Chromosome Structure I: Loop extrusion or boundary:boundary pairing?

Xinyang Bing[1,4]*, Wenfan Ke[2]*, Miki Fujioka[3]*, Amina Kurbidaeva[2,5], Sarah Levitt[2], Mike Levine[1], Paul Schedl[2], and James B. Jaynes[3]

1: Lewis Sigler Institute, Princeton University, Princeton, NJ. 08544

2: Department of Molecular Biology, Princeton University, Princeton NJ 08544

3: Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, PA 19107

4: Present address: BlueRock Therapeutics, Cambridge, Ma 20142

5: Present address: Center for Genomics and Systems Biology, New York University, NY 10003

* These authors contributed equally.

Abstract

Two different models have been proposed to explain how the endpoints of chromatin looped domains ("TADs") in eukaryotic chromosomes are determined. In the first, a cohesin complex extrudes a loop until it encounters a boundary element roadblock, generating a stem-loop (and an unanchored loop). In this model, boundaries are functionally autonomous: they have an intrinsic ability to halt the movement of incoming cohesin complexes that is independent of the properties of neighboring boundaries. In the second, loops are generated by boundary:boundary pairing. In this model, boundaries are functionally non-autonomous, and their ability to form a loop depends upon how well they match with their neighbors. Moreover, unlike the loop-extrusion model, pairing interactions can generate both stem-loops and circle-loops. We have used a combination of MicroC to analyze how TADs are organized and experimental manipulations of the *even skipped* TAD boundary, *homie*, to test the predictions of the "loop-extrusion" and the "boundary-pairing" models. Our findings are incompatible with the loop-extrusion model and instead suggest that endpoints of TADs in flies are determined by a mechanism in which boundary elements physically pair with their partners, either head-to-head, or head-to-tail, with varying degrees of specificity. How the partners find each other is not clear but is unlikely to require a loop extrusion mechanism.

Introduction

The chromosomes of multicellular animals have a regular and inheritable physical organization. This was first recognized in studies on the lampbrush chromosomes in amphibian oocytes and the polytene chromosomes of insects (Callan, 1963; Nora et al., 2020; Zhimulev and Belyaeva, 1991; Zhimulev et al., 1983). Subsequent research has shown that the architectural features inferred from analysis of lampbrush and polytene chromosomes are present in chromosomes throughout much of the animal kingdom. The key organizational principle is the subdivision of the chromatin fiber into a series of independent looped domains, commonly called "TADs" (Cavalheiro et al., 2021; Chetverina et al., 2017; Jerković et al., 2020; Matthews and White, 2019; Rowley and Corces, 2018). With important exceptions, the arrangement of TADs along a given chromosome tend to be invariant and are largely independent of the cell type or developmental stage. This regular and inheritable organization is a reflection of the underlying mechanism of TAD formation. TADs are separated from each other by special elements called boundaries or insulators. While these elements have been found in many different species, they have been most fully characterized in *Drosophila* (Cavalheiro *et al.*, 2021; Chetverina *et al.*, 2017). Fly boundaries span DNA sequences of 150 bp to 1.5 kb in length and contain one or more nucleosome-free nuclease-hypersensitive regions. These nuclease-hypersensitive regions are targets for a large collection of DNA binding proteins that have been implicated in boundary function. Some of the fly boundary factors are widely conserved (CTCF, BEN family proteins, GAF), while others appear to be restricted to insect linages (Su(Hw), Pita, Zw5, Pita, Mod(mdg4), BEAF) (Heger et al., 2013; Heger and Wiehe, 2014; Schoborg and Labrador, 2010).

Boundary elements in flies are not only responsible for organizing the chromatin fiber, they also have genetic activities. When interposed between enhancers or silencers and target

53 promoters, boundary elements block regulatory interactions (Bell et al., 2001; Chetverina et al.,
54 2014; Chetverina *et al.*, 2017). This insulating activity provides a mechanism for delimiting
55 units of independent gene activity: genes located between a pair of compatible boundaries are
56 subject to regulatory interactions with enhancers/silencers present in the same chromosomal
57 interval, while they are insulated from the effects of enhancers/silencers located beyond either
58 boundary in adjacent regulatory neighborhoods. Genetic studies suggest that the insulating
59 activity of boundary elements is a consequence of subdividing the chromosome into a series of
60 topologically independent domains (Cai and Shen, 2001; Gohl et al., 2011; Muravyova et al.,
61 2001). Organizing the chromatin fiber into looped domains enhances contacts between
62 sequences within the loop, while it suppresses contacts with sequences outside of the loop.
63 While it is not known whether a similar insulation mechanism is at play in mammals, the fact
64 that mammalian chromosomes are also subdivided into TADs by boundary-like elements
65 suggests that it might be.

66 A critical question in chromosome biology is the mechanism(s) responsible for
67 determining the endpoints of the loop domains, the TADs. In mammals, a novel loop-extrusion
68 model has been proposed to not only generate chromosomal loops but also determine the
69 endpoints of those loops (Alipour and Marko, 2012; Fudenberg et al., 2016; Guo et al., 2012;
70 Guo et al., 2015; Nuebler et al., 2018; Sanborn et al., 2015). In this model, a cohesin complex
71 initiates loop formation at a loading site within the "loop-to-be" and then extrudes a chromatin
72 loop until it bumps into CTCF-dependent roadblocks, one on each side of the extruding loop.
73 The location of these roadblocks combined with the processive action of the cohesin complex
74 determines the endpoints of each TAD (Davidson and Peters, 2021; Ghosh and Meyer, 2021;
75 Mirny and Dekker, 2022; Perea-Resa et al., 2021). A key assumption of the loop-extrusion
76 model is that mammalian boundaries are fully autonomous: they are roadblocks, and their

77 physical presence in and of itself is sufficient to define the loop endpoint, independent of the
78 functional properties of neighboring boundaries. In a more specific variant of this model, the
79 relative orientation of the paired CTCF roadblocks is also important—in order to halt cohesin-
80 mediated extrusion, the CTCF sites on each side of the loop must have a convergent orientation.
81 In this model the CTCF roadblocks are able to block an oncoming cohesin complex in only one
82 direction; however, their intrinsic blocking activity is independent of the orientation of other
83 CTCF sites in their neighborhood.

84 While loop extrusion is widely thought to be operative in mammals, the evidence in flies
85 is inconsistent with this mechanism. Genetics studies have shown that fly boundaries are
86 functionally non-autonomous and that their activities in both loop formation and gene regulation
87 depend upon their ability to engage in direct physical interactions with other boundaries
88 (Chetverina *et al.*, 2014; Chetverina *et al.*, 2017). That fly boundaries might function by
89 physical pairing was first suggested by studies which showed that the *gypsy* transposon *su(Hw)*
90 boundary and the bithorax complex (BX-C) *Mcp* boundary can mediate regulatory interactions
91 (enhancer/silencer: reporter) between transgenes inserted at sites separated by 15-32 MB or more
92 (Muller et al., 1999; Sigrist and Pirrotta, 1997; Vazquez et al., 1993). Consistent with direct
93 physical interactions, these distant transgenes co-localize *in vivo*, and in living tissue remain in
94 contact for extended periods of time (Chen et al., 2018; Li et al., 2011; Vazquez et al., 2006).

95 Some of the parameters that govern pairing interactions have been defined in two different
96 transgene assays, insulator bypass and boundary competition. In one version of the bypass
97 assay, a set of enhancers are placed upstream of two different reporters (Cai and Shen, 2001;
98 Kyrchanova et al., 2008a; Muravyova *et al.*, 2001). When a single boundary is introduced
99 between the enhancers and the closest reporter, the enhancers are unable to activate either
100 reporter. However, if the same boundary is placed downstream of the closest reporter, bypass is

101 observed. In this case the closest reporter, which is bracketed by the two boundaries, is still
102 insulated from the enhancers; however, the downstream reporter is activated (Muravyova *et al.*,
103 2001). Kyrchanova *et al.* (Kyrchanova *et al.*, 2008a) showed that bypass activity is, in most
104 cases, orientation-dependent. When the same boundary is placed in the upstream and
105 downstream position, they typically need to be introduced in the opposite orientation. The
106 reason for this is that self-pairing is head-to-head and this configuration generates a stem-loop
107 topology, bringing the upstream enhancers in close proximity to the downstream reporter. When
108 the two boundaries are introduced in the same orientation, the enhancers do not activate the
109 downstream reporter because a circle-loop rather than a stem-loop is formed (Chetverina *et al.*,
110 2017; Kyrchanova *et al.*, 2008a). Head-to-head pairing seems to be a common feature of self-
111 pairing interactions between fly boundaries and was observed for *scs*, *scs'*, *iA2* and *wari*
112 (Kyrchanova *et al.*, 2008a). A preference for head-to-head interactions is also observed for
113 heterologous pairing interactions between different boundaries in the *Abdominal-B* (*Abd-B*)
114 region of the BX-C (Kyrchanova et al., 2011; Kyrchanova et al., 2008b). There are of course
115 some exceptions. The BX-C *Fab-7* boundary pairs with itself and with its neighbor *Fab-8* in
116 both orientations (Kyrchanova *et al.*, 2011; Kyrchanova *et al.*, 2008b). Bypass activity also
117 depends upon finding the proper match. For example, when multimerized dCTCF or Zw5
118 binding sites are placed both upstream and downstream of the closest reporter, bypass is
119 observed. However, there is no bypass when multimerized dCTCF sites are tested with
120 multimerized Zw5 sites (Kyrchanova *et al.*, 2008a). Similar results have been reported for the
121 Zipic and Pita factors (Zolotarev et al., 2016). Boundary:boundary pairing preferences are also
122 observed in competition assays in which different boundaries are introduced into the same
123 transgene (Gohl *et al.*, 2011). In these assays, the insulating activity of a boundary in the
124 blocking position (between an enhancer and a reporter that is flanked by a 3' boundary) is

125 challenged by placing a heterologous boundary upstream of the enhancer. If the boundary
126 upstream of the enhancer is a better match for the boundary located downstream of the reporter,
127 then the insulating activity of the boundary in the blocking position can be compromised or lost
128 altogether. These and other experiments argue that fly boundaries are functionally non-
129 autonomous—that is, their activities depend upon their ability to physically interact with other
130 boundaries.

131 However, the differences between the two models are not limited to the mechanisms for
132 determining the endpoints of TADs. The models also differ in the possible topologies of the
133 chromatin loops that form TADs. Only two loop topologies can be generated by the cohesin-
134 extrusion mechanism. One is a stem-loop while the other is an unanchored loop. Depending
135 upon the location of the road-blocks that halt cohesin extrusion and the cohesin loading sites, the
136 chromatin fiber would be organized into a series of stem-loops. These stem-loops will be
137 separated from each other by unanchored loops (see Fig. 1B). This is the configuration that is
138 often illustrated in articles discussing the loop-extrusion model. If the relevant road-blocks are
139 very closely juxtaposed, then the unanchored connecting loops will disappear and be replaced by
140 a series of stem-loops which point in opposite orientations, as illustrated in Fig. 1C. Stem-loops
141 are also possible in the boundary-pairing model, and they will be formed when two heterologous
142 boundaries pair with each other head-to-tail (see Fig 1D) in *cis*. Like loop-extrusion, the stem-
143 loops could be separated from each other by unanchored loops (Fig. 1E). If the boundaries in the
144 neighborhood pair with each other head-to-tail, a series of connected stem-loops pointing in
145 opposite orientations will be generated (Fig. 1D). In this case, the main axis of the chromosome
146 would be a series of paired boundaries (Fig. 1F).

147 Stem-loops are not, however, the only loop topology that can be generated in the boundary-
148 pairing model. If the neighboring boundaries pair with each other head-to-head instead of head-

149 to-tail, a circle-loop structure will be generated (Fig 1G). Like stem-loops, the circle-loops could
150 be connected by unanchored loops (Fig. 1H). Alternatively, if boundaries pair with both
151 neighbors head-to-head this will generate a series of linked circle-loops oriented in (more or less)
152 the same direction (see Fig. 1I).

153 In the studies reported here we have critically evaluated these two models. We have
154 used a combination of MicroC to analyze how TADs are organized and experimental
155 manipulations to test the predictions of the "loop-extrusion" and the "boundary-pairing"
156 models. For these experimental manipulations, we have used the well-characterized *homie*
157 boundary from the *even skipped* (*eve*) locus. *homie* together with *nhomie* flank the 16 kb
158 *eve* locus, and these two elements share many of the properties of other fly boundaries
159 (Fujioka et al., 2013; Fujioka et al., 2009). Like the *gypsy* and *Mcp* boundaries, *nhomie* and
160 *homie* can mediate long-distance regulatory interactions (Fujioka et al., 2016). For
161 example, a *lacZ* reporter transgene containing *homie* (or *nhomie*) inserted on one homolog
162 can be activated by an *eve* enhancer in a *homie*- (or *nhomie*-) containing transgene inserted
163 on the other homolog over a distance of 2 Mb. Long-distance activation is also observed in
164 heterologous combinations of *nhomie* and *homie*, but not in combinations with phage
165 *lambda* DNA. Similarly, regulatory interactions are observed when a *homie*- or *nhomie*-
166 containing reporter is inserted at an attP landing site in the *hebe* gene, 142 kb upstream of
167 the *eve* locus (Chen *et al.*, 2018; Fujioka *et al.*, 2016; Fujioka *et al.*, 2009). In this case, the
168 *eve* enhancers drive reporter expression during development in a pattern that mimics the
169 stage- and tissue-specific expression of the endogenous *eve* gene. Activation also depends
170 upon the orientation of the boundary (5'→3') relative to the reporter in the transgene. For
171 *homie*, the reporter must be located "upstream" of the boundary (5'→3'), just like the *eve*
172 enhancers (and *eve* gene) are "upstream" of endogenous *homie* (see Fig. 2B). Little or no

173 activation is observed when the relative orientation is reversed so that the reporter is located
174 "downstream" of *homie*. For *nhomie*, the reporter must be located "downstream", in the
175 same relative position as that in which the *eve* enhancers and *eve* gene are located with
176 respect to endogenous *nhomie* (5'→3': Fig. 2B).
177

178 **Results**

179 *The eve TAD*

180 We used MicroC to probe the TAD organization of the *eve* locus and the genomic region
181 that extends from *eve* to just beyond the attP site at –142 kb. For this purpose, we chose nuclear
182 cycle 14 embryos. At this stage of development, most nuclei in our sample will have completed
183 replication and will be arrested in G2. As sister chromosomes in flies pair with each other in
184 register, there will be two juxtaposed copies of this region of the chromosome. The sister
185 chromosomes should be linked to each other by cohesin in preparation for the asynchronous
186 nuclear divisions at the end of NC14 (Collier and Nasmyth, 2022; Gligoris et al., 2014). While
187 homolog pairing is limited in pre-cellular blastoderm embryos, there are regions of the
188 chromosome that already show evidence of pairing (Fung et al., 1998). In fact, transvection
189 between *gypsy* insulator-containing transgenes has been observed in live imaging studies of
190 nuclear cycle 14 embryos (Lim et al., 2018). Since the NC14 nuclei are just emerging from an
191 earlier mitosis and a round of DNA synthesis, one would imagine that other key features of
192 chromosome organization might also be in the process of being established. Thus, it should be
193 possible to capture evidence of intermediates like those expected to be generated as the extruding
194 cohesin complexes move through the *eve* TAD and other nearby loci at this particular stage of
195 development.

196 Shown in Fig. 2A is a blow-up of the TAD organization of the *eve* locus. As expected, the
197 left and right boundaries of the *eve* TAD correspond to *nhomie* and *homie*, respectively, and
198 sequences containing these two boundaries converge at the apex of the *eve* interaction triangle.
199 All of the sequences within the TAD appear to come in contact with each other and define a high
200 density of internal contact domain (HDIC). These contacts are expected to be generated, at least
201 in part, by the sliding of the chromatin fiber within the TAD against itself, much like sequences

202 in circular/supercoiled DNA can interact with each other (Fig. 1E). There are also sequences
203 within the *eve* TAD that show enhanced interactions. For example, there is a darker interaction
204 triangle connecting *nhomie* and the promoter region, while another triangle connects *homie* (and
205 the PRE adjacent to *homie*) to the promoter region. These enhanced internal contacts would be
206 expected to help insulate the *eve* gene from external regulatory elements and at the same time
207 facilitate interactions between *eve* and its upstream and downstream enhancers.

208 As predicted by both the loop-extrusion and boundary-pairing models, the *eve* TAD is a
209 stem-loop, not a circle-loop (Fig. 2B). This is evident from the prominent "volcano plume" at
210 the apex of the *eve* TAD (volcano triangle). A volcano plume is observed because sequences
211 flanking *eve* are brought into close proximity when *homie* and *nhomie* are linked to each other by
212 either cohesin or by head-to-tail pairing (see Fig.1: see also Ke et al.). Fig. Supplemental 1
213 shows the hierarchy of interactions between TADs to either side of *eve* (L-M, K-M and J-M in
214 Sup. Fig. 1). This conclusion is supported by previous transgene studies. As first described in
215 the boundary bypass experiments of Cai and Shen (Cai and Shen, 2001) and Muravyova et al.
216 (Muravyova *et al.*, 2001), the formation of a stem-loop structure when two boundaries are linked
217 together brings sequences immediately beyond the linked boundaries into close proximity with
218 each other. In this assay, a stem-loop structure is required, as it enables enhancers flanking the
219 "upstream" boundary to interact with reporters flanking the downstream boundary (Kyrchanova
220 *et al.*, 2008a). In contrast, a circle-loop configuration does not bring the neighbors to either side
221 of *eve* into close proximity.

