

Chromosome-level organization of the regulatory genome in the *Drosophila* nervous system

Giriram Mohana^{1*}, Julien Dorier^{2,3*}, Xiao Li^{4*}, Marion Mougnot¹, Rebecca C. Smith⁵, H el ena Malek¹, Marion Leleu^{2,3}, Daniel Rodriguez¹, Jenisha Khadka¹, Patrycja Rosa⁶, Pascal Cousin¹, Christian Iseli^{2,3}, Simon Restrepo⁷, Nicolas Guex^{2,3}, Brian D. McCabe⁵, Aleksander Jankowski^{6**}, Michael S. Levine^{4**}, Maria Cristina Gambetta^{1**}

1-Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

2-Bioinformatics Competence Center, University of Lausanne, 1015 Lausanne, Switzerland

3-Bioinformatics Competence Center, Swiss Federal Institute of Technology Lausanne, 1015 Lausanne, Switzerland

4-Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA

5-Brain Mind Institute, Swiss Federal Institute of Technology Lausanne, 1015 Lausanne, Switzerland

6-Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, 02-097 Warsaw, Poland

7-Arcoris bio AG, L ussirainstrasse 52, 6300 Zug, Switzerland

*These authors contributed equally.

**Corresponding authors: aleksander.jankowski@uw.edu.pl, msl2@princeton.edu,
mariacristina.gambetta@unil.ch

Lead contact: mariacristina.gambetta@unil.ch

Running title: Specific TAD interactions support megabase-scale gene regulation in the fly nervous system.

SUMMARY

Past studies have identified topologically associating domains (TADs) as basic units of genome organization. Here, we present evidence for a new level of genome folding, whereby distant TAD pairs megabases apart interact to form meta-domains. Within meta-domains, gene promoters and structural intergenic elements present in distant TADs are specifically paired. The associated genes encode neuronal determinants, including those engaged in axonal guidance and adhesion. These long-range associations occur in a large fraction of neurons but support transcription in only a subset of neurons. Meta-domains are formed by diverse transcription factors that are able to pair over long and flexible distances. We present evidence that two such factors, GAF and CTCF, play direct roles in this process. The relative simplicity of higher-order meta-domain interactions in *Drosophila*, compared with those previously described in mammals, allowed the demonstration that genomes can fold into highly specialized cell type-specific scaffolds that enable megabase-scale regulatory associations.

KEYWORDS

TAD, chromosomal loop, genome architecture, genome organization, gene regulation, transcription, nervous system, neuron, *Drosophila*

INTRODUCTION

Past studies have identified diverse genome folding mechanisms at different scales. Local folding into topologically associating domains (TADs) occurs when cohesin extrudes chromosomal loops until it is stalled by DNA-bound CTCF at domain boundaries^{1,2}, or when active and inactive contiguous chromatin regions physically segregate³. TADs can further associate into higher-order assemblies⁴⁻⁶. Beyond TADs, chromosomes organize into territories and more broadly into transcriptionally active and inactive compartments⁷⁻⁹. Transcription factors and machinery shape three-dimensional (3D) genome contacts within these frameworks¹⁰⁻¹³. A basic question arises regarding the role of 3D genome folding in facilitating specialized gene expression, as regulatory interactions are to a certain degree distance-dependent¹⁴⁻¹⁶.

Functional interactions between regulatory elements and gene promoters primarily occur within TADs, spanning distances of 10 to 100 kb in flies and megabases (Mb) in mice^{17,18}. TADs guide and constrain regulatory interactions, although the impact on TAD-internal gene regulation varies between loci^{15,19-24}. While TADs may be sufficient or even dispensable for most local regulation, long-range regulation may require additional 3D structural support. In flies, tethering elements bridge distal regulatory elements and target genes within the same TAD, facilitating rapid activation of certain developmental genes²⁵ and co-regulation of distant paralogs TAD²⁶. In mammals, loop extrusion brings promoters and regulatory elements into contact over long distances^{15,27,28}. Additionally, mammalian “tethering-like elements” bridge some enhancers to their target promoters within TADs for activation²³. These TAD-level features are often constitutive features of genomes.

Recent studies explored how 3D genome organization contributes to cell type-specific transcription programs in mammalian tissues, purified cells and cell culture models. Local variations in 3D genome folding were observed within TADs, at their boundaries, and in long-range contacts and compartments^{10,29-31}. In some striking cases, distant genes and regulatory elements co-bound by common transcription factors coalesce into specialized hubs^{10,11,32,33}. However, the redundancy of such assemblies challenges the understanding of how clustering may impact co-regulation. Studies in *Drosophila* embryos identified cell type-specific chromatin folding^{34,35} but also suggested that enhancer and promoter hubs are often constant across cell types and precede gene activity^{17,36-38}. These findings underscore the uncertainty regarding the extent to which cell type-specific 3D genome folding directly supports specialized transcriptional programs.

We investigated specialized genome folding in the *Drosophila* central nervous system (CNS) to determine its role in cell type-specific transcription. Using Micro-C³⁹, we compared genome folding in the CNS and wing imaginal disc at single-nucleosome resolution and discovered that specific pairs of TADs that are megabases apart interact to form meta-domains predominantly in the CNS. Within meta-domains, 58 high-frequency contacts (“meta-loops”) connect CNS gene promoters and intergenic accessible chromatin sites. Evolutionary analysis revealed that meta-loops are remarkably robust and specific. Hi-C experiments on neuroblasts, glia and neurons showed that meta-domains are most prominent in differentiated neurons. We disrupted meta-loops by deleting anchors or mutating transcription factors, GAGA factor (GAF, encoded by the *Trl* gene) and CTCF, resulting in decreased and delayed expression of neural genes at meta-loop anchors. Meta-loops are present in a large fraction of CNS cells but support gene transcription in only a subset. Our findings unveil a previously unrecognized level of genome folding and demonstrate that meta-loops facilitate the longest-range regulatory associations reported so far.

RESULTS

Meta-domains harbor loops involving genes expressed in the nervous system

Micro-C analysis of dissected central nervous systems (CNSs) or wing imaginal discs of *Drosophila melanogaster* third instar larvae revealed specific interactions between TADs separated by millions of base pairs in CNSs (Figs. 1A-1B and S1A, S1B, S2A, Table S1). These interacting TADs, referred to here as meta-domains, harboured one or several high-frequency interactions within them, referred to here as meta-loops (Figs. 1A-1B and S1A). Unlike traditional compartmental interactions that reflect segregation of transcriptionally active (A) and inactive (B) genomic compartments, meta-domains represent selective, long-range associations between specific pairs of TADs at unexpectedly high frequencies. In larval CNSs, the normalized Micro-C counts within meta-loops were orders of magnitude higher than background compartmental interactions (Fig. 1C). Single-cell assays for transposase accessible chromatin-sequencing (scATAC-seq) conducted on larval CNSs demonstrated that all loop anchors overlapped accessible chromatin peaks (Fig. 1A). Loop anchors also overlapped published DNase-seq peaks (DHSs) identified in mid-embryogenesis neurons but less frequently in muscle cells⁴⁰ (Fig. 1A, S2B). Together, these results suggest that transcription factors bind to meta-loop anchors in the CNS.

Using a custom algorithm for automatic meta-loop identification in larval CNS contact maps and manual validation, a curated list of 58 meta-loops was identified (Fig. S1A, see Methods). Published DNase-seq peaks from mid-embryogenesis neurons⁴⁰ were then assigned to meta-loop anchors (Fig. 1B). Meta-loops ranged in size between 403 kb to 22.7 Mb (Fig. 1D), spanning a chromosome arm in one case but always restricted to individual chromosome arms. In total, 79 identified meta-loop anchors interacted with each other to form 68 loops comprising 58 meta-loops and 10 intra-TAD loops (Fig. S1, Table S2). These 68 loops were present in 28 meta-domains, with most (68%) meta-domains containing clusters of two to six loops (Figs 1E, S1A). Among meta-loop anchors, 48% engaged in pairwise interactions, 34% paired with two possible anchors, and 18% potentially interacted with three or four others (Figs. S2C). Hi-C demonstrated concurrent three-way interactions between some meta-loop anchors (Figs. S2D-S2E).

We compared CNS meta-loops to loops described in *Drosophila* Kc167 cells⁴¹ or early embryos^{25,26}. Unlike meta-loops, previously described loops were detected in different cell types^{25,26}, orders of magnitude smaller (Fig. 1C), limited to individual TADs, and used different anchors (Fig. S2F) enriched for pioneer factor and Polycomb protein binding sites^{25,26,41} (Fig. S2G).

Meta-loop anchors were categorized as promoter anchors if the underlying DNase-seq peak was ± 200 bp from a gene transcription start site (TSS) or otherwise as intergenic anchors (Fig. 1F). Of all 68 loops involving meta-loop anchors, 9 connected two promoter anchors (P-P loops), 25 connected two intergenic anchors (I-I loops), and 34 connected a promoter and an intergenic anchor (I-P loops) (Fig. 1G). 6/9 (66%) P-P loops joined paralogous gene promoters (Fig. S2H). Genes at promoter anchors are more highly transcribed in adult brains than in other tissues (Fig. 1H) and start to be transcribed in differentiating embryonic nervous systems (Fig. S2I). Notably, these genes encode extracellular or transmembrane proteins involved in neuron guidance, including immunoglobulin receptor superfamily members, or ion channels specific to neuronal subtypes⁴². Gene ontology (GO) analysis demonstrated a significant enrichment of cell adhesion and axon guidance terms meta-looping genes (Fig. S2J). In contrast, genes with TSSs closest to intergenic anchors were not enriched for these GO terms (Fig. S2J) and were not specifically expressed in the brain (Fig. 1H). This suggests that if intergenic anchors are involved in transcriptional regulation, they may affect the promoter of a gene in a distant TAD to which they loop, rather than a local promoter in cis.

Meta-loop anchors specifically pair over flexible genomic distances in evolution

We investigated the impact of chromosomal rearrangements on meta-loops during evolution. Hi-C was performed on the adult CNS of *Drosophila virilis*, with an evolutionary distance from *Drosophila melanogaster* comparable to that between humans and lizards⁴³ (Fig. S3A, Table S1). Using a similar procedure to call meta-loops as in *D. melanogaster*, 84 meta-loop anchors were discovered that formed 58 meta-loops and 6 intra-TAD loops (Table S3). We examined loop conservation between both species by searching for homologous loop anchor candidates (see Methods). Loop candidates were defined as all pairwise combinations of loop anchor candidates in the other species, for each observed loop. Finally, conserved loops were identified as loop candidates present in both Hi-C maps. Most (61 out of 79) *D. melanogaster* meta-loop anchors mapped uniquely

to corresponding *D. virilis* loop anchor candidates (Fig. S3B). This homology was not solely driven by coding sequence similarity (Fig. S3C). Overall, 37 loops involving meta-loop anchors were conserved in both species, accounting for 37 out of 68 (54%) loops in *D. melanogaster* and 37 out of 64 (58%) loops in *D. virilis* (Figs. 2A and S3D). Orthologs of genes closest to *D. melanogaster* meta-loop anchors were often close to homologous meta-loop anchors observed in *D. virilis* (Figs. S3E-S3F).

As illustrated for chromosome 3L, conserved loops persisted despite chromosomal breakpoints and inversions (Fig. 2B). The lengths of the 37 conserved loops varied between *D. melanogaster* and *D. virilis* (Fig. 2C), and loop conservation was not limited by loop length (Fig. S3G), indicating that loop anchors pair over flexible distances. Loop anchor orientation did not impact loop formation, even at shorter looping distances (Figs. 2B-C). Moreover, pairs of loop anchors faithfully looped to each other and rarely to other anchors, highlighting the highly specific pairwise interactions between loop anchors through evolution (Fig. S3H). This suggests that dedicated transcription factors bound to loop anchors mediate distinct loops, a hypothesis explored in the last two results sections.

In an exceptional case, a meta-loop observed in one species corresponded to an intra-TAD loop in the other species. Whereas the promoter of the *sidestep IV* (*side-IV*) gene involved in axon guidance⁴⁴ loops to an intergenic anchor 12 Mb away in *D. virilis* (meta-loop Dvir_L48 in Fig. 2D), it loops to a homologous intergenic anchor 89 kb away within the same TAD in *D. melanogaster* (Fig. 2D). A nearby meta-loop (Dvir_L49) connecting two intergenic anchors in *D. virilis* was not conserved in *D. melanogaster*. This unique example suggests that there might not be strong selective pressure to maintain *side-IV* and its intergenic anchor in separate TADs. Nonetheless, *side-IV* promoter in *D. melanogaster* remains distant (89 kb) from its intergenic anchor within the same TAD.

Despite the conservation of many loops involving meta-loop anchors, the reason why some loop anchor candidates failed to loop in *D. melanogaster* or *D. virilis* remains unclear. The inability of loop anchor candidates to form a loop could not be attributed to their chromosomal separation between the two species, consistent with previous observations that most chromosomal rearrangements between *D. melanogaster* and *D. virilis* occur within each chromosomal arm⁴⁵.

Meta-loops strengthen during neuronal differentiation

We asked when and where meta-loops emerge in the *D. melanogaster* CNS. Meta-loops were observed in both larval and adult CNS Micro-C maps, with certain meta-loops exhibiting greater strength in adults (Fig. S1B). Our scATAC-seq dataset in larval CNSs (Fig. 1A) and a published scATAC-seq dataset in adult CNSs⁴⁶ revealed that many meta-loop anchors were more accessible in neurons than in glia or neuroblasts (neural stem cells) (Figs. S4A-S4B). A published DNase-seq dataset spanning neuronal development in embryogenesis⁴⁰ indicated that loop anchor chromatin gradually gained accessibility during neuronal differentiation (Fig. S4C). Collectively, these chromatin accessibility analyses suggested that certain meta-loops may be stronger in or specific to differentiated neurons.

To test this hypothesis, we performed Hi-C on fluorescence activated cell sorting (FACS)-purified neurons, glia and neuroblasts from embryos at three developmental stages: neuroblast proliferation (6-8 hours of development), neural patterning and axon outgrowth (10-12 hours), and neuronal terminal differentiation (14-16 hours) (Table S1). These cell types were purified from fixed embryos using antibodies against lineage markers Elav (neurons), Repo (glia) and Wor (neuroblasts), which were labeled using Multimodal Universal Signal Enhancement (MUSE) technology (Fig. S5A). Chromosomes of undifferentiated cells, including neuroblasts from embryos of any stage and glia and neurons from 6-8 hour old embryos, displayed a Rab1 configuration where centromeres and telomeres cluster at opposite poles and chromosome arms are aligned (Fig. 3A row 1 and column 1). The Rab1 configuration was gradually replaced by organization into chromosome territories in neurons and glia from 14-16 hour old embryos (Fig. 3A column 3). This demonstrates significant genome reorganization during cell differentiation.

Closer data inspection revealed that some meta-loops identified in larval CNSs were present in all CNS cell types throughout embryogenesis, while other meta-loops specifically emerged in differentiated neurons (Figs. 3A-3B). Although several meta-loops were not exclusive to neurons, most exhibited maximal strength in differentiated neurons (Fig. 3B). This trend is exemplified by the 2.6 Mb P-P meta-loop (L35) involving the *Glutamate receptor IA* and *IB* (*GluRIA-GluRIB*) paralogs, which emerged in mature neurons near a pre-existing I-I loop within the same meta-domain (Fig. S5B). The formation of TAD boundaries coincided with the emergence

of some meta-loop anchors, consequently restricting the meta-domain borders (Figs. S5B-S5C). These findings collectively suggest that, particularly during neuronal differentiation, transcription factors occupy meta-loop anchors that exhibit increased chromatin accessibility and participate in meta-looping.

Meta-loop anchors are in enhanced proximity in a large fraction of neurons

To understand how meta-loops form and their functional significance, we examined two specific examples. One meta-loop connects the promoter of *beaten path IV* (*beat-IV*), involved in motor neuron axon guidance⁴⁴, to an intergenic anchor located 5.1 Mb away (Fig. 4A). The other connects the promoter of *no long nerve cord* (*nolo*), hypothesized to function in the extracellular matrix⁴⁷, to an intergenic anchor located 1.6 Mb away (Fig. 4B). Similar to the *GluRIA-GluRIB* meta-loop mentioned earlier, the *beat-IV* meta-loop (L54) also emerged in neurons of 14-16 hour old embryos near an earlier meta-loop (L53), coinciding with the reinforcement of TAD boundaries near meta-loop anchors (Fig. 4A). L53 meta-loop anchor chromatin was accessible earlier than L54 meta-loop anchor chromatin (Fig. 4A), mirroring the observations for *GluRIA-GluRIB* (Fig. S5B). In contrast, the meta-domain containing *nolo* features a single meta-loop (L22, Fig. 4B).

To examine the reproducibility of meta-domains across cells, physical distances between *beat-IV* or *nolo* meta-loop anchors were measured in whole-mount embryos by DNA-FISH (Figs. 4C and 4D). The proportion of nuclei with physically close *beat-IV* or *nolo* meta-loop anchors specifically increased in neuronal nuclei (immunolabeled with α -Elav) as embryos aged (Figs. 4C-4D). Meta-loop strengthening observed by Hi-C during neuronal maturation (Figs. 4A-4B) thus reflects an increasing proportion of nuclei in which *beat-IV* or *nolo* meta-loop anchors are physically close. The increased proximity of *beat-IV* and *nolo* meta-loop anchors began at a similar developmental stage, coinciding with transcriptional onset of these genes (Figs. 4C-4D). *beat-IV* and *nolo* meta-loop anchors were physically close in most nuclei of late-stage embryonic nerve cords, despite *beat-IV* and *nolo* being transcribed only in a subset of these cells (Figs. 4C and 4D). Overall, Hi-C and DNA-FISH results indicate that meta-domains are established throughout the embryonic nervous system at the onset of transcription of genes at meta-loop anchors, even though expression is restricted to a subset of cells.

To investigate whether meta-loops exist in nuclei independently of the expression of genes at promoter anchors, we compared our Micro-C data in larval CNSs and wing discs. Approximately one third of meta-loops were visible in both tissues, although certain shared meta-loops were stronger in larval CNSs compared to wing discs (Figs. S1A-S1B). Surprisingly, meta-loops involving a subset of promoter anchors of genes primarily expressed in the CNS were also observed in wing discs (Fig. S1, Table S2). The fact that the *nolo* meta-loop and meta-domain were comparable in both CNSs and wing discs (in which *nolo* is not expressed⁴⁸) indicates that this meta-loop can exist independently of *nolo* transcription.

Meta-domains are formed by meta-loop anchors and delimited by TAD boundaries

To test whether meta-loop anchors are required for meta-domain formation, we deleted 815 bp or 199 bp overlapping accessible chromatin regions at *nolo* or *beat-IV* intergenic anchors. Hi-C on *nolo* ^{Δ A23} larval CNSs revealed that deleting the anchor of the single meta-loop in the *nolo* meta-domain disrupted the entire meta-domain (Fig. 5A). In contrast, Hi-C on *beat-IV* ^{Δ A65} larval CNSs revealed that deleting an anchor of only one of a pair of meta-loops in the *beat-IV* meta-domain was not sufficient to disrupt the meta-domain, possibly tethered by the remaining intact meta-loop (L53) (Fig. 5B).

