Published in final edited form as:
Cell. 2023 August 31; 186(18): 3826–3844.e26. doi:10.1016/j.cell.2023.07.008.

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

Giriram Mohana[1,*], **Julien Dorier**[2,3,*], **Xiao Li**[4,*], **Marion Mouginot**[1], **Rebecca C. Smith**[5], **Héléna Malek**[1], **Marion Leleu**[2,3], **Daniel Rodriguez**[1], **Jenisha Khadka**[1], **Patrycja Rosa**[6], **Pascal Cousin**[1], **Christian Iseli**[2,3], **Simon Restrepo**[7], **Nicolas Guex**[2,3], **Brian D. McCabe**[5], **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, Lüssirainstrasse 52, 6300 Zug, Switzerland

SUMMARY

Past studies have identified topologically associating domains (TADs) as basic units of genome organization. We present evidence for a previously unreported 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

---

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

[*]These authors contributed equally.

Author contributions

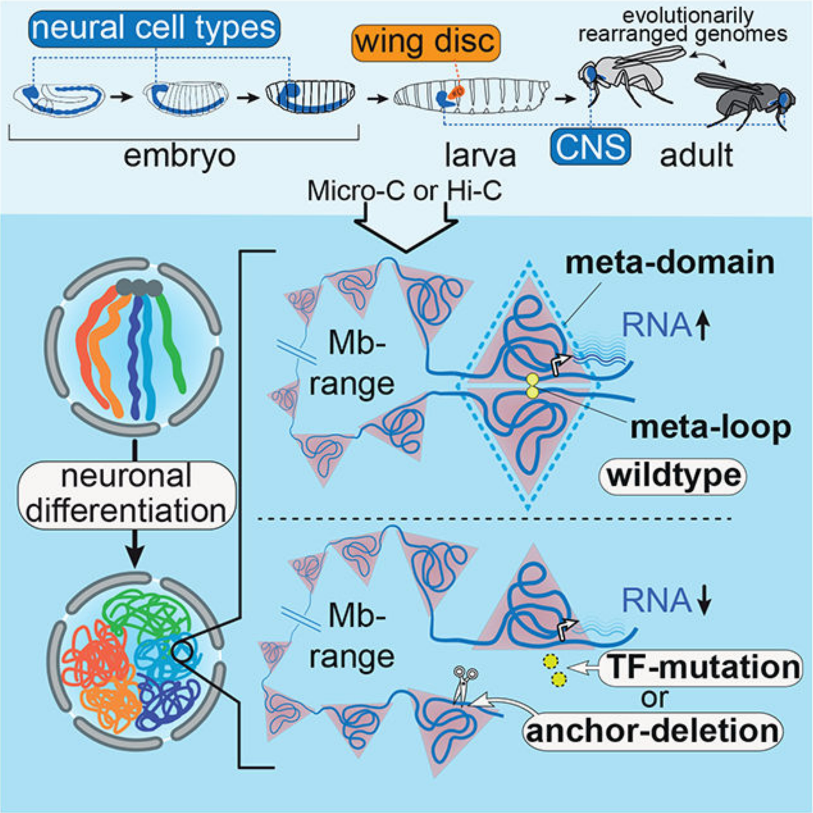
Conceptualization, A.J., M.S.L. and M.C.G.; Methodology, G.M., J.D., S.R., A.J. and M.C.G.; Investigation, G.M., X.L., M.M., R.C.S., H.M., D.R., J.K. and M.C.G.; Formal Analysis, J.D., X.L., M.L., P.R., C.I., N.G. and A.J.; Writing, N.G., A.J., M.S.L., and M.C.G.; Supervision, N.G., B.D.M., A.J., M.S.L.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

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.

Graphical Abstract



Brief statement:

Topologically associating domains that are megabases apart form meta-domains in mature fly neurons to regulate neuronal gene transcription.

Keywords

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

INTRODUCTION

Past studies 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 segregate [3]. TADs can further associate into higher-order assemblies [4-6]. Beyond TADs, chromosomes organize into territories and more broadly into transcriptionally active and inactive compartments [7-9]. Transcription factors shape three-dimensional (3D) genome contacts within these frameworks [10-13]. 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-dependent [14-16].

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 genes [25] and co-regulation of distant paralogs [26]. 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 activation [23]. 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-C [39], 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 certain neural 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 *Trithorax-like* gene) and CTCF, resulting in decreased 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 to date.

RESULTS

Meta-domains harbor loops involving genes expressed in the CNS

Micro-C analysis of CNSs or wing 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 compartmental interactions that reflect segregation of transcriptionally active and inactive genomic compartments, meta-domains represent selective, long-range associations between specific pairs of TADs at 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) in mid-embryogenesis neurons but less frequently in muscle cells [40] (Fig. 1A, S2B). 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, 58 curated meta-loops were identified (Fig. S1A, see Methods). DNase-seq peaks from mid-embryogenesis neurons [40] 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. S1A, Table S2). These 68 loops were present in 28 meta-domains, with most (68%) meta-domains containing clusters of two to six loops (Fig. 1E). 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 cells [41] or early embryos [25,26]. Unlike meta-loops, previously described loops were: detected in different cell types [25,26], orders of magnitude smaller (Fig. 1D), 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 an intergenic and a promoter 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). These genes encode extracellular or transmembrane proteins involved in axon guidance, including immunoglobulin receptor superfamily members, or ion channels specific to neuronal subtypes [42]. Genes at promoter anchors were significantly enriched for Gene Ontology terms related to cell adhesion and axon guidance (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 control transcription, they may affect a gene in a distant TAD to which they loop, rather than a local gene 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 adult CNSs of *Drosophila melanogaster* and *Drosophila virilis*, with an evolutionary distance comparable to that between humans and lizards [43] (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 54% of *D. melanogaster* loops and 58% of *D. virilis* loops (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 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 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 later in the Results.

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 guidance [44] 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 I-I meta-loop (Dvir_L49) 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 considerable loop conservation, 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 arm [45].

Meta-loops strengthen during neuronal differentiation

We asked when and where meta-loops emerge in the *D. melanogaster* CNS. Meta-loops were observed in 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 CNSs [46] 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 part of embryogenesis [40] 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 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 by fluorescence-activated cell sorting (FACS) from fixed embryos using antibodies against lineage markers Elav (neurons), Repo (glia) and Wor (neuroblasts) labeled using Multimodal Universal Signal Enhancement (MUSE) technology (Fig. S5A). Chromosomes of undifferentiated cells, including neuroblasts from any stage and glia and neurons from 6-8 hour old embryos, displayed a Rabl configuration where centromeres and telomeres cluster at opposite poles and chromosome arms are aligned (Fig. 3A row 1 and column 1). The Rabl configuration was gradually replaced by organization into chromosome territories in neurons and glia from 14-16 hour old embryos (Fig. 3A). This demonstrates significant genome reorganization during cell differentiation.

Closer 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 emergence of L35 coincided with the emergence of TAD boundaries that restricted the meta-domain borders (Figs. S5B-S5C). Similarly, the 5.1 Mb I-P meta-loop L54 involving the promoter of *beaten path IV* (*beat-IV*), involved in motor neuron axon guidance [44], emerged in neurons of 14-16 hour old embryos near an earlier I-I meta-loop (L53), whose anchor chromatin was accessible earlier than L54 anchor chromatin (Fig. 3C). L54 emergence also coincided with the reinforcement of TAD boundaries near meta-loop anchors (Fig. 3C). These findings suggest that, particularly during neuronal differentiation, transcription factors occupy meta-loop anchors and participate in meta-looping.

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

To examine the reproducibility of meta-domains across cells, physical distances between the I-P *beat-IV* meta-loop anchors were measured in whole-mount embryos by DNA-fluorescence in situ hybridization (DNA-FISH). The proportion of nuclei with physically close *beat-IV* meta-loop anchors specifically increased in neuronal nuclei (immunolabeled with α-Elav) as embryos aged (Fig. 3D). Meta-loop strengthening observed by Hi-C during neuronal maturation (Fig. 3C) thus reflects an increasing proportion of nuclei in which *beat-IV* meta-loop anchors are physically close. The increased proximity of *beat-IV* meta-loop anchors began at the developmental stage coinciding with *beat-IV* transcriptional onset (Fig. 3D). *beat-IV* meta-loop anchors were close in most nuclei of late-stage embryonic nerve cords, despite *beat-IV* being transcribed only in a subset of these cells (Fig. 3D). Overall, these 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). One example is the 1.6 Mb I-P meta-loop L22 involving *no long nerve cord* (*nolo*), hypothesized to function in the extracellular matrix [47]. The fact that the *nolo* meta-loop and meta-domain were comparable in CNSs and wing discs (in which *nolo* is not expressed [48]) 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. In contrast to most meta-domains (Fig. 1E), the *nolo* meta-domain features a single meta-loop (L22, Fig. 4A). Hi-C on $nolo^{\Delta A23}$ larval CNSs revealed that deleting the intergenic anchor of the *nolo* meta-loop disrupted the entire meta-domain (Fig. 4A). In contrast, Hi-C on $beat\text{-}IV^{\Delta A65}$ 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, that was possibly tethered by the remaining intact meta-loop (Fig. 4B).

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). *CG10651* was the only gene in the TAD containing the *nolo* intergenic meta-loop anchor (Fig. 4A). Genes around the *nolo* intergenic anchor are expressed in the male reproductive system [48] and were not reliably detected in wildtype (WT) or $nolo^{\Delta A23}$ mutant heads. mRNA levels measured in whole-bodied adult flies showed no significant change in $nolo^{\Delta A23}$ mutants (Fig. S6A). Similar observations were made for genes surrounding the deleted anchor in $beat\text{-}IV^{\Delta A65}$

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 CNSs [49], and determined 5 out of 28 (18%) meta-domains 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. 4C)[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 *CTCF*^{0} mutants did not strongly affect gene expression within the meta-domain, as assessed by RNA-sequencing (RNA-seq) on WT and *CTCF*^{0} mutant larval CNSs [49] (Fig. 4C). 48% of meta-loop anchors were within one genomic bin of a TAD boundary called by Hi-C in WT larval CNSs [49], including both promoter and intergenic anchors (Fig. S5C). Examples showed meta-loop anchors coinciding with TAD boundaries in WT (e.g. A46 in Fig. 4C), but boundary and meta-looping functions were separable in *CTCF*^{0} 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.