222 While a stem-loop is the topology expected for the *eve* TAD in both models, one feature of
223 the TAD is inconsistent with the loop-extrusion model. According to this model, loops (TADs)
224 are generated by a cohesin complex that forms a small bubble at a loading site somewhere in the
225 TAD-to-be and then extrudes both strands of the growing loop at an equal rate until it bumps into

226 the boundary road-blocks (Davidson and Peters, 2021; Ghosh and Meyer, 2021; Mirny and
227 Dekker, 2022; Perea-Resa *et al.*, 2021). If the loading site were located in the middle of the
228 TAD-to-be, and the extrusion intermediates in a population of nuclei were captured by
229 crosslinking, one should observe a crosslink-generated stripe that begins at the loading site and
230 extends perpendicular to the chromosomal DNA until it reaches the apex that links the two TAD
231 boundaries with imaginary lines of 45° (Fig. Supplemental 2B). This is the pattern that would be
232 generated by a "broken zipper" when it transiently links the zipper teeth equidistant from the
233 initial "loading site". If the loading site is off center to the left, the crosslinking of the passing
234 cohesin complex in a population of nuclei will generate a perpendicular stripe until cohesin come
235 to a halt when it encounters the roadblock on the left (Fig. Supplemental 2C). Assuming that the
236 complex continues to extrude the right strand, it will then generate a stripe at 45° degrees that
237 extends to the apex of the TAD. If the loading site corresponds to one of the CTCF roadblocks
238 or is located close to the roadblock, the cohesin complex will generate a 45° stripe that comes to
239 a halt at the apex of the triangle when it encounters the neighboring CTCF roadblock (Fig.
240 Supplemental 2D and E). If there are multiple loading sites within the TAD, there should be a
241 series of perpendicular stripes that generate and sequentially reinforce the stripe(s) at 45° (Fig.
242 Supplemental 2F). However, in spite of the fact that NC14 fly embryos should provide by far
243 the best prospect of actually capturing loop extrusion intermediates, there are no vertical stripes
244 in the *eve* TAD, nor are there stripes at 45° that extend to the apex of the *eve* TAD. Nor do we
245 observe a population of perpendicular stripes that convert into a series of reinforced 45° stripes.
246 There is also no evidence of stripes that are initially tilted to the left or right of a "loading site",
247 as might be observed if the rate of extrusion were unequal on the two strands, and then
248 subsequently convert to a 45° stripe as cohesin comes to a halt at the first boundary encountered.
249 *TADs in the eve environs*

250 The attP insertion site at –142 kb is near the end of the first exon of the *hebe* transcription
251 unit (Fig. 2C). It is located just within a ~14 kb TAD, TB, that has a high density of internal
252 contacts, and extends from the *hebe* promoter to the 3' end of the neighboring *dila* gene. In
253 between *hebe* and the *eve* locus there at least are ten distinct TADs, TC-TL (Fig. 2C). The
254 endpoints of most of these TADs correspond to sequences that are associated with one or more
255 known chromosomal architectural factors (dots in Fig. 2C).
256 These TADs correspond to the fundamental building blocks for organizing the 3D
257 architecture of this chromosomal segment. Superimposed upon these TADs are regions that
258 exhibit lower density of contacts (LDC). For example, the *mef2* gene spans two TADs, TJ and
259 TK. TK contains the *mef2* distal promoter region and extends to the internal *mef2* promoter,
260 while TJ extends to near the promoter of the divergently transcribed *Etf-QO* gene (blue arrow).
261 Both of these TADs are linked together by a rectangular LDC domain J-K (see Fig.
262 Supplemental 2). Likewise, TJ is linked to its immediate neighbor TI by the LDC domain I-J. In
263 the next level, TADs separated by a single TAD are linked together. In the region immediately
264 above TJ, the LDC domain I-K links TI to TK (Fig. Supplemental 2). Similarly, TJ is linked to
265 TL by J-L while TJ is also linked to the large TAD TM (which contains *TER94* and the
266 neighboring gene, *pka-R2*) on the other side of *eve* by J-M. The same pattern of a hierarchical
267 series of LDC domains linking TAD neighbors, TAD next-door neighbors and next-next-door
268 neighbors is observed in the DNA segment that includes the *hebe* gene.
269 --*cohesin-mediated loop extrusion:* Like *eve*, the TADs in the region between *eve* and the attP
270 site in the *hebe* locus are defined by right angle triangles (volcanos) with a high density of
271 internal contacts. However, unlike *eve,* these volcano triangles do not have a plume, and instead
272 are surrounded by a series of LDC rectangular clouds. This is not the contact pattern that one
273 might expect for either a series of stem-loops connected by unanchored loops or a series of stem-

274 loops connected to each other. Moreover, there are no perpendicular or angled crosslinking
275 stripes (broken zippers) emanating from the base of these triangles, nor are their stripes along
276 their 45° legs. Thus, in spite of the fact that the NC14 nuclei used in this analysis are just
277 emerging from an earlier mitosis and a round of DNA synthesis and should be in the process of
278 assembling TADs, the crosslinking signatures that should be generated by the embrace of
279 passing cohesin complexes are completely absent.

280 In the loop extrusion model, the LDC domains would arise in a subset of nuclei because the
281 extruding cohesin complex breaks through one or more boundary roadblocks before its
282 progression is halted (Fig. Supplemental 3A versus B, C and D) (Hsieh et al., 2020; Krietenstein
283 et al., 2020). Breakthroughs could occur at only one roadblock on one side of the extruding loop
284 (Fig. Supplemental 3B and C), or at multiple roadblocks on one or both sides (Fig. Supplemental
285 3D), giving rise to a set of overlying LDC domains like that evident in the region between *hebe*
286 and *eve*. However, while there would have to be multiple breakthrough events to account for the
287 hierarchical array of LDC domains observed in these NC14 embryos, there is no evidence of the
288 45° crosslinking stripes that would be expected to mark the legs of the LDC domains following
289 the breakthrough. For example, the I-J LDC domain (Fig. Supplemental 2) could be generated
290 by cohesin complexes that initiated extrusion in either TI or TJ and then failed to halt at the
291 TI:TJ boundary (green arrowhead). In either case, there should be a 45° stripe of crosslinking
292 that marks (at least part of) the left (TI) and/or the right (TJ) leg of the I-J LDC domain and
293 extends to the apex of the TI-TJ LDC triangle; however, there is no evidence of such stripes.
294 Nor is there evidence of crosslinking stripes marking one or both legs of the J-K or I-K LDC
295 domains (Fig. Supplemental 2).

296 --*boundary:boundary pairing:* At least two different mechanisms are expected to account for the
297 TADs and LDC domains in the region between *eve* and the attP insertion site at -142 kb. First,

298 since none of the TADs in this DNA segment are topped by volcano plumes, they could
299 correspond to circle-loops. As circle-loops are generated by head-to-head pairing, the
300 boundaries in the region between *nhomie* and *hebe* would be expected to pair with each other
301 head-to-head. (The TAD to TAD contact maps generated by circle-loops and stem-loops are
302 considered further in the accompanying paper, Ke *et al.*)

303 As illustrated in Fig. Supplemental 3E-G, the coiled circle-loop TADs, though topologically
304 independent, are expected to be in relatively close proximity to each other. This means that in
305 addition to crosslinking events within the TADs, there will be crosslinking events between
306 neighboring TADs. These crosslinking events could generate the LDCs seen in Fig. 2C and Fig.
307 Supplemental 2. Since TADs next to each other (Fig. Supplemental 3E and F; Fig. Supplemental
308 2 I-J or J-K) would typically be expected to interact more frequently than TADs separated by
309 one, two or more TADs (Fig. Supplemental 3G; Fig. Supplemental 2 I-K or J-L), there should be
310 a progressive reduction in contact frequency at each step. This is generally what is seen (see Fig.
311 Supplemental 2). While connected stem-loops generated by cohesin extrusion should also bump
312 into one another, contacts between next-door neighbors might be expected to be less frequent
313 than contacts between next-next-door neighbors (Fig. 1B, C and Fig. Supplemental 3A). (TAD-
314 to-TAD crosslinking and the impact of loop topology is further examined in Ke et al., 2023.)

315 The second mechanism would be switching and/or combining pairing partners (Fig.
316 Supplemental 3H-J). Though imaging studies suggest that pairing interactions can be of
317 relatively long duration (>30 min: Chen *et al.*, 2018; Vazquez *et al.*, 2006)), they are not
318 permanent, and other nearby boundaries can compete for pairing interactions (Gohl *et al.*, 2011).
319 If switching/combining occurs in NC14 embryos (see Fig. Supplemental 3H-J), then the precise
320 pattern of TADs and LDC domains could also be impacted by the relative avidity of potential
321 pairing partners in the neighborhood and also the distances separating potential partners. For

322 example, the rectangle linking TH and TI (H-I in Fig. Supplemental 2) could represent partner
323 switching so that the boundary between TI and TJ (green arrowhead in Fig. Supplemental 2)
324 sometimes pairs with the boundary separating TH and TI (blue arrowhead) and sometimes with
325 the boundary separating TG and TH (red arrowhead in Fig. Supplemental 1). Distant boundary
326 elements could also preferentially interact with each other, generating a supra-TAD that contains
327 multiple TAD domains. For example, at the apex of the LDC domain B-I there is an interaction
328 dot (purple arrow, Fig. Supplemental 2) that links the TA-TB boundary to the TI-TJ boundary
329 (green arrowhead).

330 *Activation of a distant reporter by the eve enhancers*

331 In previous studies we found that a minimal 367 bp *homie* fragment can orchestrate
332 regulatory interactions between the *eve* enhancers and reporters inserted at an attP site in the first
333 intron of the *hebe* transcription unit, 142 kb upstream of the *eve* gene (Fujioka *et al.*, 2016). This
334 minimal fragment lacks the single *homie* CTCF site, but has sites for several other generic
335 boundary proteins including BEAF, Zipic, Su(Hw), Pita and Ibf2. In blastoderm stage embryos,
336 *lacZ* reporter expression was observed in a 7-stripe pattern that coincided with the stripes of the
337 endogenous gene. However, at this stage only a subset of *eve*-expressing nuclei expressed the
338 reporter. Later in development, during mid-embryogenesis, reporter expression was observed in
339 cells in the dorsal mesoderm, the ventral CNS, and in the anal plate region, in a pattern that
340 recapitulated *eve* gene expression (Fujioka *et al.*, 2016). At these later stages, most of the *eve*-
341 expressing cells also appear to express the reporter. The *eve*-dependent activation of the reporter
342 in a stripe pattern has also been visualized in live imaging of pre-cellular blastoderm-stage
343 embryos (Chen *et al.*, 2018). Within the 7 *eve* stripes, the majority of the nuclei (~80-85%) did
344 not express the *lacZ* reporter. In these nuclei, the transgene was found to be at a considerable
345 average physical distance from the *eve* locus (~700 nM). In the remaining nuclei, the transgene

346 was located in closer proximity to the *eve* locus (~350 nM, on average). Within this subset, the
347 transgene was expressed in most of the nuclei. Moreover, in those cases in which the transgene
348 was active, it remained in close proximity and continued to express *lacZ* for the duration of the
349 experiment (~30 min).

350 In the two models for TAD formation, quite different mechanisms must be invoked to
351 account for activation of the reporter at –142 kb by the *eve* enhancers. In the boundary-pairing
352 model, the transgene *homie* boundary at –142 kb loops over the intervening TADs and pairs with
353 the *nhomie:homie* complex flanking the *eve* TAD (c.f., Fig. 3B and C). In the loop-extrusion
354 model, a cohesin complex initiating loop extrusion in the *eve* TAD must break through the
355 *nhomie* roadblock at the upstream end of the *eve* TAD. It must then make its way past the
356 boundaries that separate *eve* from the *attP* site in the *hebe* gene, and come to a halt at the *homie*
357 boundary associated with the *lacZ* reporter. This would generate a novel TAD, *eveMammoth*
358 (*eveMa*), that extends from the *eve homie* all the way to the *homie* fragment at –142 kb, and
359 encompasses both the reporter and the *eve* gene, including its enhancers (c.f., Fig. 3D). Of
360 course, the *eveMa* TAD could also be generated by a cohesin complex that initiated in, for
361 example, TF. However, in this case, the runaway cohesin complex would have to break through
362 the intervening boundary roadblocks in both directions. In both the boundary-pairing and loop-
363 extrusion models, the configuration of the chromatin fiber would lead to the activation of *lacZ*
364 expression by the *eve* enhancers, while the reporter would be protected from the *hebe* enhancers
365 by the *homie* boundary.

366 To test these two models, we used a transgene that has two divergently transcribed
367 reporters, *lacZ* and *gfp*, that are each under the control of an *eve* promoter (Fig. 3A). Two
368 different versions of the transgene were generated. In the first, *LhomieG*, the *eve-lacZ* reporter is
369 located "upstream" of *homie* (with respect to the 5'→3'orientation of *homie* in the *eve* locus),

370 while *eve-gfp* is located downstream (Fig. 3A: top). In the second, the orientation of *homie*
371 within the transgene is reversed, so that *gfp* is upstream of *homie*, while *lacZ* is downstream
372 (*GhomieL*) (Fig. 3A: bottom). These two transgenes were then individually inserted into the *attP*
373 site at –142 kb, in either the Z5 or Z3 orientation. In the Z5 orientation, the *eve-lacZ* reporter is
374 on the *eve* side of the *homie* boundary, so that it is closer to the *eve* locus (Fig. 3A). The *eve-gfp*
375 reporter is on the opposite side of the *homie* boundary and is farther away from the *eve* locus.
376 This places the *eve-gfp* reporter next to a series of *hebe* enhancers located farther upstream, in
377 the intron of the *hebe* gene, and thus it should be subject to regulation by these enhancers (Fig.
378 2C). In the Z3 orientation, the entire transgene is inserted in the opposite orientation so that *eve-*
379 *gfp* is on the *eve* side of *homie*, while *eve-lacZ* is next to the *hebe* enhancers (Fig. 6A). As a
380 control, we also generated a Z5 transgene in which *lambda* DNA instead of *homie* was inserted
381 between the *eve-gfp* and *eve-lacZ* reporters.

382 *LhomieG Z5 and GhomieL Z5*

383 The results for transgenes *LhomieG Z5* and the control *LlambdaG Z5* are most
384 straightforward and will be considered first. In both the loop-extrusion and boundary-pairing
385 models, the *LlambdaG Z5* control is not expected to display any regulatory interactions with the
386 *eve* locus. Fig. 4C shows that this is the case: the *eve-lacZ* and *eve-gfp* reporters in *LlambdaG Z5*
387 embryos are silent at the blastoderm stage, and transcripts are only very rarely detected. Later in
388 development, both reporters are expressed in a repeating pattern along the dorsal midline (Fig.
389 6E). This expression is driven by the *hebe* gene enhancers located just beyond the *gfp*-reporter
390 (Fujioka *et al.*, 2009).

391 In the case of the *LhomieG Z5* insert, regulatory interactions between the transgene and the
392 *eve* locus will be observed if a stem-loop is established that links the transgene to *eve*. In the
393 loop-extrusion model, the breakthrough cohesin complexes would generate a stem-loop,

394 *eveMammoth* (*eveMa*), by coming to a halt at the *homie* boundary in the transgene and the *homie*
395 boundary in *eve*, as illustrated in Fig. 3D. This will place the *lacZ* reporter in close proximity to
396 the *eve* enhancers. It will also disrupt the *eve* TAD, as *nhomie* is no longer linked to the *homie*
397 boundary by cohesin. In the boundary-pairing model, the *lacZ* reporter in the transgene is
398 "upstream" of *homie* just like the *eve* gene and its enhancers are "upstream" of *homie* in the *eve*
399 locus. Since *homie* pairs with itself head-to-head, pairing between *homie* in the transgene and
400 *homie* in the *eve* locus will generate the same stem-loop as that generated by loop-extrusion.
401 However, because *homie* pairs with *nhomie* head-to-tail, a more complex structure like that in
402 Fig. 3B would be expected if the pairing interactions with the transgene do not disrupt the *eve*
403 TAD. Consistent with a tripartite structure of the sort shown here, previous studies have shown
404 that three *Mcp* transgenes on three different chromosomes can interact with each other
405 genetically (Muller *et al.*, 1999). Subsequent direct visualization of *Mcp*-mediated pairing
406 interactions in imaginal discs by Vazquez et al. (Vazquez *et al.*, 2006) showed that four *Mcp*-
407 containing transgene separated by Mbs (at Bridges cytogenetic intervals ~65 on the left arm of
408 chromosome 3, and ~83 and ~95 on the right arm of chromosome 3) and/or located on different
409 homologs, clustered in the same nuclear foci in 94% of the nuclei.

410 As predicted from the formation of a stem-loop linking the transgene to the *eve* locus, *lacZ*
411 is expressed in blastoderm-stage embryos in 7 stripes that coincide with the 7 stripes of the
412 endogenous *eve* gene. However, unlike *eve*, not all nuclei in the seven stripes express *lacZ*
413 (compare with the *eve* control in Fig. 4A). Later, during mid-embryogenesis, *lacZ* is expressed
414 in the mesoderm, the CNS, and the anal plate in a pattern that mimics the endogenous *eve* gene.
415 At this stage almost all *eve* positive cells are also positive for *lacZ*. Unlike the *LlambdaG Z5*
416 transgene, the *hebe* enhancers do not drive *eve-lacZ* expression, as they are beyond the *homie*
417 boundary (Fig. 4E). The *eve-gfp* reporter in the *LhomieG Z5* transgene responds differently. As

418 was observed for the *LlambdaG Z5* control, the *eve* enhancers drive little if any *gfp* expression,
419 either at the blastoderm stage or later in development (Fig. 5C, E). Instead, *gfp* is expressed in
420 the midline during mid-embryogenesis under the control of the *hebe* enhancers (Fig. 5E), which
421 should be in the same TAD as *eve-gfp*. These results would be consistent with both the loop-
422 extrusion (*eveMa*) and boundary-pairing models (Fig. 3B and D).

423 *GhomieL Z5*

424 The situation is more complicated for the *GhomieL Z5* transgene. This transgene is
425 inserted in the same chromosomal orientation as *LhomieG Z5*: *eve-lacZ* is on the *eve* side of the
426 *homie* boundary and *eve-gfp* is on the *hebe* enhancer side of the boundary. However, the
427 orientation of the boundary within the transgene is switched so that *eve-gfp* rather than *eve-lacZ*
428 is "upstream" of *homie*. In the boundary-pairing model, flipping the orientation of *homie* in the
429 transgene but keeping the same orientation of the transgene in the chromosome will have two
430 consequences. First, as illustrated in Fig. 3C, the topology of the chromatin loop connecting the
431 *homie* boundary in the transgene and the *eve* locus will change from a stem-loop to a circle-loop.
432 The reason for the change in loop topology is that *homie* pairs with *homie* head-to-head, and with
433 *nhomie* head-to-tail (Fujioka *et al.*, 2016). As before, the transgene *homie* will pair with *eve*
434 *homie* head-to-head to generate a simple circle-loop; however, if, as expected, it also pairs with
435 *nhomie* so that the *eve* TAD remains intact, a more complicated loop structure would be
436 generated (Fig. 3C). Second, the reporter that is preferentially activated by the *eve* enhancers
437 will switch from *eve-lacZ* to *eve-gfp*. In order to be activated by the *eve* enhancer, the reporter
438 must be "upstream" of the transgene *homie*. This physical relationship means that the *eve-gfp*
439 reporter will be activated by the *eve* enhancers independent of the orientation of the transgene in
440 the chromosome. On the other hand, the *eve-lacZ* reporter will not be activated by the *eve*
441 enhancers. In addition, it will still be insulated from the *hebe* enhancers by the *homie* boundary.