To test whether deletion of intergenic meta-loop anchors affects transcription of nearby genes, potentially leading to 3D contact changes, mRNA levels of adjacent genes were measured by reverse transcription-quantitative PCR (RT-qPCR). Genes around the *nolo* intergenic anchor are expressed in the male reproductive system⁴⁸ and were not reliably detected in wildtype (WT) or *nolo* ^{Δ A23} mutant heads. mRNA levels measured in whole-bodied adult flies showed no significant change upon *nolo* intergenic meta-loop anchor deletion (Fig. S6A). *CG10651* was the only gene in the TAD containing the *nolo* intergenic meta-loop anchor (Fig. 5A). Similar observations were made for genes surrounding the deleted anchor in *beat-IV* ^{Δ A65} mutants (Fig. S6B). Structural changes in *nolo* and *beat-IV* meta-domains therefore occur without strong transcriptional changes of genes near the deleted meta-loop anchors.

We next asked whether meta-domain borders are established by TAD boundary-forming proteins. CTCF is required to form 10% of TAD boundaries in larval CNSs⁴⁹, and determined 5 out of 28 (18%) meta-domain boundaries. The *GluRIA-GluRIB* meta-domain's right and bottom borders were bound by CTCF in larval CNSs of WT but not *CTCF*⁰ animals lacking maternal and zygotic CTCF (Fig. 5C)⁴⁹. In *CTCF*⁰ mutants, these borders

were weakened, and the meta-domain expanded to the next TAD boundaries. Expanding the *GluRIA-GluRIB* meta-domain in *CTCF⁰* mutants did not strongly affect gene expression within the meta-domain, as assessed by RNA-sequencing (RNA-seq) on WT and *CTCF⁰* mutant larval CNSs⁴⁹ (Fig. 5C). 48% of meta-loop anchors were within one genomic bin of a TAD boundary called by Hi-C in WT larval CNSs⁴⁹, including both promoter and intergenic anchors (Fig. S5C). Examples showed meta-loop anchors coinciding with TAD boundaries in WT (e.g. A46 in Fig. 5C), but boundary and meta-looping functions were separable in *CTCF⁰* mutants, indicating dispensability of meta-looping factors for boundary formation in these cases.

These results collectively suggest that meta-domains are tethered by meta-loops and set by TAD boundaries, rather than being driven by compartmentalization.

An intergenic meta-loop anchor enables distal activation of *nolo*

To investigate whether *nolo* and *beat-IV* intergenic anchors are required for transcription of these genes, we measured their mRNA levels in adult fly heads by RT-qPCR. This revealed a 3-fold decrease in *nolo* mRNA levels in *nolo^{ΔA23}* mutants compared to WT, while *beat-IV* mRNA levels remained unchanged in *beat-IV^{ΔA65}* mutants (Figs. S6A-S6B). Deletion of anchor A23 therefore did not significantly impact the expression of nearby genes in cis but decreased *nolo* expression over 1.6 Mb (Fig. S6A). To further investigate transcriptional defects, single-cell RNA-seq (scRNA-seq) was performed on larval CNSs from WT, *nolo^{ΔA23}* and *beat-IV^{ΔA65}* mutants (Fig. 6A). *nolo* and *beat-IV* were expressed in different cell subsets in WT CNSs (11% and 9% of WT cells, respectively) (Fig. 6B). In *nolo^{ΔA23}* mutants, *nolo* was expressed in fewer cells (2%) and at lower levels, while *beat-IV* expression remained relatively stable in *beat-IV^{ΔA65}* (Fig. 6C). Differential expression analysis per cell cluster in each mutant relative to WT revealed significant differential expression (adjusted P value ≤ 0.01) of *nolo* in three *nolo^{ΔA23}* mutant clusters, whereas *beat-IV* was unaffected in any cluster or genotype (Fig. 6C, Tables S4-S5). Thus, deleting one gene's intergenic meta-loop anchor did not impact the meta-loop or the expression of the other gene (Figs. 6C and S6D), revealing their independence.

nolo RNA fluorescence in situ hybridization (RNA-FISH) signals were strongly reduced in embryonic nerve cords of *nolo^{ΔA23}* mutants compared to WT (Fig. 6D). To assess if decreased *nolo* mRNA levels in *nolo^{ΔA23}* mutants resulted from reduced *nolo* transcription, we imaged *nolo* nascent RNA and *nolo* promoter DNA by RNA-DNA-FISH (Figs. 6E and S6E). Counting *nolo* promoter DNA-FISH spots with or without co-localizing nascent *nolo* RNA-FISH signal in late stage (stage 17) embryonic nerve cords revealed that *nolo* was transcribed in 28% of WT nuclei but only in 7% (4-fold less) of *nolo^{ΔA23}* nuclei (Fig. 6E). Thus, deletion of the *nolo* intergenic anchor decreases *nolo* transcription (Figs. 6C-6E).

Intergenic anchors may facilitate neural gene transcription by serving as distal enhancers or by bridging promoters to enhancers within the distal TAD, in an analogous manner to shorter-range “tethering” loops described recently^{25,26}. We tested five intergenic loop anchors (550-999 bp) for their ability to activate reporter gene transcription in WT transgenic reporter embryos when cloned upstream of the minimal *Drosophila* synthetic core promoter (DSCP) and integrated into a common landing site (Figs. S6F-S6G). Reporter gene RNA-FISH on embryos homozygous for each transgene revealed that only the intergenic anchors of *beat-IV* (A65) and *nolo* (A23) were moderate and weak enhancers, respectively, active in the nervous system of late-stage embryos (Fig. S6G). Deletion of the *beat-IV* intergenic anchor did not visibly affect *beat-IV* expression in larval CNSs (Fig. 6C and S6B), implying that *beat-IV* transcription in larval CNSs relies on other uncharacterized enhancers. Of note, the enhancer reporter transgenes were integrated into a different chromosomal arm (chr3L) from that housing *nolo* or *beat-IV* meta-loops to prevent potential confounding looping between the transgenic and endogenous anchors.

In summary, the *nolo* intergenic anchor is critical for *nolo* transcription (Figs. 6C-6E). However, the tested I-P meta-loops are not classical enhancer-promoter loops, as most intergenic anchors show limited enhancer activity (Fig. S6G). It is possible that *beat-IV* expression in *beat-IV^{ΔA65}* mutants is sustained by distal regulatory elements within the preserved meta-domain due to the intact meta-loop (L53 in Fig. 5B). This hypothesis is tested in the last results section.

To determine whether *nolo^{ΔA23}* mutation affects *nolo* function, we compared the lethal stages of *nolo^{ΔA23}* and *nolo^{KO}* mutants with nonsense mutations after codons 47 or 48 (Figs. 6F-6G). *nolo^{KO}* and *nolo* whole-body or neuroblast-specific knock-down pupae showed impaired hatching that was not observed in *nolo^{ΔA23}* mutants (Fig. 6G). In *nolo* knock-down flies, the *nolo^{ΔA23}* mutation decreased pupal hatching further (e.g. from 66% upon neuroblast-specific *nolo* knock-down to 11% when performed in a *nolo^{ΔA23}* mutant) (Fig. 6G). We phenotyped

neuromuscular junctions of *nolo* knock-out, knock-down and *nolo*^{ΔA23} third instar larvae. α-HRP immunostaining of neuronal membranes at muscle 6/7 neuromuscular junctions revealed that the neuromuscular junction area and the number of synaptic boutons were significantly reduced in *nolo*^{KO}, *nolo*^{ΔA23} and *nolo* knock-down flies compared to WT controls (Figs. 6H-6I), suggesting that neuromuscular junctions are under-developed or impaired. Thus, *nolo*^{ΔA23} is a hypomorphic *nolo* allele with similar neuromuscular junction defects to those observed in *nolo* knock-out and knock-down animals.

GAGA-factor (GAF) is required to form two specific meta-loops

A comprehensive analysis of all meta-loop anchors did not reveal enriched transcription factor motifs. GAF binds to a subset of tethering elements in early embryos and is required to form chromosomal loops in a manner dependent on its BTB (broad-complex, tramtrack and bric-à-brac – zinc finger, also known as POZ) domain⁵⁰. To investigate GAF's potential role in meta-loop formation, we performed Micro-C on WT and *GAF*^{ΔBTB} mutant larval CNSs⁵¹. The GAF BTB domain was specifically required to form two meta-loops connecting the promoter of *sticks and stones* (*sns*) with that of *hibris* (*hbs*) (L25) or with an intergenic anchor (L26) (Figs. 7A-7B). GAF ChIP-seq data in larval tissue chromatin, comprising the CNS⁵², revealed lower enrichment of GAF binding around meta-loop anchors compared to anchors of loops detected in embryos and cultured cells (Fig. 7C). Nonetheless, GAF bound to a subset of meta-loop anchors and A29, the common anchor of both GAF-dependent meta-loops, was most strongly bound by GAF (Figs. 7A and 7C). Thus, both meta-loops destabilized in *GAF*^{ΔBTB} mutants have relatively strong GAF binding at their anchors in WT, whereas other meta-loops with anchors that are less bound or unbound by GAF are unaffected (Figs. 7B-7C). The long-range interaction between *sns* and *hbs* is not limited to the CNS (Fig. S1) and was also previously observed in whole embryo Hi-C maps, suggesting a non-CNS-specific role of GAF in forming these meta-loops^{53,54}.

Visualization of *sns* and *hbs* mRNAs by *in situ* Hybridization Chain Reaction (HCR) revealed a moderate decrease in *sns* expression in a stripe in optic lobes of *GAF*^{ΔBTB} mutants, while *hbs* expression was mostly unchanged (Fig. 7D). These observations were further supported by RT-qPCR on larval CNSs (Fig. 7E). These results indicate that GAF is required to form meta-loops involving *sns* promoter and for WT levels of *sns* transcription in larval CNSs.

CTCF is required to form a different subset of meta-loops than GAF in the CNS

Comparison of WT and *CTCF*⁰ mutant larval CNS Hi-C maps⁴⁹ identified six meta-loops that were significantly weakened in *CTCF*⁰ mutants compared to WT (Figs. 8A-8B and S7A), and these were unaffected in *GAF*^{ΔBTB} larval CNSs (compare Figs. 7B and 8B). CTCF-dependent meta-loops were identified in CNSs but not wing discs (Fig. S1). Although CTCF generally binds constitutively in different cell types⁴⁰, the chromatin of CTCF-bound meta-loop anchors was more accessible in embryonic neurons than muscle (Figs. 8C and S7A).

Surprisingly, the *beat-IV* meta-domain was weakened in *CTCF*⁰ mutant larval CNSs even though CTCF only binds to the intergenic anchors of meta-loop L53 but not L54, which connects *beat-IV* promoter to its intergenic anchor (Figs. 8A and 8D). This suggests that L53 loss in the absence of CTCF may destabilize L54 in the same meta-domain. No genes within the *beat-IV* meta-domain were significantly differentially expressed in *CTCF*⁰ mutant relative to WT larval CNSs by RNA-seq⁴⁹ (Fig. 8A). We further examined *beat-IV* expression in WT and *CTCF*⁰ embryos by RNA-FISH. In WT embryos, *beat-IV* expression is detected throughout the nerve cord at stage 15 and increases at stage 16 (Fig. 8E). In *CTCF*⁰ embryos, *beat-IV* expression was strongly reduced at stages 15-17 except for a few bright cells in the center of the nerve cord (Fig. 8E). CTCF recruits its BTB-ZnF cofactor, Centrosomal protein 190 kDa (Cp190) to CTCF binding sites⁴⁹. Consistently, Cp190 co-bound with CTCF at L53 meta-loop anchors in WT but was either reduced or absent at these anchors in *CTCF*⁰ larval CNSs (Figs. 8A and 8D)⁴⁹. Cp190 is, in contrast, largely dispensable for CTCF binding to DNA^{20,49}, and CTCF was bound at wildtype levels at L53 anchors in *Cp190*^{KO} (maternal+ zygotic-) larval CNSs (Fig. S7A). In *Cp190*⁰ (maternal- zygotic-) embryos, *beat-IV* expression was reduced to an even greater extent than in *CTCF*⁰ embryos (Fig. 8E). In contrast, in *beat-IV*^{ΔA65} mutants with a deletion of the intergenic anchor looping to *beat-IV* promoter (Fig. 5B), *beat-IV* expression was largely normal (Fig. 8E).

We also investigated the consequences of weakening of a second CTCF-dependent meta-domain containing two 2.2 Mb meta-loops normally both bound by CTCF in WT (Fig. S7B). One meta-loop (L33) connects two intergenic meta-loop anchors and the other (L34) connects the promoter of *Multiplexin* (*Mp*) to an intergenic anchor (Fig. S7A). RNA-seq analysis revealed a 12-fold upregulation of the gene *CG7509* within the

Mp meta-domain in *CTCF⁰* mutants compared to WT larval CNSs (Fig. S7A) ⁴⁹. We further examined *Mp* expression in embryos by RNA-FISH. In WT embryos, *Mp* is expressed in the posterior nerve cord at stage 15 and expands to the anterior nerve cord at stage 16 (Fig. S7C). In *CTCF⁰* embryos, *Mp* expression appeared normal at stage 15 but failed to expand anteriorly in stage 16, resembling the expression pattern of stage 15 embryos (Fig. S7C). In *Cp190⁰* embryos, *Mp* was expressed in much fewer cells at stage 15, partially recovered in the posterior nerve cord at stage 16, but remained lower than WT in the anterior nerve cord similar to *CTCF⁰* embryos (Fig. S7C).

In conclusion, the destabilization of distinct meta-loops in *GAF^{ΔBTB}* and *CTCF⁰* mutants suggests that specific transcription factors are recruited to different meta-loop anchors, ensuring interaction specificity. Loss of CTCF or Cp190 delays and reduces the transcriptional activation of neuronal genes at CTCF-dependent meta-loops (*Mp*) and meta-domains (*beat-IV*).

DISCUSSION

We studied genome folding in the developing *Drosophila* CNS and discovered long-range associations between specific pairs of TADs. We propose that meta-domains represent a new paradigm of genome folding that enables coordinated megabase-range control of gene transcription in the CNS. Below, we discuss the implications of these findings.

Specialized chromosomal folding can support cell type-specific gene expression

Recent studies have explored cell type-specific genome folding in the mammalian nervous system. Differential genome folding was observed between brain cell types and during differentiation, resulting in altered TAD boundaries, differential enhancer-promoter contacts, and increased inactive B compartment interactions^{10,30,55}. The most striking long-range contacts occurred in *cis* (over up to 50 Mb) and *trans* between heterochromatinized silent olfactory receptor (OR) gene clusters on the one hand, and between OR enhancer clusters and the single active OR allele on the other, in neurons^{11,30}. Clustering of OR genes was crucial for activating a single OR allele⁵⁶, but deletion of an OR enhancer cluster largely only affected the transcription of OR genes in *cis*¹¹. The lack of clear defects in distal OR gene activation suggests that OR enhancers act redundantly and compensate for deletions of other enhancer islands¹¹. By contrast, the specificity of *Drosophila* CNS meta-domain associations uncovered examples of megabase-scale gene regulation.

Meta-loop function

Three lines of evidence support the functional relevance of meta-loops in CNS development and function. Firstly, deleting the *nolo* intergenic loop anchor disrupted the meta-domain and reduced *nolo* transcription over 1.6 Mb (Figs. 5 and 6). Secondly, *GAF^{ΔBTB}* mutants displayed notable loss of 6.2 Mb meta-loops involving the *sns* promoter and decreased *sns* expression (Fig. 7). Thirdly, CTCF loss weakened three CNS-specific meta-domains and decreased expression of *beat-IV* and *Mp* located at meta-loop anchors (Fig. 8). Notably, the expression of *CG7509*, located between two formerly CTCF-bound intergenic anchors in the *Mp* meta-domain, was increased in larval CNSs upon CTCF loss (Fig. S7B). Thus, the expression of two genes at (*Mp*) or near (*CG7509*) opposite meta-loop anchors was oppositely affected in *CTCF⁰* mutant CNSs, potentially suggesting coordination of gene expression over 2.2 Mb (Figs. S7A and S7C). Together, these findings indicate that gene expression in the CNS is controlled over longer distances than those previously described in the mammalian genome, such as *SOX9* activation by its enhancer 1.5 Mb away in the cranial neural crest of mice and humans⁵⁷.

Our observation of reduced *beat-IV* expression in *CTCF⁰*, but not in *beat-IV^{ΔA65}*, mutants suggests that the *beat-IV* meta-domain and not the *beat-IV* meta-loop itself, is important for *beat-IV* expression (Figs. 6C and 8E). Meta-domains may facilitate the sharing of regulatory information between distant TADs. Investigating the impact of meta-domain boundary perturbations on gene expression of genes at meta-loop anchors would provide valuable insights.

Our findings suggest that meta-loops are formed through long-range focal contacts of structural elements rather than enhancer-promoter loops. Unlike enhancer-promoter interactions, which correlate with active transcription^{40,58}, meta-loops are detected in a large fraction of neurons, including those where associated genes are not actively transcribed (Figs. 4C-4D and S1). Like tethering elements that bridge promoters to distal regulatory elements in fly embryos and mammalian cells, meta-loop anchors lack robust enhancer activity (Fig. S6G) and pair in more cells than those that transcribe the associated gene^{23,25,26,59}. However, tethering element activity is generally constrained to individual TADs^{23,25,26}, whereas CNS meta-loops span tens of TADs. Meta-loops reconcile contrasting views of 3D genome folding as either invariant scaffolds^{17,25,26,36,37} or tissue-specific structures by demonstrating that specialized and relatively invariant scaffolds exist within neurons.

Mechanism of meta-loop formation

Our findings indicate that meta-domains are tethered by meta-loops, not by compartmental interactions (Figs. 1 and 5), that pair over long and flexible genomic distances with striking specificity (Fig. 2) and rely on dedicated transcription factors including GAF and CTCF (Figs. 7 and 8) for their formation. While we focused on meta-interactions, similar mechanisms may also operate at shorter distances, such as intra-TAD loops involving meta-loop anchors.

In contrast to GAF which is required to form a meta-domain that is also observed outside of the CNS (Figs. 7 and S1)^{53,54}, CTCF was required to form three CNS-specific meta-domains (Figs. 8 and S1). Our previous finding that CTCF is specifically required in neurons for fly viability⁴⁹ suggests that CTCF plays a specialized role in the CNS that includes meta-domain formation. Differentiating neurons may express additional factors that cooperate with CTCF at meta-loop anchors. It should be noted that while CTCF is required for meta-loop formation, it cannot be the sole factor responsible, as weak meta-domains persist in *CTCF⁰* CNSs despite complete loss of CTCF⁴⁹ (Figs. 8A and S7B). Cp190 may also be required, as partial meta-domain retention in *CTCF⁰* mutants correlates with partial Cp190 retention at meta-loop anchors (Figs. 8A and S7A-S7B). The severe effect of Cp190 loss on *Mp* and *beat-IV* gene expression compared to CTCF loss may further suggest an important role of Cp190 in meta-loop formation (Figs. 8E and S7C). GAF and Cp190 belong to the same BTB-ZnF protein family, indicating a possible shared mechanism for meta-loop formation. Investigating the effect of Cp190 loss on meta-domain formation will therefore be crucial. Another possibility is that CTCF or Cp190 collaborate with other essential factor(s) to form meta-domains.

