported by previously published tissue-specific Capture-C data[35], which showed significant contact between the *twist* and *CG9876* loci at 10-12 hours AEL in muscle cells, but not at 2-3 hours AEL in whole embryos or at 10-12 hours AEL in developing neurons (Fig. 5C). To further investigate the spatio-temporal dynamics of this interaction, we performed 3D DNA FISH across developmental stages. At stage 11, we observed a marked decrease in the distance between E3 and *CG9876* compared to stage 5, coinciding with the onset of *CG9876* expression (Fig. 5D).

Overall, our data suggest that E3, *twist*, *Fatp2*, and *l(2)k09913* remain in close spatial proximity throughout embryogenesis, including in cells where these genes are not expressed. This is consistent with the formation of a stable chromatin hub. In contrast, E3 forms stage and tissue-specific interactions with *CG9876* which correlate with the expression of the gene. While this may explain how CG9876 expression is regulated by E3, the mechanism that confers tissue specificity to the other E3 targets - *twist*, *Fatp2*, and *l(2)k09913* - remains to be elucidated.

Promoters impose tissue-specific constraints on pleiotropic enhancer activity

Through the use of transgenic reporter lines, we showed that E3 is an enhancer characterized by a broad and pleiotropic activity, being capable of activating transcription in mesodermal cells, including *twist* expressing cells and garland cells, as well in ectodermal cells such as hindgut and posterior spiracle cells. Despite this broad activity, each target of E3 only accesses partially the information contained with the enhancer. As tissue- and stage-specific expression of E3 target genes cannot be explained by the formation of instructive chromatin interactions, we tested whether the promoter sequence itself might encode responsiveness to E3. To this end, we generated a new series of E3 reporters in which we replaced the minimal *hsp70* promoter with the minimal promoters of *twist*, *Fatp2*, and *CG9876*. These minimal promoters were designed to encompass both the core promoter region (−50 to +50 bp from the TSS), the full 5′ untranslated region (5′UTR; 100–250 bp depending on the gene), and an upstream proximal region (50–200 bp; Supplementary Fig. 8A). The upstream boundary was defined to exclude any potential proximal enhancer elements. We confirmed that none of these minimal promoters had intrinsic enhancer activity when tested in reporter assays without E3, except for an ectopic expression in the salivary glands probably caused by an artificial TFBS created upstream of the promoter constructs (Supplementary Fig. 8B-D).

As previously shown, pairing E3 with the *hsp70* promoter results in widespread reporter expression across tissues and germ layers (Fig. 1B-E, Fig.6A, Supplementary Fig. 1). In contrast, combining E3 with the minimal promoters of its target genes revealed striking tissue specificity (Fig. 6B-D). The *twist* minimal promoter drove expression in the trunk mesoderm specifically at stages 11–12, consistent with the gene's endogenous activation window (Fig. 6C). Although the reporter signal extended more broadly than the gene's native expression, its confinement to the correct germ layer and developmental window suggests that the *CG9876* promoter provides spatial and temporal regulatory inputs. Finally, the *Fatp2* minimal promoter directed expression only in the posterior spiracles, consistent with its native