442 It is not possible to form a circle-loop in the loop extrusion model. Instead, the
443 breakthrough cohesin complex will come to a halt as before when it encounters the inverted
444 *homie* boundary at –142 kb (Fig. 3D). This means that the *eve-lacZ* reporter will be in the same
445 TAD (*eveMa*) as the *eve* enhancers and will be activated by the *eve* enhancers. In contrast, the
446 *eve-gfp* reporter will be in the neighboring TAD and will not be activated by the *eve* enhancers.
447 It will, however, be regulated by the nearby *hebe* enhancers.
448 Fig. 5 shows that the four predictions of the boundary-pairing model are correct: a) the *eve*
449 enhancers drive *eve-gfp* expression, b) the *hebe* enhancers drive *eve-gfp* expression, c) *eve-lacZ*
450 is not subject to regulation by the *eve* enhancers, and d) *eve-lacZ* is insulated from the *hebe*
451 enhancers. This would also mean that pairing of the transgene *homie* in the *GhomieL Z5* insert
452 with *homie* (and *nhomie*) in the *eve* locus generates a circle-loop, not a stem-loop, as was the
453 case for *LhomieG Z5*. In contrast, since loop-extrusion cannot generate circle-loops, the key
454 predictions of this model, namely that *eve-gfp* will be silent while *eve-lacZ* will be regulated by
455 the *eve* enhancers, are not satisfied. Instead, *eve-gfp* is activated by the *eve* enhancers, while *eve-*
456 *lacZ* is not.
457 While the results described above for the *GhomieL Z5* transgene are inconsistent with a
458 "simple" loop-extrusion model, there is one potential caveat: the 5'→3' orientation of the *homie*
459 boundary is inverted so that it is "pointing" towards the *eve* locus rather than away from it like
460 the endogenous *homie* boundary. In this orientation, it is possible that *homie* no longer functions
461 as an effective roadblock, and instead the cohesin complex emanating from the *eve* locus (or
462 from a loading site somewhere in between) slips past this *homie* fragment and continues until it
463 is blocked by a properly oriented boundary further upstream. In this case the newly formed
464 stem-loop, *eveGargantuan* (*eveGa*) (Fig. 3E), would include the *eve-gfp* reporter, and it would
465 then be activated by the *eve* enhancers. However, there are two problems with postulating this

466 novel *eveGa* TAD. First, *eve-lacZ* would be in the same *eveGa* TAD as *eve-gfp*, and it should
467 also be activated by the *eve* enhancers. Second, since the *homie* boundary was bypassed by the
468 cohesin complex to form the larger *eveGa* TAD, it should not be able to block the *hebe*
469 enhancers from activating *eve-lacZ*. Neither of these predictions is correct.

470 *GhomieL Z3* and *LhomieG Z3*

471 To confirm these conclusions, we examined the expression patterns of the *GhomieL* and
472 *LhomieG* transgenes when they are inserted in the opposite orientation (Z3) in the –142 kb attP
473 site. In the boundary-pairing model, the orientation of *homie* in the transgene determines which
474 reporter will be preferentially activated, while the orientation of the transgene *homie* in the
475 chromosome determines whether a stem-loop or a circle loop will be formed. In *GhomieL Z3* a
476 stem-loop will be formed, and *eve-gfp* will be activated by the *eve* enhancers; however, since
477 *eve-lacZ* is adjacent to the *hebe* enhancers, they will activate it rather than the *eve-gfp* reporter
478 (Fig. 6B). In *LhomieG Z3* the topology of the loop will switch to a circle-loop, and *eve-lacZ* will
479 be activated by both the *eve* and *hebe* enhancers (Fig. 6C). In the first version of the loop-
480 extrusion model, the *eve-gfp* reporter in both *GhomieL Z3* (Fig. 6D) and *LhomieG Z3* (not
481 shown), should be activated by the *eve* enhancers, since it is included in the *eveMa* TAD when
482 the transgenes are inserted in the Z3 orientation, while *eve-lacZ* will be regulated by the *hebe*
483 enhancers (Fig. 6D). In the second version of the loop-extrusion model, the functioning of the
484 *homie* roadblock depends upon its orientation relative to endogenous *homie*. In this model,
485 *GhomieL Z3* is expected to form the *eveMa* TAD (Fig. 6D), while *LhomieG Z3* will form the
486 larger *eveGa* TAD (Fig. 6E).

487 As would be predicted by all three models, the *eve* enhancers activate the *eve-gfp* reporter in
488 the *GhomieL Z3* insert (Fig. 6F and G). Likewise, *homie* blocks the *hebe* enhancers from
489 activating the *eve-gfp* reporter, while they turn on *eve-lacZ* expression instead. As predicted by

490 the boundary-pairing model, *eve*-dependent expression of the reporters switches from *gfp* to *lacZ*
491 in the *LhomieG Z3* insert (Fig. 6F and G). This would not be expected in the first version of the
492 loop-extrusion model, but it is predicted in the second version. However, as indicated in the
493 diagram in Fig. 6E, the two reporters, along with the *hebe* enhancers, are included in the larger
494 *eveGa* TAD, and for this reason both reporters should be activated equally by both the *eve* and
495 *hebe* enhancers. This is not the case. The *hebe* enhancers are blocked by the intervening *homie*
496 boundary, while there is little or no *eve*-dependent *gfp* expression.

497 *TAD formation by transgenes inserted at –142 kb*

498 We used MicroC to examine the TAD organization of the genomic region extending from
499 upstream of the attP site through the *eve* locus in 12-16 hr embryos carrying different transgene
500 insertions. At this stage, *eve* is expressed in only a small number of cells in the CNS, mesoderm
501 and anal plate, and this is also true for *hebe*. This means that most of the interactions detected by
502 MicroC are in cells in which the *eve* gene and the two reporters are inactive. In addition, while
503 the transgene co-localizes with *eve* in less than about a fifth of blastoderm stage nuclei, the
504 frequency of physical contact is expected to be much higher in 12-16 hr embryos (c.f., Vazquez
505 et al. Vazquez *et al.*, 2006)). This is because at this stage of development, cell cycles are much
506 longer, and homologs are fully paired. The key findings are summarized below.

507 *LlambdaG Z5*: While there is no evidence of contact between sequences in *eve* and
508 sequences around the attP site at –142 kb in wild-type embryos, this not true for embryos
509 carrying the dual reporter transgene. Perhaps the most unexpected finding is a weak, but clearly
510 discernable pattern of interaction linking the *LlambdaG Z5* transgene to the *eve* locus (Fig. 7A).
511 However, since the *LlambdaG Z5* transgene does not respond to the *eve* enhancers, this physical
512 interaction is not sufficient to drive detectable expression of the reporters. Though our
513 experiments do not allow us to unequivocally identify which sequences in the transgene are

514 responsible for this long-distance interaction, the most likely candidates are the two *eve*
515 promoters in the *LlambdaG Z5* transgene. As noted above, interaction triangles are observed
516 within the *eve* TAD spanning the regions between *nhomie* and *homie* (and/or the *eve* 3' PRE) and
517 the *eve* promoter. It seems possible that the factors supporting these intra-TAD interactions are
518 also responsible for connecting *LlambdaG Z5* to the *eve* locus. In this respect it is interesting to
519 note that *homie* and *nhomie* stimulate enhancer-promoter interactions in a transgene assay
520 (Yokoshi et al., 2020).

521 *LhomieG Z5*: Unlike *LlambdaG Z5*, the *eve-lacZ* reporter in the *LhomieG Z5* insert is
522 activated by the *eve* enhancers, and this regulatory interaction is perfectly paralleled by the
523 strong pattern of crosslinking between the transgene and the *eve* locus (Fig. 7B). On the *eve*
524 side, the interactions between the transgene and the *eve* locus generate a heavily populated band
525 of crosslinked sequences that span the 16 kb *eve* TAD. On the transgene side, there is an
526 unequal distribution of crosslinked sequences to either side of the transgene *homie* boundary. As
527 shown in the inset, the heaviest density of crosslinked sequences is on the *eve-lacZ* side of the
528 transgene, consistent with the activation of this reporter by *eve* enhancers. To confirm that the
529 interaction bias mimics the differences in activity of the two reporters in cells in which the
530 reporters are expressed, we analyzed the interaction profiles from viewpoints within either the
531 *eve-lacZ* or *eve-gfp* genes. Fig. 7D shows that *eve-lacZ* interacts with sequences extending
532 across the *eve* TAD, while *eve-gfp* interactions are largely restricted to the *eve homie* element.
533 While these results would be seemingly be consistent with both the boundary-pairing and loop-
534 extrusion models, there is no evidence of cohesin breaking through multiple intervening
535 boundaries into order to establish the novel *eve(Mam)* TAD. Instead, the organization of the
536 TADs and LDC domains in the interval in between the transgene and *eve* (Fig. 7B) resemble that
537 seen in wild-type NC14 embryos (see Fig. 2) (with the caveat that there are significantly fewer

538 reads in the transgene experiment). Equally important, the same is true for the *eve* TAD. In
539 order to generate the *eve*(*Mam*) TAD, the cohesin complex must break through the *nhomie*
540 boundary and, in so doing, disrupt the *eve* TAD. However, there are no obvious alterations in
541 either the *eve* volcano or plume with respect to their structure or relative contact density
542 compared to the *LlambdaG Z5* and wild type NC14 embryos. Likewise, there is no evidence of
543 novel vertical and 45° crosslinking stripes that would reflect the boundary breakthrough events
544 that are expected in the loop-extrusion model to accompany the formation of a physical
545 connection between the transgene and the *eve* locus.

546 We next used the transgene *homie* as the viewpoint to test the predictions of the
547 boundary-pairing and loop-extrusion models. Since *homie* can pair with itself (head-to-head)
548 and with *nhomie* (head-to-tail) in transvection assays (Fujioka *et al.*, 2016), a prediction of the
549 boundary-pairing model is that there will be physical interactions between the transgene *homie*
550 and <u>both</u> *nhomie* and *homie* in the *eve* locus (c.f., Vazquez et al., Vazquez *et al.*, 2006)). Also
551 based on what is known about how fly boundaries pair with each other (Chetverina *et al.*, 2014;
552 Chetverina *et al.*, 2017; Kyrchanova *et al.*, 2008a), self-interactions are expected to be stronger
553 than heterologous interactions. In contrast, in the loop-extrusion model, the cohesin complex
554 forms the *eveMa* TAD by linking *homie* in the transgene to *homie* in the *eve* locus, and this
555 requires the cohesin complex to break through *nhomie* and disrupt the connections between
556 *nhomie* and *homie* in the *eve* locus. In this case, *transgene:homie*←→*eve:homie* interactions
557 should be observed, while *transgene:homie*←→*-eve:nhomie* should not. Fig. 8A shows that, as
558 predicted by the boundary-pairing model, the transgene *homie* interacts with both *homie* and
559 *nhomie* in the *eve* locus, and the cross-linking frequency is greater for *homie:homie* interactions.

560 Since the *eve* TAD does not appear to be disturbed by the presence of the *LhomieG Z5*
561 transgene, these findings would suggest that the transgene *homie* likely pairs with both *nhomie*

562 and *homie* in a pattern somewhat like that illustrated in Fig. 3B. While the interactions with the
563 transgene are inconsistent with the loop-extrusion model, one could make the *ad hoc* assumption
564 that runaway cohesin complexes form not one but two different novel TADs. One would be
565 *eveMa*. In the other, *eveElephant* (*eveEl*), a cohesin complex that initiates in, for example, the
566 TE TAD would come to a halt on one side at the *homie* boundary in the transgene, and on the
567 other side at the *nhomie* boundary in the *eve* locus (Fig. Supplemental 4). However, this
568 configuration would not explain how the transgene reporter is activated by *eve* enhancers. Since
569 the *eve* enhancers in *eveEl* are located in a different TAD than the TADs containing *eve-lacZ* and
570 *eve-gfp*, they would not be able to activate either reporter.

571 *GhomieL Z5*: When *homie* is reversed in the *GhomieL Z5* transgene, the *eve-gfp* reporter
572 responds to the *eve* enhancers instead of *eve-lacZ*. In the boundary-pairing model, *eve-gfp* is
573 activated because the self- and heterologous-pairing interactions between *nhomie* and *homie* are
574 orientation-dependent. As a result, the topology of the loop formed between the transgene and
575 *eve* switches from a stem-loop to a circle-loop, and this brings the *eve-gfp* reporter in close
576 proximity to the *eve* enhancers instead of the *eve-lacZ* reporter. Since circle-loops cannot be
577 formed in the loop-extrusion model, activation of the distal *eve-gfp* reporter would require the
578 formation of a larger *eveGa* TAD in order to encompass the reporters in the transgene and the
579 *eve* enhancers in the same looped domain (Fig. 3E). However, as shown in Fig. 7C, we do not
580 observe the cross-linking pattern expected for a new *eveGa* TAD that includes the entire
581 transgene at –142 kb and then extends to an even more distal boundary. Nor are there any novel
582 vertical/45° crosslinking stripes or obvious perturbations in the basic TAD organization in the
583 region beyond –142 kb, or between –142 kb and the *eve* locus, as might be expected for cohesin
584 complexes breaking through multiple intervening roadblocks. Instead, as was the case for
585 *LhomieG Z5*, the *GhomieL Z5* transgene is physically linked to the *eve* locus, and there is a

586 prominent band of crosslinked transgene sequences that spans the *eve* TAD (Fig. 7C).
587 Consistent with the expression patterns of the two reporters, the heaviest density of crosslinked
588 sequences is on the *eve-gfp* side of the transgene (see inset). This is confirmed by the interaction
589 profiles generated using either the *eve-gfp* or *eve-lacZ* gene body as the viewpoint: the most
590 frequent contacts are between sequences within *eve* and the *eve-gfp* reporter (Fig. 7D). In
591 addition, as was the case for *LhomieG Z5*, when the transgene *homie* is used as the viewpoint,
592 interactions are observed between both *homie* and *nhomie* in the *eve* locus, with the most
593 frequent corresponding to self-pairing interactions (Fig. 8B). Other than postulating the *ad hoc*
594 *eveEl* TAD (Fig. Supplemental 4), this would not be expected for either version of the loop-
595 extrusion model.
596 *LhomieG Z3* and *GhomieL Z3*: To confirm the findings for transgenes inserted in the Z5
597 orientation, we used MicroC to analyze the TAD organization of *LhomieG Z3* and *GhomieL Z3*.
598 (Sup. Figs. 5 and 6). Like their Z5 counterparts, both show a strong band of interaction between
599 the transgene and sequences spanning the *eve* locus. In the case of *GhomieL Z3*, which is
600 predicted to form a stem-loop, bringing *eve-gfp* in contact with the *eve* enhancers, the *eve-gfp*
601 sequences interact more frequently with the *eve locus* (see inset in Fig. Supplemental 5A).
602 *homie* is in the opposite orientation in the transgene and in the chromosome in *LhomieG Z3*, and
603 in the boundary pairing model is expected to generate a circle-loop. While the sample
604 preparation for MicroC was of poorer quality than the others, the inset in Sup. Fig.6A shows that,
605 as expected, *eve-lacZ* interacts more frequently with sequences in the *eve* locus than *eve-gfp*. On
606 the other hand, there is no evidence of a larger *evaGa* TAD even though *eve-lacZ* is activated by
607 the *eve* enhancers. These interactions are confirmed when either *eve-lacZ* or *eve-GFP* are used
608 as viewpoints to plot the physical contacts between the *eve* TAD and the two reporters in
609 *LhomieG Z3* and *GhomieL Z3* (Sup. Figs. 5B and 6B). In addition, as was the case for the two

610 Z5 transgenes, when the transgene *homie* is used as the viewpoint, interactions are observed
611 between *homie* in the transgene and *both homie* and *nhomie* in the *eve* locus (Fig. 8A and B).
612 Again, the most frequent interactions are self-pairing interactions between *homie* in the transgene
613 and the *eve homie*.
614 *Loop topology and interactions with neighbors:* The results in the previous sections demonstrate
615 that the topology of the loop formed between the transgene and the *eve* locus depends upon the
616 orientation of *homie* in the transgene *and* the orientation of the transgene in the chromosome.
617 When *homie* in the transgene is pointing away from the *eve* locus, the topology is a stem-loop.
618 When *homie* in the transgene is pointing towards the *eve* locus, the topology is a circle-loop.
619 Changing the topology also impacts how sequences flanking the transgene and flanking *eve*
620 interact with each other. With the important caveat that the depth of our MicroC sequencing is
621 limited, the interaction profiles reflect the topology of the transgene-induced loop. For *homie*
622 pointing away from *eve* (*LhomieG Z5* and *GhomieL Z3*), sequences to the *TER94* side of *eve* and
623 to the *hebe* enhancer side of the transgene are more frequently crosslinked with each other (Fig.
624 Supplemental 7). These interactions would be expected for a stem-loop structure (Fig. 3B). For
625 *homie* pointing towards *eve*, sequences on the *TER94* side of *eve* interact with sequences on the
626 *eve* side of the transgene. This would be expected for a circle-loop structure (Fig. 3C).

627 **Discussion**

628 Two different mechanisms have been proposed to explain how TADs are formed and their
629 endpoints determined. In the first, a cohesin complex first induces a small loop at a loading site
630 within the TAD-to-be and then progressively extrudes one or both strands until the complex
631 encounters boundary roadblocks. If the loading site is located asymmetrically within the TAD-
632 to-be, extrusion will continue on the unblocked strand until a roadblock is encountered. The
633 endpoints of the loops were initially thought to be determined by the location of the first two
634 roadblocks encountered by the cohesin complex (Alipour and Marko, 2012). In a refinement of
635 this model, the orientation of a given roadblock relative to the direction of movement of the
636 cohesin complex determines whether it will come to a halt (Davidson and Peters, 2021;
637 Fudenberg *et al.*, 2016; Guo *et al.*, 2012; Guo *et al.*, 2015; Sanborn *et al.*, 2015). It has also been
638 suggested that cohesin complexes can break through a roadblock and come to a halt at more
639 distant roadblocks, so that in a population of cells analyzed by Hi-C or MicroC, the basic units of
640 organization, the TADs, are overlaid by a hierarchy of LDC domains (Hsieh *et al.*, 2020;
641 Krietenstein *et al.*, 2020). However, the only loop topologies that are possible in the loop-
642 extrusion model are a stem-loop and an unanchored loop.

643 In the other mechanism, TAD formation is governed by boundary:boundary pairing
644 interactions. In this case, the critical factors are the pairing and orientation preferences of
645 potential partners, together with proximity. Since boundary pairing interactions often exhibit an
646 orientation preference, the topology of the loop formed by paired boundaries can be either a
647 stem-loop or a circle-loop, depending on their relative orientation in the chromosome. In this
648 respect it differs from the loop-extrusion model in that switching the orientation of the boundary
649 does not preclude the formation of a TAD, whereas it can in some versions of the loop-extrusion
650 model.