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*$^{\Delta A23}$ mutants compared to WT, while *beat-IV* mRNA levels remained unchanged in *beat-IV*$^{\Delta A65}$ 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*$^{\Delta A23}$ and *beat-IV*$^{\Delta A65}$ mutants (Figs. 5A and S6C). *nolo* and *beat-IV* were expressed in different cell subsets in WT CNSs (11% and 9% of WT cells, respectively) (Fig. 5B). In *nolo*$^{\Delta A23}$ mutants, *nolo* was expressed in fewer cells (2%) and at lower levels, while *beat-IV* expression remained relatively stable in *beat-IV*^{A65} (Fig. 5C). 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*$^{\Delta A23}$ mutant clusters, whereas *beat-IV* was unaffected in any cluster or genotype (Fig. 5C, 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. 5C and S6D), revealing their independence.

nolo RNA-FISH signals were strongly reduced in embryonic nerve cords of *nolo*$^{\Delta A23}$ mutants compared to WT (Fig. 5D). To assess if decreased *nolo* mRNA levels in *nolo*$^{\Delta A23}$ mutants resulted from reduced *nolo* transcription, we imaged *nolo* nascent RNA and *nolo* promoter DNA by RNA-DNA-FISH (Figs. 5E 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*$^{\Delta A23}$ nuclei (Fig. 5E). Thus, deletion of the *nolo* intergenic anchor decreases *nolo* transcription.

To investigated whether intergenic anchors may facilitate neural gene transcription by serving as distal enhancers, 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. 5C 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. 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*$^{\Delta A65}$ mutants is sustained by distal regulatory elements within the preserved meta-domain due to the intact meta-loop (Fig. 4B). This hypothesis is tested in the last Results section.

To determine whether *nolo*$^{\Delta A23}$ mutation affects *nolo* function, we compared the lethal stages of *nolo*$^{\Delta A23}$ and *nolo*^{KO} mutants with nonsense mutations after codons 47 or 48 (Figs. 5F-5G). *nolo*^{KO} and *nolo* whole-body or neuroblast-specific knock-down pupae showed impaired hatching that was not observed in *nolo*$^{\Delta A23}$ mutants (Fig. 5G). In *nolo* knock-down flies, the *nolo*$^{\Delta A23}$ mutation decreased pupal hatching further (e.g. from 66% upon neuroblast-specific *nolo* knock-down to 11% when performed in a *nolo*$^{\Delta A23}$ mutant) (Fig. 5G). We analyzed neuromuscular junctions (NMJs) of *nolo* knock-out, knock-down and *nolo*$^{\Delta A23}$ third instar larvae. α-HRP immunostaining of neuronal membranes at muscle 6/7 NMJs revealed that the NMJ area and the number of synaptic boutons were significantly reduced in *nolo*^{KO}, *nolo*$^{\Delta A23}$ and *nolo* knock-down flies compared to WT controls (Figs. 5H-5I), suggesting that NMJs are under-developed or impaired. Thus, *nolo*$^{\Delta A23}$ is a hypomorphic *nolo* allele with similar NMJ 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 is required to form a subset of chromosomal loops between tethering elements in embryos in a manner dependent on its BTB (broad-complex, tramtrack and bric-à-brac – zinc finger, also known as POZ) domain [50]. To investigate GAF's potential role in meta-loop formation, we performed Micro-C on WT and *GAF*$^{\Delta BTB}$ mutant larval CNSs [51]. The GAF BTB domain was specifically required to form two ~6 Mb meta-loops connecting the promoter of *sticks and stones* (*sns*) with that of *hibris* (*hbs*) (L25) or with an intergenic anchor (L26) (Figs. 6A-6B). GAF ChIP-seq data in larval tissues, comprising the

CNS [52], revealed lower enrichment of GAF binding around meta-loop anchors compared to anchors of loops detected in embryos and cultured cells (Fig. 6C). 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. 6A and 6C). Thus, both meta-loops destabilized in *GAF*$^{\Delta BTB}$ 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. 6B-6C). The long-range interaction between *sns* and *hbs* is not limited to the CNS (Fig. S1) and was also previously observed in whole embryos [53,54], suggesting a non-CNS-specific role of GAF in forming these meta-loops.

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*$^{\Delta BTB}$ mutants, while *hbs* expression was mostly unchanged (Fig. 6D). These observations were further supported by RT-qPCR on larval CNSs (Fig. 6E). 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 maps [49] identified six meta-loops that were significantly weakened in *CTCF*^{0} mutants compared to WT (Figs. 7A-7B and S7A), and these were unaffected in *GAF*$^{\Delta BTB}$ larval CNSs (Fig. 6B). CTCF-dependent meta-loops were identified in CNSs but not wing discs (Fig. S1). Although CTCF generally binds constitutively in different cell types [40], the chromatin of CTCF-bound meta-loop anchors was more accessible in embryonic neurons than muscle (Figs. 7C and S7A).

Surprisingly, the *beat-IV* meta-domain was weakened in *CTCF*^{0} mutant larval CNSs even though CTCF only binds to the anchors of meta-loop L53 but not L54, which connects *beat-IV* promoter to its intergenic anchor (Figs. 7A and 7D). 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-seq [49] (Fig. 7A). 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. 7E). 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. 7E). CTCF recruits its BTB-ZnF cofactor, Centrosomal protein 190 kDa (Cp190) to CTCF binding sites [49]. Consistently, Cp190 co-bound with CTCF at L53 anchors in WT but was either reduced or absent at these anchors in *CTCF*^{0} larval CNSs (Figs. 7A and 7D) [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. S7B). In *Cp190*^{0} (maternal− zygotic−) embryos, *beat-IV* expression was reduced to an even greater extent than in *CTCF*^{0} embryos (Fig. 7E). In contrast, in *beat-IV*$^{\Delta A65}$ mutants with a deletion of the intergenic anchor looping to *beat-IV* promoter (Fig. 4B), *beat-IV* expression was largely normal (Fig. 7E).

We also investigated a second CTCF-dependent meta-domain containing two 2.2 Mb meta-loops normally both bound by CTCF in WT (Fig. S7B). L33 connects two intergenic anchors and L34 connects the promoter of *Multiplexin* (*Mp*) to an intergenic anchor (Fig. S7A). RNA-seq analysis revealed an 11-fold upregulation of the gene *CG7509* within the *Mp* meta-domain in *CTCF*^{0} 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 *CTCF*^{0} 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 *Cp190*^{0} 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 *CTCF*^{0} embryos (Fig. S7C).

In conclusion, the destabilization of distinct meta-loops in *GAF*$^{\Delta BTB}$ and *CTCF*^{0} 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 previously unrecognized 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 allele [56], 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 clusters [11]. 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. 4A and 5). Secondly, *GAF*$^{\Delta BTB}$ mutants displayed notable loss of 6.2 Mb meta-loops involving the *sns* promoter and decreased *sns* expression (Fig. 6). Thirdly, CTCF loss weakened three CNS-specific meta-domains and decreased expression of *beat-IV* and *Mp* located at meta-loop anchors (Figs. 7 and S7). Of note, the expression of two genes at (*Mp*) or near (*CG7509*) opposite meta-loop anchors was oppositely affected in *CTCF*^{0} 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 humans [57].

Our observation of reduced *beat-IV* expression in *CTCF*^{0}, but not *beat-IV*$^{\Delta A65}$, mutants suggests that the *beat-IV* meta-domain and not the *beat-IV* meta-loop itself, is important for *beat-IV* expression (Figs. 5C and 7E). 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], and 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 (Figs. 3D and S1) [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 4), that pair over long and flexible genomic distances with striking specificity (Fig. 2) and rely on dedicated transcription factors including GAF and CTCF (Figs. 6 and 7) for their formation.

In contrast to GAF which is required to form a meta-domain that is also observed outside of the CNS (Figs. 6 and S1) [53,54], CTCF was required to form three CNS-specific meta-domains (Figs. 7 and S1). Our previous finding that CTCF is specifically required in neurons for fly viability [49] 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-locp 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 *CTCF*^{0} CNSs despite complete loss of CTCF [49] (Figs. 7A and S7A). Cp190 may also be required, as partial meta-domain retention in *CTCF*^{0} mutants correlates with partial Cp190 retention at meta-loop anchors. 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. 7E and S7C). GAF and Cp190 belong to the same BTB-ZnF protein family and possibly share a 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 factors to form meta-domains.

Most meta-domains contain clustered loops (Fig. 1E). In *beat-1V*(Fig. 3C) and *GluRIA-GluRIB* (Fig. S5B) meta-domains, an I-I meta-loop emerged earlier in neuronal differentiation than the I-P or P-P meta-loop. Deleting the intergenic anchor connected to *beat-IV* promoter did not disrupt the *beat-IV* meta-domain (Fig. 4B). However, destabilizing the earlier CTCF-bound I-I meta-loop by CTCF loss, also destabilized the I-P meta-loop (not directly bound by CTCF) and the entire *beat-IV* meta-domain (Fig. 7). 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 previously unrecognized interplay between certain meta-loop anchors and TAD boundaries which appear to emerge concomitantly during neural differentiation, though the temporal resolution of our time-course Hi-C experiments is limited (Figs. 3C 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. 4C).

Meta-loops differ from traditional models of chromosome folding, which involve cohesin-mediated loop extrusion [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 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 clusters [11].

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. Determining whether meta-loops confer advantages for regulating specific CNS genes in a given species will require identifying the relevant neurons where gene expression depends on meta-loops and understanding how this regulation confers species-specific traits, ideally through studies conducted in natural habitats.

STAR METHODS

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 plasmids and fly strains generated in this study are available from the lead contact without restriction.

Data and Code Availability

- All sequencing data were deposited at NCBI Gene Expression Omnibus (GEO) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- All original code is publicly available in Github. Links to Github and all software used as described in the Star Methods are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental Model and Study Participant Details

Drosophila—*Drosophila melanogaster* and *virilis* were cultured under standard laboratory conditions at 25°C. 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. All fly strains are listed in the key resources table.