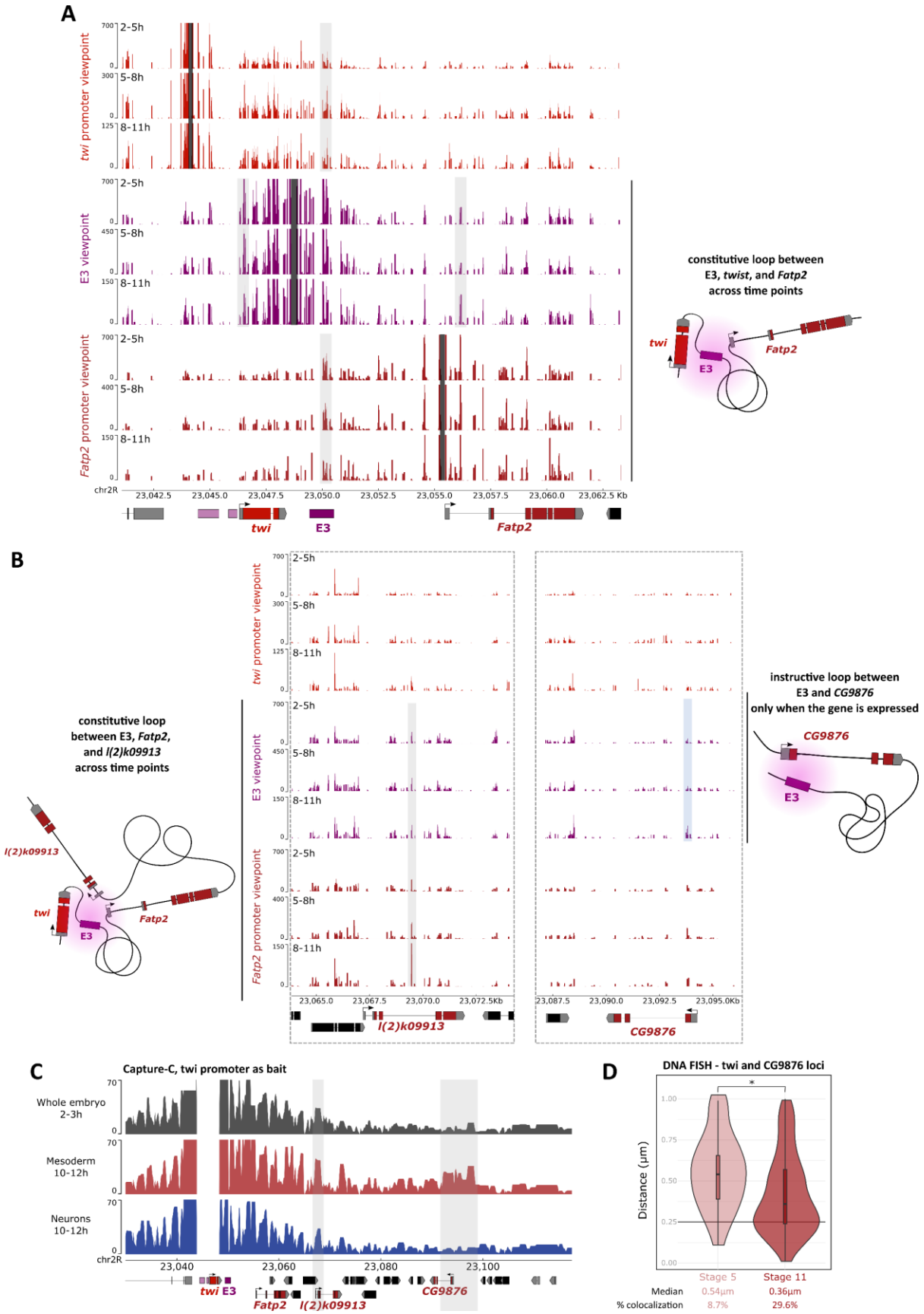

A
twi promoter viewpoint
E3 viewpoint
Fatp2 promoter viewpoint
2-5h
5-8h
8-11h
chr2R
23,042.5
23,045.0
23,047.5
23,050.0
23,052.5
23,055.0
23,057.5
23,060.0
23,062.5 Kb
twi
E3
Fatp2
constitutive loop between
E3, twist, and Fatp2
across time points
B
constitutive loop
between E3, Fatp2,
and l(2)k09913
across time points
l(2)k09913
instructive loop between
E3 and CG9876
only when the gene is expressed
CG9876
23,065.0
23,067.5
23,070.0
23,072.5Kb
23,087.5
23,090.0
23,092.5
23,095.0Kb
C
Capture-C, twi promoter as bait
Whole embryo
2-3h
Mesoderm
10-12h
Neurons
10-12h
23,040
23,060
23,080
23,100
twi E3
Fatp2
l(2)k09913
D
DNA FISH - twi and CG9876 loci
Distance (µm)
Stage 5
Stage 11
Median
0.54µm
0.36µm
% colocalization
8.7%
29.6%

Figure 5. Chromatin interactions between the E3 enhancer and its target genes.
(A-B) 4C-seq interaction maps using the *twist* (*twi*) promoter, the E3 enhancer and the *Fatp2* promoter as viewpoints in wild-type embryos at 2-5h, 5-8h and 8-11h AEL. Vertical black bars mark viewpoint locations; vertical grey and blue bars indicate interactions of interest at the E3 enhancer (A), *Fatp2* (A), *l(2)k09913* (B) and *CG9876* (B) genes. One representative experiment is shown. A schematics summary of the results is shown on the side.
(C) Normalized Capture-C counts with the *twi* promoter as bait in 2-3h AEL whole embryos, 6-8h and 10-12h AEL mesodermal cells (sorted for Mef2), and 10-12h AEL neuronal cells (sorted for elav)[35]. Vertical grey bars indicate interactions of interest at the *l(2)k09913* and *CG9876* promoter.
(D) Violin plots representing 3D DNA FISH distances between a probe located at the *twist* locus and the *CG9876* promoter in mesodermal cells of stage 5 and stage 11 embryos. At least 25 nuclei from two biologically independent embryos were analyzed. A non-parametric two-sample Kolmogorov-Smirnov test was used to assess the significant difference between DNA FISH distributions ($p = 0.03$, * indicates $p \leq 0.05$). Boxplots within the violin plots show median, edges are 25th and 75th percentiles, whiskers extend to non-outlier data points. The median distance and percentage of colocalization (defined as percentage of probe pairs with a distance ≤ 0.25µm) is indicated below each condition.

expression pattern (Fig. 6D, Supplementary Fig. 9).

Taken together, these findings demonstrate that, despite sharing the same enhancer input, each promoter is characterized by a distinct transcriptional output. This supports a model in which pleiotropic enhancers drive different tissue- and stage-specific expression of target genes through selective restriction by their target promoters. This restriction is likely mediated by sequence features in the proximal promoter region. In this framework, enhancer–promoter interfaces encode the regulatory logic required to interpret complex regulatory signals and convert them into precise spatial and temporal patterns. This mechanism enables functional specificity without the need for physical segregation through chromatin interactions.