651 In the studies reported here, we have used MicroC to analyze the TAD organization of a
652 ~150 kb chromosomal segment flanking the *eve* locus, and have experimentally tested the
653 predictions of these two models for the mechanisms of TAD formation. Our studies do not
654 support a loop-extrusion mechanism, and are inconsistent with this model on multiple levels.
655 *Manipulating "TAD" formation:* In our experimental paradigm, a transgene containing two
656 divergently transcribed reporters and the *eve* boundary *homie* is inserted 142 kb upstream of the
657 *eve* TAD. Depending upon the orientation of the *homie* boundary within the transgene, either
658 *eve-gfp* or *eve-lacZ* is activated by developmentally regulated enhancers in the *eve* TAD. Since
659 there are a series of well-defined TADs and their associated boundary elements separating the
660 transgene from *eve*, a version of the loop-extrusion model in which the cohesin complex
661 invariably arrests upon encountering *nhomie* and *homie* at the ends of the *eve* TAD cannot
662 account for this long-distance regulation. Instead, the extruding cohesin complex must be able to
663 break through the boundaries located between *homie* in the transgene and *homie* in the *eve* TAD.
664 This would generate a novel TAD, *eveMa*, with endpoints corresponding to the two *homie*
665 boundaries (Fig. 3D). In this case, the reporter on the *eve* side of the transgene *homie* would be
666 in the same domain as the *eve* enhancers, and thus potentially subject to regulation. While
667 *eveMa* could account for reporter expression in inserts in which the transgene *homie* is pointing
668 away from *eve* (and also away from the reporter on the *eve* side of the transgene insert), it does
669 not explain reporter expression when the transgene *homie* is pointing towards *eve* (and in this
670 case towards the reporter on the *eve* side of the transgene insert). In the latter case, the reporter
671 activated by the *eve* enhancers lies outside of the *eveMa* (cf., Fig. 5D), and thus is not included in
672 the same TAD as the *eve* enhancers. In order to account for activation of the distal reporter, one
673 must imagine that the cohesin complex breaks through the *homie* boundary in the transgene and
674 continues on until extrusion is halted by an even more distal boundary. In this revision of the

675 loop-extrusion model, both reporters would be included in the same TAD, *eveGa* (c.f., Fig. 5E),
676 as the *eve* enhancers, and both reporters should be activated. However, this is not observed.
677 *TAD organization: loop-extrusion versus boundary pairing:* Analysis of the TAD organization
678 in the region spanning the transgene insertion site and *eve* also does not fit with the expectations
679 of the loop-extrusion model. To account for reporter activation, the cohesin complex must be
680 able to break through the *nhomie* boundary. This breakthrough could be engineered by a cohesin
681 complex that initiated loop extrusion either from within the *eve* TAD or from one of the TADs
682 located between the *nhomie* boundary and the transgene at –142 kb. However, in the absence of
683 the transgene, there is no indication that *nhomie* is particularly prone to break-through events. In
684 fact, the *eve* TAD differs from most of the TADs in between *eve* and the transgene in that it is
685 incorporated into larger LDC interaction triangles much less frequently. There is also no reason
686 to imagine that the presence of a *homie* (or *nhomie*: see Fujioka *et al.*, 2016)) transgene at –142
687 kb would somehow induce *nhomie* break-through events.
688 As expected from the regulatory interactions, *homie*-containing transgenes at –142 kb
689 generate a strong band of crosslinking events that physically link the transgene to the *eve* locus.
690 In the loop-extrusion model, this strong band of interaction should generate a novel interaction
691 triangle (LDC or HDIC) with endpoints corresponding to the *homie* elements in the transgene
692 and *eve*. Contrary to this expectation, a triangle of increased interaction frequencies spanning the
693 entire region between –142 kb and *homie* is not observed for any of the transgene inserts. In
694 viewpoints from the *homie* element in the *LhomieG Z5* and *GhomieL Z3* transgenes, there are
695 two peaks in the *eve* TAD. The more prominent peak maps to *homie*, while the less prominent
696 peak maps to *nhomie*. To account for these interaction peaks, the loop-extrusion model would
697 have to assume that the presence of the *LhomieG Z5* and *GhomieL Z3* transgenes induces the
698 formation of two novel TADs, the *eveMa* (*homie:homie*) TAD in Fig. 5 and another TAD,

699 *eveElephant*, that links the transgene *homie* to the *nhomie* element in the *eve* locus (Fig.
700 Supplemental 4). However, while the *eveMa* TAD could potentially explain the activation of
701 reporters on the proximal side of the transgene *homie* by the *eve* enhancers, these TADs do not
702 explain how the *eve* enhancers are able to activate reporters on the distal (*hebe* enhancer) side of
703 the *GhomieL Z5* and *LhomieG Z3* transgene. This would require the formation of yet another
704 TAD, *eveGa*, that encompasses *eve*, the transgene, and one or more TADs beyond the *hebe* gene
705 (Fig. 3E). However, there is no indication of physical interactions between the *homie* boundary
706 in *eve* and a boundary located beyond the *hebe* gene. Moreover, like *eveMa*, *eveGa* would also
707 have to be accompanied by the formation of *eveEl*, since the *homie* element in *GhomieL Z5* and
708 *LhomieG Z3* interacts with both *nhomie* and *homie*.

709 While these results are inconsistent with the expectations of the loop-extrusion model,
710 they dovetail nicely with many of the predictions of the boundary-pairing model. First, in all
711 four transgenes, *homie* interacts with both *homie* and *nhomie* in the *eve* locus, but does not
712 interact with the multiple boundary elements located in between. This is consistent with other
713 studies which show that boundaries often have strong partner preferences (Gohl *et al.*, 2011;
714 Kyrchanova *et al.*, 2011; Kyrchanova *et al.*, 2008b). Second, interaction frequencies are greater
715 for self-pairing than for heterologous pairing. This fits with what is known about how partner
716 preferences are determined (Blanton et al., 2003; Chetverina *et al.*, 2017; Erokhin et al., 2021;
717 Kyrchanova *et al.*, 2008a; Zolotarev *et al.*, 2016). Third, the pattern of reporter activation in the
718 four transgenes is precisely as expected from previous studies on the orientation dependence of
719 *homie* pairing with itself (head-to-head) and with *nhomie* (head-to-tail) (Chen *et al.*, 2018;
720 Fujioka *et al.*, 2016). This orientation dependence means that the reporter located "upstream" of
721 *homie* is the one that is activated, independent of the orientation of the transgene in the
722 chromosome. Fourth, the pattern of physical interactions between the *eve* locus and the

723 transgenes recapitulates the pattern of reporter activation. That is, when *eve-lacZ* sequences are
724 preferentially crosslinked to sequences in the *eve* locus, then the *eve-lacZ* reporter is activated by
725 the *eve* enhancers. Likewise, *eve-gfp* is preferentially crosslinked to sequences in *eve* when it is
726 activated by the *eve* enhancers. Again this is dependent on the orientation of *homie* in the
727 transgene, but not on the orientation of the transgene in the chromosome. Fifth, unlike the loop-
728 extrusion model where only stem-loops or disconnected loops can be generated, both stem-loops
729 and circle-loops can be, and, as we have shown here, actually are generated by
730 boundary:boundary pairing. The topology of the loop formed between the transgene and *eve*
731 depends on the orientation of the transgene *homie* in the chromosome. When *homie* in the
732 transgene insert is pointing away from the *eve* locus, a stem-loop is generated. When *homie* in
733 the transgene insert is pointing towards the *eve* locus, a circle-loop is generated. That both loop
734 topologies can be generated by boundary:boundary pairing is confirmed by not only by the
735 pattern of expression of the two reporters in the four transgene inserts but also by their pattern of
736 physical interactions evident in our MicroC experiments.
737 *Are the pairing interactions of homie and nhomie unusual?* There are two reasons to think that
738 the properties of most fly boundaries are similar to those of *homie* and *nhomie*. First, all of the
739 fly boundaries that have been tested in assays that are expected to require pairing do have this
740 activity (Blanton *et al.*, 2003; Cai and Shen, 2001; Erokhin et al., 2011; Gohl *et al.*, 2011;
741 Kyrchanova *et al.*, 2008a; Kyrchanova *et al.*, 2011; Kyrchanova et al., 2007; Kyrchanova *et al.*,
742 2008b; Muravyova *et al.*, 2001). Second, multimerized binding sites for generic polydactyl zinc
743 finger proteins like Su(Hw), Pita, dCTCF, Zw5, and Zipic that are found associated with many
744 fly boundaries are not only able to block enhancer/silencer action but can also mediate pairing
745 interactions in insulator bypass and transvection assays (Erokhin *et al.*, 2021; Kyrchanova *et al.*,
746 2008a; Zolotarev *et al.*, 2016). Moreover, as is the case for *homie* and *nhomie*, pairing

747 interactions require the appropriate partners: multimerized sites for Zw5 do not pair with
748 multimerized sites for dCTCF (Erokhin *et al.*, 2021; Kyrchanova *et al.*, 2008a; Zolotarev *et al.*,
749 2016). This makes good sense, since these proteins form dimers/multimers, as do other fly
750 boundary factors like BEAF, Mod(mdg4), the BEN-domain protein Insensitive and the Elba
751 complex (Avva and Hart, 2016; Bonchuk et al., 2021; Bonchuk et al., 2011; Bonchuk et al.,
752 2020; Fedotova et al., 2018; Fedotova et al., 2019; Fedotova et al., 2017; Hart et al., 1997). The
753 ability to multimerize means that boundaries that share binding sites for the same protein, i.e., for
754 Pita, Zipic or CTCF, can be physically linked to each other by a single protein complex.
755 *Boundary pairing orientation and loop topology:* Our analysis of the TAD organization in the
756 *eve* neighborhood suggests that the predominant loop topology may be a circle-loop, not a stem-
757 loop. Only the *eve* TAD (and perhaps the very small TAD, TA) is clearly a stem-loop. It is
758 topped by a distinctive volcano plume and is not overlaid by a hierarchy of LDC domains,
759 indicating that it must be "insulated" from crosslinkable interactions with other nearby TADs.
760 By contrast, none of the TADs in the region spanning *eve* and the *hebe* gene have volcano
761 plumes at their apex. Instead, all of these TADs are overlaid by a hierarchal organization of
762 LDC domains (clouds) that diminish in crosslinking intensity as the number of in-between TADs
763 increases. As discussed above, the contact pattern of the TADs to the left of *eve* are likely to
764 arise from partner switching and/or head-to-head boundary pairing interactions that generate
765 circle-loops. Pairing orientation versus loop topology is considered further in the accompanying
766 paper (Ke *et al.*).
767 *Orientation and long-distance regulatory interactions:* Our experiments highlight an important
768 constraint on long-distance regulatory interactions that are mediated not only by boundary
769 elements like *homie* and *nhomie* but also by the GAF-dependent tethering and bypass elements
770 (Batut et al., 2022; Kyrchanova et al., 2023; Kyrchanova et al., 2018; Kyrchanova et al., 2019a;

771 Kyrchanova et al., 2019b; Levo et al., 2022; Li et al., 2023; Mohana et al., 2023), namely the
772 orientation of the physical interactions that bring distant enhancers (or silencers) into close
773 proximity with their gene targets. As we have shown here, the orientation of the physical
774 interactions that bring distant sequences close together determines whether the enhancer will be
775 able to activate its target gene. In principle the effects of orientation could be mitigated by
776 increasing the distance between the boundary/tethering elements and either the enhancers
777 (silencers) or the promoter. However, distance can be a critical if not a rate limiting step in
778 transcriptional activation. For example, live imaging experiments have shown that the frequency
779 and amplitude of transcriptional bursts depend on the distance between the enhancer and target
780 promoter. Nearly continuous bursting and high amplitude transcription was observed when the
781 enhancer was located close the promoter (Fukaya et al., 2016). However, at distances as little as
782 6 kb, transcription became discontinuous with discrete bursts, while at 9 kb, the bursts were
783 much less frequent, and the amplitude/output was substantially reduced (Yokoshi *et al.*, 2020).
784 If similar distance-dependent reductions in transcriptional efficiency are common, then it would
785 not be possible to overcome the topological constraints on regulatory interactions by simply
786 increasing the distance between the boundary/tethering elements and the enhancers/promoters.
787 While our studies do not directly address how chromosome architecture impacts regulatory
788 interactions in vertebrates, it seems clear that long-distance regulatory interactions in vertebrates
789 (c.f., Tan et al., 2023) will be subject to the same types of topological constraints as those that
790 apply in flies.
791
792 **Acknowledgements:** We would like to thank the Gordon Grey for running the fly food
793 facility at Princeton, members of the Lewis Sigler Genomics Core facility for their invaluable
794 assistance with DNA sequencing and Sergey Ryabichko for his help with smFISH probes. We

would also like to thank members of MOL431 for creative input. Special thanks to Olga Kyrchanova, Daria Chetverina, Maksim Erokhin, Pavel Georigev, Tsutomu Aoki, Girish Deshpande, Airat Ibragimov, Sergey Ryabichko, Yuri Pritykin and Alex Ostrin for stimulating discussions and sharing unpublished data.

802 # Materials and methods

803

804 **Key resources table**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
n-Heptane	Fisher Chemical	O3008-4
Paraformaldehyde 20% solution, EM Grade	Electron Microscopy Sciences	15713-S
Formaldehyde, 16%, methanol free, Ultra Pure	Polysciences Inc	18814-10
PBS - Phosphate-Buffered Saline (10X) pH 7.4, RNase-free	Thermal Fisher	AM9624
Tween 20	Sigma	P1379
Triton X-100	Bio-Rad	161-0407
Tris base	Sigma	11814273001
Methanol	Fisher Chemical	203403
SSC, 20X	Thermal Fisher	15557044
Formamide	Thermal Fisher	17899
Dextran sulfate	Sigma	D8906
Salmon Sperm DNA	Thermal Fisher	AM9680
Ribonucleoside Vanadyl Complex	NEB	S1402S
Nuclease-free BSA	Sigma	126609
Triethylammonium Acetate	Sigma	625718
dGTP (100 MM)	VWR	76510-208
dTTP (100 MM)	VWR	76510-224
Lonza NuSieve 3:1 Agarose	Thermal Fisher	BMA50090
T4 DNA ligase	NEB	M0202L
Biotin-11-dCTP	Jen Bioscience	NU-809-BIOX

Biotin-14-dATP	Jen Bioscience	NU-835-BIO14
Qubit dsDNA HS Assay Kit	Life Technologies Corporation	Q32851
Atto 633 NHS ester	Sigma	01464
Phase Lock Gel™, QuantaBio - 2302830, Phase Lock Gel Heavy	VMR	10847-802
NEBNext Ultra II DNA Library Prep Kit for Illumina	NEB	E7645S
Ampure Xp 5ml Kit	Thermal Fisher	NC9959336
Hifi Hotstart Ready Mix	Thermal Fisher	501965217
Dynabeads MyOne Streptavidin C1	Life Technologies Corporation	65001
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Sigma	11873580001
N,N-Dimethylformamide	Sigma	227056
Potassium acetate solution	Sigma	95843
DSG (disuccinimidyl glutarate)	Thermal Fisher	PI20593
T4 Polynucleotide Kinase - 500 units	NEB	M0201S
DNA Polymerase I, Large (Klenow) Fragment - 1000 units	NEB	M0210L
End-it DNA End Repair Kit	Thermal Fisher	NC0105678
Proteinase K recomb. 100 mg	Sigma	3115879001
Nuclease Micrococcal (s7)	Thermal Fisher	NC9391488
EGS (ethylene glycol bis(succinimidyl succinate))	Thermal Fisher	PI21565
Atto 565 NHS ester	Sigma	72464
Oligonucleotides		
WK_eve_probeset Sequence see “Supplemental Data S1”	Biosearch Technologies	

WK_LacZ_probeset Sequence see “Supplemental Data S1”	Biosearch Technologies	
WK_GFP_probeset Sequence see “Supplemental Data S1”	Biosearch Technologies	
NEBNext Multiplex Oligos for Illumina	NEB	E7335S
Software and algorithms		
Fiji (ImageJ)	Schindelin et al., 2012	fiji.sc
NIS element	Nikon	microscope.healthcare.nikon.com/products/software/nis-elements
Graphpad Prism 8	GraphPad Software	graphpad.com
HiGlass	Kerpedjiev et al., 2018	higlass.io
bwa	Li H. and Durbin R., 2009	bio-bwa.sourceforge.net/
samtools	Github/open source	samtools.github.io
pairsamtools	Github/open source	github.com/mirnylab/pairsamtools
pairix	Github/open source	github.com/4dn-dcic/pairix
cooler	Abdennur, N., and Mirny, L., 2019	github.com/mirnylab/cooler
Miniconda	Anaconda	docs.conda.io/en/latest/miniconda.html
Snakemake	Github/open source	snakemake.github.io

805

806 **Plasmid construction and transgenic lines**

807 The dual reporters construct was described previously (Fujioka *et al.*, 2016). In short,
808 both reporters contain the *eve* basal promoter (-275 to +106 bp relative to *eve* start site), the *lacZ*
809 (*eve-lacZ*) or *EGFP* (*eve-GFP*) coding region, and the *eve* 3' UTR (+1300 to +1525 bp). These
810 two reporters are divergently transcribed. The 367 bp *homie* fragment, or a 500 bp fragment
811 from *lambda* phage DNA, was placed between the two reporters. The *homie* fragment was
812 placed in both orientations: in *LhomieG*, the *eve-lacZ* reporter is located upstream of the *homie*
813 fragment, while in *GhomieL*, *eve-GFP* is located upstream of the *homie* fragment (see Fig 2A).

814 The -142 kb attP landing site was described previously (Fujioka *et al.*, 2009). The -142
815 kb site contains two attP target sites for phiC31 recombinase-mediated cassette exchange
816 (RMCE) (Bateman et al., 2006) and *mini-white* as a marker. RMCE can result in either
817 orientation of an inserting transgene (Z5 and Z3; see text). RMCE events were identified by loss
818 of *mini-white*, and the orientation of each insert was determined by PCR.

819 **smFISH probe preparation**

820 The sequences of target genes were obtained from Flybase (flybase.org)(Gramates et al.,
821 2022). To design probes, the target gene sequences were submitted to the Biosearch
822 Technologies Stellaris RNA FISH probe designer tool (biosearchtech.com), with parameters set
823 to 20nt probes, 48 probes per gene, and mask level 5 for *Drosophila melanogaster*. The sequence
824 of the probes used are listed in Supplemental Data S1. The probes with 3' MdC(TEG-Amino)
825 modifications were then ordered from Biosearch Technologies in 96-well plate format, cartridge
826 purification and salt free conditions and in 5 nmol scale. To prepare the fluorophore for probe
827 labeling, the Sigma Atto NHS ester fluorophore (Atto 565 NHS or Atto 633 NHS) were
828 dissolved in dimethylformamide at 5mg/mL as the stock solution. For probe labeling, 1 nmol of
829 each probe was combined and mixed with Atto NHS ester fluorophore at 1mg/mL in 500µL of

830 0.1M sodium bicarbonate (pH 8.0-8.5) for 2 hours, and the probes were then precipitated with
831 0.5M potassium acetate (pH 5.2) in 75% EtOH. After that, the probes were dried and resuspend
832 in 50µL of DePC-treated H_2O. From there, the probes were injected into HPLC in which the
833 percentage 0.1M triethylammonium acetate varied from 7% to 30% with a flow rate of 1mL/min.
834 The detector to monitor coupled probes during the HPLC purification were set to 260nm for
835 DNA and the absorption wavelength of the coupled fluorophore. The coupled probes were
836 collected when both 260nm and the absorption wavelength channel were detected. The coupled
837 probes were dried and diluted into appropriate concentration in DePC-treated H_2O for smFISH
838 experiments.

839 **smFISH**

840 smFISH methods were adopted and modified from previous studies (Little and Gregor,
841 2018; Trcek et al., 2017). 100-200 flies were placed in a cage with an apple juice plate at the
842 bottom of the cage. Early stages of embryos were for 7hours, while for later stages embryos,
843 collection was overnight. Embryos from each plates were washed into collection mesh and
844 dechorionated in bleach for 2min, then fixed in 5mL of 4% paraformaldehyde in 1X PBS and
845 5mL of heptane for 15min with horizontal shaking. The paraformaldehyde were then removed
846 and replaced with 5mL methanol. The embryos were then devitellinized by vortexing for 30s,
847 and washed in 1mL of methanol twice. Methanol were then removed and replaced by PTw(1X
848 PBS with 0.1% Tween-20) through serial dilution as 7:3, 1:1, and 3:7 methanol:PTw. The
849 embryos were washed twice in 1mL of PTw and then twice in 1mL smFISH wash buffer (4X
850 SSC, 35% formamide and 0.1% Tween-20). Embryos were incubated with ~5nM coupled
851 smFISH probes in hybridization buffer (0.1g/mL dextran sulfate, 0.1mg/mL salmon sperm
852 ssDNA, 2mM ribonucleoside vanadyl complex, 20µg/mL RNAse-free BSA, 4X SSC, 1%
853 Tween-20 and 35% formamide) for 12-16h. Embryos were then washed twice for 2 hours in

854 1mL smFISH wash buffer, followed by 4X washing in 1mL PTw. For DAPI/Hoechst staining,
855 the embryos were stained with 1ug/mL DAPI or Hoechst in PTw, and washed 3X with 1mL
856 PTw. Finally, the embryos were mounted on microscope slides with Aqua PolyMount and #1.5
857 coverslip for imaging.