METHOD DETAILS

Micro-C—Wing imaginal discs and larval brains were dissected from *y w* or $GAF^{\Delta BTB}$ 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 μL 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 μL of 2 M Tris-HCl pH 7.5 and washed with PBST 3 times for 5 min. The tissue was resuspended in 500 μL buffer MB1 (50 mM NaCl, 10 mM Tris, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 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 μL MB1 with 1.5 μL (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 $MgCl_2$), 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 µL buffer containing 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$. Biotin labeling was performed in 75 µL 50 mM NaCl, 10 mM Tris, 10 mM $MgCl_2$, 100 µg/mL BSA, 2 mM ATP and 5 mM DTT and 67 µM 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 µL MB3 (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$) 3 times, proximity ligation was performed in 250 µL ligation mix (50 mM Tris, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, 100 µg/mL BSA) with 5000 U T4 ligase. Unligated biotin-nucleotides were removed with 500 U Exonuclease III in 100 µL of NEB1 (10 mM Bis-Tris-Propane-HCl pH 7, 10 mM $MgCl_2$, 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 µl of 10 mM Tris-HCl pH 8.5, treated with 2 µl 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 µl, 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 µL EB buffer. 3.5 µL of End Prep Reaction Buffer and 1.5 µL 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 µL, Ligation Master Mix 15 µL, Ligation Enhancer 0.5 µL. After incubating at 20°C for 30 min, 15 µL 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 µL EB buffer. 25 µl of 2X KAPA HiFi Hot Start Mix, 0.5 µl of 10 µM universal reverse primer and 0.5 µL 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 [72] (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 [68].

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 µm 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 μm 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 μm 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 μl 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, dmel_r6.36 (dm6) for *Drosophila melanogaster*, and GSE120751_Dvir_HiC.fa.gz downloaded from ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE120nnn/GSE120751/suppl/ [81] for *Drosophila virilis*. For each sample,unique pairs of reads with high mapping quality scores on both sides (MAPQ ≥ 30) from all biological replicates were aggregated into contact matrices using cooler [60] 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 procedure [64] implemented in cooler. Hi-C maps were visualized in HiGlass [67] 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. 2021 [49]) 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 GSE146752 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146752). 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(mean count) vs log(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 median-2*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 EBImage [69] 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. 2020 [40]. 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 lops 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* embryos [64]. 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. 6B and 7B, 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 DESeq2

[73] 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 CNSs [49] to infer information on higher-order chromatin interactions. We relied on the fact that Hi-C experiments were performed using innucleus 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 mem [82] 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 suite [72] 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. [83] 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 package [80] 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 Na2CO$_3$, 40 mM NaHCO$_3$ 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 [77].

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 [77]. 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. 6E, 15 larval CNSs of WT (*y w*) or $GAF^{\Delta BTB}$ 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 (w^{1118}), $nolo^{\Delta A23}$ or $beat\text{-}IV^{\Delta A65}$ 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 $GAF^{\Delta BTB}$ mutants) or to *RpL15* (for $nolo^{\Delta A23}$ or $beat\text{-}IV^{\Delta A65}$ 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—$beat\text{-}IV^{\Delta A65}$ and $nolo^{\Delta A23}$ 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 $beat\text{-}IV^{\Delta A65}$, or 814 bp (dm6 coordinates chr2L:20,120,764-20,121,578) for $nolo^{\Delta A23}$. 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 promoters [84], then assembled into a single transgene and integrated by site-specific recombination into the ZH86Fb landing site [85]. All *Drosophila* injections were performed by FlyORF, Zurich. Males expressing both sgRNAs and Cas9 in their germline (*nanos-Cas9*) [84] 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* [84] 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 w^{111} 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 w^{1118} 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-vGlut [66] 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 recombination [86]. 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 $MgCl_2$, 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 $MgCl_2$, 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, dmel_r6.36 (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 cellrangeratac count, using the default settings. The custom peak calling algorithm implemented in Single Cell ATAC fits a probabilistic model to a smoothened 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*$^{\Delta A65}$, 157 million raw reads for *nolo*$^{\Delta A23}$) 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) [78]. 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*$^{\Delta A65}$, and 5871 cells and 8698 genes for *nolo*$^{\Delta A23}$.

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. 5A based on the expressions of selected markers shown in Fig. 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 package [75] complemented with UCSC Kent's utilities (http://hgdownload.cse.ucsc.edu/admin/exe/) [74] 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/) [76] 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 BLAT [87] 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*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/) [88], 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) [70].

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 [71]. 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 snATAC-seq data in WT adult brains, this published dataset [46] was downloaded from GEO GSE163697 (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, the fastq file of sample GSM5503528 from GEO GSE181494 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181494) [46] 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, this published dataset [64] was downloaded from ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) accessions E-MTAB-3797 and E-MTAB-9480, and remapped using the BWA tool [89] 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 [81]).

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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank David Garfield for helpful discussions and Victor Rossier for his gene family visualization interface. 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.).

REFERENCES

1. Rao SSP, Huang S-C, Hilaire BGS, Engreitz JM, Perez EM, Kieffer-Kwon K-R, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. (2017). Cohesin Loss Eliminates All Loop Domains. Cell 171, 305–309.e24. 10.1016/j.cell.2017.09.026. [PubMed: 28985562]
2. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, and Bruneau BG (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. [PubMed: 28525758]
3. Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y, and Corces VG (2017). Evolutionarily Conserved Principles Predict 3D Chromatin Organization. Molecular Cell 67, 837–852. 10.1016/j.molcel.2017.07.022. [PubMed: 28826674]
4. Paulsen J, Ali TML, 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. [PubMed: 30578416]
5. Szabo Q, Jost D, Chang J-M, Cattoni DI, Papadopoulos GL, 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. [PubMed: 29503869]
6. Beagrie RA, Scialdone A, Schueler M, Kraemer DCA, Chotalia M, Xie SQ, Barbieri M, de Santiago I, Lavitas L-M, Branco MR, et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. Nature 543, 519–524. 10.1038/nature21411. [PubMed: 28273065]
7. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293. 10.1126/science.1181369. [PubMed: 19815776]
8. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, 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. [PubMed: 25497547]
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. [PubMed: 33854237]
10. Bonev B, Cohen NM, Szabo Q, Fritsch L, Papadopoulos GL, 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. [PubMed: 29053968]
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. [PubMed: 30626972]
12. Hug CB, Grimaldi AG, Kruse K, and Vaquerizas JM (2017). Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. Cell 169, 216–228.e19. 10.1016/j.cell.2017.03.024. [PubMed: 28388407]
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. [PubMed: 37012454]
14. Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, and Gregor T (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nat. Genet 50, 1–1303. 10.1038/s41588-018-0175-z. [PubMed: 29273803]
15. Kane L, Williamson I, Flyamer IM, Kumar Y, Hill RE Lettice LA, and Bickmore WA (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. [PubMed: 35046575]

16. Fulco CP, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR, Anyoha R, Doughty BR, Patwardhan TA, 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. [PubMed: 31784727]
17. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, and Furlong EEM (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512, 96–100. 10.1038/nature13417. [PubMed: 25043061]
18. Symmons O, Uslu VV, 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. [PubMed: 24398455]
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. [PubMed: 27867070]
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. [PubMed: 35418676]
22. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, 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. [PubMed: 25959774]
23. Pachano T, Sánchez-Gaya V, Ealo T, Mariner-Faulí M, Bleckwehl T, Asenjo HG, Respuela P, Cruz-Molina S, Martín MM-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. [PubMed: 34183853]
24. Hafner A, Park M, Berger SE, Murphy SE, Nora EP, and Boettiger AN (2023). Loop stacking organizes genome folding from TADs to chromosomes. Mol. Cell 83, 1377–1392.e6. 10.1016/j.molcel.2023.04.008. [PubMed: 37146570]
25. Batut PJ, Bing XY, Sisco Z, Raimundo J, Levo M, and Levine MS (2022). Genome organization controls transcriptional dynamics during development. Science 375, 566–570. 10.1126/science.abi7178. [PubMed: 35113722]
26. Levo M, Raimundo J, Bing XY, Sisco Z, Batut PJ, Ryabichko S, Gregor T, and Levine MS (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 JD, 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. [PubMed: 33398174]
28. Rinzema NJ, Sofiadis K, Tjalsma SJD, Verstegen MJAM, Oz Y, Valdes-Quezada C, Felder A-K, Filipovska T, van der Elst S, de A. dos Ramos Z, 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. [PubMed: 35710842]
29. Dixon JR, Jung I, Selvaraj S, Shen Y Antosiewicz-Bourget JE, Lee AY, 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. [PubMed: 25693564]
30. Winick-Ng W, Kukalev A, Harabula I, Zea-Redondo L, Szabò D, Meijer M, Serebreni L, Zhang Y, Bianco S, Chiariello AM, et al. (2021). Cell-type specialization is encoded by specific chromatin topologies. Nature 599, 684–691. 10.1038/s41586-021-04081-2. [PubMed: 34789882]
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. [PubMed: 23744951]

32. Dotson GA, 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. [PubMed: 36127324]
33. Allahyar A, Vermeulen C, Bouwman BAM, Krijger PHL, Verstegen MJAM, Geeven G, van Kranenburg M, Pieterse M, Straver R, Haarhuis JHI, et al. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. Nature Genetics 50, 1151–1160. 10.1038/s41588-018-0161-5. [PubMed: 29988121]
34. Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, and Boettiger AN (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature 167, 1–54. 10.1038/s41586-019-1035-4.
35. Gizzi AMC, Cattoni DI, Fiche J-B, Espinola SM, Gurgo J, Messina O, Houbron C, Ogiyama Y, Papadopoulos GL, 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. [PubMed: 30951649]
36. Ing-Simmons E, Vaid R, Bing XY, Levine M, Mannervik M, and Vaquerizas JM (2021). Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning. Nat Genet 53, 487–99. 10.1038/s41588-021-00799-x. [PubMed: 33795866]
37. Espinola SM, Gotz M, Bellec M, Messina O, Fiche J-B, Houbron C, Dejean M, Reim I, Gizzi AMC, Lagha M, et al. (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. Nat Genet 53, 477–486. 10.1038/s41588-021-00816-z. [PubMed: 33795867]
38. Stadler MR, Haines JE, 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. [PubMed: 29148971]
39. Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, 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 JP, Garfield DA, Sigalova OM, Calviello AK, Marco-Ferreres R, Girardot C, Viales RR, Degner JF, Ohler U, and Furlong EEM (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. [PubMed: 33171098]
41. Eagen KP, Aiden EL, and Kornberg RD (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. [PubMed: 28765367]
42. Özel MN, Simon F, Jafari S, Holguera I, Chen Y-C, Benhra N, El-Danaf RN, Kapuralin K, Malin JA, 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. [PubMed: 33149298]
43. Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, et al. (2007). Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature 450, 219–232. 10.1038/nature06340. [PubMed: 17994088]
44. Özkan E, Carrillo RA, Eastman CL, Weiszmann R, Waghray D, Johnson KG, Zinn K, Celniker SE, and Garcia KC (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 AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, et al. (2007). Evolution of genes and genomes on the Drosophila phylogeny. Nature 450, 203–218. 10.1038/nature06341. [PubMed: 17994087]
46. Janssens J, Aibar S, Taskiran II, Ismail JN, Gomez AE, Aughey G, Spanier KI, Rop FVD, González-Blas CB, Dionne M, et al. (2022). Decoding gene regulation in the fly brain. Nature 601, 630–636. 10.1038/s41586-021-04262-z. [PubMed: 34987221]
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 SA, Overend G, Dow JAT, and Leader DP (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. [PubMed: 33579945]