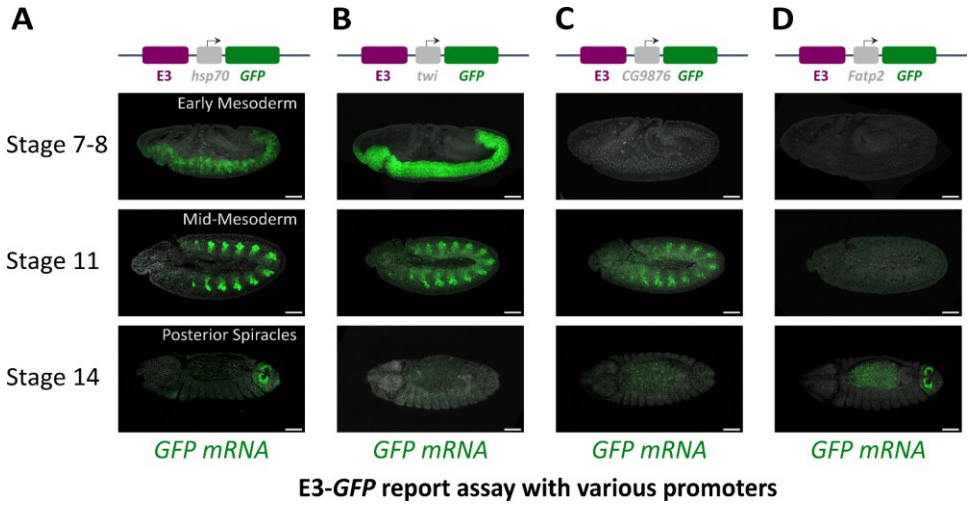


Figure 6. Tissue-specific activity of E3 is restricted by the proximal promoter of its target genes.
HCR RNA *in situ* hybridization of *GFP* (green) mRNA in E3-*hsp70-GFP* (A), E3-*twi-GFP* (B), E3-*CG9876-GFP* (C), and E3-*Fatp2(isoform1)-GFP* (D) embryos. Representative embryos are shown at stages 7-8, 11, and 14. A schematic of the reporter construct is shown on the top. Scale bar, 50 µm. The intensity of GFP signal was adjusted for each construct.

DISCUSSION

Pleiotropic enhancers are often defined as regulatory elements that drive the expression of a single pleiotropic gene[6]. Our work broadens this definition. We show that the *Drosophila* E3 enhancer, originally described as mesoderm-specific, regulates multiple genes expressed in different germ layers and at different embryonic stages. While enhancer sharing is well documented in both mammals and *Drosophila*, shared enhancers typically regulate the

expression of paralogous or functionally related genes[32,36–39]. In contrast, E3 constitutes an example of a shared pleiotropic enhancer that drives distinct spatiotemporal outputs for functionally diverse target genes. We propose that such broad form of enhancer pleiotropy is not exceptional, but rather a widespread and underappreciated feature of gene regulation.

Strikingly, despite E3's complex activity, each of its target gene is expressed in a precise and non-overlapping spatiotemporal window, which poses the question of how this specificity is achieved. We demonstrate that specificity does not arise from the enhancer itself or from the formation of tissue- and stage-specific enhancer-promoter interactions, but is instead mediated by the promoters of the target genes, which selectively restrict E3's regulatory input. Our data suggests that this restrictive activity is controlled by the proximal promoter region, a ~100 bp region upstream of the core promoter, which lacks enhancer activity on its own, but acts instead as a "gatekeeper". Similar promoter-proximal elements have been implicated in the context of co-regulated genes. In the *Drosophila* wing disc, a silencer within the *pdm2* promoter restricts input from an enhancer shared with its paralog *nub*[38]. The differential responses of the *Drosophila* genes *gooseberry* (*gsb*) and *gooseberry-neuro* (*gsbn*) to their cognate enhancers[40], or the differential expression of the *Drosophila brinker* (*brk*) gene under the control of two different enhancers[41] have also been attributed to the genes proximal promoters. Here, however, we show that proximal promoter regions can have a much more important and possibly more widespread role. They do not simply act as co-evolved switches that control the differential expression of two paralogous genes. Instead, they are required to process a complex cis-regulatory input that can be shared by several genes.

Our work, together with complementary work on the *Sox9* locus in a mouse cell culture system (see accompanying paper by Majchrzycka *et al*.), suggests that such gatekeeping elements are likely not specific to the E3 enhancer, but could be a widespread regulatory mechanism across various genes and species. More generally, promoter-proximal elements may constitute a critical mechanism for ensuring enhancer–promoter specificity. Their action explains why neighbouring genes exposed to the same enhancer are not expressed identically. By filtering enhancer input, promoter-proximal elements provide an additional layer of regulation that sharpens transcriptional output and ensures precision in gene expression.

The division of labour between enhancers and promoters is part of a broader paradigm shift in our understanding of gene regulation, from individual regulatory elements contributing additively to gene expression, to combinatorial interactions between multiple elements[42]. Together with "facilitators"[43], "range extenders"[44], and "tethering elements"[45], promoter-proximal gatekeepers act as noncanonical regulatory sequences that do not activate transcription on their own but help enhancers function correctly. These sequences may all belong to the same emerging class of regulatory elements: sequences that cooperate with enhancers to ensure specificity, timing, and precision of gene expression.

Promoters are thus not a passive recipient of enhancer signals, but an active interpreter, defining when and where a complex enhancer's inputs are converted into productive transcription. This additional layer of regulation may help explain how enhancers achieve flexibility: a pleiotropic enhancer like E3 can be reused across multiple developmental contexts because promoter proximal elements constrain its output to ensure fidelity. Such mechanisms would allow new regulatory connections to evolve without widespread misexpression, providing a substrate for both developmental innovation and evolutionary change.

More broadly, our findings support a view of the genome as a *cis*-regulatory network in which transcription specificity emerges not from single regulatory elements acting in isolation, but from combinatorial interactions between various types of regulatory elements. Understanding these interactions will be key to deciphering how gene expression programs are built, maintained,

and modified across development, disease and evolution. In this context, promoter-proximal gatekeepers may help explain why structural rearrangements observed in cancers and genetic disorders differentially impact neighbouring genes, as responsiveness to the same enhancer can vary depending on promoter identity.