68% of meta-domains contain clustered loops (Fig. 1E). In *beat-IV* (Fig. 4A) and *GluRIA-GluRIB* (Fig. S5B) meta-domains, an I-I meta-loop emerged earlier in neuronal differentiation than the I-P meta-loop involving *beat-IV* or *GluRIA/GluRIB* promoters. Deleting the intergenic anchor connected to *beat-IV* promoter did not disrupt the *beat-IV* meta-domain (Fig. 5B). However, genetically ablating CTCF, which destabilizes the earlier CTCF-bound I-I meta-loop in the *beat-IV* meta-domain, also destabilized the I-P meta-loop (not directly bound by CTCF) and the entire *beat-IV* meta-domain (Figs. 8A-8B). These results suggest a hierarchy in long-range associations, where a meta-loop can facilitate the emergence of another meta-loop in the same meta-domain. Moreover, we observe a novel interplay between certain meta-loop anchors and TAD boundaries which appear to emerge concomitantly during neural differentiation, but the temporal resolution of our time-course Hi-C experiments is limited (Figs. 4A and S5B-S5C). It will be important to determine whether TAD boundaries may facilitate meta-domain emergence at specific loci, although TAD boundaries at other loci can be shifted without affecting meta-domain formation (Fig. 5C).

Meta-loops differ from traditional models of chromosome folding, which involve cohesin-mediated loop extrusion within a TAD^{1,2}, clustering of transcription factor-bound elements into complex “hubs” or “factories”^{10,32,60,61}, or compartmentalization^{11,56,62}. While the specificity of CTCF-dependent loops in mammalian cells is explained by a tracking mechanism, the specificity of meta-loop interactions invokes a distinct mechanism. CTCF-dependent meta-loops are unlikely to be extruded due to their large (up to over 16 Mb) size and spanning of hundreds of TAD boundaries in between. Furthermore, in contrast to meta-domains, the compartmentalization of heterochromatic OR clusters relies not only on the binding of a common set of transcription factors to local OR enhancer clusters, but primarily on heterochromatin-driven compartmentalization of OR clusters¹¹.

In conclusion, the striking specificity of meta-loop interactions is arguably their most intriguing property. The mechanisms by which GAF or CTCF enable specific pairwise interactions, despite binding to over a thousand loci^{49,63}, remains unclear. Meta-loops provide a powerful and relevant paradigm to explore the emergence of specific long-range regulatory interactions in genomes and their role in establishing tissue-specific gene activity.

Limitations of the study

Why meta-loops are enriched in the CNS remains unresolved. Meta-loops may be most prominent in neurons because neurons are post-mitotic. However, our data suggest that meta-loops are functionally significant in the CNS. It remains unclear whether meta-loops represent a mere adaptation or an evolutionary advantage. Meta-loops may compensate for chance breakpoints (illustrated in Fig. 2B) that separate CNS genes from their regulatory elements. Alternatively, meta-loops may prevent premature regulation of genes by distal regulatory elements until 3D genome organization in the CNS matures (Fig. 3A) and stabilizes meta-loops. Differentiating between these models requires identifying putative distal regulatory elements and positioning them close to target genes to bypass long-range regulation. Another possibility is that meta-loops evolved to generate novelty in the brain. The competitive advantage of distal regulation may only become apparent over evolutionary timescales, in species studied in the wild. To determine if meta-loops confer advantages for regulating specific CNS genes in a given species, it is necessary to identify the relevant neurons where gene expression depends on meta-loops and understand how this regulation confers species-specific traits, ideally through studies conducted in natural habitats.

Acknowledgements

We thank David Garfield for helpful discussions, Victor Rossier for his gene family visualization interface, and Michael Taschner for the graphical abstract. H.M. was enrolled in the University of Paris Cité European Master Program in Genetics. R.C.S. received funding from the Marie Skłodowska-Curie grant agreement No. 945363. This work was supported by the Polish National Agency for Academic Exchange (Polish Returns 2019 to A.J.), National Institutes of Health grant R35 GM118147 and U01 DK127429 (to M.S.L.), the Swiss National Science Foundation grant no. 184715 (to M.C.G.), the Pierre Mercier Science Foundation (to M.C.G.), and the University of Lausanne (to M.C.G.).

Author contributions

Conceptualization, A.J., M.S.L. and M.C.G.; Methodology, G.M., J.D., S.R., A.J. and M.C.G.; Investigation, G.M., X.L., M.M., R.C.S., H.M., D.R., J.K. and M.C.G.; Formal Analysis, J.D., X.L., M.L., P.R., C.I., N.G. and A.J.; Writing, N.G., A.J., M.S.L., and M.C.G.; Supervision, N.G., B.D.M., A.J., M.S.L. and M.C.G; Funding Acquisition, A.J., M.S.L., and M.C.G.

Declaration of interests

S.R. developed MUSE technology marketed by arcoris bio AG, has shares and sits on the board. The authors declare no other competing interests.

FIGURE LEGENDS

Figure 1. Meta-domains harbor loops involving nervous system-expressed genes. See also Figures S1-S2.

- A. (Top) Third instar larva wing disc (left) or CNS (right) Micro-C map of chromosome 3L. Dotted squares mark meta-domains. (Bottom) Zoom-in on the left (in x) and right (in y) anchors of a meta-domain (center) harboring meta-loops L35 and L36. Pseudo-bulk scATAC-seq on larval CNSs, DNase-seq on embryonic neurons and muscles⁴⁰, mRNA-seq on larval CNSs⁴⁹ and gene tracks (with longest isoforms colored according to transcription direction) are shown. 10 kb regions around each meta-loop anchor are shown in dotted rectangles.
- B. Cartoon representation of panel A column 2 (see main text for details).
- C. Normalized larval CNS Micro-C count distribution (at 6.4 kb resolution) versus genomic distance of all pairwise contacts on all chromosomes (except Y). Red dots show meta-loops, purple dots show intra-TAD loops involving meta-loop anchors, light and dark grey curves respectively show 1-99 or 25-75 percentile ranges of all contact pairs, and the black line shows the median of all contact pairs.
- D. Distribution of sizes (in bp, in y) of all 58 meta-loops (points) called in larval CNSs.
- E. Number of meta-domains (in y) each harboring the indicated numbers of loops involving meta-loop anchors (in x).
- F. Distribution of distances (in bp, in y) of all 79 meta-loop anchors (points) to the nearest protein-coding gene transcription start site. A line at 200 bp marks the cut-off for defining promoter versus intergenic anchors.
- G. Classification of all 58 meta-loops called in larval CNSs as intergenic-promoter (I-P), intergenic-intergenic (I-I) and promoter-promoter (P-P).
- H. RNA-seq on adult fly tissues from FlyAtlas v2⁴⁸. All indicated tissues are from adult females except for testis and accessory gland.

Figure 2. Meta-loop anchors specifically pair over flexible genomic distances in evolution. See also Figure S3.

- A. Loops observed by Hi-C in CNSs of adult *D. melanogaster* (*D. mel.*, left) and *D. virilis* (*D. vir.*, right). For each chromosome arm, conserved loops are uniquely colored (but colors are re-used between chromosome arms) whereas loops present only in *D. mel.* or *D. vir.* are gray. Scale is in Mb.
- B. Dot plot of aligned regions of *D. mel.* (in x) and *D. vir.* (in y) chromosome 3L. Aligned regions in the same orientation are in blue, those that are inverted in one species relative to the other are in red. Conserved loops are colored as in A and loop anchors are highlighted (in x and y).
- C. The sizes of all 37 conserved loops involving meta-loop anchors (colored symbols) are plotted in *D. vir.* (in y) versus *D. mel.* (in x). Different symbols represent conserved loops whose anchors have one of three possible relative orientations in *D. mel.* versus *D. vir.* The “complex” case is illustrated in Figs. S3I-S3J.
- D. Hi-C maps showing a *D. vir.* meta-loop (circled Dvir_L48, right) that is homologous to an intra-TAD loop in *D. mel.* (circled, left). These loops connect the promoters of *side-IV* orthologs to an intergenic anchor. The Dvir_L49 meta-loop was not conserved even though Dvir_L49 anchors aligned to the *D. mel.* genome. Chromatin accessibility tracks in each species are shown: bulk ATAC-seq in WT *D. mel.* adult CNSs from Janssens et al. 2022⁴⁶ or DNase-seq on whole *D. vir.* embryos of the indicated hours of development from Liu et al. 2021⁶⁴. Anchors of meta-loops called in *D. vir.* were assigned to underlying DNase-seq peaks in old (25-28h) *D. vir.* embryos.

Figure 3. Meta-loops strengthen during neuronal differentiation. See also Figures S4-S5.

- A. (Top) Embryogenesis timeline (in hours), embryonic stages and major events in neurogenesis. Whole-genome Hi-C maps in neuroblasts (row 1), glia (row 2) or neurons (row 3) FACS-purified respectively with α -Wor, Repo or Elav antibodies from embryos of the indicated hours of development. Dotted squares show a zoomed-in view of a portion of chromosome 3R harboring meta-loops L53-L54 and L55-L60 (marked by arrowheads). 3D genome configurations deduced by Hi-C are shown in a cartoon below and each bounded by a differently colored box.

- B. For each loop involving a meta-loop anchor observed in WT larval CNSs (rows), loop strength (color, measured as the observed-over-expected normalized Hi-C count) is shown in each sample (columns). Intra-TAD loops and meta-loops (that loop between TADs) are indicated. I-I, I-P and P-P loops are color-coded on the right.

Figure 4. Meta-loop anchors are in enhanced proximity in a large fraction of neurons.

- A. Hi-C maps, DNase-seq profiles in neurons FACS-purified from staged embryos from Reddington et al. 2020⁴⁰, and gene tracks of the left (in x) and right (in y) anchors of a meta-domain (center) harboring meta-loop L54 involving *beat-IV* promoter, in neurons of 6-8 (column 1), 10-12 (column 2) or 14-16 (column 3) hour old embryos. Arrowheads mark emerging TAD boundaries coinciding with meta-loop anchors. DNA-FISH probes used in panel C are highlighted in green and red in the gene tracks. Zoomed-in views of 10 kb around each meta-loop anchor are shown in dotted boxes.
- B. Same as C but for *nolo*.
- C. Violin plot of physical distances (in y, in μm) measured by two color DNA-FISH with 5 kb probes overlapping *beat-IV* and an intergenic meta-loop anchor, in Elav-immunolabeled (pink) or neighboring (blue) nuclei of embryos of the indicated stages (in x). Below, lateral views of embryos of the indicated stages stained by immuno-RNA-FISH with a probe antisense to *beat-IV* mRNA (green), α -Elav (pink) and DAPI (anterior, left; posterior, right). *P* values and *D*-statistics from two-sided Kolmogorov-Smirnov tests are indicated; n.s., not significant, *, significant ($p < 0.01$).
- D. Same as C but for *nolo*.

Figure 5. Meta-domains are formed by meta-loop anchors and delimited by TAD boundaries. See also Figure S6.

- A. Hi-C maps, pseudo-bulk scATAC-seq profiles in WT and gene tracks are shown for the left (in x) and right (in y) anchors of a meta-domain (center) harboring meta-loop L22 that connects *nolo* promoter to an intergenic anchor in larval CNSs of WT (column 1) or *nolo* ^{$\Delta A23$} animals (column 2). Filled/empty arrowheads point to L22 presence/absence in the respective genotype. Zoomed-in views of 10 kb around each meta-loop anchor are shown in dotted boxes and mark the extents of meta-loop anchor A23 deletion or testing for enhancer activity in Fig. S6G.
- B. Same as A but for the meta-domain harboring meta-loops L53 and L54 which involves *beat-IV* promoter, in WT (column 1) or *beat-IV* ^{$\Delta A65$} mutants (column 2).
- C. Hi-C maps, pseudo-bulk scATAC-seq profiles in WT, CTCF ChIP-seq profiles with called peaks and gene tracks are shown for the meta-domain harboring L35 (connecting *GluRIA* and *GluRIB* promoters) and L36 in larval CNSs of WT (left) or *CTCF*⁰ (middle) mutants. Filled/empty arrowheads mark normal/weakened TAD boundaries in the respective genotypes. A differential (*CTCF*⁰ minus WT) Hi-C map is shown on the right (scale is \log_2 fold change of contacts). Genes determined to be differentially expressed by RNA-seq in *CTCF*⁰ mutant relative to WT larval CNSs ($\text{padj} < 0.05$ and $|\text{fold change}| > 1.5$)⁴⁹ are highlighted in pink and labeled with a letter on the gene track in the middle, and the fold changes are indicated in the table.

Figure 6. An intergenic anchor enables distal activation of *nolo*. See also Figure S6.

- A. Combined scRNA-seq datasets generated in WT, *nolo* ^{$\Delta A23$} and *beat-IV* ^{$\Delta A65$} larval CNSs. Cell clusters are numbered, and major cell types are color-coded.
- B. Same as A but coloring cells expressing *nolo* (red) or *beat-IV* (blue).
- C. For each cell cluster numbered as in A and B, the percentage (circle size) of cells expressing the indicated genes (columns) and the average scaled expression (color) are shown in WT (1), *nolo* ^{$\Delta A23$} (2) or *beat-IV* ^{$\Delta A65$} (3) larval CNSs. Cells from cluster 11 were variably recovered between samples and thus excluded from this plot. *pros*, *repo* and *wor* are respectively neuronal, glial and neuroblast marker genes. Asterisks mark cell clusters that significantly ($\text{padj} \leq 0.01$) differentially expressed the indicated gene in *nolo* ^{$\Delta A23$} (red asterisk) or *beat-IV* ^{$\Delta A65$} (orange asterisk) mutants relative to WT.
- D. RNA-FISH with a probe antisense to *nolo* mRNA (green) in dissected nervous systems of stage 16 WT and *nolo* ^{$\Delta A23$} mutant embryos immunostained with α -Elav (pink, to label neurons) and labeled with DAPI (anterior, left; posterior, right). Reduced *nolo* expression shown here was observed in 100% of *nolo* ^{$\Delta A23$} mutants (≥ 10 embryos imaged per genotype).

- E. Percentage (in y) of nuclei with *nolo* promoter DNA-FISH signal that overlapped nascent *nolo* RNA-FISH signal in nerve cords of individual stage 17 WT (left) or *nolo*^{ΔA23} (right) mutant embryos (dots). Horizontal lines show means. N, number of embryos; n, number of nuclei with DNA-FISH signal scored.
- F. Schematic of the protein encoded by the longest Nolo isoform (Uniprot accession number Q0E8N3). Protein domains and the positions of premature translation termination codons induced by two independent CRISPR/Cas9-mediated 5 bp deletions are marked.
- G. Percentages of indicated genotypes (with/without *nolo* mutations and *nolo* tissue-specific knock-down) that successfully developed from pupae to fully hatched adults in three independent batches of 30-40 animals (dots). Horizontal lines show means.
- H. α-vGlut immunostaining of synaptic boutons and α-HRP immunostaining of the whole neuromuscular junction in muscle 6/7 terminals in WT and the indicated *nolo* mutant third instar larvae. Representative images of each genotype are shown, and the results from several images are shown in panel I.
- I. Quantification of neuromuscular junction (NMJ) area (left) and numbers of synaptic boutons (right) in images such as those shown in panel H from n ≥ 8 independent larvae (data points) of each of the indicated genotypes. P values from one-way ANOVA (for wildtype versus mutant comparisons) or unpaired two-sided t tests (for *Luciferase* to *nolo* knock-down comparisons) are indicated.

Figure 7: The GAF BTB domain is required to form the *sns-hbs* meta-domain.

- A. Micro-C map of a meta-domain (center) harboring meta-loops L25 and L26 that connect *sns* promoter to that of *hbs* or to an intergenic anchor, respectively, in larval CNSs of WT (column) and *GAF*^{ΔBTB} animals (column 2). Filled/empty arrowheads point to wildtype/reduced meta-loop strength in the respective genotype. Pseudo-bulk scATAC-seq in larval CNSs and published GAF ChIP-seq profiles in larval tissues including CNS⁵² are also shown. Dotted rectangles show zoomed-in views of 10 kb regions around meta-loop anchors.
- B. Differential analysis of strengths of loops involving meta-loop anchors (colored dots) measured by Micro-C in larval CNSs of WT versus *GAF*^{ΔBTB} mutants. Labeled loops (pink dots) were significantly weaker (padj ≤ 0.01) in *GAF*^{ΔBTB} compared to WT.
- C. Distribution of GAF ChIP-seq signal in larval tissues including CNS from Tsai et al. 2016⁵² in 1-kb bins ± 25 kb around all anchors of meta-loops (lane 1) or loops previously defined in tissue culture cells (lane 2) or embryos (lanes 3 and 4) ranked by GAF ChIP occupancy. Average signals over all loop anchors are shown above.
- D. Hybridization chain reaction (HCR) RNA-FISH with co-hybridized probes antisense to *sns* (magenta) and *hbs* (yellow) mRNAs in larval CNSs of WT (left) and *GAF*^{ΔBTB} animals (right). Filled/empty arrowheads mark wildtype/reduced *sns* expression in a specific stripe in the respective genotypes. This phenotype was observed in 100% of *GAF*^{ΔBTB} mutants (≥ 10 larval CNSs imaged per genotype).
- E. Fold change of *hbs* and *sns* mRNA levels measured by RT-qPCR with two independent primer pairs each, in *GAF*^{ΔBTB} relative to WT larval CNSs. Dots represent individual biological triplicate values, thick horizontal bars represent average values, and error bars indicate standard deviation. P values from two-sided unpaired t tests are shown.

Figure 8: CTCF is required to form a different set of meta-loops than GAF. See also Figure S7.

- A. Hi-C map of meta-loops L53 and L54 in WT (left) and *CTCF*⁰ mutant (right) larval CNSs. L54 connects *beat-IV* promoter to an intergenic anchor located 5.1 Mb away. Filled/empty arrowheads point to wildtype/reduced meta-loop strength in the respective genotypes. Pseudo-bulk scATAC-seq and ChIP-seq profiles of CTCF and Cp190 in larval CNSs are also shown. Hi-C and ChIP-seq data were published by Kaushal et al.⁴⁹. Genes determined to be differentially expressed (DE) by RNA-seq in *CTCF*⁰ mutant relative to WT larval CNSs (padj < 0.05 and |fold change| > 1.5)⁴⁹ are highlighted in pink and labeled with a letter on the gene track on the right, and the fold changes are indicated in the table below. Dotted rectangles show zoomed-in views of 10 kb regions around meta-loop anchors.
- B. Differential analysis of strengths of loops involving meta-loop anchors (colored dots) measured by Hi-C in larval CNSs of WT versus *CTCF*⁰ mutants. Labeled meta-loops (pink dots) were significantly weaker (padj ≤ 0.01) in *CTCF*⁰ compared to WT.