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. [PubMed: 37003261]

51. Tang X, Li T, Liu S, Wisniewski J, Zheng Q, Rong Y, Lavis LD, 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. [PubMed: 35835866]

52. Tsai S-Y, Chang Y-L, Swamy KBS, 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, 458–72. 10.1016/j.cell.2012.01.010. [PubMed: 22265598]

54. Ghavi-Helm Y, Jankowski A, Meiers S, Viales RR, Korbel JO, and Furlong EEM (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. Nature Genetics 51, 1272–1282. 10.1038/s41588-019-0462-3. [PubMed: 31308546]

55. Song M, Pebworth M-P, Yang X, Abnousi A, Fan C, Wen J, Rosen JD, Choudhary MNK, Cui X, Jones IR, et al. (2020). Cell-type-specific 3D epigenomes in the developing human cortex. Nature, 1–6. 10.1038/S41586-020-2825-4.

56. Clowney EJ, LeGros MA, Mosley CP, Clowney FG, Markenskoff-Papadimitriou EC, Myllys M, Barnea G, Larabell CA, 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. [PubMed: 23141535]

57. Long HK, Osterwalder M, Welsh IC, Hansen K, Davies JOJ, Liu YE, Koska M, Adams AT, 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. [PubMed: 32991838]

58. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. (2018). The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature 555, 538–542. 10.1038/nature25981. [PubMed: 29539636]

59. Fujioka M, Mistry H, Schedl P, and Jaynes JB (2016). Determinants of Chromosome Architecture: Insulator Pairing in cis and in trans. PLoS Genet. 12, e1005889. 10.1371/journal.pgen.1005889. [PubMed: 26910731]

60. Tsai A, Muthusamy AK, Alves MR, Lavis LD, Singer RH, Stern DL, 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 II (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science (New York, N.Y.) 361, 412–415. 10.1126/science.aar4199. [PubMed: 29930094]

62. Rosencrance CD, Ammouri HN, Yu Q, Ge T, Rendleman EJ, Marshall SA, and Eagen KP (2020). Chromatin Hyperacetylation Impacts Chromosome Folding by Forming a Nuclear Subcompartment. Mol Cell 78, 112–126.e12. 10.1016/j.molcel.2020.03.018. [PubMed: 32243828]

63. Gaskill MM, Gibson TJ, Larson ED, and Harrison MM (2021). GAF is essential for zygotic genome activation and chromatin accessibility in the early Drosophila embryo. Elife 10, e66668. 10.7554/elife.66668. [PubMed: 33720012]

64. Liu J, Viales RR, Khoueiry P, Reddington JP, 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 GL, 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. [PubMed: 31950077]
66. Banerjee S, Vernon S, Jiao W, Choi BJ, Ruchti E, Asadzadeh J, Burri O, Stowers RS, and McCabe BD (2021). Miniature neurotransmission is required to maintain Drosophila synaptic structures during ageing. Nat Commun 12, 4399. 10.1038/s41467-021-24490-1. [PubMed: 34285221]
67. Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, 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. [PubMed: 30143029]
68. Abdennur N, and Mirny LA (2019). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics 36, 311–316. 10.1093/bioinformatics/btz540.
69. 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. [PubMed: 20338898]
70. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, et al. (2006). The UCSC Genome Browser Database: update 2006. Nucleic Acids Res 34, D590–D598. 10.1093/nar/gkj144. [PubMed: 16381938]
71. Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. 10.1093/bioinformatics/btq033. [PubMed: 20110278]
72. Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, and Venev SV (2023). Pairtools: from sequencing data to chromosome contacts. Biorxiv, 2023.02.13.528389. 10.1101/2023.02.13.528389.
73. Love MI, 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. [PubMed: 25516281]
74. Kuhn RM, Haussler D, and Kent WJ, (2013). The UCSC genome browser and associated tools. Brief Bioinform 14, 144–161. 10.1093/bib/bbs038. [PubMed: 22908213]
75. 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. [PubMed: 31449516]
76. Harris RS (2007). Improved pairwise alignment of genomic DNA. Ph.D. Thesis, The Pennsylvania State University.
77. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. 10.1038/nmeth.2019. [PubMed: 22743772]
78. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 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. [PubMed: 31178118]
79. Stuart T, Srivastava A, Madad S, Lareau CA, and Satija R (2021). Single-cell chromatin state analysis with Signac. Nat Methods 18, 1333–1341. 10.1038/s41592-021-01282-5. [PubMed: 34725479]
80. 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.
81. Renschler G, Richard G, Valsecchi CIK, Toscano S, Arrigoni L, Ramírez F, and Akhtar A (2019). Hi-C guided assemblies reveal conserved regulatory topologies on X and autosomes despite extensive genome shuffling. Genes & Development 33, 1591–1612. 10.1101/gad.328971.119. [PubMed: 31601616]
82. Li H (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Arxiv. 10.48550/arxiv.1303.3997.

83. Bintu B, Mateo LJ, Su J-H, Sinnott-Armstrong NA, Parker M, Kinrot S, Yamaya K, Boettiger AN, 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. [PubMed: 30361340]

84. Port F, Chen H-M, Lee T, and Bullock SL (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. [PubMed: 25002478]

85. Bischof J, Maeda RK, 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. [PubMed: 17360644]

86. Venken KJT, He Y, Hoskins RA, and Bellen HJ (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. [PubMed: 17138868]

87. Kent WJ (2002). BLAT—The BLAST-Like Alignment Tool. Genome Res 12, 656–664. 10.1101/gr.229202. [PubMed: 11932250]

88. Zdobnov EM, Kuznetsov D, Tegenfeldt F, Manni M, Berkeley M, and Kriventseva EV (2020). OrthoDB in 2020: evolutionary and functional annotations of orthologs. Nucleic Acids Res 49, gkaa1009-. 10.1093/nar/gkaa1009.

89. Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. 10.1093/bioinformatics/btp324. [PubMed: 19451168]

Highlights

- Meta-domains are specific associations of distant TADs in mature fly neurons.
- In meta-domains, meta-loops tether neuronal gene promoters and intergenic elements.
- Meta-loops enable megabase-range regulation of neuronal gene transcription.
- Dedicated transcription factors such as CTCF and GAF form individual meta-loops.

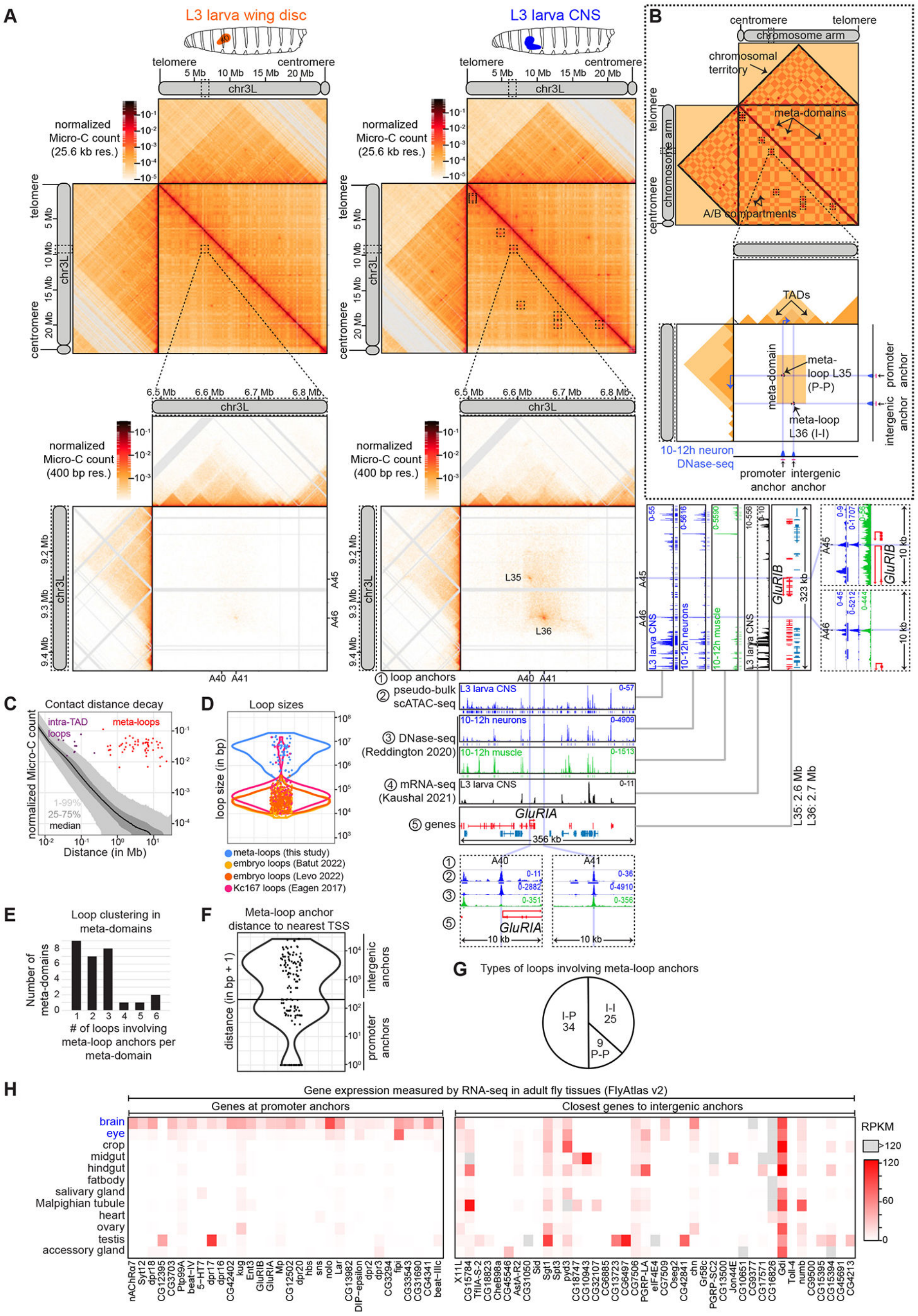


Figure 1. Meta-domains harbor loops involving CNS-expressed genes. See also Figures S1-S2.