METHODS

Plasmid construction and transgenesis

All plasmids were assembled using standard molecular cloning techniques, employing either restriction enzyme digestion and ligation with T4 DNA ligase (New England Biolabs), or the NEBuilder HiFi DNA Assembly Kit (New England Biolabs). All constructs were verified by Sanger sequencing. A comprehensive list of all primer sequences is provided in Table 1.

Reporter constructs were integrated into the *Drosophila* genome via ΦC31-mediated site-specific recombination (ref), using the *nos-ϕC31\int.NLS; attP40* fly line[46]. Transgenic lines were subsequently balanced, and homozygous stocks were established for downstream analyses. All fly lines used in this study are listed in Table 2.

Unless indicated otherwise, all experiments were conducted using the *yw* (y^1 w^{1118}) strain (BDSC_6598) as the wild-type background, except for all the reporter assays where wild-type controls refer to the E3-*hsp70-GFP* line. Flies were reared on standard medium at 25°C under controlled conditions.

<u>E3 mutants:</u>

Twelve distinct sub-fragments of the E3 enhancer (chr2R: 23,049,440–23,050,529) were designed based on the multiple alignment and conservation track across 124 insect genomes (https://genome.ucsc.edu/cgi-bin/hgTrackUi?db=dm6&g=cons124way) and on DNAseq hypersensitivity data[18]. These fragments were amplified by PCR using the Q5 High-Fidelity DNA Polymerase kit (NEB E0555), and inserted by SphI-mediated digestion into the pBID vector backbone (Addgene #35190) upstream of the *hsp70* minimal promoter and of the codon optimized *mGFPmut2* reporter gene[47] (pBID-*hsp70-mGFPmut2*).

In addition to these sub-fragments, we generated more targeted mutations within the E3 enhancer. Transcription factor binding motifs on the E3 enhancer sequence were identified using JASPAR 2024's CORE PWM collection for *Drosophila melanogaster* and its Scan tool (relative profile score threshold = 80%)[48]. Two Twist binding sites (16 and 21 bp in length; chr2R: 23,049,883-23,049,899 and chr2R: 23,049,987-23,050,008), identified by ChIP-Nexus analysis[49], were deleted using overlap extension PCR[50]. The mutated enhancer fragment was subsequently cloned into the SphI-digested pBID-*hsp70-mGFPmut2* vector. Another 111-bp region (chr2R: 23,049,639-23,049,751), containing transcription factor motifs associated with spiracle and garland cell specificity, was also deleted. This mutant E3 fragment was cloned into the pBID-E3-*hsp70-mGFPmut2* vector following digestion with XhoI and NsiI-HF restriction enzymes.

<u>Promoter reporters:</u>

Minimal promoters for *Fatp2*, *twist*, *CG9876*, and *l(2)k09913* were defined using annotations from the Ensembl genome browser (*Drosophila melanogaster* BDGP6 genome assembly, Ensembl release 109). For each gene, we identified the cDNA corresponding to the primary transcript and mapped the 5′ untranslated region (5′ UTR). Promoters were defined as the sequence immediately upstream of the annotated transcription start site (TSS), including the full 5′ UTR and an upstream region ranging from 62 to 128 bp in length (Supplementary Fig. 7A). To refine TSS positions and confirm promoter isoform usage, we incorporated data from published PRO-cap datasets[51].

Promoter sequences were synthesized by Thermo Fisher Scientific (Table 1), flanked by SphI and XbaI restriction sites. Each promoter was cloned into the SphI/XbaI-digested pBID-*hsp70-mGFPmut2* vector, replacing the *hsp70* minimal promoter. Control constructs lacking the E3 enhancer were generated in parallel to assess the basal activity of each promoter in isolation.

Identification of E3 candidate target genes

To identify candidate genes potentially regulated by the E3 enhancer, we scanned a genomic region spanning 100 kb upstream and downstream of the E3 locus. Given the compact genome of *Drosophila melanogaster* and previous observations that enhancers often regulate nearby genes, sometimes skipping one or more intervening genes[52], we extracted gene annotations from this region using the Ensembl BioMart tool[53,54].

Gene expression data were obtained from the BDGP *in situ* hybridization database[25,26], filtering for expression in mesodermal tissues, garland cells, and posterior spiracles across multiple developmental stages. Gene identifiers were standardized to FlyBase FBgn IDs, and candidate genes were defined as those present within the E3 genomic interval and expressed in at least one of the tissues of interest. For *Fatp2*, where BDGP data was unavailable, expression information was retrieved from the FlyFISH database[55,56].

Embryo collections

Embryo collection was performed as previously described[16]. 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 45 min for stage-specific collections), dechorionated using 50% bleach, and washed alternately with water and PBS containing 0.1% Triton X-100. The embryos used for RNA and DNA FISH were covalently crosslinked in 4% formaldehyde for 25 min and stored at -20°C in methanol. The embryos used for 4C-seq were covalently crosslinked in 1.8% formaldehyde for 15 min, quenched with 2M Tris-HCl pH7.5 for 5 min, washed in cold PBS containing 0.1% Triton X-100, and stored at -80°C.

Hybridization chain reaction (HCR) RNA Fluorescent *in situ* hybridization (FISH) and HCR RNA-FISH combined with immunofluorescence (IF)

All probes were synthesized and provided by Molecular Instruments (Table 3). HCR *in situ* hybridization was performed according to the manufacturer's protocol.