- C. Distribution of DNase-seq signal (data from Reddington et al. 2020⁴⁰) around CTCF-bound anchors of meta-loops weakened in *CTCF⁰* mutant larval CNSs in 500-bp bins ± 25 kb around these anchors ordered by genomic position, in 10-12 hour old embryonic neurons (top) or muscle (bottom).
- D. Distribution of CTCF ChIP-seq peaks and differential Cp190 ChIP-seq occupancy (colored by average best.logFC) in *CTCF⁰* versus WT larval CNSs (data from Kaushal et al. 2021²⁰) in 1-kb bins ± 25 kb around all meta-loop anchors ordered by genomic position. Percentages of meta-loop anchors with at least one CTCF ChIP peak or differential Cp190 binding region in WT versus *CTCF⁰* mutants over all anchors are shown on top. Pink filled/empty arrowheads indicate anchors of meta-loops weakened in *CTCF⁰* that are bound/not bound by CTCF. These anchors are labeled on the left and meta-loops are labeled on the right.
- E. Ventral views of embryos of indicated genotypes (rows) and stages (columns) stained by immuno-RNA-FISH with a probe antisense to *beat-IV* mRNA (green), α -Elav (pink) and DAPI (anterior, left; posterior, right; merged images on top). Specific RNA-FISH signal appears as dots; string-like and broad green non-specific signal is visible in trachea and epidermis. Filled/empty arrowheads point to wildtype/reduced *beat-IV* RNA-FISH signal in the respective genotypes. The phenotypes shown here (described in the results) were observed in 100% of embryos (≥ 6 embryos imaged per genotype and stage).

SUPPLEMENTAL INFORMATION

Figure S1. Comparison of meta-domains in larval wing discs, CNSs and adult CNSs. Related to Figure 1.

- A. Micro-C maps (400 bp resolution) generated in the indicated tissues (columns) are shown separately for each meta-domain. Loops involving meta-loop anchors are marked in larval CNSs (in which they were defined). The color scales show normalized Micro-C counts.
- B. For each loop involving meta-loop anchors observed in WT larval CNSs (columns), loop strength (color, measured as the observed-over-expected normalized Micro-C count) is shown in each sample (rows). Intra-TAD loops and meta-loops (that loop between TADs) are indicated. Intergenic-intergenic (I-I), intergenic-promoter (I-P) and promoter-promoter (P-P) loops are color-coded below.

Figure S2. Meta-loops involve genes expressed in the nervous system. Related to Figure 1.

- A. Average normalized Micro-C (left) and Hi-C (right) counts per pairs of 1.6 kb (for Micro-C) or 2 kb (for Hi-C) bins (in y) as a function of genomic distance (using exponentially increasing bins in x) per genotype (merged biological duplicates or triplicates are shown for separately colored genotypes).
- B. Distribution of DNase-seq signal (blue, average per 500 bp) from the indicated published datasets (columns, data from Reddington et al 2020⁴⁰) in 500-bp bins ± 25 kb around all 79 meta-loop anchors defined in larval CNSs (in rows, ordered by genomic position). The datasets are FACS-purified mesodermal cells (Meso) or neurons (Neuro) from 10-12 hour old embryos. Meta-loop anchors were originally defined by Hi-C and subsequently intersected with underlying DNase-seq peaks in FACS-purified neurons from 10-12 hour old embryos; hence the distance from the loop anchor is the distance from the summit of the DNase-seq peak associated with that anchor measured in the 5'-to-3' direction (as described in the Methods). For each column, the upper plot summarizes the average signal over all loop anchors. This shows that neurons of 10-12 hour old embryos have DNase-seq peaks enriched at loop anchors, and only some of these peaks are also present in muscle.
- C. The number of anchors (in x) that each meta-loop anchor pairs with is shown (in y).
- D. Pseudo-bulk scATAC-seq data (this study) and virtual 4C-seq tracks derived from Hi-C data (at 2 kb resolution, data from Kaushal et al. 2021⁴⁹) in WT larval CNSs are shown. The 4C-seq tracks correspond to reads involving meta-loop anchor A16 (row 1), both A16 and A22 (row 2) and both A16 and A25 (row 3). Viewpoints are marked by red triangles. Regions with numbers of reads exceeding the y axis scale are marked by a red line. 955/39,716 (2.4%) of all Hi-C reads involving A22 also involve A25, whereas 25/480 (5.2%) of reads involving A16 and A22 also involve A25. This corresponds to the triplet cooperativity factor of 2.2, and shows that A16, A22 and A25 meta-loop anchors can engage in concurrent three-way interactions.
- E. Comparison of the distributions of triplet cooperativity factor values (points, defined in the Methods) for 21 looping triplets consisting of three meta-loop anchors forming three observed loops, and for 42 control triplets. The control triplets consist of two anchors from a looping triplet and a third non-looping anchor at a comparable genomic distance as the third looping anchor, but in the opposite genomic direction. The indicated p value is from a Wilcoxon test.
- F. Upset plot showing the overlap between anchors of meta-loops reported in larval CNSs in this study and loops defined by Hi-C in Kc167 cells⁴¹ or by Micro-C in early embryos^{25,26} (rows). This shows that most meta-loop anchors found in the CNS are unique, i.e. are not used in embryos or Kc167 cells.
- G. Distribution of GAGA-factor (GAF) (data from Tsai et al. 2016⁵²) and subunits of Polycomb Repressive 1 comprising Polycomb (Pc), Polyhomeotic (Ph), and Posterior sex combs (Psc) (data from Loubière et al. 2020⁶⁵) ChIP-seq signal (blue, average per 1 kb) in 1-kb bins ± 25 kb around loop anchors defined by Hi-C in Kc167 cells (data from Eagen et al. 2017⁴¹), by Micro-C in early (nuclear cycle 14) embryos (data from Batut et al. 2022²⁵ or Levo et al. 2022²⁶), or reported in larval CNSs in this study. Loop anchors are shown in rows, ordered by genomic location. For published loop anchors, the distance from the center of the loop anchor measured in the 5' to 3' direction is shown in the x-axis. Larval CNS meta-loop anchors from this study were originally defined by Hi-C and subsequently intersected with underlying DNase-seq peaks in FACS-purified neurons from 10-12 hour old embryos; hence the distance from the loop anchor is the distance from the summit of the DNase-seq peak associated with that anchor measured in the 5' to 3' direction (as described in the Methods). For each column, the upper plot shows

the average signal over all loop anchors as a function of distance from the loop anchors, per 1 kb. GAF ChIP-seq was performed on larval tissues comprising CNSs⁵², while Polycomb ChIP-seq was performed on larval eye-antenna discs composed in large part of neurons⁶⁵.

- H. Average number of PANTHER orthologs for each gene, or pair of paralogous genes, present at promoter anchors (rows) found in the indicated species or species group (columns). Vertical lines indicate whether genes at promoter anchors only loop to an intergenic anchor (in I-P meta-loops) or also loop to another promoter anchor (in P-P meta-loops). If a gene at a promoter anchor engages in both I-P and P-P meta-loops, it was annotated as a P-P meta-loop in this figure. Another vertical line further indicates whether a P-P meta-loop involved two paralogous gene promoter anchors (“paralogous promoter meta-loops”) or not.
- I. Lateral views of embryos of the indicated stages (columns) stained by immuno-RNA-FISH with a probe antisense to the indicated genes at meta-loop promoter anchors (rows, in green), α -Elav (pink, to label neurons) and DAPI (anterior, left; posterior, right; merged images on the right). This shows that all tested genes at meta-loop promoter anchors start to be expressed in the nervous system at stage 14 as neurons differentiate.
- J. GO term enrichment analysis of genes at promoter anchors (left) and of genes whose transcription start sites (TSSs) are closest to intergenic anchors. GO terms (in x) are ranked by gene enrichment ratios (in y). Dot sizes indicate the number of genes associated with a given GO term and are colored according to false discovery rate-adjusted p value.

Figure S3. Meta-loop conservation over evolution. Related to Figure 2.

- A. Whole-genome views of Hi-C maps of *D. mel.* (left) and *D. vir.* (right) adult CNSs. Dotted circles show examples of meta-domains.
- B. For each *D. mel.* meta-loop anchor, the number of homologous loci in *D. vir.* is counted (labeled as “candidate *D. mel.* meta-loop anchors in *D. vir.*”).
- C. Violin plots showing exon coverage at all 40 anchors of the 37 loops involving meta-loop anchors conserved in *D. mel.* and *D. vir.* For each anchor, we considered all the aligned nucleotides of the respective alignment chain and calculated the fraction of these nucleotides overlapped by gene exons. The overlaid boxplot shows the median (thick vertical line) and the 25th and 75th percentiles (hinges).
- D. For each of the 68 loops involving meta-loop anchors in *D. mel.*, the number of pairs of loci found in *D. vir.* that are homologous to the respective pair of *D. mel.* meta-loop anchors is shown (labeled as “candidate loops in *D. vir.*”) in x. Candidate loops in *D. vir.* for a given *D. mel.* meta-loop are colored in gray when none of these were observed to loop by Hi-C in *D. vir.* (labeled “not conserved”), or in red when at least one of these candidates were observed to loop in *D. vir.* (labeled “conserved”). Note that a few *D. mel.* meta-loops had up to 2 or 4 candidate matches in *D. vir.* due to anchor duplication in *D. vir.* Nevertheless, each *D. mel.* meta-loop had at most one loop conserved in *D. vir.* (i.e. the remaining 1 or 3 candidate meta-loops in *D. vir.* did not loop).
- E. All 79 meta-loop anchors in *D. mel.* (rows) are sorted by their distance to the closest indicated gene TSS (left). For each *D. mel.* meta-loop anchor, dots indicate whether (dot present) or not (dot absent) the anchor is a promoter anchor, has a candidate match in *D. vir.*, and has an ortholog in *D. vir.* of the gene that it is closest to (middle). Colors indicate whether the candidate anchor in *D. vir.* loops (dark blue) or not (light blue) according to *D. vir.* CNS Hi-C maps. For anchors with both a candidate match and an ortholog in *D. vir.* present on the same chromosome, the closest distance between the match and the putative TSS of the ortholog is shown (right).
- F. Out of 79 *D. mel.* meta-loop anchors, 52 have both a candidate anchor in *D. vir.* and a *D. vir.* ortholog of the gene closest to the meta-loop anchor in *D. mel.*. For each of these 52 loop anchor candidates in *D. vir.* the distance to the ortholog is plotted (in y, in kb). Anchor candidates are categorized as those that loop in *D. vir.* (left) and those that do not (right). This plot shows that loop anchor candidates that loop in *D. vir.* are not any closer to the ortholog genes than loop anchor candidates that do not loop ($p=0.09$, Wilcoxon test).
- G. (Top) For each of all 68 loops involving meta-loop anchors called in *D. mel.*, those with candidate loops in *D. vir.* that are conserved (i.e. loop, red) or not (i.e. do not loop, orange) in *D. vir.*, and those that did not have a candidate match in *D. vir.* (grey) are shown as dots arranged according to their sizes (in x). (Bottom) For each of all 64 loops involving meta-loop anchors called in *D. vir.*, those with candidate loops

- in *D. mel.* that are conserved (red) or not (orange) in *D. mel.*, and those that did not have a candidate match in *D. mel.* (grey) are shown as dots arranged according to their sizes (in x).
- H. Matching of randomly permuted loops in *D. mel.* to observed loops in *D. vir.* The distribution (gray) is derived from 1000 random pairings of *D. mel.* loop anchors. The average number of matches of randomly permuted loops (dashed vertical line) is 4.86 times lower than the observed number of loops conserved in *D. vir.* (vertical orange line).
 - I. Illustration of the “complex” loop referenced in Fig. 2C that had rearrangement junctions at both anchors. Alignment chains of conserved sequence fragments in concordant (blue) and discordant (red) orientation. The four colored arrows on the sides mark the roughly conserved regions. Green and cyan stripes indicate anchors of loops conserved in *D. mel.* and *D. vir.*; gray stripes indicate exons.
 - J. The gray box indicates a possible chromosomal rearrangement scenario that may have occurred during evolution to result in the observed meta-loop anchors in *D. mel.* (top, white background) and *D. vir.* (bottom, white background).

Figure S4. Chromatin accessibility at meta-loop anchors. Related to Figure 3.

- A. Meta-loop anchor chromatin accessibility in L3 larval CNSs determined by scATAC-seq. An “accessibility score” was defined as the percentage (color) of single cells in which chromatin was accessible in a 20 bp region around each meta-loop anchor. Accessibility scores are shown for each meta-loop anchor (columns, labeled A1-A79 and ordered by genomic position) in each cell type (rows) defined as neuron, neuroblast or glia (as described in the Methods) https://app.readcube.com/library/971d43d2-3a7a-4850-b3a4-305df315b977/all?uuid=2775183052286859&item_ids=971d43d2-3a7a-4850-b3a4-305df315b977:52cd7a85-868a-4beb-9bd1-78c37c111538. Meta-loop anchors were classified as intergenic or promoter (top row). For each cell type (rows), the side boxplots summarize the average accessibility scores plotted at every loop anchor (each data point is a loop anchor). The significance of the observed relative to the expected result is also shown. For each meta-loop anchor (columns), the top boxplots summarize the average accessibility score in every cell type (each data point is a cell type).
- B. Same as A but showing meta-loop anchor chromatin accessibility in adult fly CNSs determined by scATAC-seq by Janssens et al. 2022⁴⁶. Cell types (rows) were defined by that study.
- C. Distribution of DNase-seq signal (blue, average per 500 bp) from the indicated published datasets (columns, data from Reddington et al 2020⁴⁰) in 500-bp bins ± 25 kb around all 79 meta-loop anchors defined in larval CNSs (in rows, ordered by genomic position). The datasets are FACS-purified mesodermal cells (Meso), neurons (Neuro), non-mesodermal and non-neuronal cells (other) from embryos or whole embryos (WE) of the indicated hours of development. Meta-loop anchors were originally defined by Hi-C and subsequently intersected with underlying DNase-seq peaks in FACS-purified neurons from 10-12 hour old embryos (shown in column 8); hence the distance from the loop anchor is the distance from the summit of the DNase-seq peak associated with that anchor measured in the 5'-to-3' direction (as described in the Methods). For each column, the upper plot summarizes the average signal over all loop anchors. This shows that neurons of 10-12 hour old embryos have DNase-seq peaks enriched at loop anchors, and only some of these peaks are present in younger neurons or in other embryo samples.

Figure S5. Developmental time-course of meta-loop formation in CNS cell types. Related to Figure 3.

- A. (Top) Embryos of three developmental stages (at 6-8, 10-12 or 14-16 hours of development) were fixed and dissociated into nuclei, then co-stained with directly labeled antibodies against lineage markers of neurons (α -Elav), glia (α -Repo) or neuroblasts (α -Wor). (Bottom) Each cell type was FACS-purified with the indicated gates. Note that the proportion of neurons and glia progressively increased in older embryos, whereas the inverse was true for neuroblasts.
- B. Hi-C maps, DNase-seq tracks in neurons purified from 6-8h or 10-12h embryos from Reddington et al.⁴⁰, and gene tracks of the left (in x) and right (in y) anchors of a meta-domain harboring meta-loops L35 and L36 in neuroblasts (row 1), glia (row 2) or neurons (row 3) FACS-purified from 6-8 (column 1), 10-12 (column 2) or 14-16 (column 3) hour old embryos. The meta-domain is shown in grid view in the center. This shows that L36 is detectable in all samples, whereas L35 connecting the promoters of *GluRIA* and *GluRIB* is only clearly distinguishable in neurons from 14-16 hour old embryos. The

arrowhead marks an example of a TAD boundary that appears close to meta-loop L35 anchor A40 concomitantly with looping.

- C. Distribution of the presence of TAD boundaries called in WT larval CNS Hi-C maps by Kaushal 2021⁴⁹ in 1-kb bins ± 25 kb around all 79 meta-loop anchors. The summarized percentage of loci with a TAD boundary present over all loop anchors is shown on top. Anchors are classified as intergenic or promoter on the right. This shows that 48% of meta-loop anchors are within ± 2 kb (i.e. within ± 1 binsize at which TAD boundaries were called) of a TAD boundary.

Figure S6. Effect of meta-loop anchors on gene expression. Related to Figures 5 and 6.

- A. Fold change (in y) of mRNA levels of indicated genes (in x) in WT versus *nolo* ^{$\Delta A23$} adult whole-bodied flies (left) or heads (right) quantified by RT-qPCR. Dots represent individual biological triplicate values, thick horizontal bars represent average values, and error bars indicate standard deviation. P values from two-sided unpaired heteroscedastic t tests are shown. Gene distance to the deleted A23 intergenic meta-loop anchor is indicated below. *CTCF* is not expected to be affected by *nolo* ^{$\Delta A23$} and serves as negative control. This shows that *nolo* mRNA levels are reduced by 2- (in whole bodies) to 3-fold (in heads) in *nolo* ^{$\Delta A23$} mutants relative to WT.
- B. Same as A but quantifying the expression of *beat-IV* and genes around the deleted A65 intergenic meta-loop anchor in WT versus *beat-IV* ^{$\Delta A65$} mutants. This shows that *beat-IV* mRNA levels are not significantly changed in *beat-IV* ^{$\Delta A65$} mutants relative to WT.
- C. Marker gene expression used to annotate broad cell types in the combined WT, *beat-IV* ^{$\Delta A65$} and *nolo* ^{$\Delta A23$} larval CNS scRNA-seq data presented in Fig. 6A.
- D. For each meta-loop called in WT larval CNSs (columns, ordered by hierarchical clustering), loop strength (color, measured as the observed-over-expected normalized Hi-C count) is shown in larval CNSs of each indicated genotype (rows).
- E. Nascent RNA-DNA-FISH images of nerve cords of stage 17 WT (row 1) and *nolo* ^{$\Delta A23$} mutant (row 2) embryos with antisense probes against *nolo* intronic RNA (red) and *nolo* promoter DNA (green). *nolo* nascent RNA-FISH signal was weaker in *nolo* ^{$\Delta A23$} mutants compared to WT. We therefore imaged RNA-FISH signal in the mutant with higher laser intensity than in WT to detect nuclei with even low RNA-FISH signal. Because of the resulting higher background signal, less *nolo* nascent RNA-FISH spots colocalized with *nolo* promoter DNA-FISH spots in *nolo* ^{$\Delta A23$} mutants than in WT. To distinguish real *nolo* nascent RNA-FISH signal from background in *nolo* ^{$\Delta A23$} mutants, *nolo* promoter DNA-FISH spots with or without overlapping *nolo* nascent RNA-FISH signal were counted in biological triplicates and plotted in Fig. 6E.
- F. List of intergenic anchors (IA, 550-999 bp) tested for their ability to drive reporter gene transcription when cloned upstream of the minimal *Drosophila* Synthetic Core Promoter (DSCP) transcribing a reporter gene, as schematized below. A transgene without an intergenic anchor was included as negative control to reveal the activities of local enhancers present around the transgene integration site. Each transgene was each stably integrated into the VK33 genomic landing site on chromosome 3L of WT embryos. For each tested anchor, its name and genomic coordinates and the observed effect it had on reporter gene expression relative to the negative control is reported.
- G. Lateral views of embryos homozygous for the indicated transgenes (rows, described in panel F) at the indicated developmental stages (columns) stained by immuno-RNA-FISH with a probe antisense to reporter gene mRNA (green), α -Elav (pink) and DAPI (anterior, left; posterior, right; merged images on the right). The negative control (row 1) shows that local enhancers around the VK33 transgene integration site drive reporter gene expression in dynamic spatio-temporal patterns, but not in the embryonic nervous system (labeled by α -Elav) at any stage. Of four intergenic anchors tested, only those of the *nolo* or *beat-IV* meta-loops weakly or moderately drove reporter gene expression in the nervous system of late-stage embryos, respectively. Empty/filled arrowheads mark absence/presence of reporter gene mRNA in stage 16-17 embryonic nerve cords in each genotype. The shown reporter gene expression patterns were observed in 100% of embryos (≥ 5 embryos imaged per genotype and stage).