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 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 muscles [40], mRNA-seq on larval CNSs [49] and gene tracks (with longest isoforms colored according to transcription direction) are shown. Zoom-ins of 10 kb regions around each meta-loop anchor are shown in dotted rectangles.

B. Cartoon representation of panel A column 2.

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.

D. Sizes (in bp, in y) of all indicated loop types (points).

E. Number of meta-domains (in y) each harboring the indicated numbers of loops involving meta-loop anchors (in x).

F. Distances (in bp, in y) of all 79 meta-loop anchors (points) to the nearest protein-coding gene TSS. The 200 bp line marks the cut-off for defining promoter or intergenic anchors.

G. Classification of all 58 meta-loops called in larval CNSs (I: intergenic, P: promoter).

H. RNA-seq on adult fly tissues from FlyAtlas v2 [48].

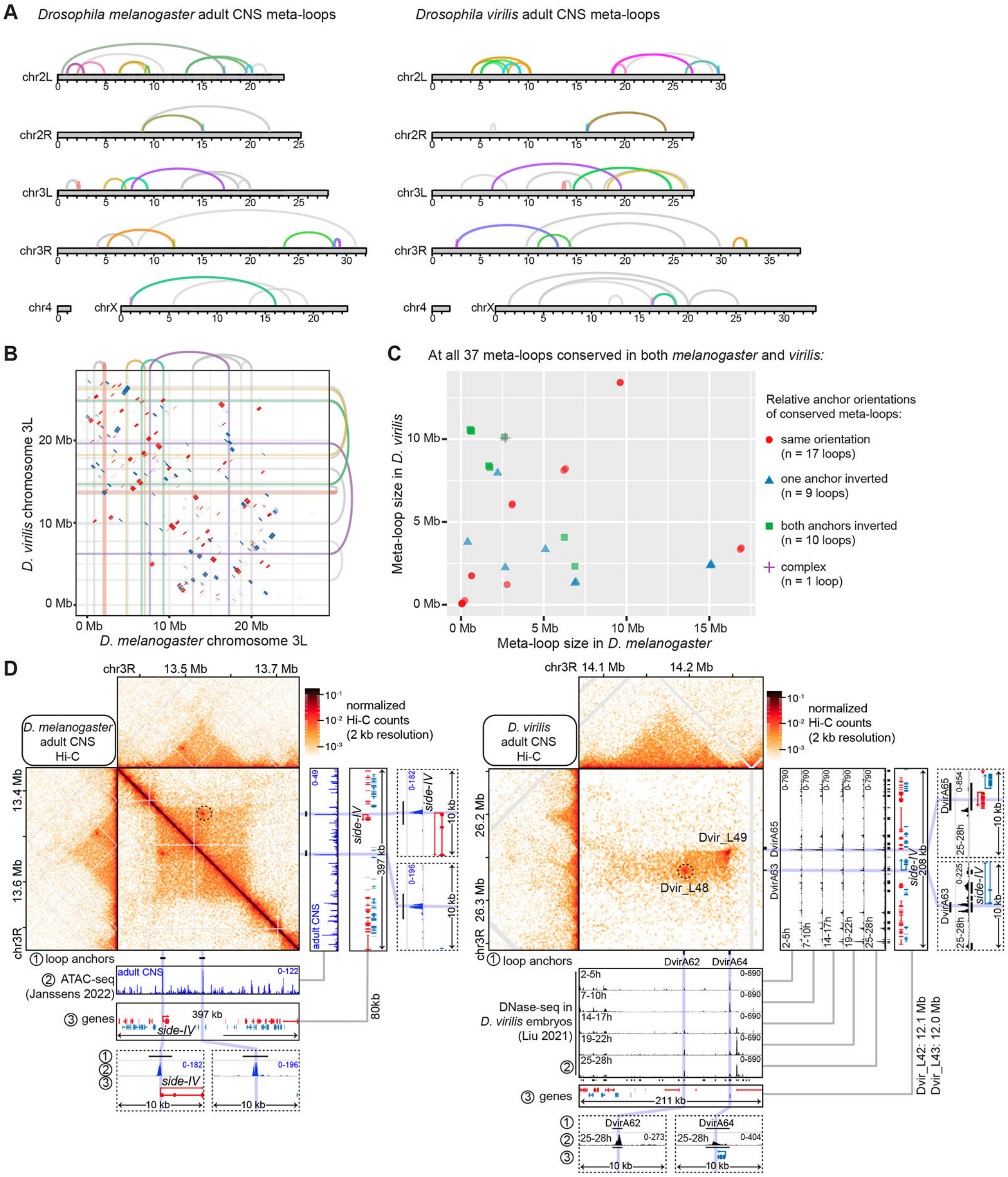


Figure 2. Meta-loop anchors specifically pair over flexible distances in evolution. See also Figure S3.

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 blue, those that are inverted in one species are red. Conserved loops are colored as in A and loop anchors are highlighted (in x and y).

C. Sizes of all 37 conserved loops involving meta-loop anchors (colored symbols) plotted in *D. vir.* (in y) versus *D. mel.* (in x). Different symbols represent conserved loops with different relative anchor orientations. The "complex" case is illustrated in Figs. S3I-S3J.

D. Similar to Fig. 1A but showing Hi-C maps of a *D. vir*. meta-loop (Dvir_L48) homologous to an intra-TAD loop in *D. mel*. (circled). *D. mel*. adult CNS ATAC-seq from Janssens et al. 2022 [46] or whole *D. vir*. embryo DNase-seq at the indicated hours of development from Liu et al. 2021 [64] are shown below. Anchors of meta-loops called in *D. vir*. were assigned to underlying DNase-seq peaks in old (25-28h) *D. vir*. embryos.