858 **Imaging, image analysis and statistics**

859 Embryos from smFISH were imaged by using a Nikon A1 confocal microscope system,
860 Plan Apo 20X/0.75 DIC objective. Z-stack images were taken with interval of 2µm, 4X average,
861 1024X1024 resolution, and the appropriate laser power and gain were set for 405, 561, and 640
862 channels to avoid overexposure. Images were processed by ImageJ and the maximum projection
863 was applied to each of the stack images. To measure the stripe intensity of early embryos, multi-
864 channel images were first split into single channels and the stripe signal was highlighted and
865 selected by the MaxEntropy thresholding method. For APR intensity measurement, the ROI tool
866 was used to crop out the APR region of late-stage embryos. The cells with APR signal were also
867 highlighted and selected by MaxEntropy thresholding. The particle measurement tool was used
868 to measure the average intensity of all cells that have the signal. At the same time, the
869 background signal (average intensity) was measured on cells without a signal in the same
870 embryo. The relative intensity (signal to background) for each embryo was calculated by using
871 the stripe signal and the background from the same embryo. To make comparisons between
872 independent biological replicates, the average background signal of all embryos from each
873 replicate was calculated. The relative intensity of each embryo from each replicate were
874 normalized based on the average background signal of all embryos from same replicates.
875 GraphPad Prism was used for data visualization and statistical analysis. To compare intensity
876 from embryos in different groups, different signals from the same embryo (e.g., *lacZ* and *GFP*)

877 were paired and paired two-tailed t-tests were used to calculate p-values. All raw measurements
878 and normalized data are included in Supplemental Data S2.

879 **MicroC library construction**

880 Embryos were collected on yeasted apple juice plates in population cages for 4 hours,
881 incubated for 12 hours at 25°C, then subjected to fixation as follows. Embryos were
882 dechorionated for 2 mins in 3% sodium hypochlorite, rinsed with deionized water, and
883 transferred to glass vials containing 5 mL PBST (0.1% Triton-X100 in PBS), 7.5 mL n-heptane,
884 and 1.5mL fresh 16% formaldehyde. Crosslinking was carried out at room temperature for
885 exactly 15 mins on an orbital shaker at 250rpm, followed by addition of 3.7 mL 2M Tris-HCl
886 pH7.5 and shaking for 5 mins to quench the reaction. Embryos were washed twice with 15 mL
887 PBST and subjected to secondary crosslinking. Secondary crosslinking was done in 10mL of
888 freshly prepared 3mM final DSG and ESG in PBST for 45 mins at room temperature with
889 passive mixing. The reaction was quenched by addition of 3.7mL 2M Tris-HCl pH7.5 for 5
890 mins, washed twice with PBST, snap-frozen, and stored at -80°C until library construction.
891 Micro-C libraries were prepared as previously described (Batut *et al.*, 2022) with the following
892 modifications: 50uL of 12-16hr embryos were used for each biological replicate. 60U of MNase
893 was used for each reaction to digest chromatin to a mononucleosome:dinucleosome ratio of 4.
894 Libraries were barcoded, pooled and subjected to paired-end sequencing on an Illumina Novaseq
895 S1 100 nt Flowcell (read length 50 bases per mate, 6-base index read).

896 **Micro-C data processing**

897 MicroC data for *D. melanogaster* were aligned to custom genomes edited from the
898 Berkeley Drosophila Genome Project (BDGP) Release 6 reference assembly (dos Santos et al.,
899 2015) with BWA-MEM (Li and Durbin, 2009) using parameters **-S -P -5 -M**. Briefly, the
900 custom genomes are simply insertions of the transgenic sequence into the –142kb integration

901 site, as predicted from perfect integration. These events were confirmed using PCR post-
902 integration. The resultant BAM files were parsed, sorted, de-duplicated, filtered, and split with
903 Pairtools (https://github.com/mirnylab/pairtools). We removed pairs where only half of the pair
904 could be mapped, or where the MAPQ score was less than three. The resultant files were
905 indexed with Pairix (https://github.com/4dn-dcic/pairix). The files from replicates were merged
906 with Pairtools before generating 100bp contact matrices using Cooler (Abdennur and Mirny,
907 2020). Finally, balancing and Mcool file generation was performed with Cooler's Zoomify tool.
908 Virtual 4C profiles were extracted from individual replicates using FAN-C (Kruse et al.,
909 2020) at 400bp resolution. The values were summed across replicates and smoothed across three
910 bins (1.2kb). Viewpoints were determined based on the most informative region for
911 interpretation. Ultimately we decided to use 800bp regions in the gene body of either *GFP* or
912 *lacZ*, moving downstream from the *eve* promoter. However, similar results were obtained using
913 viewpoints between transgenic *homie* and the promoters of either gene. For *homie* viewpoints,
914 due to transgenic *homie* being masked, some loss of signal is expected at the viewpoint (Figure
915 6).

916

Figure Legends

Figure 1. Diagram of the possible loop topologies generated by loop-extrusion and boundary:boundary pairing. A) Cohesin embraces a loop at a loading site somewhere within the TAD-to-be and then extrudes a loop at an equal rate on both strands. Extrusion continues until cohesin encounters boundary roadblocks on both strands. B) In one model, the orientation of the roadblock is important. As a consequence, the chromatin fiber will be organized into a series of stem-loops separated from each other by unanchored loops. C) If the presence of a boundary roadblock, but not its orientation, is sufficient to halt extrusion, the chromatin fiber will be organized into a series of linked stem-loops. The main axis of the chromosome will be defined by the cohesin:CTCF roadblocks. D) The pairing of two boundaries head-to-tail generates a stem-loop. E) If boundary pairing interactions are strictly pairwise, head-to-tail pairing will generate a series of stem-loops separated from each other by unanchored loops. F) If boundaries can engage in multiple head-to-tail pairing interaction, the chromosome will be organized into a series of linked stem-loops. The main axis of the chromosome will be defined by a series paired boundaries. G) The pairing of two boundaries head-to-head generates a circle-loop. H) If boundary pairing interactions are strictly pairwise there will be a series of circle-loops separated from each other by unanchored loops. I) If boundaries can engage in multiple head-to-head pairing interactions, the chromatin fiber will be organized into a series of circle-loos connected to each other at their base. Pairing interactions between boundaries #1 and #2 need not be in register with pairing of boundaries #2 and #3. In this case, the main axis of the chromosome may bend and twist and this could impact the relative orientation of the circle-loops.

Figure 2. TAD organization of the *even-skipped* locus and neighboring sequences. A) The *eve* skipped TAD is a volcano with a plume that is anchored by *nhomie* and *homie*. ChIP-seq data below the MicroC map indicate many of the known fly chromosomal architectural proteins proteins are associated with the two *eve* boundary elements *in vivo* (Cuartero et al., 2014; Duan et al., 2021; Gaskill et al., 2021; Li et al., 2015; Sun et al., 2015; Ueberschär et al., 2019; Zolotarev *et al.*, 2016). B) The *eve* locus forms a stem-loop structure. In this illustration *nhomie* pairs with *homie* head-to-tail, and this configuration forms a stem-loop which brings sequences upstream *nhomie* and downstream of *homie* in contact as is observed in insulator bypass assays (Kyrchanova *et al.*, 2008a). The *eve* locus is shown assembled into a coiled "30 nM" chromatin fiber. C) MicroC contact pattern for the chromosomal region spanning the attP site at –142 kb and the *eve* locus. Like the *eve* volcano, this contact pattern was generated using aggregated previously published NC14 data (Batut *et al.*, 2022; Levo *et al.*, 2022). A black arrow indicates the –142kb locus where transgenes are integrated into the genome. A blue arrow indicates position of the *Etf-QO* gene. Note the numerous TADs separating the –142kb *hebe* locus from the *eve* "volcano TAD". Individual TADs are labeled TA—TM. Black boxes indicate positions of sequences that, based on ChIP experiments, are bound by one or more known insulator proteins *in vivo*.

Figure 3. Schematics of boundary pairing and loop extrusion. A) *LhomieG Z5* and *GhomieL Z5* transgenes and the *eve* locus on a linear map. The same color codes are used throughout. In the *eve* locus, *nhomie* (blue arrow) and *homie* (red arrow) are oriented, by convention, so that they are pointing towards the right. This convention is maintained in the two transgenes. In *LhomieG Z5 homie* (top) is in the opposite orientation from *homie* in the *eve* locus is pointing away from the *eve* locus. In *GhomieL Z5 homie* (bottom) *homie* is in the same orientation as the

964 *homie* in the *eve* locus and is pointing towards the *eve* locus. B) Predicted boundary pairing
965 interactions between *LhomieG Z5* and the *eve* locus. *homie* in the transgene pairs with *homie* in
966 the *eve* locus head-to-head. Since *homie* in the transgene is pointing in the opposite orientation
967 from *homie* in the *eve* locus, a stem-loop will be generated. If *homie* in the transgene also pairs
968 with *nhomie* in the *eve* locus head-to-tail, a loop structure like that shown in B will be generated.
969 In this topology, *eve-lacZ* is in close proximity to *eve* enhancers and *eve-gfp* is in contact with
970 the *hebe* enhancer. C) Predicted boundary pairing between *GhomieL Z5* and the *eve* locus. The
971 *GhomieL Z5* transgene is inserted in the same chromosomal orientation as *LhomieG Z5*;
972 however, the *homie* boundary is inverted so that it is pointing towards the *eve* in the same
973 orientation as the *homie* in the *eve* locus. *Homie* in the transgene will pair with *homie* in the *eve*
974 locus head-to-head, and this generates a circle-loop. If *homie* in the transgene also pairs with
975 *nhomie* in the *eve* loucs a loop structure like that shown in C will be generated. In this topology
976 *eve-gfp* will be activated by both the *eve* and *hebe* enhancers. D) Loop extrusion model for
977 *LhomieG Z5*. Transgene *homie* and endogenous *homie* determine the endpoints of the extruded
978 *eveMammoth* (*eveMa*) loop. In this topology, *eve-lacZ* is in close proximity to the *eve*
979 enhancers, while *eve-gfp* is close to the *hebe* enhancers. E) Loop extrusion model for *GhomieL*
980 *Z5*. The cohesin complex bypasses transgenic *homie* and is stopped at an upstream boundary X
981 to generate a novel *eveGarantuan* (*eveGa*) loop. Both *eve-gfp* and *eve-lacZ* are located within
982 the same *eveGa* TAD, and thus should interact with *eve*. Since the *hebe* enhancers are within
983 this TAD as well, they would also be able to activate both reporters.
984
985 **Figure 4. The *lacZ* reporter is activated by *eve* enhancers in the *LhomieG Z5* transgene.** A)
986 Expression patterns of endogenous *eve* in early (stage 4-5) and late (stage 14-16) embryos by
987 smFISH. *eve* (Atto 633) probes were used to hybridize with *eve* mRNA. *eve* is shown in yellow
988 and DAPI in blue. N>3. B) Schematics of transgenes. C) Expression patterns of transgenic
989 reporters in early stages embryos. Top: *LlambdaG Z5*. Bottom: *LhomieG Z5*. *GFP* is in green,
990 *lacZ* is in orange, and DAPI is in blue, here and in E. D) Quantification of normalized stripe
991 signals for transgenes shown in C. The reporter smFISH intensity (signal to background), here
992 and in F, was measured, normalized, and plotted as described in Methods. N>3. n=27 for
993 *LlambdaG Z5* and n=42 for *LhomieG Z5*. E) Expression patterns of transgenic reporters in late
994 stages embryos. Top: *LlambdaG Z5*. Bottom: *LhomieG Z5*. F) Quantification of normalized
995 stripe signals for transgenes shown in E. N>3. n=27 for *LlambdaG Z5* and n=56 for *LhomieG*
996 *Z5*. Scale bars = 200μm. N = # independent biological replicates. n = # animals. The paired
997 two-tailed t-test were used for statistical analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p <
998 0.0001, ns: not significant. Raw measurements are available in the Source Data files.
999
1000 **Figure 5. The *GFP* reporter is activated by both *eve* and *hebe* enhancers in the *LhomieG***
1001 ***Z5* transgene.** A) Schematics of transgenes. B) Expression patterns of transgenic reporters in
1002 early stage embryos. Top: *LlambdaG Z5*. Bottom: *GhomieL Z5*. *GFP* is in green, *lacZ* is in
1003 orange, and DAPI is in blue, here and in D. C) Quantification of normalized stripes signal as
1004 represented in B. The reporter smFISH intensity (signal to background) was measured,
1005 normalized, and plotted as described in Methods, here and in E. N>3. n=27 for *LlambdaG Z5,*
1006 and n=46 for *LhomieG Z5*. D) Expression patterns of transgenic reporters in late stage embryos.
1007 Top: *LlambdaG Z5*. Bottom: *GhomieL Z5*. E) Quantification of normalized stripe signals for
1008 transgenes shown in D. N>3. n=27 for *LlambdaG Z5*, and n=59 for *LhomieG Z5*. Scale bars =
1009 200μm. N = # independent biological replicates. n = # animals. The paired two-tailed t-test
1010 were used for statistical analysis. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, ns: not
1011 significant. Raw measurements are available in the Source Data files.

1012
1013 **Figure 6. Expression of reporters in Z3 transgenes.** A) Schematics of Z3 transgenes. B)
1014 Boundary pairing for *GhomieL Z3* transgene long-range interactions. In this topology, *GFP* is in
1015 close proximity to *eve* enhancers, and *lacZ* is in contact with the *hebe* enhancer region. C)
1016 Boundary pairing for the *LhomieG Z3* transgene. In this topology, *lacZ* is in close proximity to
1017 both the *eve* enhancers and the *hebe* enhancers. D) Loop extrusion model for *GhomieL Z3*.
1018 Transgenic *homie* and endogenous *homie* flank the *eveMammoth* loop. In this topology, *GFP* is
1019 in close proximity to *eve* enhancers, and *lacZ* is close to the *hebe* enhancers. E) Loop extrusion
1020 model for the *LhomieG Z3* transgene. The cohesin complex bypasses transgenic *homie* and is
1021 stopped at an upstream boundary X to create the *eveGarantuan* loop. Both *GFP* and *lacZ* are at
1022 a similar physical distance from *eve* enhancers and *hebe* enhancers in this topology. F)
1023 Expression patterns of transgenic reporters in early stage embryos. Top: *GhomieL Z3*. Bottom:
1024 *LhomieG Z3*. *GFP* is in green, *lacZ* is in orange, and DAPI is in blue. N=3, n=24 for both
1025 *GhomieL Z3* and *LhomieG Z3*. **G)** Expression patterns of transgenic reporters in late stage
1026 embryos. Top: *GhomieL Z3*. Bottom: *LhomieG Z3*. N=3, n=24 for both *GhomieL Z3* and
1027 *LhomieG Z3*. Scale bars = 200µm, N = # independent biological replicates, n = # animals.
1028
1029 **Figure 7. *homie* in Z5 transgene *LhomieG* and *GhomieL* Z3 mediates long-range**
1030 **interactions with the *eve* locus** *Small dark blue arrows* – insulators upstream and downstream
1031 of *hebe* that appear important to demarcate interaction domains between –142kb and *eve*. *Large*
1032 *blue arrow* – endogenous *nhomie*. *Large red arrow – homie*, either endogenous or in the
1033 transgene. The direction of arrows follow established convention on *nhomie/homie*, and do not
1034 reflect orientation of insulator protein binding motifs *per se*. *Gray boxes* – enhancers. A)
1035 MicroC map of the control line *LlambdaG Z5*. Scaled cartoons of the two loci of interest are
1036 shown directly below the MicroC map, and an unscaled blowup of the elements of interest at
1037 each locus is provided. Within the MicroC map of the entire locus, a zoom-in of the off-diagonal
1038 interaction between the –142kb insertion cassette and the endogenous *eve* locus is shown. Note a
1039 slight increase in interaction frequency (compare to Figure 1A). B, C**)** MicroC map of *LhomieG*
1040 *Z5* and *GhomieL Z5*, respectively. The only difference between the two lines is the orientation of
1041 transgenic *homie*, as indicate below each MicroC map. Blow up of the off-diagonal interaction
1042 between –142kb and *eve*, including scaled cartoons denoting features of interest in the top right
1043 corner. Note the changed pattern of interaction with the endogenous *eve* locus due to the
1044 orientation switch. D) "Virtual 4C" maps obtained from MicroC maps of "B" (top panel), and
1045 "C" (bottom panel). Viewpoints are shown in both panels from either the *lacZ* gene (orange) or
1046 the *GFP* gene (green). Note in each cassette the altered interaction of either gene with
1047 endogenous *eve*, depending on the change of orientation of *homie*.
1048
1049 **Figure 8. *homie* in the transgene interacts with *eve homie* and *nhomie.*** "Virtual 4C" maps
1050 of the *LhomieG* (A) or *GhomieL* (B) transgenes, for the Z5 (blue) and Z3 (orange) orientations.
1051 Viewpoints are taken from the transgenic *homie* sequence. Importantly, note that transgenic
1052 *homie* preferentially interacts with endogenous *homie* no matter the orientation. Intriguingly, the
1053 orientation switch of transgenic *homie* leads to an apparent change in its degree of preference for
1054 endogenous *homie* over *nhomie*.
1055
1056
1057
1058
1059