Figure S7: CTCF loss destabilizes the *Mp* meta-domain. Related to Figure 8.

- A. The indicated pseudo-bulk scATAC-seq or ChIP-seq profiles generated in larval CNSs of the indicated genotypes are shown for anchors A63 and A66 of meta-loop L53 (shown in Fig. 8A), A38 and A43 of

meta-loop L33 (shown in Fig. S7A), and A39 and A42 of meta-loop L34 (shown in Fig. S7A). Note that CTCF binding is largely unaffected at these meta-loop anchors in *Cp190^{KO}* mutant larval CNSs.

- B. Hi-C map of meta-loops L33 and L34 in WT (left) and *CTCF⁰* mutant (right) larval CNSs. L34 connects the downstream promoter of *Mp* to an intergenic anchor located 2.2 Mb away. Filled/empty arrowheads point to wildtype/reduced meta-loop strength in *CTCF⁰* mutants (in the meta-domain), or to wildtype/weakened TAD boundaries in *CTCF⁰* mutants (in x and y). Pseudo-bulk scATAC-seq and ChIP-seq profiles of CTCF and Cp190 in larval CNSs are also shown. Hi-C and ChIP-seq data were published by Kaushal et al. ⁴⁹. Genes determined to be differentially expressed (DE) by RNA-seq in *CTCF⁰* mutant relative to WT larval CNSs ($p_{adj} < 0.05$ and $|\text{fold change}| > 1.5$) ⁴⁹ are highlighted in pink and labeled with a letter on the gene track on the right, and the fold changes are indicated in the table below. Dotted rectangles show zoomed-in views of 10 kb regions around meta-loop anchors. Note that the scale for ChIP-seq tracks is kept constant between WT (left) and *CTCF⁰* mutants (right).
- C. Ventral views of embryos of indicated genotypes (rows) and stages (columns) stained by immuno-RNA-FISH with a probe antisense to *Mp* mRNA (green), α -Elav (pink) and DAPI (anterior, left; posterior, right; merged images on top). Embryos were staged according to developmental landmarks indicated above (as in Fig. 8E). Specific RNA-FISH signal appears as dots; string-like and broad green non-specific signal is visible in trachea and epidermis. Filled/empty arrowheads point to wildtype/reduced *Mp* RNA-FISH signal. The phenotypes shown here (described in the results) were observed in 100% of embryos (≥ 6 embryos imaged per genotype and stage).

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Maria Cristina Gambetta (mariacristina.gambetta@unil.ch).

Materials Availability

All sequencing data (Micro-C, Hi-C, scATAC-seq, scRNA-seq) that support the findings of this study were deposited in Gene Expression Omnibus with accession code GSE214707 (for Hi-C, scATAC-seq, scRNA-seq) [reviewer login: link = <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214707>, reviewer token = qhqzyccebpexbab] and GSE228095 [reviewer login: link = <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE228095>, reviewer token = whkngugoxhgnfid] (for Micro-C). All other relevant data needed to evaluate the conclusions in this study are present in the article and/or its Supplemental Figures and Supplemental Tables.

Data and Code Availability

Custom code used for the analysis is available at <https://github.com/gambettalab/meta-loops-2022>.

All software used as described in the Methods to map, visualize and analyze data is published open source and freely available for download following the links in Key Resources Table.

REMOVE UNTIL THE END OF PARAGRAPH: in the following links: “HiGlass v1.11.8 [<https://higlass.io/>]”; “distiller v0.3.3 [<https://github.com/open2c/distiller-nf>]”; “cooler v0.8.11 [<https://github.com/open2c/cooler>]”; “R v4.2.0 [[https:// www.R-project.org/](https://www.R-project.org/)]” with packages “ggplot2 v3.3.6 [<https://ggplot2.tidyverse.org/>]”, “data.table v1.14.2 [<https://CRAN.R-project.org/package=data.table>]”, “EBImage v4.38.0 [<http://bioconductor.org/packages/EBImage>]”, “patchwork v1.1.1 [<https://CRAN.R-project.org/package=patchwork>]”, “rhdf5 v2.40.0 [<https://github.com/grimbough/rhdf5>]”, “fastcluster v1.2.3 [<https://CRAN.R-project.org/package=fastcluster>]”; “LiftOver [<http://genome.ucsc.edu/cgi-bin/hgLiftOver>]”; “bedtools merge [<https://bedtools.readthedocs.io/en/latest/content/tools/merge.html>]”; “pairtools [<https://doi.org/10.5281/zenodo.2649383>]”; “3D-Distance-Tool [https://imagej.nih.gov/ij/macros/tools/3D_Distance_Tool.txt]”; “CNEr Bioconductor package [<http://hgdownload.cse.ucsc.edu/admin/exe/>]”; “LASTZ [<https://www.bx.psu.edu/~rsharris/lastz/>]”.

Experimental Model and Study Participant Details

Drosophila melanogaster and *virilis* were cultured under standard laboratory conditions. Samples were prepared as described in the Methods Details. Embryos and larvae were not sexed, but adults were analyzed in 50:50 pools of females:males.

METHOD DETAILS

Micro-C

Wing imaginal discs and larval brains were dissected from *y w* or *GAF^{ΔBTB}* homozygous third instar larvae, and adult brains were dissected from *y w* 5-day old adult animals in PBST (0.1% Triton-X in PBS). Tissues were fixed in 1 mL of 1% formaldehyde in PBST for 15 min at room temperature. The fixation was quenched by adding 370 μ L of 2 M Tris-HCl pH 7.5 and washed with PBST 3 times. 60 discs or 30 brains were collected for each biological replicate. The tissue was then further fixed with 1 mL of 3 mM DSG and 3 mM EGS at room temperature for 45 min. The fixation was stopped by adding 370 μ L of 2 M Tris-HCl pH 7.5 and washed with PBST 3 times for 5 min. The tissue was resuspended in 500 μ L buffer MB1 (50 mM NaCl, 10 mM Tris, 5 mM MgCl₂, 1 mM CaCl₂, 0.2% NP-40, 1x protease inhibitor cocktail) and permeabilized on ice for 20 min. The tissue was rinsed twice in 1 mL MB1 and digested in 500 μ L MB1 with 1.5 μ L (30 U) MNase at 37 °C for 10 min. The digestion was terminated by 4 mM EGTA and incubating at 65 °C for 10 min. After two washes in MB2 (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂), the tissues were treated with 25 U T4 PNK at 37 °C for 15 min and 25 U Klenow Fragment at 37 °C for 15 min, for chewing 3' ends for biotin labeling in 50 μ L buffer containing 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂. Biotin labeling was performed in 75 μ L 50 mM NaCl, 10 mM Tris, 10 mM MgCl₂, 100 μ g/mL BSA, 2 mM ATP and 5 mM DTT and 67 μ M of each nucleotide (dATP and dCTP were biotinylated) at 25°C for 45 min and terminated with 30 mM EDTA incubating at 65 °C for 20 min. After rinsing in 200 μ L MB3 (50 mM Tris-HCl pH 7.5, 10 mM MgCl₂) 3 times, proximity ligation was performed in 250 μ L ligation mix (50 mM Tris, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, 100 μ g/mL BSA) with 5000 U T4 ligase. Unligated biotin-nucleotides were removed with 500 U Exonuclease III in 100 μ L of NEB1 (10 mM Bis-Tris-Propane-HCl pH 7, 10 mM MgCl₂, 1 mM DTT). The sample was then reverse crosslinked in 1 mg/mL Proteinase K and 1% SDS at 65 °C overnight. DNA was extracted with Phenol:Chloroform:Isoamylalcohol, ethanol-precipitated, resuspended in 50 μ L of 10 mM Tris-HCl pH 8.5, treated with 2 μ L RNase A at 37 °C for 30 min and purified on a ZymoClean spin column. Biotin-labeled DNA was bound to Dynabeads MyOne Streptavidin C1 beads, incubated with BW buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl) at room temperature for 20 min in a total volume of 100 μ L, and washed twice with BW at 55 °C for 5 min. The sample was rinsed once in EB buffer (10 mM Tris-HCl pH 7.5) and resuspended in 25 μ L EB buffer. 3.5 μ L of End Prep Reaction Buffer and 1.5 μ L End Prep Enzyme Mix were added and the reaction was incubated at 20 °C for 30 min for end repair, then heat-inactivated at 65°C for 30 min. The following components were added for adapter ligation: Adapter for Illumina 0.5 μ L, Ligation Master Mix 15 μ L, Ligation Enhancer 0.5 μ L. After incubating at 20°C for 30 min, 15 μ L of USER enzyme was added and the reaction was incubated 37°C for 15 min. The reaction was washed in BW buffer once at 55 °C for 5 min and resuspended in 24 μ L EB buffer. 25 μ L of 2X KAPA HiFi Hot Start Mix, 0.5 μ L of 10 μ M universal reverse primer and 0.5 μ L of 10 μ M PE2.0 primer were added to the beads and PCR-amplified for 12 cycles. The library was finally purified with 0.9x Ampure XP beads. Micro-C libraries were sequenced on a NovaSeq S1 100nt Flowcell v1.5 by the Princeton University Genomics Core Facility.

Micro-C analysis

The paired reads were mapped to the dm6 reference by *bwa* with *mem -SP5M -t8*. Aligned reads were parsed, sorted and deduplicated by *pairtools* (<https://doi.org/10.5281/zenodo.2649383>) with default settings. The deduplicated pairs were then selected with '(pair_type == "UU") or (pair_type == "UR") or (pair_type == "RU")' by *pairtools*. After being split and indexed by *pairtools* with default parameters, the output pairs were processed and normalized into multi-resolution matrices with default settings by *Cooler* ⁶⁶.

FACS of embryonic neurons, glia, neuroblasts for Hi-C

Embryos were collected on yeasted apple juice agar plates from 8 cages each with 25 grams of flies for 2 hours, then aged at 25°C until the embryos were 6-8, 10-12 or 14-16 hours old. Embryos were then washed off the collection plates, dechorionated by stirring in 50% bleach for 2.5 minutes, extensively rinsed with water, dried by blotting embryos in 125 μ m Nitex membrane squares on paper towels, transferred into 50 ml Falcons containing 10 ml of crosslinking solution (50 mM Hepes pH 8, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl) and 30 ml of heptane, and fixed by vigorous shaking for 15 minutes at room temperature. Falcons were centrifuged at 500 g for 1 minute, and crosslinking solution and heptane were decanted. Formaldehyde crosslinking was quenched by vigorously shaking embryos for 1 minute in 50 ml of 125 mM glycine in PBS with 0.1% Triton X-

100. Falcons were centrifuged again, and the buffer was decanted. Embryos were rinsed in embryo wash (PBS, 120 mM NaCl, 0.1% Triton X-100), dried by blotting in 125 μ m Nitex membrane squares on paper towels, transferred to cryovials, weighed, snap-frozen in liquid nitrogen and stored at -80°C . An aliquot of each staged collection was devitilized by shaking in eppendorfs containing 0.5 ml of heptane and 0.5 ml of methanol, washed in methanol, rehydrated in 0.5 ml of methanol and 0.5 ml of PBT, rinsed with PBT, stained with DAPI, and mounted on coverslips for observation on a fluorescent microscope with a 20x objective. The stages of at least 100 embryos were counted to estimate the proportions of each stage present in the collection.

For FACS, 0.5 g aliquots of 2 independent biological replicates of 6-8 or 10-12 or 14-16 hour old embryos prepared as above were processed. Frozen embryos were dounce-homogenized in 10 ml of cold homogenization buffer (15 mM Tris-HCl pH 7.4, 15 mM NaCl 60 mM KCl, 340 mM sucrose, 0.2 mM EDTA pH 8, 0.2 mM EGTA pH 8, 1x complete protease inhibitor cocktail) with 20 loose pestle strokes and 10 tight pestle strokes. The lysate was filtered through 2 layers of miracloth rotated by 90° to each other's grain. Nuclei were pelleted by centrifuging at 3'000 g for 10 minutes at 4°C , then blocked in 1 ml of MUSE incubation buffer with 0.1% NP-40 in 2 ml eppendorfs by rotating for 1 hour at 4°C . Primary antibodies against Worniu (Abcam ab196362), Repo (Developmental Studies Hybridoma Bank mouse monoclonal 8D12), or Elav (Developmental Studies Hybridoma Bank mouse monoclonal 9F8A9) conjugated to MUSE oligos were added at a 1 in 1'000 dilution and incubated with nuclei overnight at 4°C . The next day, immunostained nuclei were washed twice in MUSE wash buffer, incubated with MUSE primary amplifiers for 1 hour, washed twice in MUSE wash buffer, incubated with MUSE secondary amplifiers for 1 hour, washed twice in MUSE wash buffer, incubated with MUSE read-out probes for 30 minutes, washed twice in 2xSCCT (2x SCC pH 7, 0.1% Triton X-100) and incubated in 2xSSCT with DAPI overnight at 4°C . The next day, nuclei were passed up and down 10 times through a 20G needle and 10 times through a 22G needle, filtered through a 35 μ m nylon mesh, and FAC-sorted into PBT (PBS, 0.1% Triton X-100, 1x complete protease inhibitor cocktail) on a Beckman Coulter Astrios EQ instrument. Aliquots of 1 million nuclei were transferred to eppendorfs and pelleted by centrifuging 5'000 g 5 minutes at 4°C in a swinging bucket rotor. PBT was removed leaving approximately 20 μ l in the tube, and nuclei were snap-frozen and stored at -80°C .

Hi-C

About 10 dissected adult CNSs or 30 dissected third instar larval CNSs per biological replicate (2-3 biological replicates per genotype, except for the neuroblasts from 6-8 hour old embryo sample for which only a single biological replicate was used) were crushed in RPMI supplemented with 10% fetal bovine serum using a micro-pestle. Nuclei were fixed in 1% (v/v) formaldehyde for 10 minutes at room temperature. Crosslinking was stopped by adding 200 mM glycine, then nuclei were washed in PBS and snap-frozen for -80°C storage.

For Hi-C, CNS nuclei prepared above from dissected CNSs or 1 million FAC-sorted embryonic nuclei per replicate were restricted with MseI and Csp6I, restricted ends were marked with biotin, and then ligated. DNA was purified by proteinase K digestion and reverse crosslinking at 65°C for 6 hours, then sonicated in AFA microtubes in a Covaris S220 sonicator, and purified on SPRIselect beads (Beckman Coulter B23319). DNA was end-repaired, A-tailed and ligated to barcoded adapters using the NEBNext Ultra II DNA Library Prep kit for Illumina (NEB E764S), then enriched for pairwise DNA junctions by biotin pull-down using Dynabeads MyOne Streptavidin T1 (Invitrogen 65601) beads following the manufacturer's instructions. Libraries were amplified using KAPA HiFi HotStart Ready Mix (Roche KK2501) and purified on SPRIselect beads. 4 nM equimolar pools of multiplexed Hi-C libraries were subjected to 150 bp paired-end sequencing on HiSeq4000 instruments.

Hi-C analysis

FASTQ files were mapped to the reference genome using distiller v0.3.3 (<https://github.com/open2c/distiller-nf>). The following reference genomes were used: FlyBase version FB2020_05, dm6 (dm6) for *Drosophila melanogaster*, and GSE120751_Dvir_HiC.fa.gz downloaded from <ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE120nnn/GSE120751/suppl/>⁶³ for *Drosophila virilis*. For each sample, unique pairs of reads with high mapping quality scores on both sides (MAPQ \geq 30) from all biological replicates were aggregated into contact matrices using cooler⁶⁰ v0.8.11 (<https://github.com/open2c/cooler>) and stored into multiresolution cooler files (1kb, 2kb, 4kb, 8kb, 16kb, 32kb, 64kb, 128kb, 256kb, 512kb). All contact matrices

were normalized using the iterative correction procedure⁶⁴ implemented in cooler. Hi-C maps were visualized in HiGlass⁶⁷ and R.

Meta-loop calling

Meta-loops were called on WT *D. melanogaster* third instar larval CNS Hi-C maps (data from Kaushal et al. 2021⁴⁹) or on WT *D. virilis* adult CNS Hi-C maps. Numbers of Hi-C read pairs between pairs of restriction fragments for the 3 biological replicates of WT larval CNSs were downloaded from Gene Expression Omnibus accession GSE146750 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146750>). Using cooler v0.8.11 (<https://github.com/open2c/cooler>), read pairs from all replicates were pooled, binned at 2 kb (for *D. melanogaster*) or 4 kb (for *D. virilis*) and normalized using iterative correction. The per chromosome expected number of Hi-C contacts as a function of genomic distance was obtained by evaluating the mean number of normalized Hi-C contacts (considering only pairs of bins within the selected chromosome arm) as a function of genomic distance and smoothing the resulting curve with a spline [R function `smooth.spline()` applied to $\log(\text{mean count})$ vs $\log(\text{distance})$]. The observed-over-expected matrix was then generated by dividing the normalized Hi-C contact matrix by the expected number of Hi-C contacts at the corresponding genomic distance. Loop calling on the observed-over-expected Hi-C matrix occurred in two main steps.

In step 1, regions with high Hi-C contacts were identified. Pairs of bins satisfying all of the following conditions were selected: (1) Observed-over-expected matrix entry for this pair of bins >30 (i.e. contacts are 30-fold higher than expected for pairs of bins at this distance). (2) Normalized Hi-C count matrix entry for this pair of bins is larger than the per chromosome expected number of Hi-C contacts at distance 100 kb (i.e. the absolute normalized count is at least as high as the normalized count between bins separated by 100 kb). (3) Both bins are not in a low coverage region. Low coverage regions were determined by evaluating the total number of raw Hi-C counts per 128 kb bin for all 128 kb bins in the genome and computing their median and mad (median absolute deviation). Low coverage 128 kb bins are defined as those with counts lower than $\text{median} - 2 * \text{mad}$. Low coverage 2 kb (for *D. mel.*) or 4 kb (for *D. vir.*) bins are those within a low coverage 128 kb bin. Retained pairs of 2 kb (for *D. mel.*) or 4 kb (for *D. vir.*) bins were then clustered using single linkage hierarchical clustering with Euclidean distance in the plane (bin1 position, bin2 position) and using a cutting distance of 5 bins. Only clusters satisfying all of the following criteria were retained (other clusters were considered as noise): (1) The cluster contains at least 10 selected pairs of bins. (2) The cluster bounding box [in the (bin1 position, bin2 position) plane] measures at least 5 bins in both directions.