To visualize gene expression in the *twi*$^{\Delta E3}$*/CyO-hb-lacZ* fly line, we carried out HCR RNA FISH combined with immunofluorescence (IF) following the manufacturer's guidelines. A mouse anti-β-Galactosidase primary antibody (Promega, Ref: Z3781) was used at a 1:100 dilution. A donkey anti-mouse secondary antibody conjugated to the B4 hairpin amplifier was obtained from Molecular Instruments. Homozygous *twi*$^{\Delta E3}$ embryos were identified by the absence of β-Galactosidase staining.

Embryos were mounted on glass slides using ProLong Gold antifade mounting medium with DAPI (Invitrogen) and imaged with a Leica SP8 confocal microscope. All images were acquired using a 20× glycerol objective (or 40x glycerol for the zoom ins) under consistent settings (laser power, detector gain, and pinhole size) within each experimental batch. For comparative analyses between wild-type and mutant E3 reporter lines samples were imaged using identical acquisition parameters to ensure comparability of signal intensity and spatial resolution. Z-stacks were collected at 1 µm or 1.5 µm intervals and processed using Fiji[57]. When necessary, minor adjustments to brightness and contrast were applied uniformly across samples for clarity.

Stage-specific 4C-sequencing in *Drosophila* embryos

Nuclei dissociation

About 5000 fixed and frozen embryos from the *yw* fly line were resuspended in 300 µl of ice-cold cell lysis buffer (10mM Tris-HCl, pH8; 10 mM NaCl; 0.2% IGEPAL, protease inhibitors) in an Eppendorf tube. The embryos were grinded using an electric grinder (BIOSPEC) for about 90 seconds until the embryos were completely lysed. After a 15 minutes incubation on ice, the tubes were centrifuged for 5 minutes at 3000xg at 4°C, and the pellet was re-suspended in 500 µl of ice-cold cell lysis buffer.

4C template preparation

4C-seq was performed as previously described[16]. Briefly, nuclei were digested using MboI and

NlaIII (New England Biolabs) as the first and second restriction enzymes, respectively. 4C templates were amplified from 320 ng of DNA using the following primers:

- E3_FW: CTGTATCTGTTCGCTTCTTCCCA E3_RV: CTTTTGGGGGCCTGGAATCG, corresponding to a viewpoint located upstream for the E3 enhancer (chr2R:23,048,600-23,048,866)
- Fatp2_FW: CAATTCTGCTGTGTTTGCACGT Fatp2_RV: GCTCCGGCAACAACGGAT, corresponding to a viewpoint located upstream of the *Fatp2* promoter (chr2R:23,055,259-23,055,469)
- twi_FW: TACGTGCACCAAAAGTTTCTT twi_RV: AAAATGGTCGTCAAAGCGC, corresponding to a viewpoint located upstream for the twist promoter (chr2R:23,044,043–23,044,500)

To ensure sufficient base diversity at the start of each read following multiplexing, an additional 1–5 nucleotide 'shift' sequence was included at the 5' end of each primer. PCR products were then purified using SPRIselect magnetic beads (Beckman Coulter), and 100 ng of purified DNA per sample was used for library preparation. Libraries were constructed using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs), with indexing performed using the NEBNext Multiplex Oligos kit (New England Biolabs) to enable sample multiplexing. In total, 18 libraries were prepared, representing two independent biological replicates per condition. Sequencing was carried out on an Illumina MiSeq sequencer at the IGFL sequencing facility, using 75-bp paired-end reads, resulting in approximately 1 million reads per library.

<u>4C data analysis</u>

4C data analysis was performed as described previously[16]. Briefly, data quality check was performed using the FastQC software (http://www.bioinformatics.babraham.ac. uk/ projects/ fastqc). Adapter and primer sequences were trimmed using TrimGalore (https://www.bioinformatics.babraham.ac. uk/ projects/ trim _ galore/) and Cutadapt version 2.8[58]. The reads were then aligned to the dm6 reference genome using bowtie version 1.3.1[59]. Library read counts were normalized using the reads-per-million (RPM) method and converted into coverage bedgraph files.

Next-Generation Sequencing data visualization

All genome browser tracks were visualized using Integrative Genomics Viewer 2.12.2[60,61] and plotted using pyGenomeTracks (v3.9)[62,63]. When required, datasets were converted from dm3 to dm6 using CrossMap 0.7.0[64]. Bigwig datasets were merged for the scATAC-Seq clusters using the bigwigAverage tool of deeptools 3.5.5[65]. Visualization of the 4C-seq data was performed with pyGenomeTracks (v3.9).

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

3D DNA FISH was performed as previously described[16] in embryos of the *yw* fly line. Two probe sets were designed, mapping to regions of genomic DNA overlapping the *twi* locus (including the promoter and the E3 enhancer) and the *CG9876* promoter:

- chr2R:23040112-23051659 for the *twi* locus
- chr2R:23090690-23095383 for the *CG9876* promoter

Each probe set consisted of either four or six PCR fragments ranging from 900bp to 1.5 kb in length, labeled using the FISH Tag DNA Multicolor Kit (Life Technologies), with Alexa Fluor 647 used for the *twi* locus and Alexa Fluor 594 for the *CG9876* promoter. Mesodermal cells were visualized using an anti-Twist antibody. Embryos were mounted in ProLong Gold antifade reagent with DAPI (Life Technologies) and imaged using an Abberior confocal microscope equipped with a 40x objective. Multiple Z-stacks were collected for each embryo, with a section thickness of approximately 10 µm. At least 25 nuclei from two biologically independent embryos were analyzed. Distances between FISH signal centers were quantified using Imaris software (Bitplane). Statistical comparisons of distance distributions between samples were performed using a non-parametric two-sample Kolmogorov–Smirnov test. Signals were

considered co-localized if their centers were separated by less than 0.25 µm.