1060
1061 **Supplemental Figures:**
1062
1063
1064 **Figure Sup. 1. TAD organization of the *even-skipped* locus and neighboring sequences**
1065 MicroC contact pattern for the chromosomal region spanning the attP site at –142 kb and the *eve*
1066 locus. Like the *eve* volcano, this contact pattern was generated using aggregated previously
1067 published NC14 data (Batut *et al.*, 2022; Levo *et al.*, 2022). Arrow indicates the –142kb locus
1068 where transgenes are integrated into the genome. LDC domains are labeled according to the
1069 interacting TADs: for example, H-I indicates crosslinking events generated by physical
1070 interactions between sequences in two neighboring TADs, TH and TI, while I-J indicates
1071 interactions between sequences in TI and TJ. Likewise, G-I indicates crosslinking events
1072 generated by physical interactions between sequences TADs TG and TI, which are separated
1073 from each other by TAD TH. Note that unlike the TADs to the left of *eve*, the *eve* TAD does not
1074 interact frequently with its neighbors. Purple arrow indicates potential interaction dot between
1075 the boundary marking the border of TA and TB and the boundary marking the border of TI and
1076 TJ. Red, blue and green arrowheads mark boundaries that could engage in partner switching: see
1077 text.
1078
1079 **Figure Sup. 2. Trouble ahead, trouble behind.** A) Illustration of a broken zipper—zipping in
1080 front but not closing behind. B) Crosslinking pattern expected for cohesin loading at a site very
1081 close to the middle of the TAD-to-be: The vertical line of linked DNA is generated when the two
1082 chromatin fibers in the TAD-to-be are transiently linked together as cohesin passes. When
1083 cohesin bumps into the CTCF roadblocks (BE) at the boundaries of the TAD, the vertical line
1084 stops. C) Cohesin loading site is offset from center of TAD-to-be to the left. The crosslinking
1085 initiates at the loading site and extends vertically until the cohesin complex bumps into the
1086 CTCF roadblock on the left. At that point extrusion of the left chromatin fiber halts, while it
1087 continues to extrude the chromatin fiber on the right. The line of crosslinked sequences then
1088 extends at a 45° angle until cohesin comes to a halt at the CTCF roadblock on the right. D and
1089 E) The loading site is at or close to the boundary on the left (D) or the right (E). In this case
1090 extrusion will be unidirectional and the DNAs crosslinked as cohesin passes will form a line that
1091 extends at a 45° angle until cohesin comes to a halt at the CTCF roadblock. F). Cohesin is able
1092 to load at multiple sites in the TAD-to-be. In this case there will be a series of vertical lines
1093 whose intensity will be correlated with the relative frequency that a give loading site is used. The
1094 vertical lines will extend in each case until cohesin comes to a halt at the closest CTCF
1095 roadblock. At that point, the line of crosslinked DNA will extend at a 45° angle until that
1096 particular cohesin complex comes to halt at the CTCF roadblock that is farther away. The 45°
1097 line of crosslinked DNA will become progressive more intense as it approaches the apex of the
1098 triangle.
1099
1100
1101
1102 **Figure Sup. 3. Models for the formation of LDC domains.** Several different mechanisms
1103 could explain the LDC domains. A-D). In the loop extrusion model, the LDC domains arise
1104 because cohesin sometimes breaks through a CTCF roadblock. In A) there are three primary
1105 TADs, TAD1, TAD2, and TAD3. B and C) Occasional breakthrough events can generate a
1106 secondary TAD, LDC1-2 (TAD1+TAD2) or LDC2-3 (TAD2+TAD3). D) Even less frequent
1107 breakthrough events can generate a tripartite TAD (TAD1+TAD2+TAD3) that will correspond

1108 to LDC1-3. E-G). LDC domains arise because TADs sometimes bump into their neighbors. In
1109 E, TAD1 bumps into TAD2 to give LDC1-2, while in F TAD2 bumps into TAD3 to give
1110 LDC2:3. G) Less frequently TAD1 bumps into TAD3 to give LDC1:3. H-J) LDC domains arise
1111 because boundaries can switch partners. H and I) In H boundary A (dark green) skips boundary
1112 B (red) and pairs with boundary C (tan). This generates a new TAD, TAD1+TAD2 which
1113 corresponds to LDC domain 1-2. In I, boundary B (red) skips boundary C (tan) and pairs with D
1114 (blue). This generates a new TAD, TAD2+TAD3 which corresponds to LDC domain 2-3. J) In
1115 J, boundary A (dark green) skips boundary B (red) and boundary C (tan) and pairs with boundary
1116 D (blue). This generates a new TAD, TAD1+TAD2+TAD3 which corresponds to LDC domain
1117 1-3.
1118
1119
1120 **Figure Sup. 4. Loop extrusion: *eveElephant*.** A) Loop extrusion orientation dependent model
1121 for LhomieG Z5 transgene as the sketched in Figure 2E. This configuration generates *eveMa*. B)
1122 Instead running through *nhomie*, the cohesin complex comes to a halt at the *nhomie* boundary,
1123 generating *eveElephant*. A second cohesin complex generates the *eve* TAD by loading at a site
1124 within the *eve* locus. Both *eveElephant* and the *eve* TAD are stem-loops. However, while this
1125 double stem-loop configuration would explain why the transgene *homie* makes contact with both
1126 *nhomie* and *homie*, the *eve* enhancers are isolated in their own loop and would not make contact
1127 with the two reporters in the transgene.
1128
1129 **Figure Sup. 5. MicroC of *GhomieL Z3*.** A) MicroC map of *GhomieL Z3*. Gray lines are
1130 masked regions due to repeated sequences. In this case, focusing on the zoom-in map from top
1131 to bottom, most likely P-element remnants 3' of the *lacZ* gene, transgenic *homie*, and P-element
1132 remnants 3' of the *GFP* gene. Compare directly to *GhomieL Z5* for changes in interaction
1133 preference due to change of *homie* orientation (Figure 4C). B) "Virtual 4C" map from either
1134 *lacZ* or *GFP* gene body viewpoints.
1135
1136 **Figure Sup. 6. Micro C of *LhomieG Z3*.** A) MicroC map of *LhomieG Z3*. Gray lines are
1137 masked regions due to repeated sequences. In this case, most possibly *P-element* remnants 3' of
1138 the *lacZ* gene. Compare directly to *GhomieL Z3* for changes in interaction preference due to
1139 change of homie orientation (Figure 4B). B) "Virtual 4C" map from either *lacZ* or *GFP* gene
1140 body viewpoints.
1141
1142
1143
1144 **Figure Sup. 7. Blowup of interactions generated by the transgene with sequences**
1145 **neighboring –142 kb and the *eve* TAD.** The interaction pattern indicates that the loop topology
1146 changes as a result of the orientation of transgenic *homie*. Blow-up of off-diagonal interactions,
1147 focused on the interaction between the endogenous *eve* locus and the –142kb transgenic insert
1148 with A) *LhomieG Z3,* B) *LhomieG Z5,* C) *GhomieL Z3,* or D) *GhomieL Z5*. Note the switch in
1149 preferential interaction between regions adjacent to transgenic *homie* and endogenous *homie*
1150 (indicated by arrow brackets). In cassettes where transgenic *homie* is in the same orientation in
1151 the genome as endogenous *homie* (A and D), regions upstream of transgenic *homie* preferentially
1152 interact with regions downstream of endogenous *homie*. In contrast, when transgenic *homie* is in
1153 the opposite orientation in the genome as endogenous *homie* (B and C), regions downstream of
1154 transgenic *homie* preferentially interact with regions downstream of endogenous *homie*. This

suggests a different loop topology when the orientation of the transgene *homie* element in the chromosome is flipped.

REFERENCES

Abdennur, N., and Mirny, L.A. (2020). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics *36*, 311-316. 10.1093/bioinformatics/btz540.

Alipour, E., and Marko, J.F. (2012). Self-organization of domain structures by DNA-loop-extruding enzymes. Nucleic Acids Res. *40*, 11202-11212. 10.1093/nar/gks925.

Avva, S.V., and Hart, C.M. (2016). Characterization of the Drosophila BEAF-32A and BEAF-32B Insulator Proteins. PLoS One *11*, e0162906. 10.1371/journal.pone.0162906.

Bateman, J.R., Lee, A.M., and Wu, C.T. (2006). Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. Genetics *173*, 769-777.

Batut, P.J., Bing, X.Y., Sisco, Z., Raimundo, J., Levo, M., and Levine, M.S. (2022). Genome organization controls transcriptional dynamics during development. Science *375*, 566-570. 10.1126/science.abi7178.

Bell, A.C., West, A.G., and Felsenfeld, G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science *291*, 447-450. 10.1126/science.291.5503.447.

Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes Dev. *17*, 664-675.

Bonchuk, A., Boyko, K., Fedotova, A., Nikolaeva, A., Lushchekina, S., Khrustaleva, A., Popov, V., and Georgiev, P. (2021). Structural basis of diversity and homodimerization specificity of zinc-finger-associated domains in Drosophila. Nucleic Acids Res. *49*, 2375-2389. 10.1093/nar/gkab061.

Bonchuk, A., Denisov, S., Georgiev, P., and Maksimenko, O. (2011). Drosophila BTB/POZ domains of "ttk group" can form multimers and selectively interact with each other. J. Mol. Biol. *412*, 423-436. 10.1016/j.jmb.2011.07.052.

Bonchuk, A., Kamalyan, S., Mariasina, S., Boyko, K., Popov, V., Maksimenko, O., and Georgiev, P. (2020). N-terminal domain of the architectural protein CTCF has similar structural organization and ability to self-association in bilaterian organisms. Sci Rep *10*, 2677. 10.1038/s41598-020-59459-5.

Cai, H.N., and Shen, P. (2001). Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. Science *291*, 493-495.

Callan, H.G. (1963). THE NATURE OF LAMPBRUSH CHROMOSOMES. Int. Rev. Cytol. *15*, 1-34. 10.1016/s0074-7696(08)61114-6.

Cavalheiro, G.R., Pollex, T., and Furlong, E.E. (2021). To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? Curr. Opin. Genet. Dev. *67*, 119-129. 10.1016/j.gde.2020.12.015.

Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nat. Genet. *50*, 1296.

Chetverina, D., Aoki, T., Erokhin, M., Georgiev, P., and Schedl, P. (2014). Making connections: insulators organize eukaryotic chromosomes into independent cis-regulatory networks. Bioessays *36*, 163-172.

1199 Chetverina, D., Fujioka, M., Erokhin, M., Georgiev, P., Jaynes, J.B., and Schedl, P. (2017).
1200 Boundaries of loop domains (insulators): Determinants of chromosome form and function in
1201 multicellular eukaryotes. Bioessays *39*, 1600233.
1202 Collier, J.E., and Nasmyth, K.A. (2022). DNA passes through cohesin's hinge as well as its
1203 Smc3-kleisin interface. Elife *11*. 10.7554/eLife.80310.
1204 Cuartero, S., Fresan, U., Reina, O., Planet, E., and Espinas, M.L. (2014). Ibf1 and Ibf2 are novel
1205 CP190-interacting proteins required for insulator function. EMBO J. *33*, 637-647.
1206 Davidson, I.F., and Peters, J.M. (2021). Genome folding through loop extrusion by SMC
1207 complexes. Nat Rev Mol Cell Biol *22*, 445-464. 10.1038/s41580-021-00349-7.
1208 dos Santos, G., Schroeder, A.J., Goodman, J.L., Strelets, V.B., Crosby, M.A., Thurmond, J.,
1209 Emmert, D.B., Gelbart, W.M., and FlyBase, C. (2015). FlyBase: introduction of the Drosophila
1210 melanogaster Release 6 reference genome assembly and large-scale migration of genome
1211 annotations. Nucleic Acids Res. *43*, D690-697.
1212 Duan, J., Rieder, L., Colonnetta, M.M., Huang, A., McKenney, M., Watters, S., Deshpande, G.,
1213 Jordan, W., Fawzi, N., and Larschan, E. (2021). CLAMP and Zelda function together to promote
1214 Drosophila zygotic genome activation. Elife *10*. 10.7554/eLife.69937.
1215 Erokhin, M., Davydova, A., Kyrchanova, O., Parshikov, A., Georgiev, P., and Chetverina, D.
1216 (2011). Insulators form gene loops by interacting with promoters in Drosophila. Development
1217 *138*, 4097-4106.
1218 Erokhin, M., Gorbenko, F., Lomaev, D., Mazina, M.Y., Mikhailova, A., Garaev, A.K.,
1219 Parshikov, A., Vorobyeva, N.E., Georgiev, P., Schedl, P., and Chetverina, D. (2021). Boundaries
1220 potentiate polycomb response element-mediated silencing. BMC Biol. *19*, 113. 10.1186/s12915-
1221 021-01047-8.
1222 Fedotova, A., Aoki, T., Rossier, M., Mishra, R.K., Clendinen, C., Kyrchanova, O., Wolle, D.,
1223 Bonchuk, A., Maeda, R.K., and Mutero, A. (2018). The BEN domain protein insensitive binds to
1224 the Fab-7 chromatin boundary to establish proper segmental identity in Drosophila. Genetics
1225 *210*, 573-585.
1226 Fedotova, A., Clendinen, C., Bonchuk, A., Mogila, V., Aoki, T., Georgiev, P., and Schedl, P.
1227 (2019). Functional dissection of the developmentally restricted BEN domain chromatin boundary
1228 factor Insensitive. Epigenetics & chromatin *12*, 2.
1229 Fedotova, A.A., Bonchuk, A.N., Mogila, V.A., and Georgiev, P.G. (2017). C2H2 Zinc Finger
1230 Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors.
1231 Acta Naturae *9*, 47-58.
1232 Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016).
1233 Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.
1234 10.1016/j.celrep.2016.04.085.
1235 Fujioka, M., Mistry, H., Schedl, P., and Jaynes, J.B. (2016). Determinants of Chromosome
1236 Architecture: Insulator Pairing in *cis* and in *trans*. PLoS Genet. *12*, e1005889.
1237 Fujioka, M., Sun, G., and Jaynes, J.B. (2013). The Drosophila *eve* insulator Homie promotes *eve*
1238 expression and protects the adjacent gene from repression by Polycomb spreading. PLoS Genet.
1239 *9*, e1003883.
1240 Fujioka, M., Wu, X., and Jaynes, J.B. (2009). A chromatin insulator mediates transgene homing
1241 and very long-range enhancer-promoter communication. Development *136*, 3077-3087.
1242 Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer control of transcriptional bursting. Cell
1243 *166*, 358-368.
1244 Fung, J.C., Marshall, W.F., Dernburg, A., Agard, D.A., and Sedat, J.W. (1998). Homologous
1245 chromosome pairing in Drosophila melanogaster proceeds through multiple independent
1246 initiations. J. Cell Biol. *141*, 5-20.

1247 Gaskill, M.M., Gibson, T.J., Larson, E.D., and Harrison, M.M. (2021). GAF is essential for
1248 zygotic genome activation and chromatin accessibility in the early Drosophila embryo. Elife *10*.
1249 10.7554/eLife.66668.
1250 Ghosh, R.P., and Meyer, B.J. (2021). Spatial Organization of Chromatin: Emergence of
1251 Chromatin Structure During Development. Annu. Rev. Cell. Dev. Biol. *37*, 199-232.
1252 10.1146/annurev-cellbio-032321-035734.
1253 Gligoris, T.G., Scheinost, J.C., Bürmann, F., Petela, N., Chan, K.L., Uluocak, P., Beckouët, F.,
1254 Gruber, S., Nasmyth, K., and Löwe, J. (2014). Closing the cohesin ring: structure and function of
1255 its Smc3-kleisin interface. Science *346*, 963-967. 10.1126/science.1256917.
1256 Gohl, D., Aoki, T., Blanton, J., Shanower, G., Kappes, G., and Schedl, P. (2011). Mechanism of
1257 chromosomal boundary action: roadblock, sink, or loop? Genetics *187*, 731-748.
1258 Gramates, L.S., Agapite, J., Attrill, H., Calvi, B.R., Crosby, M.A., Dos Santos, G., Goodman,
1259 J.L., Goutte-Gattat, D., Jenkins, V.K., Kaufman, T., et al. (2022). FlyBase: a guided tour of
1260 highlighted features. Genetics *220*. 10.1093/genetics/iyac035.
1261 Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K.E., Li, W., Myers, R.M., Maniatis, T., and
1262 Wu, Q. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter
1263 choice. Proc Natl Acad Sci USA *109*, 21081-21086. 10.1073/pnas.1219280110.
1264 Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., and Tang,
1265 Y. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter
1266 function. Cell *162*, 900-910.
1267 Hart, C.M., Zhao, K., and Laemmli, U.K. (1997). The scs' boundary element: characterization of
1268 boundary element-associated factors. Mol Cell Biol. *17*, 999-1009.
1269 Heger, P., George, R., and Wiehe, T. (2013). Successive gain of insulator proteins in arthropod
1270 evolution. Evolution *67*, 2945-2956.
1271 Heger, P., and Wiehe, T. (2014). New tools in the box: an evolutionary synopsis of chromatin
1272 insulators. Trends Genet. *30*, 161-171.
1273 Hsieh, T.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq,
1274 X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin
1275 Folding. Mol. Cell *78*, 539-553.e538. 10.1016/j.molcel.2020.03.002.
1276 Jerković, I., Szabo, Q., Bantignies, F., and Cavalli, G. (2020). Higher-Order Chromosomal
1277 Structures Mediate Genome Function. J. Mol. Biol. *432*, 676-681. 10.1016/j.jmb.2019.10.014.
1278 Ke, W., Fujioka, M., Schdle, P., and Jaynes, J.B. Chromosome Structure II: Stem-loops and
1279 circle-loops.
1280 Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.S., Parsi, K.M.,
1281 Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of Mammalian
1282 Chromosome Architecture. Mol. Cell *78*, 554-565.e557. 10.1016/j.molcel.2020.03.003.
1283 Kruse, K., Hug, C.B., and Vaquerizas, J.M. (2020). FAN-C: a feature-rich framework for the
1284 analysis and visualisation of chromosome conformation capture data. Genome Biol *21*, 303.
1285 10.1186/s13059-020-02215-9.
1286 Kyrchanova, O., Chetverina, D., Maksimenko, O., Kullyev, A., and Georgiev, P. (2008a).
1287 Orientation-dependent interaction between Drosophila insulators is a property of this class of
1288 regulatory elements. Nucleic Acids Res. *36*, 7019-7028.
1289 Kyrchanova, O., Ibragimov, A., Postika, N., Georgiev, P., and Schedl, P. (2023). Boundary
1290 bypass activity in the abdominal-B region of the Drosophila bithorax complex is position
1291 dependent and regulated. Open Biol *13*, 230035. 10.1098/rsob.230035.
1292 Kyrchanova, O., Ivlieva, T., Toshchakov, S., Parshikov, A., Maksimenko, O., and Georgiev, P.
1293 (2011). Selective interactions of boundaries with upstream region of Abd-B promoter in