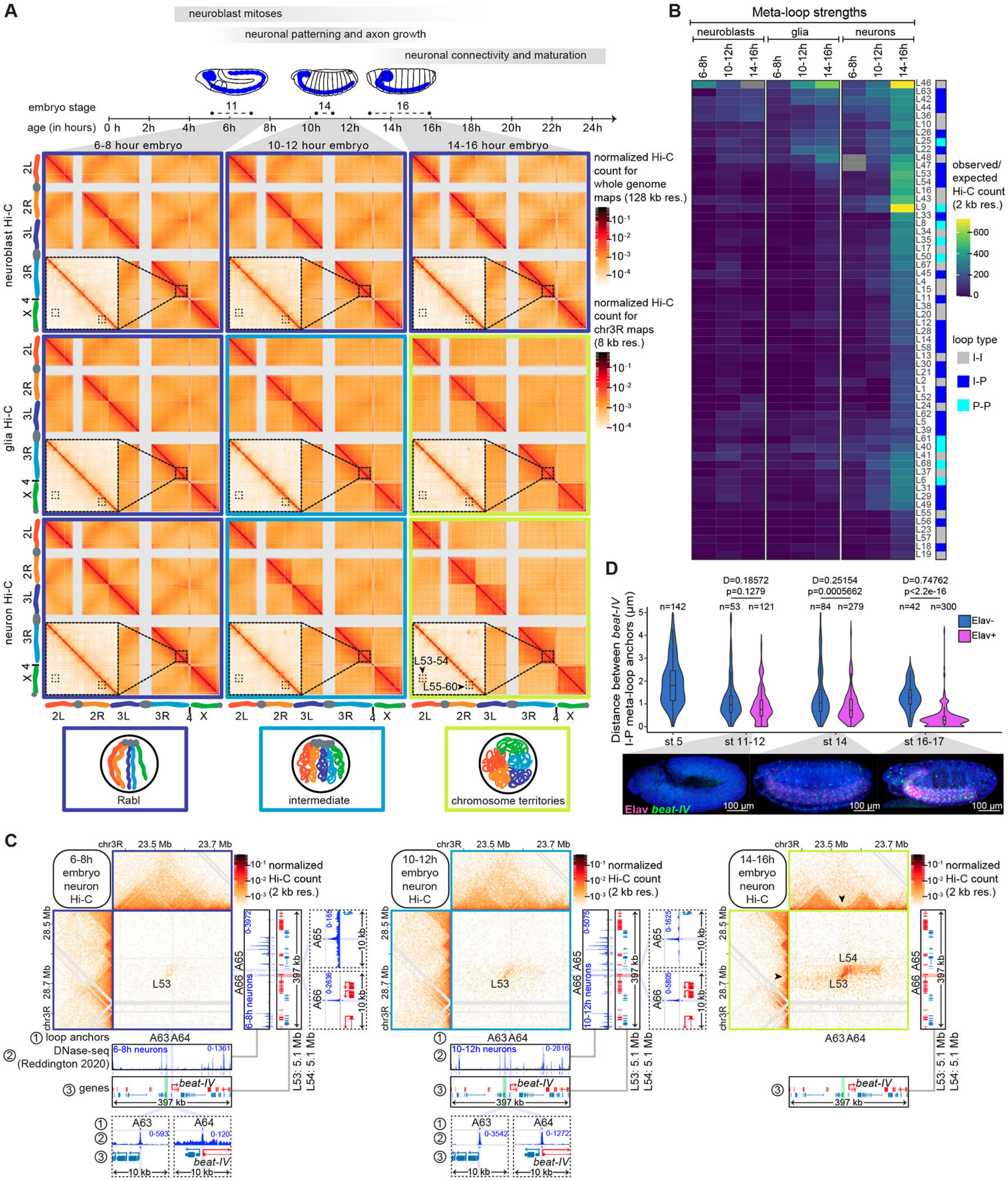


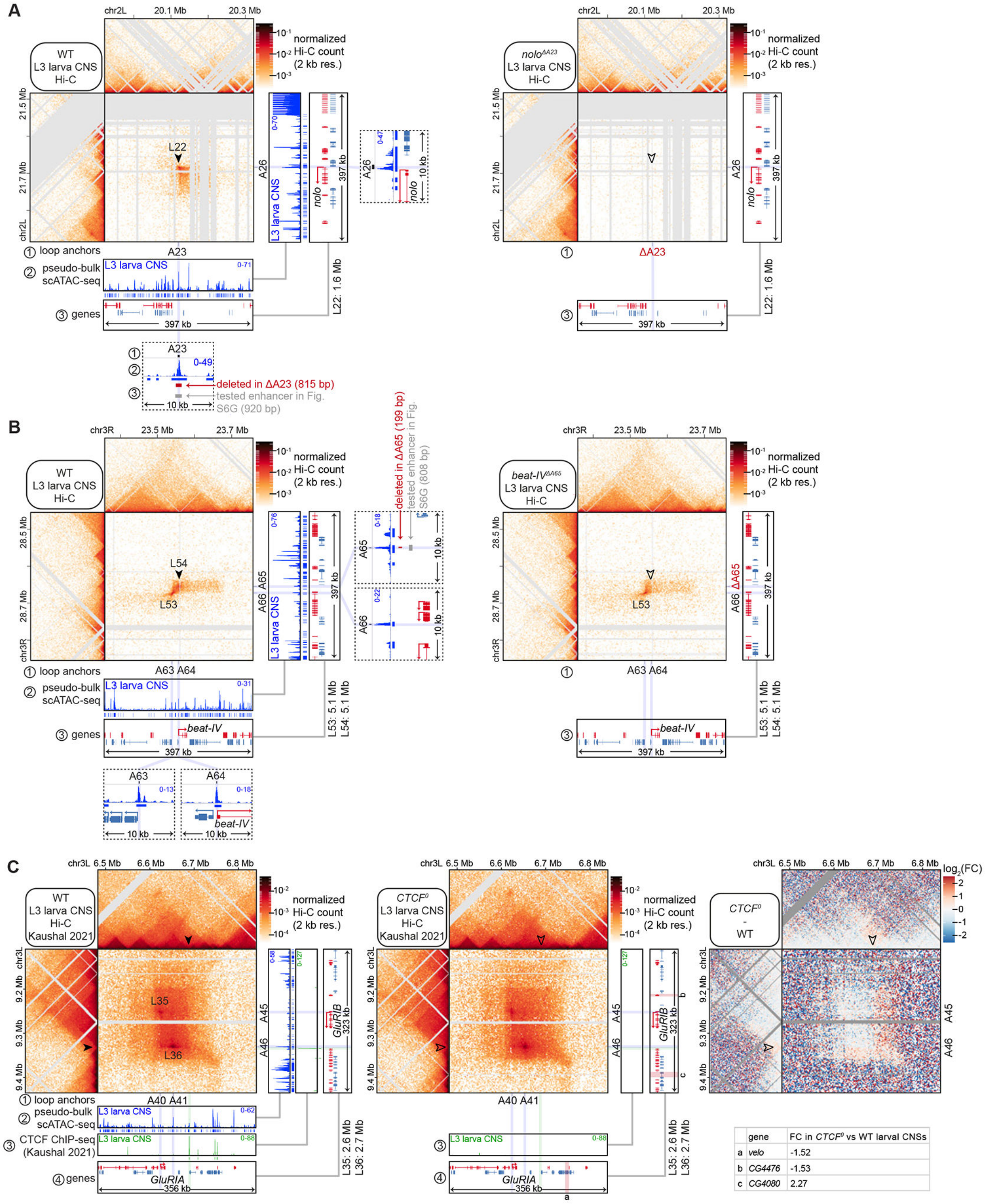
Figure 3. Meta-loop anchors come into enhanced proximity in a large fraction of neurons during neuronal differentiation. See also Figures S4-S5.

A. Top: Hours in embryogenesis, embryonic stages and major events in neurogenesis. Bottom: Whole-genome Hi-C maps in neuroblasts (row 1), glia (row 2) or neurons (row 3) FACS-purified from embryos of the indicated ages (columns). Dotted squares show zoom-ins around example meta-loops. 3D genome configurations deduced by Hi-C are schematized below and bounded by colored boxes.

B. For each meta-loop 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). Loop types are color-coded on the right.

C. Similar to Fig. 1A but showing Hi-C maps and DNase-seq profiles in neurons isolated from staged embryos by Reddington et al. 2020 [40] at the *beat-IV* meta-domain, in neurons of 6-8 (left), 10-12 (middle) or 14-16 (right) hour old embryos. Arrowheads mark emerging TAD boundaries near meta-loop anchors. DNA-FISH probes used in panel D are highlighted in green or red in the gene tracks.

D. Violin plot of distances (in y, in μm) measured by DNA-FISH with probes against *beat-IV* and its 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 with a probe against *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.



	gene	FC in *CTCF^0* vs WT larval CNSs
a	*velo*	-1.52
b	*CG4476*	-1.53
c	*CG4080*	2.27

Figure 4. Meta-domains are formed by meta-loop anchors and delimited by TAD boundaries.

A. Similar to Fig. 1A but showing Hi-C maps of the *nolo* meta-domain in WT (left) or *nolo*$^{\Delta A23}$ (right) larval CNSs. Filled/empty arrowheads point to meta-loop L22 presence/ absence in each genotype. The extents of meta-loop anchor A23 deletion or testing for enhancer activity in Fig. S6G are indicated below.

B. Same as A but for the *beat-IV* meta-domain in WT (left) or *beat-IV*$^{\Delta A65}$ mutants (right).

C. Similar to A but also showing CTCF ChIP-seq profiles and called peaks at the *GluRIA-GluRIB* meta-domain in WT (left) or *CTCF*^{0} (middle) larval CNSs [49]. Filled/ empty arrowheads mark normal/weakened TAD boundaries in each genotype. A differential

(*CTCF*^{0} minus WT) Hi-C map is shown on the right (scale is log2 fold change of contacts). Differentially expressed genes defined by RNA-seq in *CTCF*^{0} 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 in the middle, and the fold changes are indicated in the table.

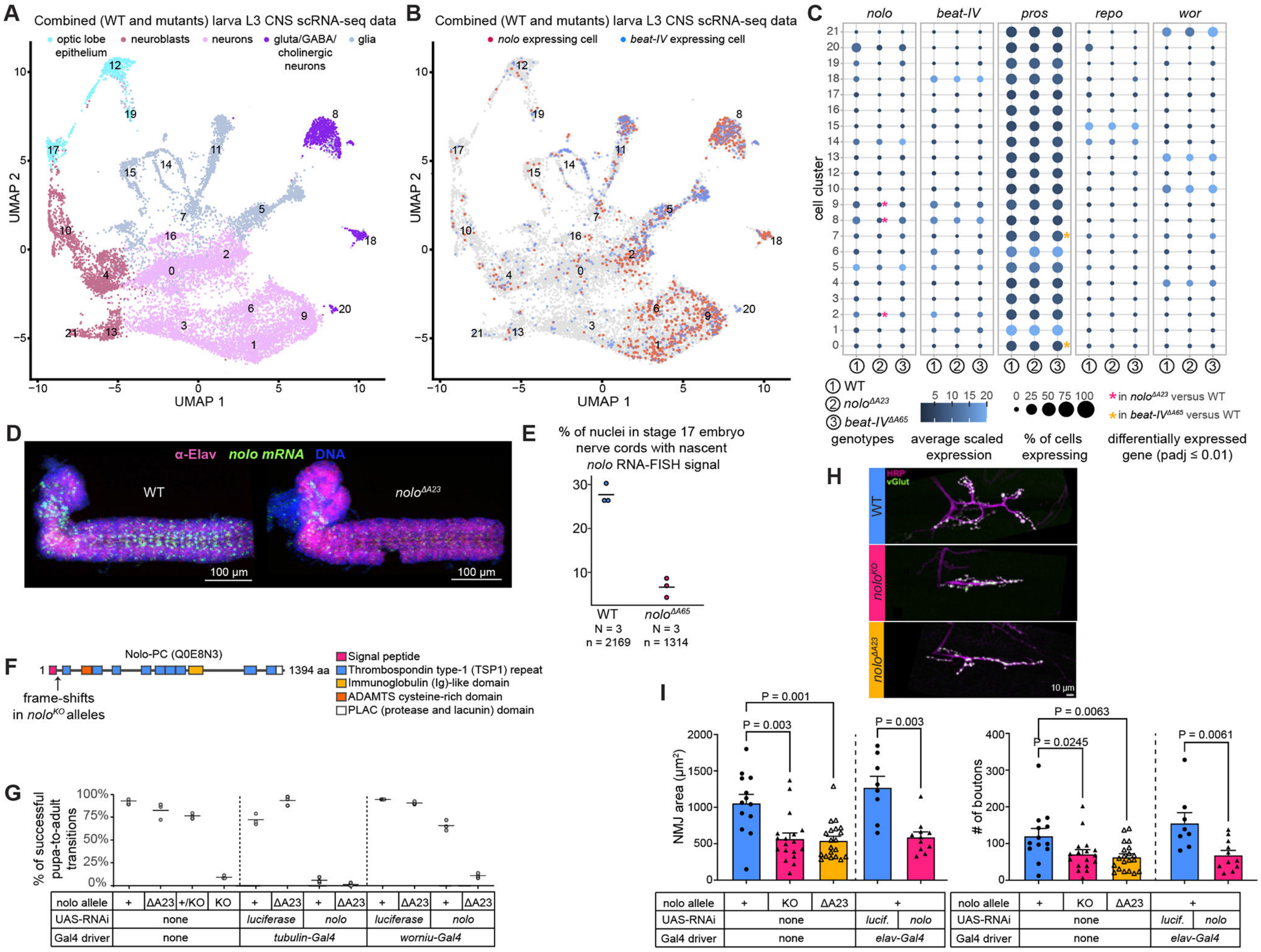


Figure 5. An intergenic anchor enables distal activation of *nolo*. See also Figure S6.

A. Combined scRNA-seq datasets generated in WT, *nolo*$^{\Delta A23}$ and *beat-IV*$^{\Delta A65}$ third instar larval CNSs. Cell clusters are numbered, and major cell types color-coded.

B. Same as A but coloring cells expressing *nolo* (red) or *beat-IV* (blue).

C. For each cell cluster, 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, *pros, repo* and *wor* are respectively neuronal, glial and neuroblast markers. Asterisks mark cell clusters that significantly (padj≤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 against *nolo* mRNA (green) in dissected nerve cords of stage 16 WT and *nolo*$^{\Delta A23}$ mutant embryos immunostained with α-Elav (pink) 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 stage 17 WT (left) or *nolo*$^{\Delta A23}$ (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 longest Nolo protein isoform (Uniprot accession Q0E8N3). Protein domains and positions of premature stop codons present in *nolo*^{KO} alleles are marked.