DATA AVAILABILITY

All raw 4C data were submitted to ArrayExpress (https: // www.ebi.ac.uk/ arrayexpress/ browse.html) under accession number E-MTAB-15531. The following publicly available databases and datasets were used: FlyBase r6.40 (https:// flybase.org/) using the dm6 reference genome; single-cell ATAC (GSE190149), Capture-C (E-MTAB-13271), DNase-seq (SRA:SRP002474), ChIP-seq for H3K27ac (GSE15292), Hi-C (GSE103625).

ACKNOWLEDGMENTS

We are very grateful to Daniel Ibrahim, Michalis Averof, Francois Leulier, Arnaud Krebs, and Guillaume Junion for their comments on the manuscript. We thank all members of the Ghavi-Helm lab for discussions and comments on the manuscript. This work was technically supported by IGFL sequencing facility (PSI), the IGFL microscopy facility, the PLATIM imaging facility, the Arthro-tools facility and Protein Sciences Facility of the Lyon SFR Biosciences (UAR3444 / US8. This work was also supported by the EquipEx+ Spatial-Cell-ID under the "Investissements d'avenir" program (ANR-21-ESRE-00016). This work was financially supported by an FRM starting grant (AJE20161236686), an ERC starting grant Enhancer3D (759708), and an ANR PRC grant DevLoop (ANR-24-CE12-3117) to Y.G-H., by the Fondation ARC pour la recherche sur le cancer (ARCDOC42024010007733 doctoral fellowship to M.M and PDF20171206672 postdoctoral fellowship to C.M.), an FRM doctoral fellowship (ECO202206015555 and FDT202504020230) to D.B.

AUTHOR INFORMATION

Affiliations

Institut de Génomique Fonctionnelle de Lyon, Univ Lyon, CNRS UMR 5242, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, 46 allée d'Italie F-69364 Lyon, France

Contributions

Y.G-H. conceived and supervised the study. P.B.P helped co-supervise the study. M.M. performed all the imaging experiments together with two master-level students (M.L. and L.C.). D.B. performed the 4C-seq and DNA FISH experiments together with C.A.A. M.M., M.L, C.M., H.T., D.L., C.C-R., J.M, and S.V. generated the transgenic lines. All of the authors discussed the results and implications and commented on the manuscript at all stages. Y.G-H., M.M., D.B., and C.M. acquired funding.

Corresponding author

Correspondence and requests for materials should be addressed to Yad Ghavi-Helm (yad.ghavi-helm@ens-lyon.fr)

ETHICS INTERESTS

Competing interests

The authors declare no competing financial interests.

REFERENCES

1. Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science *361*, 1341–1345. https://doi.org/10.1126/science.aau0320.
2. Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T., and Jaynes, J.B. (1999). Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients*. Development *126*, 2527–2538. https://doi.org/10.1242/dev.126.11.2527.
3. Small, S., Arnosti, D.N., and Levine, M. (1993). Spacing ensures autonomous expression of different stripe enhancers in the even-skipped promoter. Development *119*, 767–772. https://doi.org/10.1242/dev.119.3.767.
4. Small, S., Blair, A., and Levine, M. (1996). Regulation of Two Pair-Rule Stripes by a Single Enhancer in the *Drosophila* Embryo. Developmental Biology *175*, 314–324. https://doi.org/10.1006/dbio.1996.0117.