1294 Drosophila bithorax complex and role of dCTCF in this process. Nucleic Acids Res. *39*, 3042-
1295 3052.
1296 Kyrchanova, O., Kurbidaeva, A., Sabirov, M., Postika, N., Wolle, D., Aoki, T., Maksimenko, O.,
1297 Mogila, V., Schedl, P., and Georgiev, P. (2018). The bithorax complex iab-7 Polycomb response
1298 element has a novel role in the functioning of the Fab-7 chromatin boundary. PLoS Genet. *14*,
1299 e1007442.
1300 Kyrchanova, O., Sabirov, M., Mogila, V., Kurbidaeva, A., Postika, N., Maksimenko, O., Schedl,
1301 P., and Georgiev, P. (2019a). Complete reconstitution of bypass and blocking functions in a
1302 minimal artificial Fab-7 insulator from Drosophila bithorax complex. Proceedings of the
1303 National Academy of Sciences, 201907190.
1304 Kyrchanova, O., Toshchakov, S., Parshikov, A., and Georgiev, P. (2007). Study of the functional
1305 interaction between Mcp insulators from the Drosophila bithorax complex: effects of insulator
1306 pairing on enhancer-promoter communication. Mol. Cell. Biol. *27*, 3035-3043.
1307 Kyrchanova, O., Toshchakov, S., Podstreshnaya, Y., Parshikov, A., and Georgiev, P. (2008b).
1308 Functional interaction between the Fab-7 and Fab-8 boundaries and the upstream promoter
1309 region in the Drosophila Abd-B gene. Mol. Cell. Biol. *28*, 4188-4195.
1310 Kyrchanova, O., Wolle, D., Sabirov, M., Kurbidaeva, A., Aoki, T., Maksimenko, O.,
1311 Kyrchanova, M., Georgiev, P., and Schedl, P. (2019b). Distinct Elements Confer the Blocking
1312 and Bypass Functions of the Bithorax Fab-8 Boundary. Genetics, genetics. 302694.302019.
1313 Levo, M., Raimundo, J., Bing, X.Y., Sisco, Z., Batut, P.J., Ryabichko, S., Gregor, T., and
1314 Levine, M.S. (2022). Transcriptional coupling of distant regulatory genes in living embryos.
1315 Nature *605*, 754-760. 10.1038/s41586-022-04680-7.
1316 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
1317 transform. Bioinformatics *25*, 1754-1760. 10.1093/bioinformatics/btp324.
1318 Li, H.-B., Muller, M., Bahechar, I.A., Kyrchanova, O., Ohno, K., Georgiev, P., and Pirrotta, V.
1319 (2011). Insulators, not Polycomb response elements, are required for long-range interactions
1320 between Polycomb targets in Drosophila melanogaster. Mol. Cell. Biol. *31*, 616-625.
1321 Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.T., Cubeñas-Potts, C., Hu, M.,
1322 Lei, E.P., Bosco, G., et al. (2015). Widespread rearrangement of 3D chromatin organization
1323 underlies polycomb-mediated stress-induced silencing. Mol. Cell *58*, 216-231.
1324 10.1016/j.molcel.2015.02.023.
1325 Li, X., Tang, X., Bing, X., Catalano, C., Li, T., Dolsten, G., Wu, C., and Levine, M. (2023).
1326 GAGA-associated factor fosters loop formation in the Drosophila genome. Mol. Cell *83*, 1519-
1327 1526.e1514. 10.1016/j.molcel.2023.03.011.
1328 Lim, B., Heist, T., Levine, M., and Fukaya, T. (2018). Visualization of transvection in living
1329 Drosophila embryos. Mol. Cell *70*, 287-296. e286.
1330 Little, S.C., and Gregor, T. (2018). Single mRNA Molecule Detection in Drosophila. Methods
1331 Mol Biol *1649*, 127-142. 10.1007/978-1-4939-7213-5_8.
1332 Matthews, N.E., and White, R. (2019). Chromatin Architecture in the Fly: Living without
1333 CTCF/Cohesin Loop Extrusion?
1334 Mirny, L., and Dekker, J. (2022). Mechanisms of chromosome folding and nuclear organization:
1335 their interplay and open questions. Cold Spring Harbor Perspectives in Biology *14*, a040147.
1336 Mohana, G., Dorier, J., Li, X., Mouginot, M., Smith, R.C., Malek, H., Leleu, M., Rodriguez, D.,
1337 Khadka, J., Rosa, P., et al. (2023). Chromosome-level organization of the regulatory genome in
1338 the Drosophila nervous system. Cell *186*, 3826-3844.e3826. 10.1016/j.cell.2023.07.008.
1339 Muller, M., Hagstrom, K., Gyurkovics, H., Pirrotta, V., and Schedl, P. (1999). The mcp element
1340 from the Drosophila melanogaster bithorax complex mediates long-distance regulatory
1341 interactions. Genetics *153*, 1333-1356.

1342 Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., and
1343 Georgiev, P. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. Science
1344 *291*, 495-498.
1345 Nora, E.P., Caccianini, L., Fudenberg, G., So, K., Kameswaran, V., Nagle, A., Uebersohn, A.,
1346 Hajj, B., Saux, A.L., Coulon, A., et al. (2020). Molecular basis of CTCF binding polarity in
1347 genome folding. Nat Commun *11*, 5612. 10.1038/s41467-020-19283-x.
1348 Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N., and Mirny, L.A. (2018). Chromatin
1349 organization by an interplay of loop extrusion and compartmental segregation. Proc Natl Acad
1350 Sci USA *115*, E6697-e6706. 10.1073/pnas.1717730115.
1351 Perea-Resa, C., Wattendorf, L., Marzouk, S., and Blower, M.D. (2021). Cohesin: behind
1352 dynamic genome topology and gene expression reprogramming. Trends Cell Biol. *31*, 760-773.
1353 10.1016/j.tcb.2021.03.005.
1354 Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture.
1355 Nature Reviews Genetics *19*, 789-800.
1356 Sanborn, A.L., Rao, S.S., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov,
1357 I.D., Chinnappan, D., Cutkosky, A., and Li, J. (2015). Chromatin extrusion explains key features
1358 of loop and domain formation in wild-type and engineered genomes. Proceedings of the National
1359 Academy of Sciences *112*, E6456-E6465.
1360 Schoborg, T.A., and Labrador, M. (2010). The phylogenetic distribution of non-CTCF insulator
1361 proteins is limited to insects and reveals that BEAF-32 is Drosophila lineage specific. J. Mol.
1362 Evol. *70*, 74-84. 10.1007/s00239-009-9310-x.
1363 Sigrist, C.J., and Pirrotta, V. (1997). Chromatin insulator elements block the silencing of a target
1364 gene by the Drosophila polycomb response element (PRE) but allow trans interactions between
1365 PREs on different chromosomes. Genetics *147*, 209-221.
1366 Sun, Y., Nien, C.Y., Chen, K., Liu, H.Y., Johnston, J., Zeitlinger, J., and Rushlow, C. (2015).
1367 Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic
1368 genome activation. Genome Res. *25*, 1703-1714. 10.1101/gr.192542.115.
1369 Tan, L., Shi, J., Moghadami, S., Parasar, B., Wright, C.P., Seo, Y., Vallejo, K., Cobos, I.,
1370 Duncan, L., Chen, R., and Deisseroth, K. (2023). Lifelong restructuring of 3D genome
1371 architecture in cerebellar granule cells. Science *381*, 1112-1119. 10.1126/science.adh3253.
1372 Trcek, T., Lionnet, T., Shroff, H., and Lehmann, R. (2017). mRNA quantification using single-
1373 molecule FISH in Drosophila embryos. Nat Protoc *12*, 1326-1348. 10.1038/nprot.2017.030.
1374 Ueberschär, M., Wang, H., Zhang, C., Kondo, S., Aoki, T., Schedl, P., Lai, E.C., Wen, J., and
1375 Dai, Q. (2019). BEN-solo factors partition active chromatin to ensure proper gene activation in
1376 Drosophila. Nat Commun *10*, 5700. 10.1038/s41467-019-13558-8.
1377 Vazquez, J., Farkas, G., Gaszner, M., Udvardy, A., Muller, M., Hagstrom, K., Gyurkovics, H.,
1378 Sipos, L., Gausz, J., Galloni, M., and et al. (1993). Genetic and molecular analysis of chromatin
1379 domains. Cold Spring Harb Symp Quant Biol *58*, 45-54. 10.1101/sqb.1993.058.01.008.
1380 Vazquez, J., Muller, M., Pirrotta, V., and Sedat, J.W. (2006). The Mcp element mediates stable
1381 long-range chromosome-chromosome interactions in Drosophila. Molecular Biology of the Cell
1382 *17*, 2158-2165.
1383 Yokoshi, M., Segawa, K., and Fukaya, T. (2020). Visualizing the Role of Boundary Elements in
1384 Enhancer-Promoter Communication. Mol. Cell *78*, 224-235.e225. 10.1016/j.molcel.2020.02.007.
1385 Zhimulev, I.F., and Belyaeva, E.S. (1991). Chromomeric organization of polytene chromosomes.
1386 Genetica *85*, 65-72. 10.1007/bf00056107.
1387 Zhimulev, I.F., Pokholkova, G.V., Bgatov, A.V., Semeshin, V.F., Umbetova, G.H., and
1388 Belyaeva, E.S. (1983). Genetic interpretation of polytene chromosomes banding pattern. Mol.
1389 Biol. Rep. *9*, 19-23. 10.1007/bf00777469.

1390 Zolotarev, N., Fedotova, A., Kyrchanova, O., Bonchuk, A., Penin, A.A., Lando, A.S., Eliseeva,
1391 I.A., Kulakovskiy, I.V., Maksimenko, O., and Georgiev, P. (2016). Architectural proteins Pita,
1392 Zw5,and ZIPIC contain homodimerization domain and support specific long-range interactions
1393 in Drosophila. Nucleic Acids Res. *44*, 7228-7241. 10.1093/nar/gkw371.
1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

1409

1410

1411

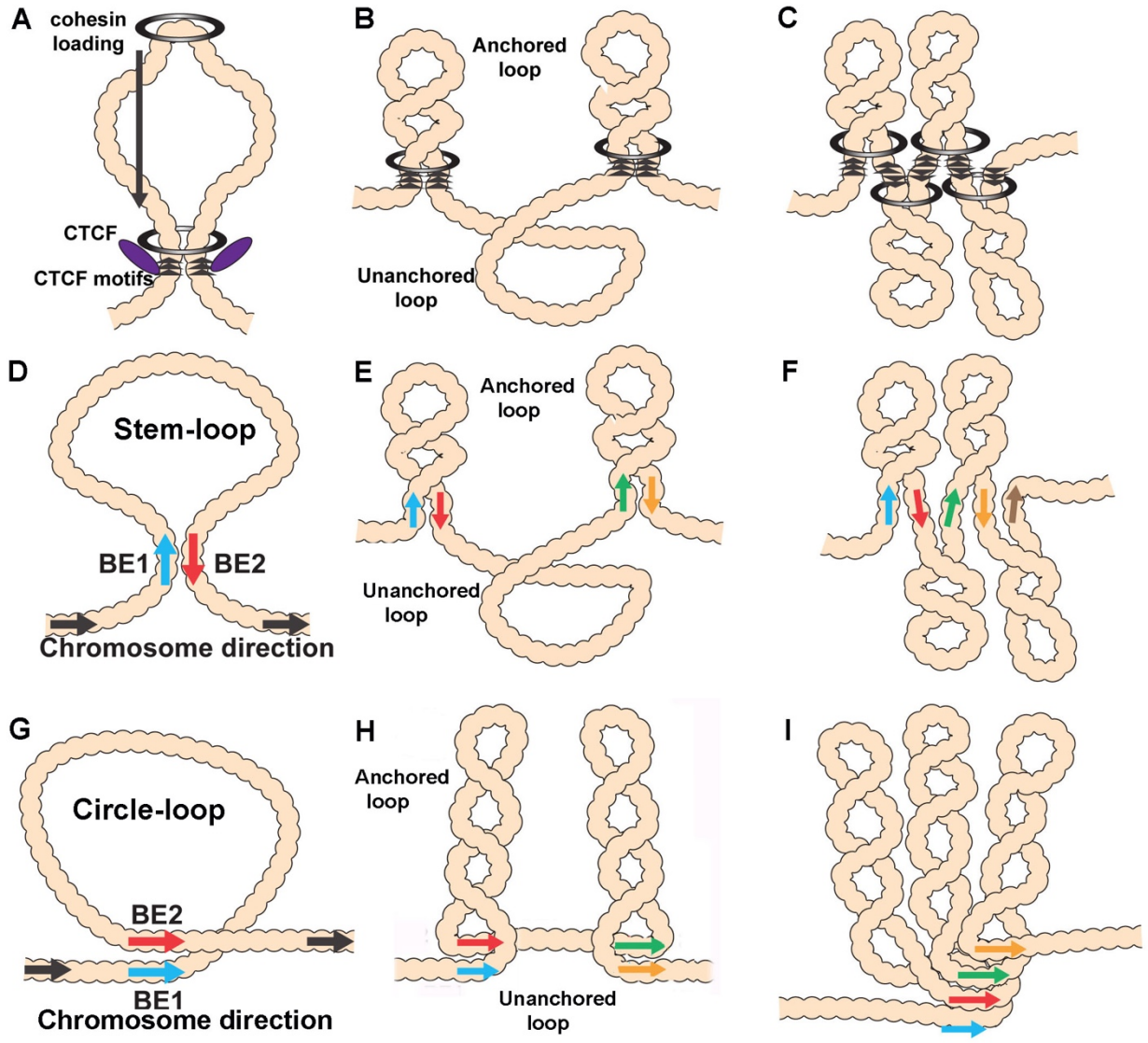
1412

1413

1414

1415

1416 **FIGURES**
1417



1418
1419 **Figure 1. Diagram of the possible loop topologies generated by loop-extrusion and**
1420 **boundary:boundary pairing.**
1421
1422
1423

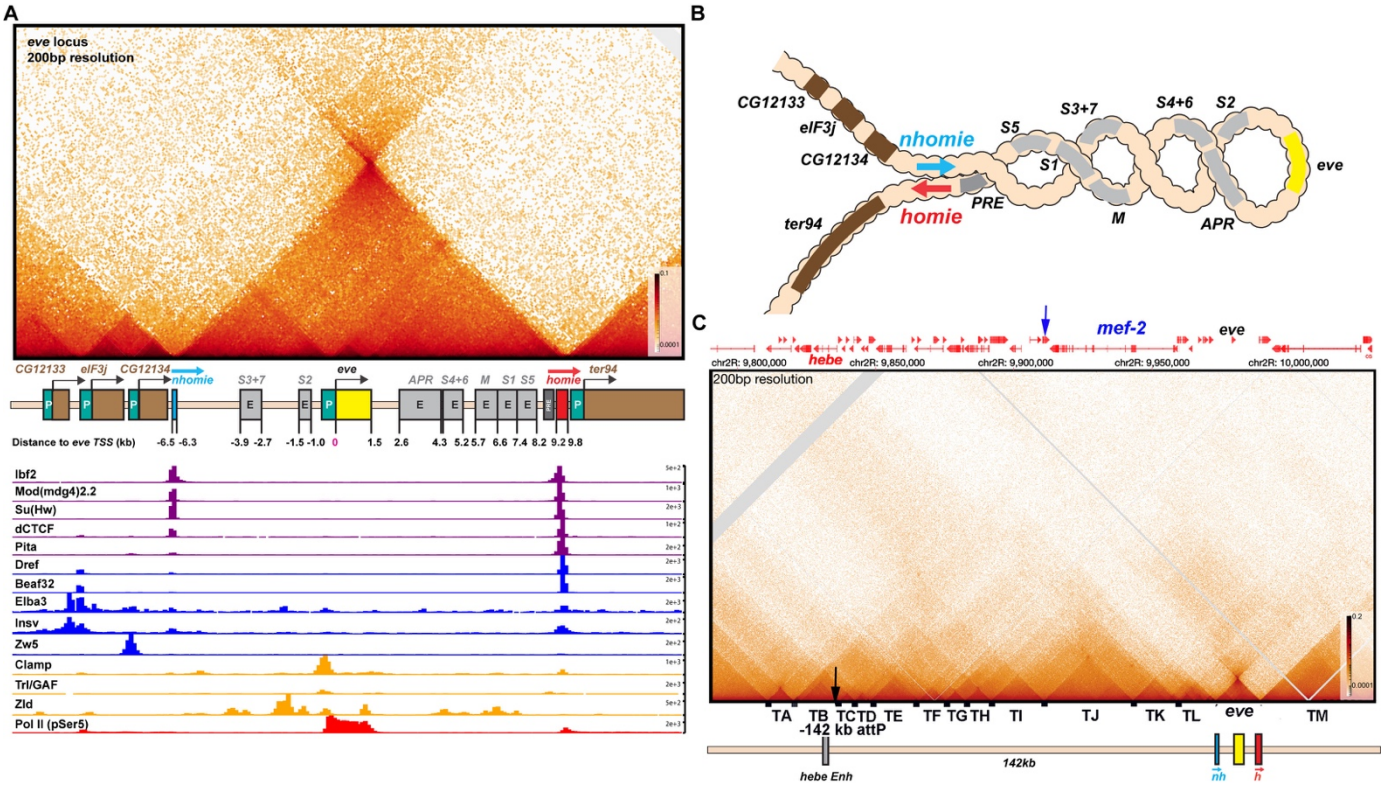


Figure 2. TAD organization of the *even-skipped* locus and neighboring sequences.

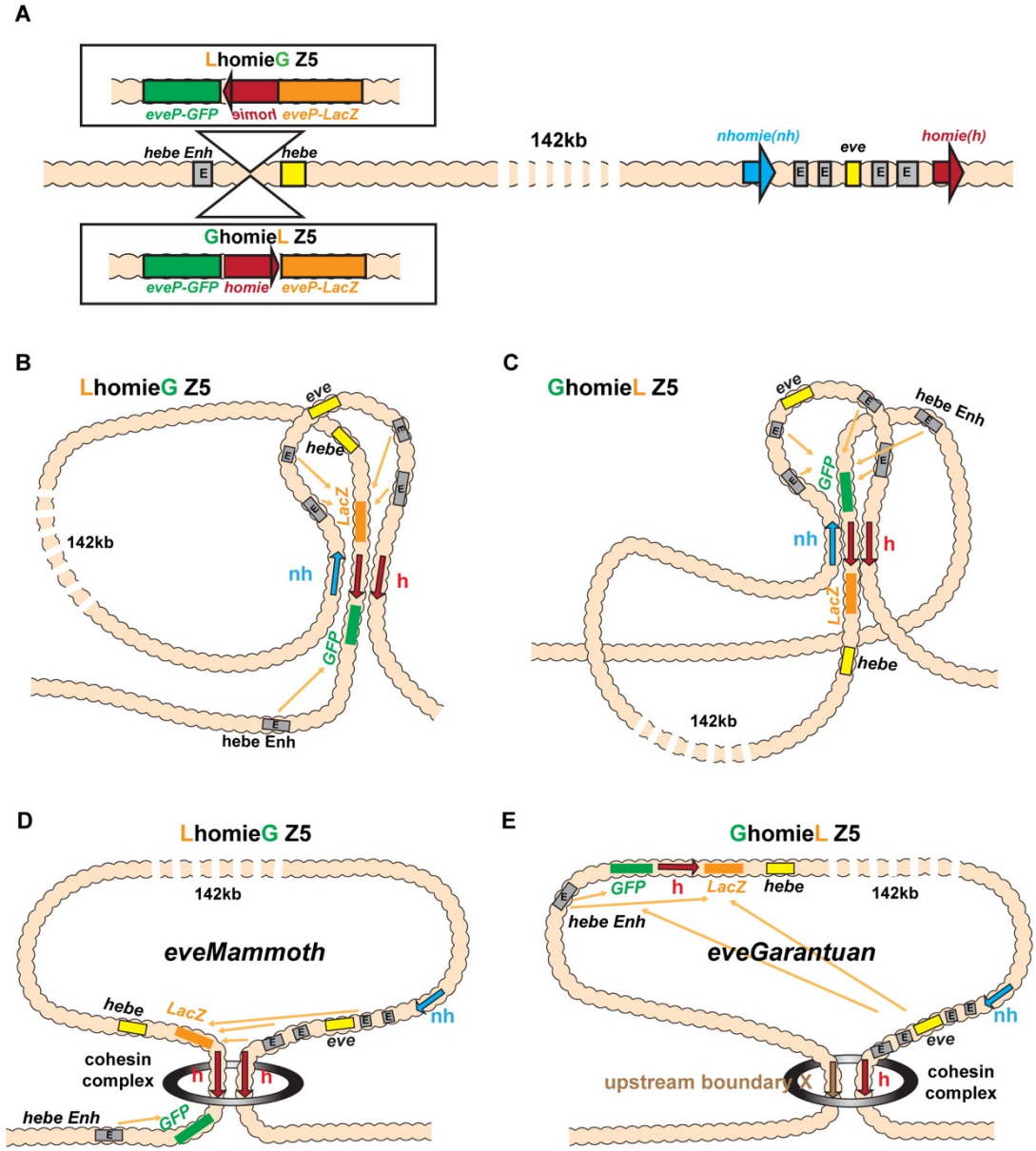


Figure 3. Schematics of boundary pairing and loop extrusion.