In step 2, loops were identified in each region of high Hi-C contacts. For each cluster retained in step 1, we used image analysis methods to detect peaks (i.e. meta-loops) in the observed-over-expected matrix. An observed-over-expected image was created from the observed-over-expected matrix by extracting all pairs of bins within the bounding box of clusters extended by 20 bins on each side. Invalid bins (i.e. bins flagged as low coverage and ignored during normalization by cooler) were excluded from the image. A second image with normalized Hi-C counts was also created in the same way. To reduce noise, both images were blurred using a Gaussian filter with a standard deviation of 1 pixel (2 kb bin for *D. mel.* or 4 kb for *D. vir.*). Local maxima in the blurred observed-over-expected image were found as the global maxima within each of the regions obtained by watershed segmentation (using EImage⁶⁸ function `watershed()` with flooding threshold 5). To mitigate the loss of precision due to blurring, the position of each local maximum was further adjusted to the pixel with the highest non-blurred observed-over-expected value within a 3x3 pixel neighborhood. Local maxima satisfying all of the following conditions were retained whereas others were considered as noise: (1) The value of the blurred observed-over-expected image at the position of the local maximum is >30 (i.e. 30-fold contact increase compared to expected). (2) The value of the blurred observed-over-expected image at the position of the local maximum is larger than half of the global maximum in this image. (3) The value of the blurred normalized Hi-C count image at the position of the local maximum was larger than the per chromosome expected number of Hi-C contacts at distance 100 kb. (4) The position of the local maximum is inside the bounding box of the cluster (i.e. not within the 20 bins added around the bounding box).

The resulting meta-loop anchors (i.e. local maxima retained after step 2) in *D. melanogaster* were associated with the presence of an accessible chromatin peak defined by DNase-seq in FACS-purified neurons (immunostained with anti-Elav antibody) from 10-12 hour old embryos from Reddington et al. 2020⁴⁰. Therefore, the coordinates (chr, start, end, summit) of the nearest 10-12h neuron DNase hypersensitivity site (DHS) were attributed to each loop anchor (the distance was measured between the center of the loop anchor 2 kb bin and

the DHS summit). If more than one DHS summit overlapped the 2 kb bin of the loop anchor, then all overlapping DHSs were conservatively kept, and the start, end and summit positions were taken as the minimum, maximum and mean of all overlapping DHS start, end and summit positions respectively.

Meta-loops automatically called in *D. melanogaster* (58 total) with the algorithm described above were then further manually curated. 51/58 (88%) automatically called meta-loops were validated by manual curation and 7/58 were considered false-positives. False positive meta-loops appeared to be artifacts arising by mapping or chromosomal rearrangements rather than pairwise associations of distal TADs and lacked a chromatin accessibility peak underlying the Hi-C interaction peak. 7 manually curated meta-loops were not automatically called, considered false-negatives, and added to our final list of 58 curated meta-loops. False negatives were missed when interaction peaks were “drowned” in regions of high interaction frequency, for example within a meta-domain with particularly high interactions or when located close to another meta-loop. Indeed, 7/7 (100%) meta-loops that we classified as false-negatives are present in the same meta-domain as an automatically called meta-loop. Each curated loop anchor was manually matched to an overlapping 10-12h neuron DHS (or several DHSs in case of doubt) as for the automatically called loops. Finally, 10 intra-TAD loops involving meta-loop anchors were manually annotated and added to the final list of 58 meta-loops + 10 intra-TAD loops = 68 loops involving meta-loop anchors shown in Fig. S1 and listed in Table S2. Note that we never observed a meta-loop anchor to loop to another locus than a meta-loop anchor.

Using a similar procedure as described above for manual curation of meta-loops automatically called in *D. virilis*, 51/54 (94%) automatically called loops were validated and 3/54 were considered false-positives (because they were likely artifacts or because one was arguably an intra-TAD loop). 4 meta-loops were considered false-negatives, leading to a final list of 58 curated meta-loops in *D. virilis*. The coordinates of called meta-loop anchors were then aligned to overlapping DNase-seq peaks from old (25-28 hours after egg-lay) *D. virilis* embryos⁶⁴. Specifically, the minimum and maximum of start and end coordinates of overlapping DNase-seq peaks were taken as the anchor coordinates. Finally, 6 intra-TAD loops involving meta-loop anchors were then added to obtain a total of 64 considered loops involving meta-loop anchors listed in Table S3.

Loop anchor classification

Promoter-proximal loop anchors were defined as those within ± 200 bp of a protein coding gene TSS (FlyBase version 6.36 for *D. melanogaster*, GSE120751 transcript annotations in GTF format for *D. virilis*). When an anchor was within ± 200 bp of >1 protein coding gene TSS, all these genes were attributed to the anchor. The remaining loop anchors were defined as intergenic, and the closest TSS was attributed to each intergenic anchor.

Differential analysis of Micro-C or Hi-C interactions

For the differential analyses presented in Figs. 7B and 8B, meta-loop coordinates were standardized to 5 kb in a bedpe format. Cooler files were dumped into the interaction matrix at 100 bp resolution using the dump function from cooler. The dumped matrix was then intersected with the meta-loop coordinates using the pairToPair function from bedtools with “both” mode. Total counts of each meta-loop were summed. The count lists for individual samples were summarized as a count table in R. The count table was used for differential analysis using the DESeq2⁶⁹ package in R. Statistically significantly changed loops were filtered using an adjusted P value ≤ 0.01 . The volcano plot was generated with ggplot2.

Analysis of three-way interactions between meta-loop anchors

To determine if multiple interactions involving a single meta-loop anchor occur concurrently, we re-analyzed sequencing reads from our published Hi-C experiments in WT larval CNSs⁴⁹ to infer information on higher-order chromatin interactions. We relied on the fact that Hi-C experiments were performed using in-nucleus ligation and each sequenced DNA molecule possibly obtained via several ligations, representing concurrent interactions in a single nucleus. We used a local sequence aligner, bwa mem⁷⁰ version 0.7.17 with options -E50 -L0 -5 for mapping. This aligner was particularly suitable for chimeric reads that had more than one alignment that each uniquely mapped to a distinct region in the genome. To extract the alignments, we used pairtools parse from the pairtools suite⁷¹ version 1.0.2. We detected cases in which the same DNA fragment was sequenced in both reads forming the paired-end read, and further merged such alignments. The specific

option used, --walks-policy all --add-columns pos5,pos3 --no-flip, ensured that all alignments are reported with their full genomic coordinates in the order in which they were sequenced.

We merged all ~597M processed paired-end reads from the three replicates of WT larval CNS Hi-C data. ~62M (10.7%) of them had three or more alignments, allowing for quantification of three-way interactions. Only ~7.5M (1.3%) of paired-end reads had more than three alignments in total, hence we did not pursue the analysis of 4-way or higher-order interactions. We further followed an approach proposed by Bintu et al.⁷² to quantify the cooperativity of three-way interactions. We focused on 21 triplets of meta-loop anchors, called looping triplets, such that all three pairs of anchors within the triplet form a loop. Since all the loops are constrained to a single chromosome arm, all the looping triplets are too.

Let us consider one such looping triplet, and denote its three anchors by A, B, and C, according to their chromosomal order. We denote by n_A the number of paired-end reads involving anchor A, i.e. paired-end reads having an alignment that maps to the genome no more than 2 kb (i.e. the resolution of the Hi-C maps) away from anchor A. Similarly, we denote by n_{AB} the number of paired-end reads involving both anchors A and B, and we generalize this notation for any anchors and their higher number. To test whether locus A facilitates or inhibits the interaction between loci B and C, we first calculated $f(C|B) = n_{BC} / n_B$ as the fraction of reads involving B that also involve C. Similarly, we can calculate $f(C|AB) = n_{ABC} / n_{AB}$ as the fraction of reads involving A and B that also involve C. The fold change between these two fractions, namely the triplet cooperativity factor $f(C|AB)/f(C|B)$, quantifies the impact of A on the interaction between B and C. The above-described measure has two important properties. First, because we consider A, B, and C in their chromosomal order, the anchor A is not between B and C, and the effect of A is beyond polymeric nature of chromatin fiber. Second, it algebraically follows that $f(C|AB) / f(C|B) = (n_{ABC} \cdot n_B) / (n_{AB} \cdot n_{BC}) = f(A|BC) / f(A|B)$, and the cooperativity factor is independent of the direction of traversing the chromosome.

Positive values of the triplet cooperativity factor indicate that the two extreme anchors (A and C) facilitate interactions between the two other anchors. We systematically compared the values of the triplet cooperativity factor for all 21 looping triplets of meta-loop anchors, and for a set of control triplets of genomic regions defined as follows. For each looping triplet of anchors A, B, and C (in this chromosomal order), we defined anchor C' located to the left of A and positioned such that the distance between B and C is equal to the distance between C' and A. The control triplet of anchors C', A, and B is characterized by the same inter-anchor distances as the looping triplet A, B, and C, but typically involved very few alignments on C' side. Hence, we defined the anchor C'' by extending the anchor C' by 100 kb on both sides, noting that in no case did the anchors C'' and A overlap. From each looping triplet of anchors A, B, and C, we derived two control triplets of regions, namely C'', A, and B, as well as B, C, and A'' obtained correspondingly.

Loop strength measurement

Strengths of loops involving meta-loop anchors (plotted in Figs. 3B, S1B and S6D) were measured as the normalized Micro-C or Hi-C count divided by the average normalized Micro-C or Hi-C count at the same genomic distance and on the same chromosome as the pair of bins containing the loop anchors. The normalized Micro-C or Hi-C count was estimated as the average normalized Micro-C or Hi-C count over 3x3 pairs of bins centered on the loop (i.e. the bin containing the loop anchors and the nearest neighbor bins). The 3x3 neighborhood was used to mitigate the problem of missing data (i.e. bins flagged as low coverage and ignored during normalization by cooler), in particular for samples in which the loop is weak or absent. All measurements were done using 1.6 kb Micro-C resolution matrices and 2 kb resolution Hi-C matrices.

GO term enrichment analysis

For Gene Ontology (GO) term enrichment analysis, we extracted FlyBase IDs of 35 genes attributed to meta-loop promoter anchors. We also considered the set of 62 genes attributed as the closest to intergenic anchors. Each of the input gene sets was compared to the universe of all *Drosophila melanogaster* coding genes, using enrichGO function from the clusterProfiler R package⁷³ v3.14.0 and gene annotations from the org.Dm.eg.db Bioconductor package version 3.10.0. In our over-representation analysis, we considered terms from all three GO sub-ontologies (Biological Process, Molecular Function, and Cellular Component), and identified significantly enriched terms with at FDR of 0.05, adjusting for multiple hypothesis testing using Benjamini-Hochberg method. The enrichment was plotted as a dotplot graph from the clusterProfiler package.

Terms were ordered by gene enrichment ratio, i.e. the fraction of genes in the input set associated with the given GO term. Up to 20 terms with the highest gene enrichment ratio are shown.

RNA-FISH

DNA templates for *5-HT7*, *beat-IV*, *DIP-ε*, *Mp*, *nolo* or a reporter gene used in the transgenic enhancer assay were PCR-amplified from genomic DNA or a template plasmid, respectively, using primer sequences in Table S6. These inserts were cloned upstream of a T7 RNA polymerase promoter oriented such that it transcribes an antisense transcript to the mRNA. Labeled RNA probes were then generated from template DNA by in vitro transcription with Dig-UTP labeling mix (Roche 11277073910) and T7 RNA polymerase (Roche 10881767001). After DNase I digestion for 20 minutes at 37°C, probes were fragmented by incubating 20 minutes at 65°C in 60 mM Na₂CO₃, 40 mM NaHCO₃ pH 10.2, precipitated in 300 mM sodium acetate pH 5.2, 1.25 M LiCl, 50 mg/ml tRNA and 80% EtOH, resuspended in 50% formamide, 75 mM sodium citrate pH 5, 750 mM NaCl, 100 µg/ml salmon sperm DNA, 50 µg/ml heparin and 0.1% Tween20, and stored at -20°C. Embryos were fixed in 4% paraformaldehyde for 30 minutes at room temperature, washed, and then stored in 100% MeOH at -20°C. Samples were rehydrated in PBS with 0.1% Tween20, post-fixed in 4% paraformaldehyde for 20 minutes at room temperature, progressively equilibrated to hybridization buffer (50% formamide, 75 mM sodium citrate pH 5, 750 mM NaCl) and heated to 65°C. RNA probes were diluted 1:50 in hybridization buffer, denatured at 80°C for 10 min then placed on ice, and added to the samples overnight shaking at 65°C. Samples were washed 6 times 10 minutes in hybridization buffer at 65°C, then progressively equilibrated to PBS with 0.1% Triton X-100. Samples were incubated overnight at 4°C in anti-Dig peroxidase (Roche 11207733910) diluted 1:2000 in PBS, 0.1% Triton X-100, 1x Western blocking reagent (Sigma 1921673). Samples were washed 6 times 10 minutes in PBS with 0.1% Tween20, labeled with Cyanine 3 tyramide in the TSA Plus kit (Perkin Elmer NEL753001KT) for 3 minutes at room temperature, washed 6 times 10 minutes in PBS with 0.1% Tween20, and finally mounted with DAPI to stain DNA. Images were acquired on a Zeiss LSM 880 microscope with a 20x objective and visualized with Fiji software v2.1.0/1.53c.

DNA-FISH

5 kb PCRs from genomic DNA using primers indicated in Table S6 were cloned into a plasmid. 2 µg of miniprep plasmid DNA was nick-translated (Abbott Molecular 07J100-001) following the manufacturer's protocol in the presence of 10 µM aminoallyl-dUTP-XX-ATTO-488 (Jena Bioscience NU-803-XX-488-S) or aminoallyl-dUTP-ATTO-550 (Jena Bioscience NU-803-550-S) at 15°C for 5 hours, and stored at -20°C. Prior to DNA-FISH, both probes per two-color DNA-FISH experiment were co-precipitated in the presence of salmon sperm DNA, and resuspended in 100% formamide. 0-24 hour old WT embryos were dechorionated and fixed in 4% formaldehyde for 20 minutes at room temperature, washed, and then stored in 100% MeOH at -20°C. Samples were rehydrated in 2xSSCT (30 mM sodium citrate dihydrate pH 7.4, 300 mM NaCl, 0.1% Tween20), treated with RNase A for 30 minutes at room temperature, washed in 2xSSCT, progressively equilibrated to 50% formamide in 2xSSCT, and incubated for 2 hours at 37°C. DNA-FISH probes were mixed with the embryos in 25 µl total of 2xSSC, 10% (w/v) dextran sulfate, 50% formamide, heated to 80°C for 10 minutes, then incubated at 37°C overnight in the dark. Embryos were washed in 50% formamide in 2xSSCT, then 20% formamide in 2xSSCT, then 2xSSCT at 37°C. Embryos were then immunostained overnight at 4°C with anti-Elav rat monoclonal antibody (clone 7E8A10 from Developmental Studies Hybridoma Bank) diluted 1:50 in PBS, 0.1% Tween20, 1x Western blocking reagent (Sigma 1921673). Samples were washed, then incubated in the dark for 2 hours at room temperature with Alexa 647 anti-rat IgG (Thermo Fisher A21247) diluted 1:200 in PBS, 0.1% Tween 20, 1x Western blocking reagent (Sigma 1921673). Samples were washed in PBS with 0.1% Tween20, and finally mounted with DAPI to stain DNA. Images were acquired on a Zeiss LSM 880 microscope with a 63x oil objective and visualized with Fiji software v2.1.0/1.53c. 3D distances between two color DNA-FISH spots were counted with the 3D distance tool Fiji plugin downloaded from https://imagej.nih.gov/ij/macros/tools/3D_Distance_Tool.txt.

RNA-DNA-FISH

For RNA-DNA-FISH experiments, RNA-FISH was performed first, RNAs were then digested by RNase A (the RNA-FISH signal perdures because it arises from fluorescent labeling of tyrosine residues of proximal proteins), and DNA-FISH was subsequently performed as described above.

HCR-RNA-FISH

The third instar larval tissue was dissected in PBST (0.1% Triton-X in PBS) and fixed in 4% formaldehyde diluted in PBST for 20 min at room temperature. The tissue was briefly rinsed in PBST 2 times, treated with 100% methanol, 75% methanol in PBST, 50% methanol in PBST and 25% methanol in PBST for 5 min at room temperature each time, then washed in PBST 3 times for 5 min at room temperature. The tissue was further processed following the protocol described in <https://files.molecularinstruments.com/MI-Protocol-RNAFISH-GenericSolution-Rev8.pdf> with the following modifications. The hybridization buffer, wash buffer and amplification buffer used in the protocol were purchased from Molecular Instruments. The probe stocks were designed and ordered from Molecular Instruments. For hybridization, oligos were mixed at 16 nM final concentration in the hybridization buffer. For amplification, the incubation time was reduced to 1.5 hours compared to the original protocol.

RT-qPCR on larval CNSs, adult heads or whole-bodied flies

For RT-qPCR on larval CNSs shown in Fig. 7E, 15 larval CNSs of WT (*y w*) or *GAF^{ΔBTB}* mutants were dissected in biological triplicates and incubated in 500 μl TRIzol (Invitrogen 15596026) at room temperature for 10 minutes. RNA was extracted following the manufacturer's instructions and dissolved in 10 μl of 10 mM Tris-HCl pH 8. 1 μg of RNA was reverse transcribed with SuperScript IV (Invitrogen 18091050) with an oligo dT primer following the manufacturer's instructions. qPCR was performed with SYBR green (Invitrogen 4367659) on a ViiA 7 Real-Time PCR System (ThermoFisher).

For RT-qPCR on adult flies or adult heads shown in Figs. S6A-S6B, total RNA was extracted from 10 heads (5 female, 5 male) or 4 flies (2 females, 2 males) of WT (*w¹¹¹⁸*), *nolo^{ΔA23}* or *beat-IV^{ΔA65}* mutants in biological triplicates with TRIzol as described above. 500 ng of RNA were reverse transcribed with GoScript (Promega A2801) with random hexamers following the manufacturer's instructions. qPCR was performed with GoTaq (Promega A600A) on a QuanStudio 6 Flex Real-Time PCR System (ThermoFisher).

Fold changes of gene expression in mutants relative to WT controls were calculated using the delta-delta Ct method, referring to the *GAPDH* signal (for *GAF^{ΔBTB}* mutants) or to *RpL15* (for *nolo^{ΔA23}* or *beat-IV^{ΔA65}* mutants) and normalizing by the mean expression level of WT for each primer pair. A two-sided unpaired heteroscedastic t test was performed to test the significance of differential expression in knock-out versus WT backgrounds.