5. Small, S., Blair, A., and Levine, M. (1992). Regulation of even-skipped stripe 2 in the Drosophila embryo. EMBO J *11*, 4047–4057. https://doi.org/10.1002/j.1460-2075.1992.tb05498.x.
6. Sabarís, G., Laiker, I., Preger-Ben Noon, E., and Frankel, N. (2019). Actors with Multiple Roles: Pleiotropic Enhancers and the Paradigm of Enhancer Modularity. Trends in Genetics *35*, 423–433. https://doi.org/10.1016/j.tig.2019.03.006.
7. Infante, C.R., Mihala, A.G., Park, S., Wang, J.S., Johnson, K.K., Lauderdale, J.D., and Menke, D.B. (2015). Shared Enhancer Activity in the Limbs and Phallus and Functional Divergence of a Limb-Genital *cis*-Regulatory Element in Snakes. Developmental Cell *35*, 107–119. https://doi.org/10.1016/j.devcel.2015.09.003.
8. Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science *346*, 1004–1006. https://doi.org/10.1126/science.1257493.
9. Singh, D., and Yi, S.V. (2021). Enhancer Pleiotropy, Gene Expression, and the Architecture of Human Enhancer–Gene Interactions. Molecular Biology and Evolution *38*, 3898–3909. https://doi.org/10.1093/molbev/msab085.
10. Preger-Ben Noon, E., Sabarís, G., Ortiz, D.M., Sager, J., Liebowitz, A., Stern, D.L., and Frankel, N. (2018). Comprehensive Analysis of a *cis*-Regulatory Region Reveals Pleiotropy in Enhancer Function. Cell Reports *22*, 3021–3031. https://doi.org/10.1016/j.celrep.2018.02.073.
11. Laiker, I., and Frankel, N. (2022). Pleiotropic Enhancers are Ubiquitous Regulatory Elements in the Human Genome. Genome Biology and Evolution *14*, evac071. https://doi.org/10.1093/gbe/evac071.
12. Fuqua, T., Jordan, J., van Breugel, M.E., Halavatyi, A., Tischer, C., Polidoro, P., Abe, N., Tsai, A., Mann, R.S., Stern, D.L., et al. (2020). Dense and pleiotropic regulatory information in a developmental enhancer. Nature *587*, 235–239. https://doi.org/10.1038/s41586-020-2816-5.
13. Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and Pervasive Changes in Genome-wide Enhancer Usage during Mammalian Development. Cell *155*, 1521–1531. https://doi.org/10.1016/j.cell.2013.11.033.
14. Huang, Y., Verstegen, M.J.A.M., Tjalsma, S.J.D., Krijger, P.H.L., Gupta, K., Park, M., Boettiger, A., and Laat, W. de (2025). Two unrelated distal genes activated by a shared enhancer benefit from localizing inside the same small topological domain. Genes Dev. *39*, 348–363. https://doi.org/10.1101/gad.352235.124.
15. Chang, T.H.-T., Primig, M., Hadchouel, J., Tajbakhsh, S., Rocancourt, D., Fernandez, A., Kappler, R., Scherthan, H., and Buckingham, M. (2004). An enhancer directs differential expression of the linked *Mrf4* and *Myf5* myogenic regulatory genes in the mouse. Developmental Biology *269*, 595–608. https://doi.org/10.1016/j.ydbio.2004.02.013.
16. Balasubramanian, D., Borges Pinto, P., Grasso, A., Vincent, S., Tarayre, H., Lajoignie, D., and Ghavi-Helm, Y. (2024). Enhancer–promoter interactions can form independently of genomic distance and be functional across TAD boundaries. Nucleic Acids Research *52*, 1702–1719. https://doi.org/10.1093/nar/gkad1183.
17. Erickson, P.A., Cleves, P.A., Ellis, N.A., Schwalbach, K.T., Hart, J.C., and Miller, C.T. (2015). A 190 base pair, TGF-β responsive tooth and fin enhancer is required for stickleback *Bmp6* expression. Developmental Biology *401*, 310–323. https://doi.org/10.1016/j.ydbio.2015.02.006.
18. 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 Biology *12*, R43. https://doi.org/10.1186/gb-2011-12-5-r43.

19. Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the Drosophila genome. Nature *471*, 527–531. https://doi.org/10.1038/nature09990.
20. Rowley, A.J., Square, T.A., and Miller, C.T. (2022). Site pleiotropy of a stickleback *Bmp6* enhancer. Developmental Biology *492*, 111–118. https://doi.org/10.1016/j.ydbio.2022.09.012.
21. González, A., and Paul, P. (2024). Pleiotropic expression quantitative trait loci are enriched in enhancers and transcription factor binding sites and impact more genes. Comput Struct Biotechnol J *23*, 4260–4270. https://doi.org/10.1016/j.csbj.2024.11.019.
22. Lovegrove, B., Simões, S., Rivas, M.L., Sotillos, S., Johnson, K., Knust, E., Jacinto, A., and Hombría, J.C.-G. (2006). Coordinated Control of Cell Adhesion, Polarity, and Cytoskeleton Underlies Hox-Induced Organogenesis in Drosophila. Current Biology *16*, 2206–2216. https://doi.org/10.1016/j.cub.2006.09.029.
23. Museridze, M., Ceolin, S., Mühling, B., Ramanathan, S., Barmina, O., Sekhar, P.S., and Gompel, N. (2024). Entangled and non-modular enhancer sequences producing independent spatial activities. Science Advances *10*, eadr9856. https://doi.org/10.1126/sciadv.adr9856.
24. 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. https://doi.org/10.1093/genetics/iyac035.
25. 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 Biology *14*, R140. https://doi.org/10.1186/gb-2013-14-12-r140.
26. Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2007). Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biology *8*, R145. https://doi.org/10.1186/gb-2007-8-7-r145.
27. 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. https://doi.org/10.1186/gb-2002-3-12-research0088.
28. Quintero-Cadena, P., and Sternberg, P.W. (2016). Enhancer Sharing Promotes Neighborhoods of Transcriptional Regulation Across Eukaryotes. G3 (Bethesda) *6*, 4167–4174. https://doi.org/10.1534/g3.116.036228.
29. 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. Molecular Cell *71*, 73-88.e5. https://doi.org/10.1016/j.molcel.2018.05.032.
30. Calderon, D., Blecher-Gonen, R., Huang, X., Secchia, S., Kentro, J., Daza, R.M., Martin, B., Dulja, A., Schaub, C., Trapnell, C., et al. (2022). The continuum of Drosophila embryonic development at single-cell resolution. Science *377*, eabn5800. https://doi.org/10.1126/science.abn5800.
31. Lagha, M., Bothma, J.P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D.S., Zeitlinger, J., et al. (2013). Paused Pol II Coordinates Tissue