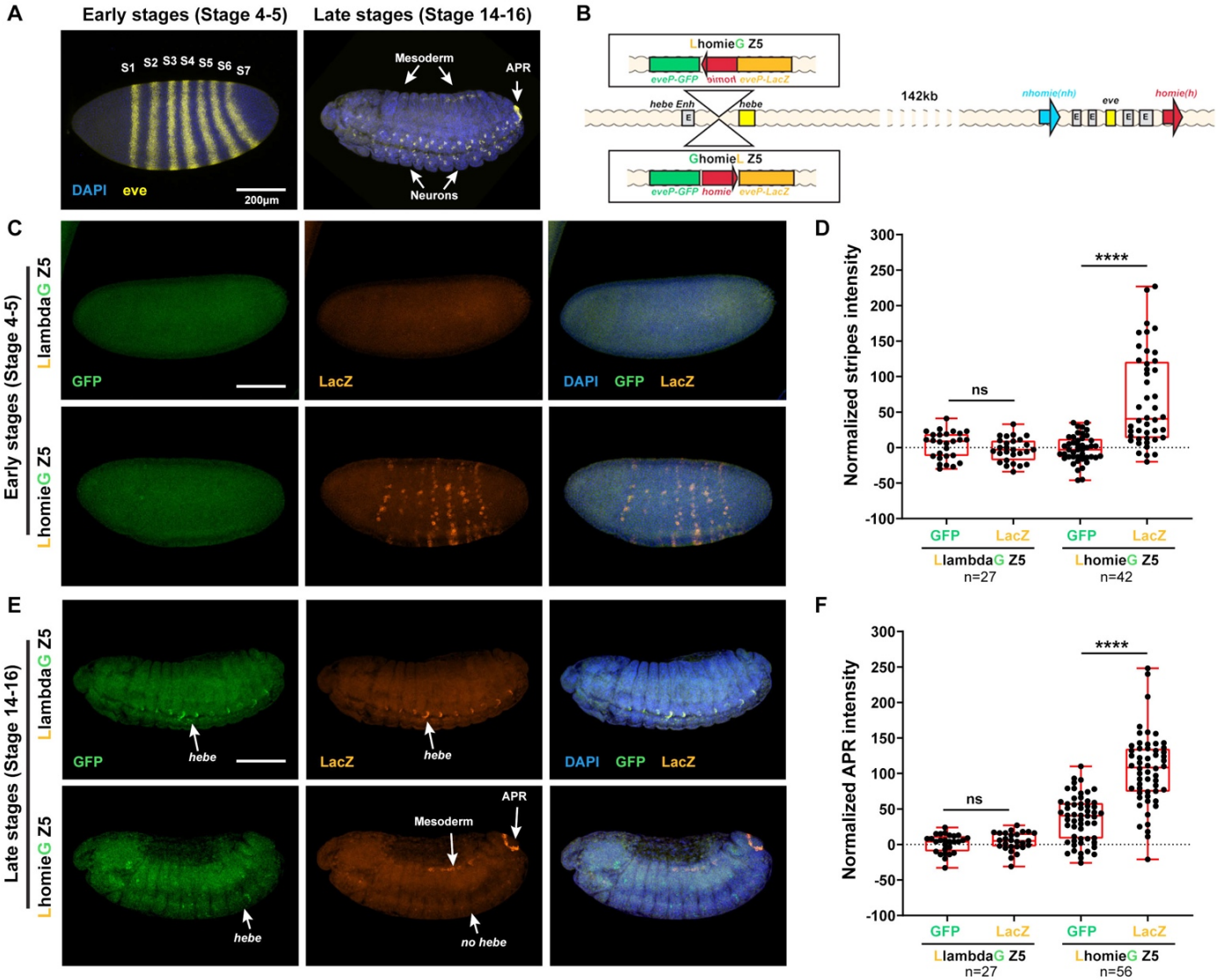


Figure 4. The *lacZ* reporter is activated by *eve* enhancers in the *LhomieG Z5* transgene.

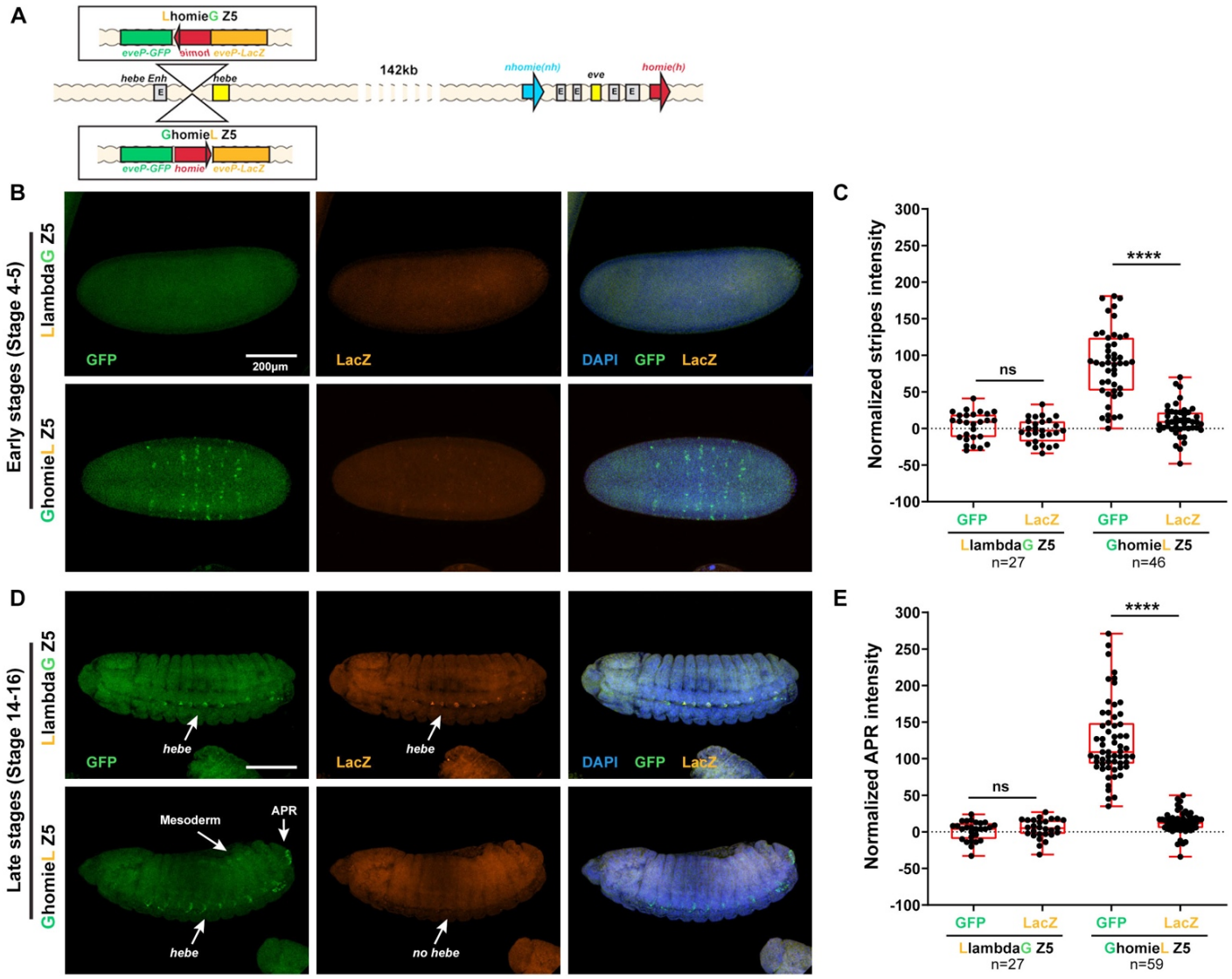


Figure 5. The *GFP* reporter is activated by both *eve* and *hebe* enhancers in the *LhomieG Z5* transgene.

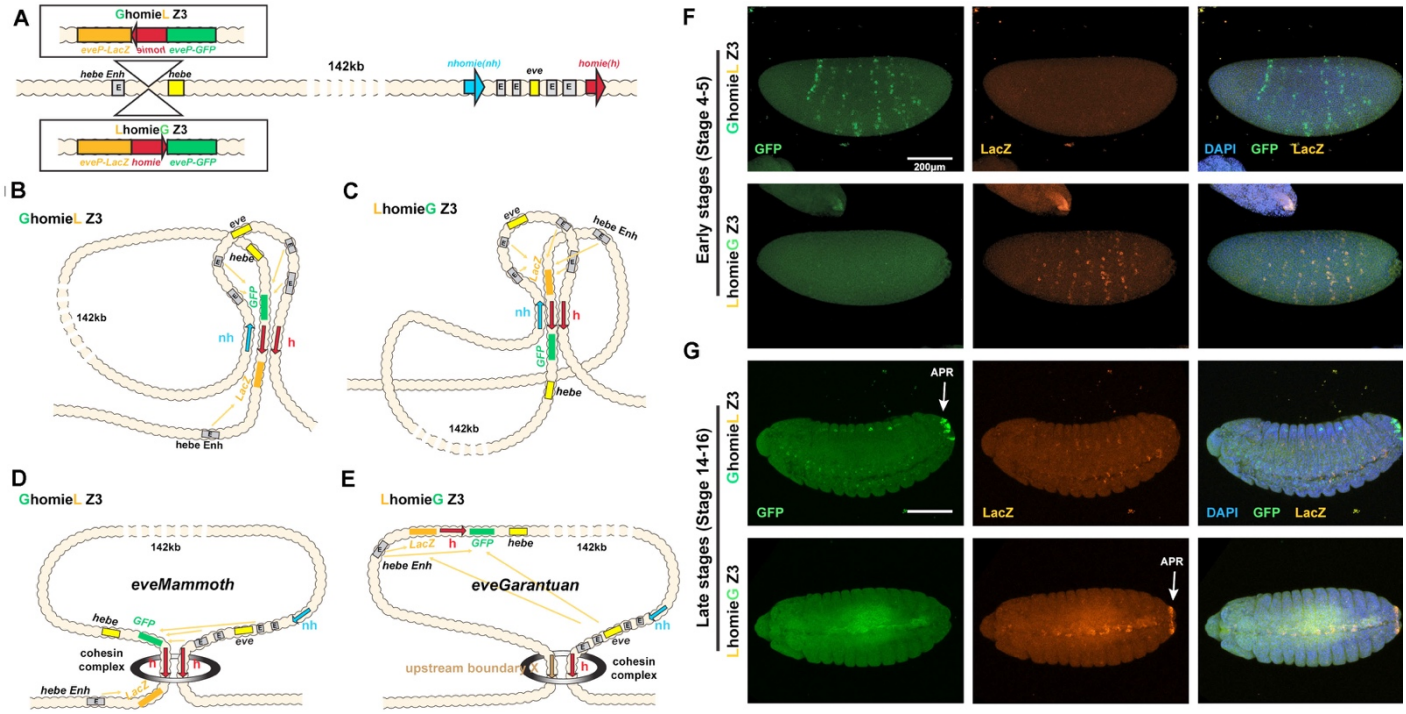


Figure 6. Expression of reporters in Z3 transgenes.

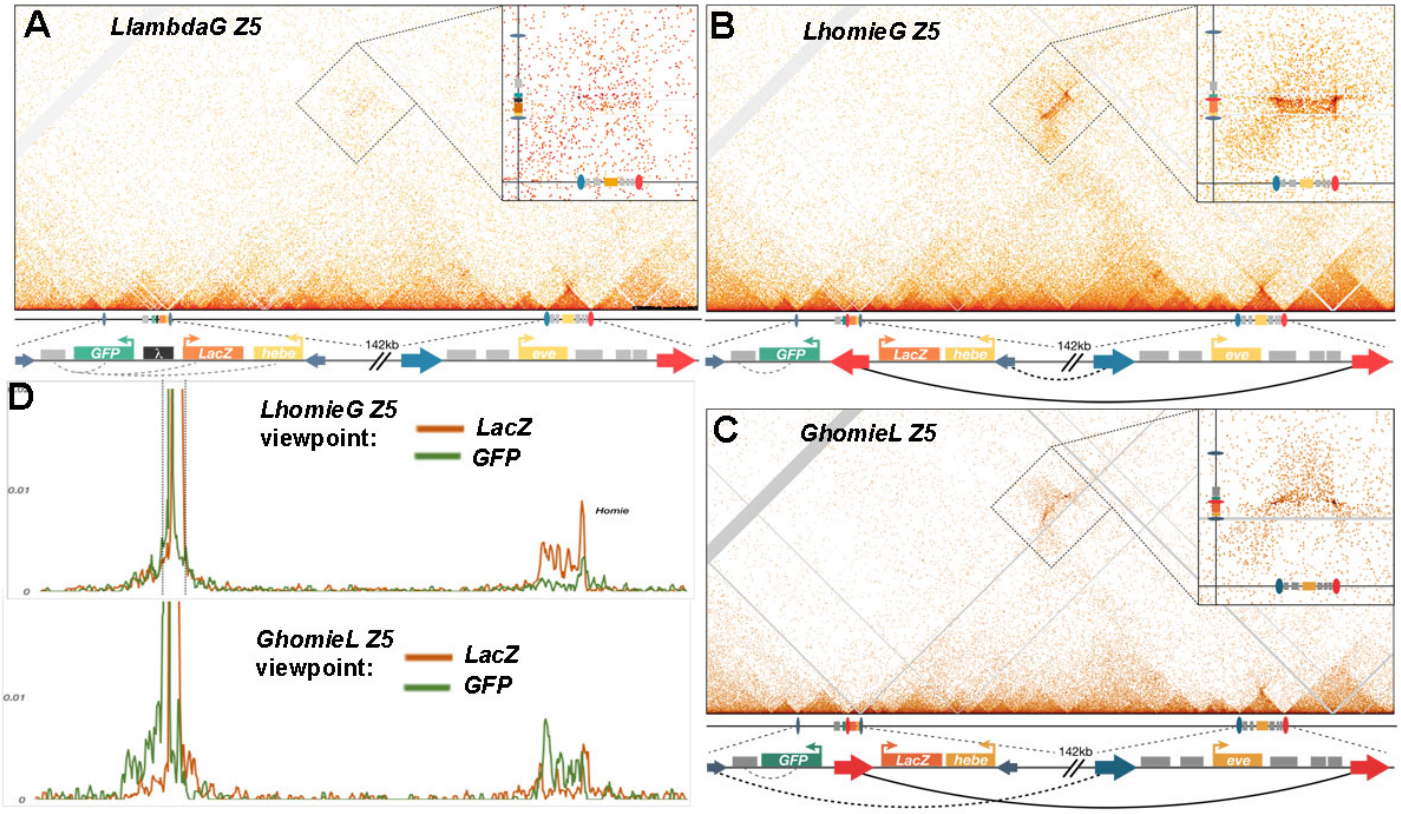
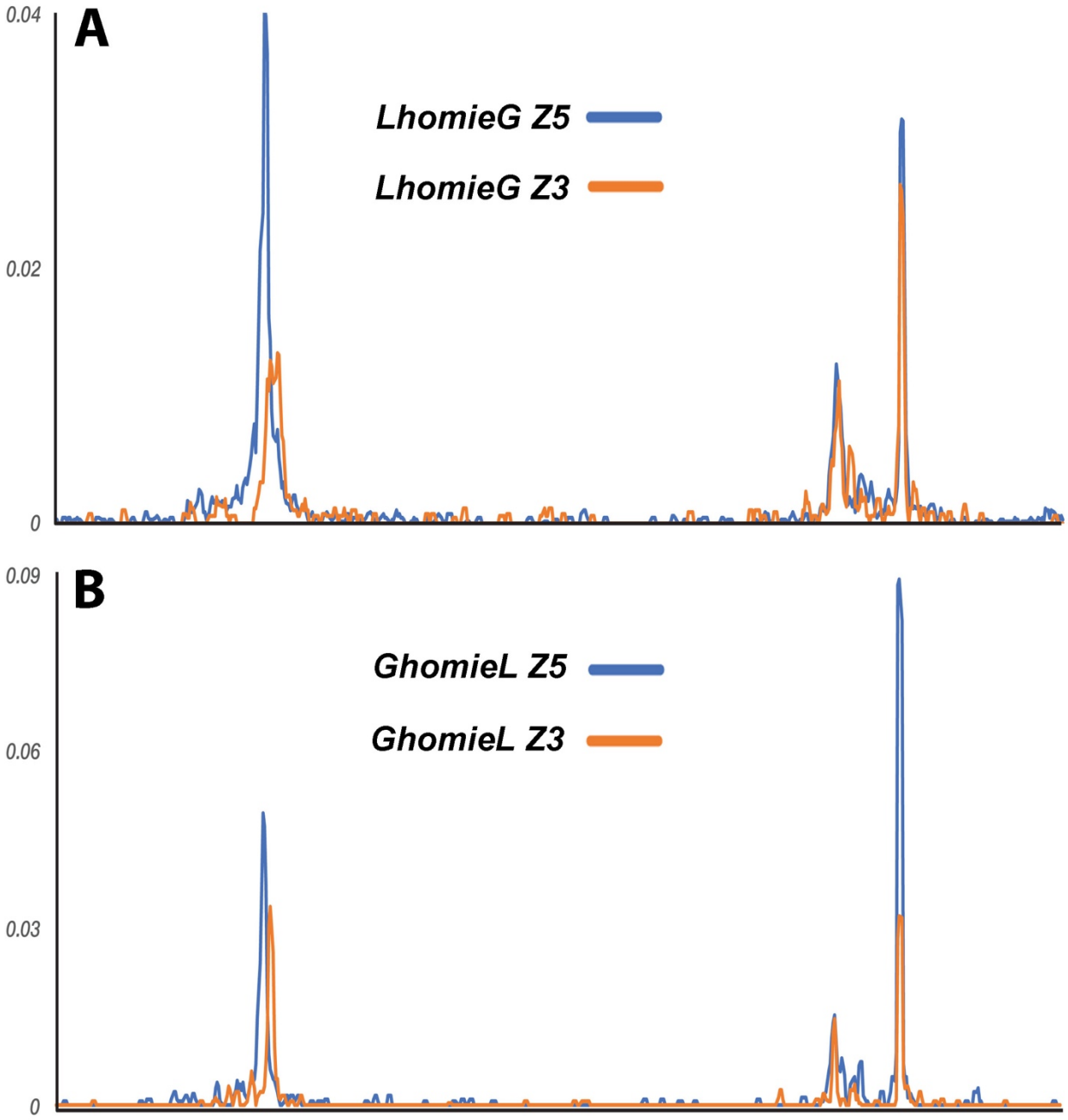


Figure 7. *homie* in Z5 transgene *LhomieG* and *GhomieL* Z3 mediates long-range interactions with the *eve* locus.



1498
1499 **Figure 8. *homie* in the transgene interacts with *eve homie* and *nhomie*.**
1500