CRISPR/Cas9-mediated intergenic anchor deletions

beat-IV^{ΔA65} and *nolo^{ΔA23}* mutants were generated by CRISPR-Cas9-mediated genome editing using small guide RNAs (sgRNAs) flanking the regions chosen for deletion: 199 bp (dm6 coordinates chr3R:28,654,618-28,654,816) for *beat-IV^{ΔA65}*, or 814 bp (dm6 coordinates chr2L:20,120,764-20,121,578) for *nolo^{ΔA23}*. Up to six sgRNAs per targeted intergenic anchor were cloned (using primers listed in Table S6) downstream of U6:1, U6:2 and U6:3 promoters⁷⁴, then assembled into a single transgene and integrated by site-specific recombination into the ZH86Fb landing site⁷⁵. All *Drosophila* injections were performed by FlyORF, Zurich. Males expressing both sgRNAs and Cas9 in their germline (*nanos-Cas9*)⁷⁴ were crossed to balancer females, and crosses of single sons were set up with balancer females. After three days, the males were sacrificed and genotyped by PCR. Males and females harboring independently isolated CRISPR knock-out alleles were crossed and resulting trans-heterozygous larvae were analyzed by Hi-C and scRNA-seq.

nolo CRISPR/Cas9-mediated knock-outs and RNAi

nolo^{KO} alleles were generated by crossing mothers expressing *nanos-Cas9*⁷⁴ to fathers transgenically expressing two *nolo* sgRNAs targeting exon 2 (Bloomington stock 83042). Transheterozygotes for two independently isolated *nolo^{KO}* alleles were used for phenotypic analysis, and *w¹¹¹⁸* animals were used as WT control. For *nolo* tissue-specific knockdowns, transgenic *UAS-Luciferase RNAi* (Bloomington stock 35788) or *UAS-nolo RNAi* mothers (Bloomington stock 55296) were crossed to fathers with transgenic Gal4 drivers expressed in neuroblasts [*worniu-Gal4* (Bloomington stock 56553)] or ubiquitously [*tubulin-Gal4* (Bloomington stock 5138)].

Drosophila viability tests

Three or four sets of 30-40 third instar larvae of desired genotypes were transferred into separate vials and the number of fully hatched adults was recorded. Viability defects in *nolo* knock-out and knock-down animals were only observed at the pupa-to-adult transitions, whereas larva-to-pupa transitions were similar to wildtype.

Analysis of neuromuscular junctions

Third instar larvae were dissected in chilled PBS and fixed for 8 minutes in Bouin's fixative. WT controls were the *w¹¹¹⁸* line. The dissected preparations were permeabilized in PBS with 0.3% Triton X-100 (PBT) for 45 minutes at room temperature, blocked in PBT with 3% goat serum for 1 hour, incubated overnight with mouse anti-vGlut⁷⁶ at 4°C, and finally incubated with secondary antibody and cellular dyes (Alexa 647 goat anti-HRP, Jackson 123-605-021 1:400; Alexa 555 anti-mouse IgG, Thermo Scientific A-21422 1:200) for 2 h at room temperature. Samples were mounted in Vectashield and imaged with LASX acquisition software (Leica) on a Leica SP8 confocal microscope using a 63x/1.40 NA oil HCPLAPOC52 objective with a 0.14 mm working distance. Maximum intensity projections of confocal stack images and image cropping were performed in Fiji ImageJ and images saved as .tif files. Thresholding of the resulting images was used to select regions of interest (ROIs) for area quantification, using the Default method of auto thresholding available in Fiji ImageJ. The selected ROIs were used to measure area values, which were recorded in a .csv file. One neuromuscular junction in each of ≥8 larvae were imaged per genotype.

Transgenic enhancer assays

Intergenic anchors were PCR-amplified from genomic DNA with primers from Table S6 and cloned upstream of a minimal *Drosophila* synthetic core promoter driving transcription of a reporter gene. Transgenes were inserted into the VK33 landing site on chromosome 3 by attP-attB recombination⁷⁷. Embryos homozygous for the reporter transgene were analyzed in Fig. S6G.

scATAC-seq on third instar larval central nervous systems

20 wandering third instar larvae were dissected in ice-cold Schneider's medium supplemented with 10% fetal bovine serum, washed in Rinaldini solution (0.36 mM NaH₂PO₄*H₂O, 11.9 mM NaHCO₃, 137 mM NaCl, 2.7 mM KCl, 5.6 mM glucose*H₂O), dissociated in 0.5 µg/µl collagenase I (Sigma C2674) for 1 hour at room temperature, washed in Schneider's medium, and dissociated into single cells by pipetting up and down several times with wide-bore tips. The cell suspension was passed through a 40 µm cell strainer, washed in PBS with 0.04% (w/v) BSA, and counted. Nuclei were extracted in 0.1x lysis buffer (1 mM Tris-HCl pH 7.4, 1 mM NaCl, 0.3 mM MgCl₂, 0.01% Tween20, 0.01% NP-40, 0.001% digitonin, 0.1% BSA) for 5 minutes on ice, washed in cold wash buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% Tween20, 1% BSA), resuspended in 1x nuclei buffer (provided by the manufacturer), counted, and diluted to 2500 nuclei/µl. 20000 nuclei were then loaded and processed according to the instructions of 10x Genomics Chromium single cell ATAC reagent kits solution v1.0 (CG000168 revB January 2019). scATAC-seq libraries were sequenced on an Illumina HiSeq 4000 using HiSeq 3000/4000 SBS kit reagents according to 10X Genomics recommendations (50 cycles read1, 8 cycles i7 index read, 16 cycles i5 index read, 50 cycles read2). Sequencing data were demultiplexed with bcl2fastq Conversion software (v 2.20, Illumina).

scRNA-seq on third instar larval central nervous systems

15 wandering third instar larvae were dissected in ice-cold Schneider's medium supplemented with 10% fetal bovine serum, washed in Rinaldini solution (0.36 mM NaH₂PO₄*H₂O, 11.9 mM NaHCO₃, 137 mM NaCl, 2.7 mM KCl, 5.6 mM glucose*H₂O), dissociated in 2 µg/µl collagenase I (Sigma C2674) and 2 µg/µl papain (Sigma P4762) for 50 minutes at room temperature, and dissociated into single cells by pipetting up and down several times. The cell suspension was washed in PBS with 1% BSA and RNase inhibitor (Sigma 3335399001), passed through a 35 µm cell strainer, resuspended in PBS with 1% BSA and RNase inhibitor, filtered again through a 0.35 µm cell strainer, counted and assessed for viability on a Moxi V instrument (Witec MXV102), and diluted to 1000 cells/µl. 20 µl (20000 cells) were processed according to the instructions of 10x Genomics Chromium Next GEM sc 3' reagent kits v3.1 dual index (CG000315 RevC August 2021). scRNA-seq libraries were sequenced on an Illumina NovaSeq 6000 using SBS cartridge 100C v 1.5 kit reagents according to 10X Genomics recommendations (28 cycles read1, 10 cycles i7 and i5 index reads, 90 cycles read2). Sequencing data were demultiplexed with bcl2fastq Conversion software (v 2.20, Illumina).

scATAC-seq analysis

Raw sequencing data (305 million read pairs, 14,957 cells) were processed using the Cell Ranger ATAC software v1.2.0, using a custom *D. melanogaster* reference based on FlyBase version FB2020_05, dm6 genome assembly. The read filtering and alignment, identification of transposase cut sites, detection of accessible chromatin peaks, cell calling and count matrix generation for peaks were performed by cellranger-atac count, using the default settings. The custom peak calling algorithm implemented in Single Cell ATAC fits a probabilistic model to a smoothed cut site distribution and sets a global signal threshold at the specificity level of 95%. The resulting 28877 accessible chromatin peaks, filtered peak-barcode matrix in HDF5 format, filtered fragments and per-barcode cell calling table were further analyzed in R (v3.6.3) using the R packages Seurat (v3.1.5)⁷⁸ and Signac (v0.2.5)⁷⁹.

For each cell, we computed the following quality metrics: number of fragments overlapping peaks, percent reads in peaks, TSS enrichment score and short-to-mononucleosomal reads ratio. Cells with the number of fragments overlapping peaks outside the 5%–95% quartile (672–18,642) were excluded from further analysis. The resulting count matrix of 28,877 peaks and 13,461 cells was used for latent semantic indexing (LSI). First, term frequency-inverse document frequency (TF-IDF) matrix was computed using the Signac function RunTFIDF and default implementation method. The resulting TF-IDF matrix was then decomposed using partial singular value decomposition implemented in the RunSVD function in Signac.

To obtain two-dimensional projection, we performed Uniform Manifold Approximation and Projection (UMAP) dimensional reduction, applying Seurat function RunUMAP and using LSI components 2 to 30. The first LSI component was found to be highly correlated with the total number of counts per cell, and hence excluded from the analysis as capturing technical variation (sequencing depth). Cells were clustered into 19 cell clusters by Seurat function FindClusters, using LSI components 2 to 30 and running the smart local moving (SLM) algorithm at the default resolution of 0.8. The 19 clusters were manually associated to one of three broad cell types (neuronal, neuroblast and glia) based on the accessibility of a 2 kb region upstream of the transcription start site of selected marker genes, respectively: *embryonic lethal abnormal vision (elav)*, *Notch (N)* and *reversed polarity (repo)*.

scRNA-seq analysis

Raw sequencing reads (143 million raw reads for WT, 192 million raw reads for *beat-IV^{ΔA65}*, 157 million raw reads for *nolo^{ΔA23}*) were processed with CellRanger v6.1.2 software using the *Drosophila melanogaster* reference genome BDGP6.28.101. Resulting UMI count matrices (h5 files filtered_feature_bc_matrix.h5) were imported into and analyzed in R (v. 4.0.0) using the R package Seurat (v3.1.5). Cells with less than 1000 or more than 30000 UMI counts, expressing less than 200 or more than 7000 genes, and detecting more than 3% of mitochondrial genes, were removed from the analysis. This resulted in the following total numbers of cells and genes remaining for the analysis: 4460 cells and 8905 genes in WT, 4431 cells and 8905 genes for *beat-IV^{ΔA65}*, and 5871 cells and 8698 genes for *nolo^{ΔA23}*.

All samples were first processed independently following the standard Seurat processing workflow with default settings unless specified: Log-normalization of the raw counts (NormalizeData with scale factor set to 10000), identification of the most variable genes (FindVariableFeatures, with nFeature set to top 2000 and selecting the *vst* method), and scaling the data (ScaleData, setting the center parameter to FALSE). After processing, objects were then integrated using the Seurat Canonical Correlation Analysis (FindIntegrationAnchors using CCA dimensions 1-20 and IntegrateData), and further processed applying the same steps as described above.

22 clusters were identified at a resolution of 0.5 (FindNeighbors and FindClusters, running the Louvain algorithm). Clusters were manually associated to one of five broad cell-types shown in Fig. 6A based on the expressions of selected markers shown in Figure S6C. For visualization, a 2-dimensional UMAP was generated from the first 10 principal components (RunPCA and runUMAP).

The differential expression analysis in either mutant relative to WT was performed from the integrated Seurat object for each of the 22 cell clusters (FindMarker using a Wilcoxon Rank Sum test).

Comparative analysis of meta-loops in *D. mel.* and *D. vir.* adult central nervous systems

We used CNEr Bioconductor package⁸⁰ complemented with UCSC Kent's utilities (<http://hgdownload.cse.ucsc.edu/admin/exe/>) to identify conserved elements. First, we converted *D. melanogaster* and *D. virilis* soft-masked genome sequences from FASTA format to twoBit format, and used the LASTZ sequence alignment program (<https://www.bx.psu.edu/~rsharris/lastz/>)⁸¹ with the default (HOXD70) scoring matrix to obtain their pairwise alignment. We converted the LASTZ output from .lav to .psl format using lavToPsl, and joined matching alignments located next to each other into chains using axtChain with the default (medium) linear gap costs. The chains were further sorted using chainMergeSort, filtered using chainPreNet to remove isolated chains, and combined to alignment nets using chainNetSyntenic. Final .net.axt files were obtained from the previous net and chain files using netToAxt.

We further combined the two .net.axt files obtained by taking one of the *Drosophila* genomes as query and the other one as reference, using the CNE function from the CNEr package. Conserved elements were identified by scanning the two sets of alignments using ceScan for regions with at least 80% identity over 50 alignment columns (nucleotide positions). We merged the conserved elements using cneMerge, as we perform their detection with each genome as reference, leaving intact the elements that overlap only in one of the genomes. As some of the conserved elements might be unannotated repeats which were not filtered out by soft-masking, we realigned each element sequence against the respective genome using BLAT⁸² and used blatCNE to discard the elements with more than four BLAT matches.

The filtered conserved elements were chained by placing edges between concordant elements within 10 kb distance in both genomes, as described below. We iteratively looped over pairs of elements, in increasing order of distances between them. For this purpose, the larger of the distances calculated in both genomes was taken. We chained the elements only if each of them was either unchained or at one of the ends of its respective chain, and only if the resulting chaining order would be in line with the strandedness of these elements. We kept only the chains containing at least two conserved elements.

For mapping the loop anchors across species, we extended the anchors by 1 kb in both directions to account for possible uncertainty in calling them. We then overlapped them with both the chained conserved elements and the edges between these elements. The coordinates in the other genome of these conserved elements and edges were reduced by joining them within each chain to define the homologous loci for the loop anchors. For a pair of anchors defining a loop, all possible combinations of homologous loci for both anchors were considered as candidate loops. We called a loop “conserved” if at least one such combination matched (overlapped) a loop defined in the other species.

To assess the contribution of coding sequences to loop anchor homology (Fig. S3C), we calculated gene exon coverage as follows. For each chain of conserved elements, we overlapped the conserved elements (i.e. the aligned sequences) with all annotated gene exons, and divided the total length of overlap by the total length of the conserved elements. This ratio was calculated separately in both species.

For the permutations of *D. melanogaster* loop anchors performed in Fig. S3H, we randomly paired $2 \times 68 = 136$ loop anchors rather than 79 unique anchors, to make the random pairing more realistic as some loop anchors are used multiple times. The random pairing was done within each chromosome arm separately.

To compare the identity of genes at meta-loop anchors, we used ortholog annotations from OrthoDB v10.1 (<https://www.orthodb.org/>)⁸³, taking only protein-coding genes and their orthologs within the clade Metazoa. For each *D. melanogaster* meta-loop anchor, only the gene with the closest TSS to the anchor was considered for the ortholog analysis. Further, when analyzing the *D. virilis* orthologs of these genes, we searched for all possible orthologs and considered the closest one (in terms of distance to TSS) to any of the homologous loci of the anchor. We note that this procedure, when applied to *D. melanogaster* anchor A22 with the same anchor-to-TSS distance for its two closest protein-coding genes (*Lar* and *CG46244*), led to the choice of the same ortholog in *D. virilis* (*GJ13567*). The choice of the ortholog also determined the homologous locus of the anchor.

Reuse of published datasets

FlyAtlas RNA-seq v2 data were downloaded from http://ftp.flybase.org/releases/FB2022_04/precomputed_files/genes/gene_rpkm_matrix_fb_2022_04.tsv.gz, and expression in adults (males for all tissues except ovary from females) were plotted in Fig. 1H.

For visualizing DNase hypersensitive sites (DHSs) mapped by DNase-seq in WT *D. melanogaster* embryos ⁴⁰, this published dataset was lifted over from dm3 to dm6 coordinates using the LiftOver tool (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>).

To quantify the overlap between meta-loop anchors described in this study with those of loops reported in Kc167 cells by Eagen et al. ⁴¹ or in nuclear cycle 14 embryos by Batut et al. ²⁵ and Levo et al. ²⁶, we extended the [start; stop] coordinates of each set of anchors by ± 200 bp and quantified the number of overlaps between independent studies by ≥ 1 bp in Fig. S2F. Prior to this, loop anchors published by Eagen et al. were first lifted over from dm3 to dm6 coordinates using the LiftOver tool, and separate anchors published by Eagen et al. that overlapped by ≥ 1 bp were merged to avoid overestimating loop anchor overlaps, using the “merge” command from bedtools ⁸⁴. Loop sizes plotted in Fig. 1D show distances between the summits of DNase-seq peaks under meta-loop anchors (from this study) or distances between the centers of published loop anchor coordinates ^{25,26,41}.

For visualizing scATAC-seq data in WT adult brains ⁴⁶, this published dataset was downloaded from GEO (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163697>, supplementary file GSE163697_cisTopicObject_Adult.Rds.gz). For visualizing bulk ATAC-seq data in WT adult brains ⁴⁶, the fastq file of sample GSM5503528 was downloaded from the sra repository (<https://sra-pub-run-odp.s3.amazonaws.com/sra/SRR15347354/SRR15347354>) and mapped the dm6 genome assembly with the BWA-MEM aligner.

For visualizing DNase-seq data in *D. virilis* embryos ⁶⁴, this published dataset was downloaded from ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) accessions E-MTAB-3797 and E-MTAB-9480, and remapped using the BWA tool ⁸⁵ as previously described ⁶⁴. Sizes of *D. vir.* loops involving meta-loop anchors shown in Fig. 2 represent distances between the midpoints of the DNase-seq peaks in 25-28 hour old *D. vir.* embryos that were assigned to meta-loop anchors. The same *D. vir.* genome assembly was used throughout this study (GSE120751 downloaded from <https://ftp.ncbi.nlm.nih.gov/geo/series/GSE120nnn/GSE120751/suppl/GSE120751%5FDvir%5FHiC%2Efa%2Egz>).

For visualizing Hi-C and CTCF and Cp190 ChIP-seq data in larval CNSs ⁴⁹, these published datasets were downloaded from GEO (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146752>).

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses and software used were indicated in the respective Methods Details section and figure legends. P values are indicated in the respective figures. All described replicate experiments are biological replicates. Samples were grouped according to genotype (wildtype or various mutants). The investigators were not blinded during data collection as the biological groups (genotypes) were well defined and handled in parallel. Computational analysis was performed by data scientists different from the researchers who collected the data. No data were excluded from the analyses.

For RNA-FISH experiments, approximately 20 embryos were examined per genotype over two independent experiments and only representative expression patterns observed in all animals are shown. These numbers were chosen because they revealed that phenotypes were reproducibly detected in all animals and because sample collection beyond this scale was rate-limiting. For Micro-C and Hi-C experiments, 10 adult or 30 third instar larval brains were dissected per replicate because these numbers allowed sufficient material to be amplified for next-generation sequencing library preparation with a limited number of PCR cycles to avoid over-amplification. This number was sufficient because all biological replicates were well correlated.