Morphogenesis in the *Drosophila* Embryo. Cell *153*, 976–987. https://doi.org/10.1016/j.cell.2013.04.045.
32. Levine, M. (2011). Paused RNA Polymerase II as a Developmental Checkpoint. Cell *145*, 502–511. https://doi.org/10.1016/j.cell.2011.04.021.
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. https://doi.org/10.1038/nature13417.
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. Nature Genetics *53*, 487–499. https://doi.org/10.1038/s41588-021-00799-x.
35. 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. https://doi.org/10.1038/s41588-021-00816-z.
36. Pollex, T., Rabinowitz, A., Gambetta, M.C., Marco-Ferreres, R., Viales, R.R., Jankowski, A., Schaub, C., and Furlong, E.E.M. (2024). Enhancer–promoter interactions become more instructive in the transition from cell-fate specification to tissue differentiation. Nat Genet, 1–11. https://doi.org/10.1038/s41588-024-01678-x.
37. Pollex, T., Marco-Ferreres, R., Ciglar, L., Ghavi-Helm, Y., Rabinowitz, A., Viales, R.R., Schaub, C., Jankowski, A., Girardot, C., and Furlong, E.E.M. (2023). Chromatin gene-gene loops support the cross-regulation of genes with related function. Molecular Cell. https://doi.org/10.1016/j.molcel.2023.12.023.
38. Levo, M., Raimundo, J., Bing, X.Y., Sisco, Z., Batut, P.J., Ryabichko, S., Gregor, T., and Levine, M.S. (2022). Transcriptional coupling of distant regulatory genes in living embryos. Nature *605*, 754–760. https://doi.org/10.1038/s41586-022-04680-7.
39. Loker, R., and Mann, R.S. (2022). Divergent expression of paralogous genes by modification of shared enhancer activity through a promoter-proximal silencer. Current Biology *32*, 3545-3555.e4. https://doi.org/10.1016/j.cub.2022.06.069.
40. Heist, T., Fukaya, T., and Levine, M. (2019). Large distances separate coregulated genes in living Drosophila embryos. Proceedings of the National Academy of Sciences *116*, 15062–15067. https://doi.org/10.1073/pnas.1908962116.
41. Li, X., and Noll, M. (1994). Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the Drosophila embryo. EMBO J *13*, 400–406. https://doi.org/10.1002/j.1460-2075.1994.tb06274.x.
42. Dunipace, L., Saunders, A., Ashe, H.L., and Stathopoulos, A. (2013). Autoregulatory Feedback Controls Sequential Action of *cis*-Regulatory Modules at the *brinker* Locus. Developmental Cell *26*, 536–543. https://doi.org/10.1016/j.devcel.2013.08.010.
43. Thomas, H.F., Feng, S., Haslhofer, F., Huber, M., García Gallardo, M., Loubiere, V., Vanina, D., Pitasi, M., Stark, A., and Buecker, C. (2025). Enhancer cooperativity can compensate for loss of activity over large genomic distances. Molecular Cell *85*, 362-375.e9. https://doi.org/10.1016/j.molcel.2024.11.008.
44. Blayney, J.W., Francis, H., Rampasekova, A., Camellato, B., Mitchell, L., Stolper, R., Cornell, L., Babbs, C., Boeke, J.D., Higgs, D.R., et al. (2023). Super-enhancers include classical enhancers and facilitators to fully activate gene expression. Cell *186*, 5826-

5839.e18. https://doi.org/10.1016/j.cell.2023.11.030.
45. Bower, G., Hollingsworth, E.W., Jacinto, S.H., Alcantara, J.A., Clock, B., Cao, K., Liu, M., Dziulko, A., Alcaina-Caro, A., Xu, Q., et al. (2025). Range extender mediates long-distance enhancer activity. Nature *643*, 830–838. https://doi.org/10.1038/s41586-025-09221-6.
46. Calhoun, V.C., Stathopoulos, A., and Levine, M. (2002). Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. Proc Natl Acad Sci U S A *99*, 9243–9247. https://doi.org/10.1073/pnas.142291299.
47. 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. https://doi.org/10.1038/ng.101.
48. 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. https://doi.org/10.1038/nmeth.4509.
49. Rauluseviciute, I., Riudavets-Puig, R., Blanc-Mathieu, R., Castro-Mondragon, J.A., Ferenc, K., Kumar, V., Lemma, R.B., Lucas, J., Chèneby, J., Baranasic, D., et al. (2024). JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Res *52*, D174–D182. https://doi.org/10.1093/nar/gkad1059.
50. He, Q., Johnston, J., and Zeitlinger, J. (2015). ChIP-nexus enables improved detection of in vivo transcription factor binding footprints. Nat Biotechnol *33*, 395–401. https://doi.org/10.1038/nbt.3121.
51. Nelson, M.D., and Fitch, D.H.A. (2012). Overlap Extension PCR: An Efficient Method for Transgene Construction. In Molecular Methods for Evolutionary Genetics (Humana Press), pp. 459–470. https://doi.org/10.1007/978-1-61779-228-1_27.
52. Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I.E., Males, M., Viales, R.R., and Furlong, E.E.M. (2018). The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. Genes Dev. *32*, 42–57. https://doi.org/10.1101/gad.308619.117.
53. Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O., and Stark, A. (2015). Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. Nature *518*, 556–559. https://doi.org/10.1038/nature13994.
54. Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global Analysis of mRNA Localization Reveals a Prominent Role in Organizing Cellular Architecture and Function. Cell *131*, 174–187. https://doi.org/10.1016/j.cell.2007.08.003.
55. Wilk, R., Hu, J., Blotsky, D., and Krause, H.M. (2016). Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. Genes Dev. *30*, 594–609. https://doi.org/10.1101/gad.276931.115.
56. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.
57. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.
58. 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*, 1–10. https://doi.org/10.1186/gb-2009-10-3-r25.
59. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24–26. https://doi.org/10.1038/nbt.1754.

60. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178–192. https://doi.org/10.1093/bib/bbs017.
61. 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. Nature Communications *9*, 189. https://doi.org/10.1038/s41467-017-02525-w.
62. 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. https://doi.org/10.1093/bioinformatics/btaa692.
63. Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.-P., and Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics *30*, 1006–1007. https://doi.org/10.1093/bioinformatics/btt730.
64. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.