G. Percentages of indicated genotypes that successfully developed from pupae to fully hatched adults in three independent batches of 30-40 animals (dots). Horizontal lines show means.

H. Immunostaining of synaptic boutons (α-vGlut) and the whole neuromuscular junction (α-HRP) in muscle 6/7 terminals in WT and the indicated *nolo* mutant third instar larvae. Representative images are shown, and results from several images are quantified in panel I.

I. Quantification of neuromuscular junction (NMJ) area (left) and numbers of synaptic boutons (right) from n≥8 larvae (data points) 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.

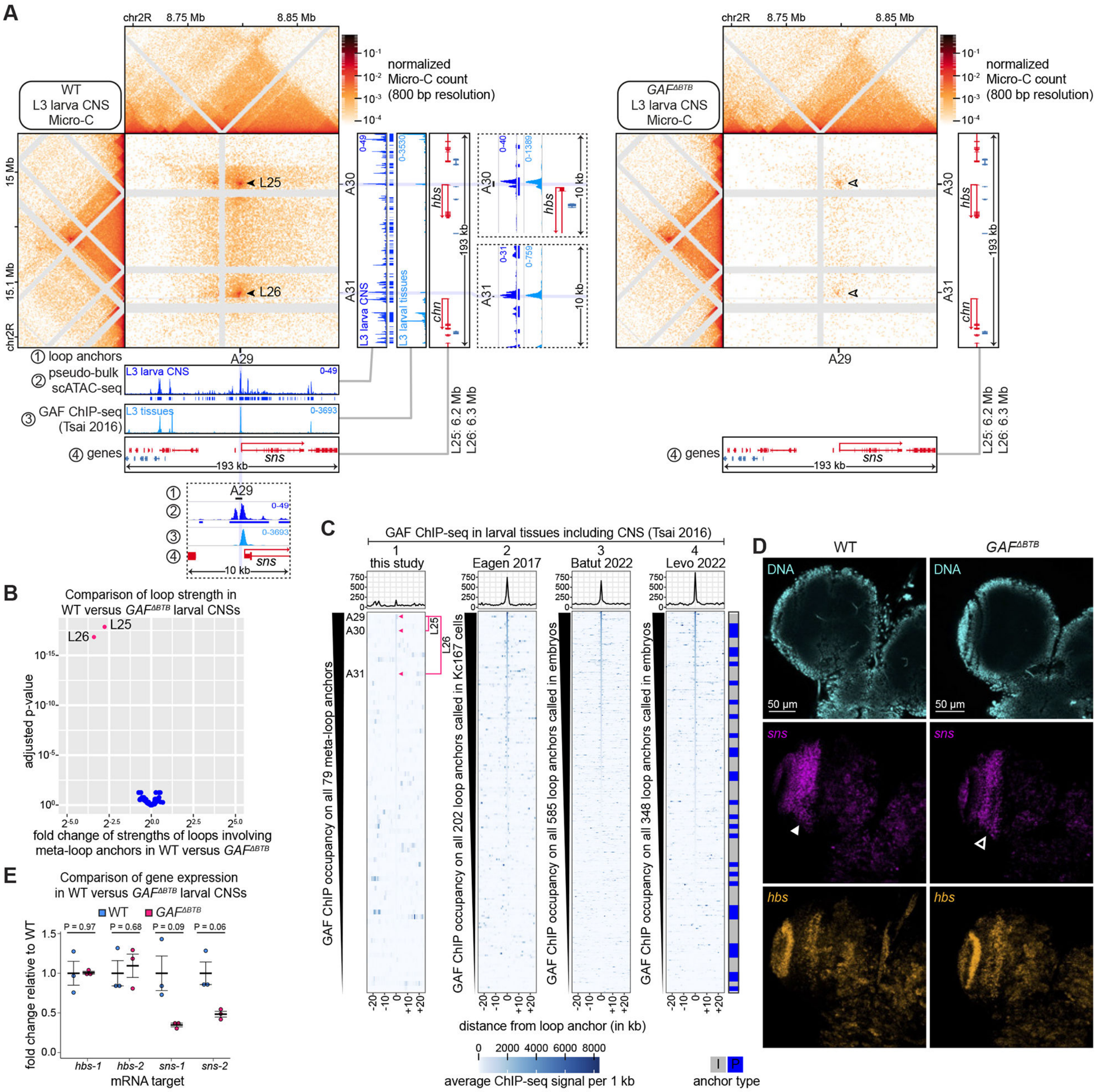


Figure 6: The GAF BTB domain is required to form the *sns-hbs* meta-domain.

A. Similar to Fig. 1A but also showing GAF ChIP-seq profiles in larval tissues including the CNS [52] at the *sns-hbs* meta-domain in WT (left) and *GAF$^{\Delta BTB}$* (right) larval CNSs. Filled/empty arrowheads point to WT/reduced meta-loop strength in each genotype. Differential analysis of strengths of loops involving meta-loop anchors (dots) measured by Micro-C in WT versus *GAF$^{\Delta BTB}$* larval CNSs. Labeled loops (pink dots) were significantly weaker (padj≤0.01) in *GAF$^{\Delta BTB}$* compared to WT.

C. Distribution of GAF ChIP-seq signal in larval tissues including CNS [52] 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. HCR RNA-FISH with co-hybridized probes against sns (magenta) and *hbs* (yellow) mRNAs in WT (left) and $GAF^{\Delta BTB}$ (right) larval CNSs. Filled/empty arrowheads mark a stripe with WT/reduced sns expression in each genotype. This phenotype was observed in 100% of $GAF^{\Delta BTB}$ 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 $GAF^{\Delta BTB}$ 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.

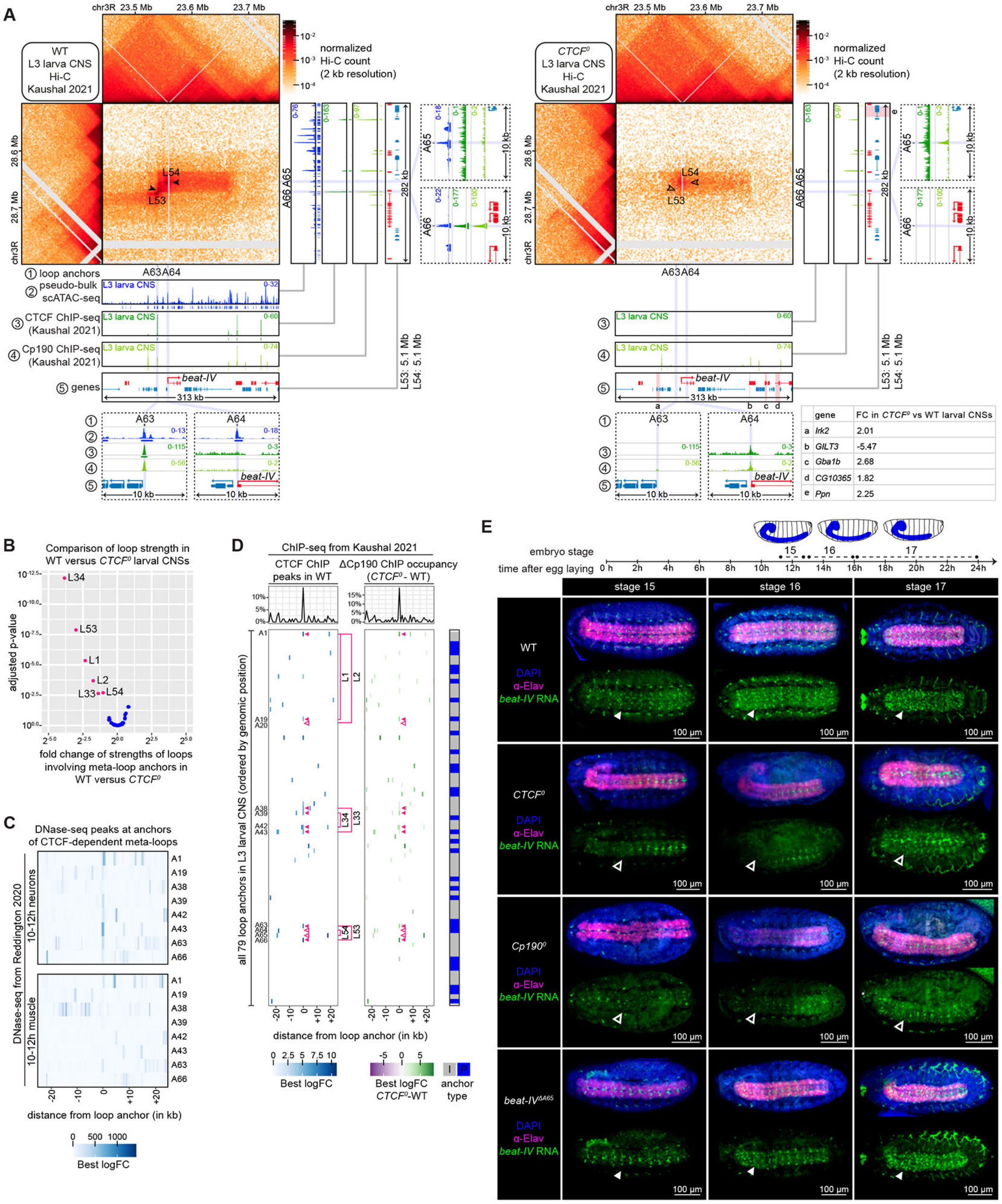


Figure 7: CTCF is required to form a different set of meta-loops than GAF. See also Figure S7.

A. Same as Fig. 4C but also showing Cp190 ChIP-seq profiles and called peaks [49] at the *beat-IV* meta-domain in WT (left) or *CTCF⁰* (right) larval CNSs.

B. Differential analysis of strengths of loops involving meta-loop anchors (dots) measured by Hi-C in WT versus *CTCF⁰* larval CNSs. Labeled loops (pink dots) were significantly weaker (padj≤0.01) in *CTCF⁰* compared to WT.