Table S2. Loops involving meta-loop anchors identified in *D. melanogaster* CNSs. Related to Figure 1.
Each loop associated with a meta-loop anchor annotated in *D. melanogaster* third instar larval CNS Hi-C maps is shown in rows, with dm6 coordinates and the number of published DNase hypersensitivity site (DHS) peaks in neurons of 10-12 hour old embryos⁴⁰ overlapping the loop anchor. Start/end coordinates of each meta-loop anchor correspond to start/end coordinates of the first/last DHS published by Reddington et al. 2020⁴⁰ that the anchor overlapped. The loop size (in bp), anchor type (intergenic or promoter), nearest annotated transcription start site (TSS), and the distance from the loop anchor to the TSS are shown. In the few cases in which two TSSs were within 200 bp of a loop anchor, both TSSs are indicated. Finally, the meta-domain in which the loop is found, whether the loop clustered together with other loops (clustered) or was unique (single) in the meta-domain, and the loop type (meta- or intra-TAD loop) are indicated.

Table S3. Loops involving meta-loop anchors identified in *D. virilis* CNSs. Related to Figure 2.
Each loop associated with a meta-loop anchor annotated in *D. virilis* adult CNS Hi-C maps is shown in rows, with genomic coordinates and the number of published DNase hypersensitivity site (DHS) peaks in old (25-28 hours after egg-lay) *D. virilis* embryos⁶⁴ overlapping the loop anchor. Start/end coordinates of each meta-loop anchor correspond to start/end coordinates of the first/last DHS that the anchor overlapped. Other annotations are provided as in Supplemental Table S2.

Table S4. Differential analysis of gene expression measured by scRNA-seq in larval CNSs of *nolo*^{ΔA23} mutants relative to WT, per cell cluster. Related to Figure 6.

Table S5. Differential analysis of gene expression measured by scRNA-seq in larval CNSs of *beat-IV*^{ΔA65} mutants relative to WT, per cell cluster. Related to Figure 6.

Table S6. Primers used for cloning, RT-qPCR and FISH experiments. Related to STAR Methods section.

REFERENCES

1. Rao, S.S.P., Huang, S.-C., Hilaire, B.G.S., 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, 305-309.e24. 10.1016/j.cell.2017.09.026.
2. 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.e22. 10.1016/j.cell.2017.05.004.
3. Rowley, M.J., Nichols, M.H., Lyu, X., Ando-Kuri, M., Rivera, I.S.M., Hermetz, K., Wang, P., Ruan, Y., and Corces, V.G. (2017). Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Molecular Cell* 67, 837–852. 10.1016/j.molcel.2017.07.022.
4. Paulsen, J., Ali, T.M.L., Nekrasov, M., Delbarre, E., Baudement, M.-O., Kurscheid, S., Tremethick, D., and Collas, P. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nature Genetics* 51, 1–843. 10.1038/s41588-019-0392-0.
5. Szabo, Q., Jost, D., Chang, J.-M., Cattoni, D.I., Papadopoulos, G.L., Bonev, B., Sexton, T., Gurgo, J., Jacquier, C., Nöllmann, M., et al. (2018). TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Science advances* 4, eaar8082. 10.1126/sciadv.aar8082.
6. Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C.A., Chotalia, M., Xie, S.Q., Barbieri, M., Santiago, I. de, Lavitas, L.-M., Branco, M.R., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519–524. 10.1038/nature21411.
7. Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. 10.1126/science.1181369.
8. 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. 10.1016/j.cell.2014.11.021.
9. Zenk, F., Zhan, Y., Kos, P., Löser, E., Atinbayeva, N., Schächtle, M., Tiana, G., Giorgetti, L., and Iovino, N. (2021). HP1 drives de novo 3D genome reorganization in early *Drosophila* embryos. *Nature* 593, 289–293. 10.1038/s41586-021-03460-z.
10. Bonev, B., Cohen, N.M., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171, 557-572.e24. 10.1016/j.cell.2017.09.043.
11. Monahan, K., Horta, A., and Lomvardas, S. (2019). LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* 565, 448–453. 10.1038/s41586-018-0845-0.
12. Hug, C.B., Grimaldi, A.G., Kruse, K., and Vaquerizas, J.M. (2017). Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* 169, 216-228.e19. 10.1016/j.cell.2017.03.024.
13. Zhang, S., Übelmesser, N., Barbieri, M., and Papantonis, A. (2023). Enhancer–promoter contact formation requires RNAPII and antagonizes loop extrusion. *Nat. Genet.* 55, 832–840. 10.1038/s41588-023-01364-4.
14. 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, 1–1303. 10.1038/s41588-018-0175-z.

15. Kane, L., Williamson, I., Flyamer, I.M., Kumar, Y., Hill, R.E., Lettice, L.A., and Bickmore, W.A. (2022). Cohesin is required for long-range enhancer action at the *Shh* locus. *Nat Struct Mol Biol*, 1–7. 10.1038/s41594-022-00821-8.
16. Fulco, C.P., Nasser, J., Jones, T.R., Munson, G., Bergman, D.T., Subramanian, V., Grossman, S.R., Anyoha, R., Doughty, B.R., Patwardhan, T.A., et al. (2019). Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat Genet* 51, 1664–1669. 10.1038/s41588-019-0538-0.
17. 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. 10.1038/nature13417.
18. 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. 10.1101/gr.163519.113.
19. Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., and Spitz, F. (2016). The *Shh* Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Dev. Cell* 39, 529–543. 10.1016/j.devcel.2016.10.015.
20. Kaushal, A., Dorier, J., Wang, B., Mohana, G., Taschner, M., Cousin, P., Waridel, P., Iseli, C., Semenova, A., Restrepo, S., et al. (2022). Essential role of Cp190 in physical and regulatory boundary formation. *Sci Adv* 8. 10.1126/sciadv.abl8834.
21. 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. 10.1038/s41586-022-04570-y.
22. 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. 10.1016/j.cell.2015.04.004.
23. Pachano, T., Sánchez-Gaya, V., Ealo, T., Mariner-Faulí, M., Bleckwehl, T., Asenjo, H.G., Respuela, P., Cruz-Molina, S., Martín, M.M.-S., Haro, E., et al. (2021). Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nat Genet* 53, 1036–1049. 10.1038/s41588-021-00888-x.
24. Hafner, A., Park, M., Berger, S.E., Murphy, S.E., Nora, E.P., and Boettiger, A.N. (2023). Loop stacking organizes genome folding from TADs to chromosomes. *Mol. Cell* 83, 1377-1392.e6. 10.1016/j.molcel.2023.04.008.
25. 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. 10.1126/science.abi7178.
26. 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*, 1–7. 10.1038/s41586-022-04680-7.
27. Kubo, N., Ishii, H., Xiong, X., Bianco, S., Meitinger, F., Hu, R., Hocker, J.D., Conte, M., Gorkin, D., Yu, M., et al. (2021). Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat Struct Mol Biol* 28, 152–161. 10.1038/s41594-020-00539-5.
28. 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. van der, Ramos, Z. de A. dos, et al. (2022). Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nat Struct Mol Biol* 29, 563–574. 10.1038/s41594-022-00787-7.
29. Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336. 10.1038/nature14222.

30. Winick-Ng, W., Kukalev, A., Harabula, I., Zea-Redondo, L., Szabó, D., Meijer, M., Serebreni, L., Zhang, Y., Bianco, S., Chiariello, A.M., et al. (2021). Cell-type specialization is encoded by specific chromatin topologies. *Nature* 599, 684–691. 10.1038/s41586-021-04081-2.
31. Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science (New York, N.Y.)* 340, 1234167–1234167. 10.1126/science.1234167.
32. Dotson, G.A., Chen, C., Lindsly, S., Cicalo, A., Dilworth, S., Ryan, C., Jeyarajan, S., Meixner, W., Stansbury, C., Pickard, J., et al. (2022). Deciphering multi-way interactions in the human genome. *Nat Commun* 13, 5498. 10.1038/s41467-022-32980-z.
33. Allahyar, A., Vermeulen, C., Bouwman, B.A.M., Krijger, P.H.L., Verstegen, M.J.A.M., Geeven, G., Kranenburg, M. van, Pieterse, M., Straver, R., Haarhuis, J.H.I., et al. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. *Nature Genetics* 50, 1151–1160. 10.1038/s41588-018-0161-5.
34. Mateo, L.J., Murphy, S.E., Hafner, A., Cinquini, I.S., Walker, C.A., and Boettiger, A.N. (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* 167, 1–54. 10.1038/s41586-019-1035-4.
35. Gizzi, A.M.C., Cattoni, D.I., Fiche, J.-B., Espinola, S.M., Gurgo, J., Messina, O., Houbron, C., Ogiyama, Y., Papadopoulos, G.L., Cavalli, G., et al. (2019). Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Molecular Cell* 74, 1–31. 10.1016/j.molcel.2019.01.011.
36. 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. 10.1038/s41588-021-00799-x.
37. Espinola, S.M., Götz, M., Bellec, M., Messina, O., Fiche, J.-B., Houbron, C., Dejean, M., Reim, I., Gizzi, A.M.C., Lagna, 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. 10.1038/s41588-021-00816-z.
38. Stadler, M.R., Haines, J.E., and Eisen, M. (2017). Convergence of topological domain boundaries, insulators, and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early *Drosophila melanogaster* embryo. *eLife* 6, e29550. 10.7554/elife.29550.
39. 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*. 10.1016/j.molcel.2020.03.002.
40. Reddington, J.P., Garfield, D.A., Sigalova, O.M., Calviello, A.K., Marco-Ferreres, R., Girardot, C., Viales, R.R., Degner, J.F., Ohler, U., and Furlong, E.E.M. (2020). Lineage-Resolved Enhancer and Promoter Usage during a Time Course of Embryogenesis. *Dev. Cell* 55, 648–664. 10.1016/j.devcel.2020.10.009.
41. Eagen, K.P., Aiden, E.L., and Kornberg, R.D. (2017). Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proceedings of the National Academy of Sciences of the United States of America* 114, 8764–8769. 10.1073/pnas.1701291114.
42. Özel, M.N., Simon, F., Jafari, S., Holguera, I., Chen, Y.-C., Benhra, N., El-Danaf, R.N., Kapuralin, K., Malin, J.A., Konstantinides, N., et al. (2021). Neuronal diversity and convergence in a visual system developmental atlas. *Nature* 589, 88–95. 10.1038/s41586-020-2879-3.
43. Stark, A., Lin, M.F., Kheradpour, P., Pedersen, J.S., Parts, L., Carlson, J.W., Crosby, M.A., Rasmussen, M.D., Roy, S., Deoras, A.N., et al. (2007). Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* 450, 219–232. 10.1038/nature06340.

44. Özkan, E., Carrillo, R.A., Eastman, C.L., Weiszmann, R., Waghray, D., Johnson, K.G., Zinn, K., Celniker, S.E., and Garcia, K.C. (2013). An Extracellular Interactome of Immunoglobulin and LRR Proteins Reveals Receptor-Ligand Networks. *Cell* 154. 10.1016/j.cell.2013.06.006.
45. Consortium, D. 12 G., Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., et al. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450, 203–218. 10.1038/nature06341.
46. Janssens, J., Aibar, S., Taskiran, I.I., Ismail, J.N., Gomez, A.E., Aughey, G., Spanier, K.I., Rop, F.V.D., González-Blas, C.B., Dionne, M., et al. (2022). Decoding gene regulation in the fly brain. *Nature* 601, 630–636. 10.1038/s41586-021-04262-z.
47. Meyer, S., Schmidt, I., and Klämbt, C. (2014). Glia ECM interactions are required to shape the *Drosophila* nervous system. *Mech Develop* 133, 105–116. 10.1016/j.mod.2014.05.003.
48. Krause, S.A., Overend, G., Dow, J.A.T., and Leader, D.P. (2021). FlyAtlas 2 in 2022: enhancements to the *Drosophila melanogaster* expression atlas. *Nucleic Acids Res* 50, D1010–D1015. 10.1093/nar/gkab971.
49. Kaushal, A., Mohana, G., Dorier, J., Özdemir, I., Omer, A., Cousin, P., Semenova, A., Taschner, M., Dergai, O., Marzetta, F., et al. (2021). CTCF loss has limited effects on global genome architecture in *Drosophila* despite critical regulatory functions. *Nat Commun* 12, 1011. 10.1038/s41467-021-21366-2.
50. Li, X., Tang, X., Bing, X., Catalano, C., Li, T., Dolsten, G., Wu, C., and Levine, M. (2023). GAGA-associated factor fosters loop formation in the *Drosophila* genome. *Mol. Cell* 83, 1519-1526.e4. 10.1016/j.molcel.2023.03.011.
51. Tang, X., Li, T., Liu, S., Wisniewski, J., Zheng, Q., Rong, Y., Lavis, L.D., and Wu, C. (2022). Kinetic principles underlying pioneer function of GAGA transcription factor in live cells. *Nat Struct Mol Biol* 29, 665–676. 10.1038/s41594-022-00800-z.
52. Tsai, S.-Y., Chang, Y.-L., Swamy, K.B.S., Chiang, R.-L., and Huang, D.-H. (2016). GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. *Epigenet Chromatin* 9, 32. 10.1186/s13072-016-0082-4.
53. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472. 10.1016/j.cell.2012.01.010.
54. Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R.R., Korb, J.O., and Furlong, E.E.M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nature Genetics* 51, 1272–1282. 10.1038/s41588-019-0462-3.
55. Song, M., Pebworth, M.-P., Yang, X., Abnoui, A., Fan, C., Wen, J., Rosen, J.D., Choudhary, M.N.K., Cui, X., Jones, I.R., et al. (2020). Cell-type-specific 3D epigenomes in the developing human cortex. *Nature*, 1–6. 10.1038/s41586-020-2825-4.
56. Clowney, E.J., LeGros, M.A., Mosley, C.P., Clowney, F.G., Markenskoff-Papadimitriou, E.C., Myllys, M., Barnea, G., Larabell, C.A., and Lomvardas, S. (2012). Nuclear Aggregation of Olfactory Receptor Genes Governs Their Monogenic Expression. *Cell* 151, 724–737. 10.1016/j.cell.2012.09.043.
57. Long, H.K., Osterwalder, M., Welsh, I.C., Hansen, K., Davies, J.O.J., Liu, Y.E., Koska, M., Adams, A.T., Aho, R., Arora, N., et al. (2020). Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder. *Cell Stem Cell* 27, 765-783.e14. 10.1016/j.stem.2020.09.001.
58. Cusanovich, D.A., Reddington, J.P., Garfield, D.A., Daza, R.M., Aghamirzaie, D., Marco-Ferreres, R., Pliner, H.A., Christiansen, L., Qiu, X., Steemers, F.J., et al. (2018). The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature* 555, 538–542. 10.1038/nature25981.
59. Fujioka, M., Mistry, H., Schedl, P., and Jaynes, J.B. (2016). Determinants of Chromosome Architecture: Insulator Pairing in cis and in trans. *PLoS Genet.* 12, e1005889. 10.1371/journal.pgen.1005889.

60. Tsai, A., Muthusamy, A.K., Alves, M.R., Lavis, L.D., Singer, R.H., Stern, D.L., and Crocker, J. (2017). Nuclear microenvironments modulate transcription from low-affinity enhancers. *eLife* 6, e1006441. 10.7554/elife.28975.
61. Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science (New York, N.Y.)* 361, 412–415. 10.1126/science.aar4199.
62. Rosencrance, C.D., Ammouri, H.N., Yu, Q., Ge, T., Rendleman, E.J., Marshall, S.A., and Eagen, K.P. (2020). Chromatin Hyperacetylation Impacts Chromosome Folding by Forming a Nuclear Subcompartment. *Mol Cell* 78, 112-126.e12. 10.1016/j.molcel.2020.03.018.
63. Gaskill, M.M., Gibson, T.J., Larson, E.D., and Harrison, M.M. (2021). GAF is essential for zygotic genome activation and chromatin accessibility in the early *Drosophila* embryo. *Elife* 10, e66668. 10.7554/elife.66668.
64. Liu, J., Viales, R.R., Khoueiry, P., Reddington, J.P., Girardot, C., Furlong, E., and Robinson-Rechavi, M. (2021). The hourglass model of evolutionary conservation during embryogenesis extends to developmental enhancers with signatures of positive selection. *Genome Res* 31, gr.275212.121. 10.1101/gr.275212.121.
65. Loubiere, V., Papadopoulos, G.L., Szabo, Q., Martinez, A.-M., and Cavalli, G. (2020). Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. *Sci Adv* 6, eaax4001. 10.1126/sciadv.aax4001.
66. Abdennur, N., and Mirny, L.A. (2019). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* 36, 311–316. 10.1093/bioinformatics/btz540.
67. Kerpedjiev, P., Abdennur, N., Lekschas, F., McCallum, C., Dinkla, K., Strobelt, H., Lubert, J.M., Ouellette, S.B., Azhir, A., Kumar, N., et al. (2018). HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol* 19, 125. 10.1186/s13059-018-1486-1.
68. Pau, G., Fuchs, F., Sklyar, O., Boutros, M., and Huber, W. (2010). EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics* 26, 979–981. 10.1093/bioinformatics/btq046.
69. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550–21. 10.1186/s13059-014-0550-8.
70. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *Arxiv*. 10.48550/arxiv.1303.3997.
71. Open2C, Abdennur, N., Fudenberg, G., Flyamer, I.M., Galitsyna, A.A., Goloborodko, A., Imakaev, M., and Venev, S.V. (2023). Pairtools: from sequencing data to chromosome contacts. *Biorxiv*, 2023.02.13.528389. 10.1101/2023.02.13.528389.
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 (New York, N.Y.)* 362, eaau1783. 10.1126/science.aau1783.
73. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics J Integr Biology* 16, 284–287. 10.1089/omi.2011.0118.
74. Port, F., Chen, H.-M., Lee, T., and Bullock, S.L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* 111, E2967-76. 10.1073/pnas.1405500111.
75. Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). Bischof_Basler2007P.pdf. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3312–3317. 10.1073/pnas.0611511104.

76. Banerjee, S., Vernon, S., Jiao, W., Choi, B.J., Ruchti, E., Asadzadeh, J., Burri, O., Stowers, R.S., and McCabe, B.D. (2021). Miniature neurotransmission is required to maintain *Drosophila* synaptic structures during ageing. *Nat Commun* 12, 4399. 10.1038/s41467-021-24490-1.
77. Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science (New York, N.Y.)* 314, 1747–1751. 10.1126/science.1134426.
78. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e21. 10.1016/j.cell.2019.05.031.
79. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. *Nat Methods* 18, 1333–1341. 10.1038/s41592-021-01282-5.
80. Tan, G., Polychronopoulos, D., and Lenhard, B. (2019). CNEr: A toolkit for exploring extreme noncoding conservation. *Plos Comput Biol* 15, e1006940. 10.1371/journal.pcbi.1006940.
81. Harris, R.S. (2007). Improved pairwise alignment of genomic DNA. Ph.D. Thesis, The Pennsylvania State University.
82. Kent, W.J. (2002). BLAT—The BLAST-Like Alignment Tool. *Genome Res* 12, 656–664. 10.1101/gr.229202.
83. Zdobnov, E.M., Kuznetsov, D., Tegenfeldt, F., Manni, M., Berkeley, M., and Kriventseva, E.V. (2020). OrthoDB in 2020: evolutionary and functional annotations of orthologs. *Nucleic Acids Res* 49, gkaa1009-. 10.1093/nar/gkaa1009.
84. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. 10.1093/bioinformatics/btq033.
85. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. 10.1093/bioinformatics/btp324.