C. Distribution of DNase-seq signal from Reddington et al. 2020 [40] around the anchors of CTCF-dependent meta-loops shown in B, 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 *CTCF*^{0} versus WT larval CNSs (data from Kaushal et al. 2021 [20]) 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 *CTCF*^{0} mutants over all anchors are shown on top. Pink filled/empty arrowheads indicate anchors of meta-loops weakened in *CTCF*^{0} 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 with a probe against *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 WT/reduced *beat-IV* RNA-FISH signal in each genotype. Phenotypes shown here were observed in 100% of embryos (≥ 6 embryos imaged per genotype and stage).

KEY RESOURCES TABLE

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
Antibodies		
sheep anti-Dig peroxidase	Roche	11207733910
mouse monoclonal anti-Elav	Developmental Studies Hybridoma Bank	9F8A9
rat monoclonal anti-Elav	Developmental Studies Hybridoma Bank	7E8A10
mouse monoclonal anti-Repo	Developmental Studies Hybridoma Bank	8D12
rat anti-Wor	Abcam	ab196362
mouse monoclonal anti-vGlut clone 10D6G	Banerjee et al. 2021 [66]	10D6G
Alexa 647 anti-rat IgG	Thermo Scientific	A21247
Alexa 647 goat anti-HRP	Jackson ImmunoResearch	123-605-021
Alexa 555 anti-mouse IgG	Thermo Scientific	A-21422
Bacterial and virus strains		
DH5alpha E. coli	New England Biolabs	C2987H
Chemicals, peptides, and recombinant proteins		
16% formaldehyde	Thermo Scientific	20593
DSG (disuccinimidyl glutarate)	Thermo Scientific	20593
EGS (ethylene glycol bis(succinimidyl succinate))	Thermo Scientific	21565
Micrococcal nuclease	New England Biolabs	M0247S
DNA polymerase I, large (Klenow) fragment	New England Biolabs	M0210
T4 polynucleotide kinase	New England Biolabs	M0201
T4 DNA ligase	New England Biolabs	M0202
Exonuclease III	New England Biolabs	M0206
Biotin-14-dATP	Jena Bioscience	NU-835-BIO14
Biotin-11-dCTP	Jena Bioscience	NU-809-BIOX
Collagenase I	Sigma	C2674
Papain	Sigma	P4762
RNase inhibitor	Sigma	3335399001
MseI	New England Biolabs	R0525S
Csp6I	Thermo Scientific	ER0211
Hoechst 33342	Sigma-Aldrich	23491-52-3
UltraPure™ SSC, 20X	Thermo Fisher Scientific	15557044
ProLong Gold Antifade Mountant	Thermo Fisher Scientific	P36934
TRIzol Reagents	Thermo Fisher Scientific	15596026
Phenol:Chloroform:Isoamylalcohol (PCI, 25:24:1)	Thermo Scientific	15593031
T7 RNA polymerase	Roche	10881767001
Dig-UTP labeling mix	Roche	11277073910
DNase I recombinant, RNase-free	Roche	04716728001

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
aminoallyl-dUTP-XX-ATTO-488	Jena Bioscience	NU-803-XX-488-S
aminoallyl-dUTP-ATTO-550	Jena Bioscience	NU-803-550-S
Critical commercial assays		
Zymoclean Gel DNA Recovery Kit	Zymo Research	D4007
TSA Plus Cyanine 3/Fluorescein System	Perkin Elmer	NEL753001KT
HCR buffer set	Molecular Instruments	www.molecularinstruments.com
HCR B1 488 amplifier	Molecular Instruments	www.molecularinstruments.com
HCR B2 546 amplifier	Molecular Instruments	www.molecularinstruments.com
Moxi V	Witec	MXV102
10x Genomics Chromium single cell ATAC reagent kits solution v1.0	10x Genomics	CG000168
10x Genomics Chromium Next GEM sc 3’ reagent kits v3.1 dual index	10x Genomics	CG000315
SPRI Select beads	Beckman Coulter	B23319
Ampure XP beads	Beckman Coulter	A63880
NEBNext Ultra II DNA Library Prep kit for Illumina	New England Biolabs	E7645S
Dynabeads MyOne Streptavidin T1 beads	Invitrogen	65601
KAPA HiFi HotStart Ready Mix	Roche	KK2501
Nick-translation kit	Abbott Molecular	07J100-001
Superscript IV	Invitrogen	18091050
GoScript	Promega	A2801
SYBR green	Invitrogen	4367659
GoTaq	Promega	A600A
Deposited data		
Raw and processed data (Hi-C, scATAC-seq, scRNA-seq)	This paper	GEO: GSE214707
Raw and processed data (Micro-C)	This paper	GEO: GSE228095
Experimental models: Organisms/strains		
D. melanogaster: nolo intergenic anchor deletion: nolo$^{\Delta A23}$	This paper	N/A
D. melanogaster: beat-IV intergenic anchor deletion: beat-IV$^{\Delta A65}$	This paper	N/A
D. melanogaster: nolo knock-out: nolo^{KO}	This paper	N/A
D. melanogaster: UAS-luciferase RNAi: y[1] v[1]; +; P{UAS-LUC.VALIUM10}attP2	Bloomington Drosophila Stock Center	35788
D. melanogaster: UAS-nolo RNAi: y[1] v[1]; +; P{UAS-nolo RNAi TRiP.MHC03983(v+)}attP2	Bloomington Drosophila Stock Center	55296
D. melanogaster: tubulin-Gal4 driver: y[1] w[*]; +; P{tubP-GAL4}LL7/TM3	Bloomington Drosophila Stock Center	5138
D. melanogaster: worniu-Gal4 driver: w[*]; P[wor.GAL4.A}2; Dr[1]/TM3	Bloomington Drosophila Stock Center	56553
D. melanogaster: enhancer assay negative control: w; +; {DSCP-Gal4}VK33	Gambetta lab	N/A
D. melanogaster: enhancer assay beat-IV intergenic anchor A65: w; +; {beat-IV A65}VK33	Gambetta lab	N/A

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
D. melanogaster: enhancer assay nolo intergenic anchor A23: w; +; {nolo A23}VK33	Gambetta lab	N/A
D. melanogaster: enhancer assay DIP-ε A12 intergenic anchor: w; +; {DIP-ε A12}VK33	Gambetta lab	N/A
D. melanogaster: enhancer assay DIP-ε A14 intergenic anchor: w; +; {DIP-ε A14}VK33	Gambetta lab	N/A
D. melanogaster: enhancer assay 5-HT7 A60 intergenic anchor: w; +; {5-HT7 A60}VK33	Gambetta lab	N/A
Drosophila melanogaster: GAF$^{\Delta BTB}$ *: Halo-Trl*$^{\Delta POZ\ 90\text{-}119}$*/ TM6B*	Tang et al., 2022 [51]	N/A
Drosophila melanogaster: CTCF^{0}	Kaushal et al., 2021 [49]	N/A
Drosophila melanogaster: Cp190^{0}	Kaushal et al., 2022 [20]	N/A
Drosophila virilis: wildtype	National Drosophila Species Stock Center	N/A
Oligonucleotides		
Oligonucleotides used for cloning, RT-qPCR and FISH experiments	Table S6	N/A
Software and algorithms		
Custom code	This paper	https://github.com/gambettalab/meta-loops-2022
HiGlass v1.11.8	Kerpedjiev et al., 2018 [67]	https://higlass.io
BWA	N/A	https://bio-bwa.sourceforge.net/
distiller v0.3.3	N/A	https://github.com/open2c/distiller-nf
cooler v0.8.11	Abdennur and Mirny, 2019 [68]	https://github.com/open2c/cooler
R V4.2.0	R Core Team	https://www.R-project.org/
ggplot2 v3.3.6	N/A	https://ggplot2.tidyverse.org/
data.table v1.14.2	N/A	https://CRAN.R-project.org/package=data.table
EBImage V4.38.0	Pau et al., 2010 [69]	http://bioconductor.org/packages/EBImage
patchwork v1.1.1	N/A	https://CRAN.Rproject.org/package=patchwork
rhdf5 V2.40.0	N/A	https://github.com/grimbough/rhdf5
igraph v1.3.5	N/A	https://igraph.org
mlpack v4.0.0	N/A	https://CRAN.R-project.org/package=mlpack
LiftOver	Hinrichs et al., 2006 [70]	http://genome.ucsc.edu/cgi-bin/hgLiftOver
bedtools merge	Quinlan and Hall, 2010 [71]	https://bedtools.readthedocs.io/en/latest/content/tools/merge.html
pairtools v1.0.2	Open2C, Abdennur et al., 2023 [72]	https://pairtools.readthedocs.io/en/1.0.2/
DESeq2	Love et al., 2014 [73]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
UCSC Kent's utilities v377	Kuhn et al., 2013 [74]	http://hgdownload.cse.ucsc.edu/admin/exe/
CNEr Bioconductor package	Tan et al., 2019 [75]	https://bioconductor.org/packages/release/bioc/html/CNEr.html
LASTZ	Harris et al., 2007 [76]	https://www.bx.psu.edu/~rsharris/lastz/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fiji software v2.1.0/1.53c	Schindelin et al., 2012 [77]	https://fiji.sc/
bcl2fastq Conversion software v2.20	Illumina	https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html
Cell Ranger ATAC software v1.2.0	10x Genomics	https://support.10xgenomics.com/single-cell-atac/software/pipelines/latest/installation
Seurat v3.1.5	Stuart et al., 2019 [78]	https://satijalab.org/seurat/
Signac v0.2.5	Stuart et al., 2021 [79]	https://stuartlab.org/signac/
CellRanger v6.1.2	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation
clusterProfiler v3.14.0	Yu et al., 2012 [80]	https://guangchuangyu.github.io/software/clusterProfiler/
3D-Distance-Tool	N/A	https://imagej.nih.gov/ij/macros/tools/3D_Distance_Tool.txt
Other		
D. virilis reference genome	Renschler et al. 2019 [81]	GEO: GSE120751
D. melanogaster third instar larva CNS Hi-C, RNA-seq, CTCF ChIP-seq, Cp190 ChIP-seq	Kaushal et al. 2021 [49]	GEO: GSE146752
D. melanogaster adult brain snATAC-seq	Janssens et al. 2022 [46]	GEO: GSE163697
D. melanogaster adult brain bulk ATAC-seq	Janssens et al. 2022 [46]	GEO: GSE181494
D. virilis whole embryo DNase-seq	Liu et al. 2021 [64]	ArrayExpress: E-MTAB-3797 and E-MTAB-9480