

SCIENTIFIC REPORTS



OPEN

Selective interactions between diverse STEs organize the ANT-C Hox cluster

Mo Li, Zhibo Ma, Sharmila Roy, Sapna K. Patel, Derrick C. Lane, Carly R. Duffy & Haini N. Cai

The three-dimensional organization of the eukaryotic genome is important for its structure and function. Recent studies indicate that hierarchies of chromatin loops underlie important aspects of both genomic organization and gene regulation. Looping between insulator or boundary elements interferes with enhancer-promoter communications and limits the spread active or repressive organized chromatin. We have used the SF1 insulator in the *Drosophila Antennapedia* homeotic gene complex (ANT-C) as a model to study the mechanism and regulation of chromatin looping events. We reported previously that SF1 tethers a transient chromatin loop in the early embryo that insulates the Hox gene *Sex comb reduce* from the neighbor non-Hox gene *fushi tarazu* for their independent regulation. To further probe the functional range and connectivity of SF1, we used high-resolution chromosomal conformation capture (3C) to search for SF1 looping partners across ANT-C. We report here the identification of three distal SF1 Tether Elements (STEs) located in the *labial*, *Deformed* and *Antennapedia* Hox gene regions, extending the range of SF1 looping network to the entire complex. These novel STEs are bound by four different combinations of insulator proteins and exhibit distinct behaviors in enhancer block, enhancer-bypass and boundary functions. Significantly, the six STEs we identified so far map to all but one of the major boundaries between repressive and active histone domains, underlining the functional relevance of these long-range chromatin loops in organizing the Hox complex. Importantly, SF1 selectively captured with only 5 STEs out of ~20 sites that display similar insulator binding profiles, indicating that presence of insulator proteins alone is not sufficient to determine looping events. These findings suggest that selective interaction among diverse STE insulators organize the *Drosophila* Hox genes in the 3D nuclear space.

The three-dimensional (3D) organization of the genome is critical for its function including transcription regulation^{1–8}. Recent studies indicate that extensive loop structures underpin the genomic architecture in mammals and *Drosophila*^{3,5,9–14}. Chromatin boundaries, also known as insulators, are DNA-protein complexes originally known to separate and insulate neighboring chromatin domains. Interactions between insulator sites tether chromatin loops, which can block or promote enhancer-promoter interactions^{15–22}. These loops can also impede the spread of silent or active chromatin^{22–25}. Multiple classes of insulator complexes, represented by SuHw, dCTCF, GAF and BEAF proteins, are known to facilitate chromatin looping in *Drosophila*^{14,17,26–34}. Genome-wide distributions of these insulator complexes are partially overlapping and the functions of the loops they tether are largely unknown. Mounting evidence suggests that interactions between insulators can be tissue- and developmental stage-specific, providing a potential mechanism for developmental gene regulation^{22,32,35–38}. However, the organization and regulation of chromatin loops, especially defined at high resolution and in the context of animal development, remain poorly understood.

The *Drosophila* Homeotic/Hox genes are activated in a tissue-specific pattern by numerous enhancers during early embryonic development. Their transcriptional status are then maintained by epigenetic mechanisms mediated by the Polycomb Group (PcG) and Trithorax Group (Trx-G) complexes in late development^{39–43}. As both enhancer- and chromatin-mediated transcriptional regulation are influences by the formation of chromatin loops, the homeotic/Hox complexes provide a good model for elucidating the mechanism and regulation of genomic loops. We have previously identified a hub of chromatin loops anchored by the SF1 insulator in the *Drosophila* Antennapedia homeotic complex (ANT-C, Fig. 1A)²². SF1, located in the regulatory region of the *Sex*

Department of Cellular Biology, University of Georgia, Athens, GA, 30602, USA. Mo Li and Zhibo Ma contributed equally. Correspondence and requests for materials should be addressed to H.N.C. (email: hcai@uga.edu)

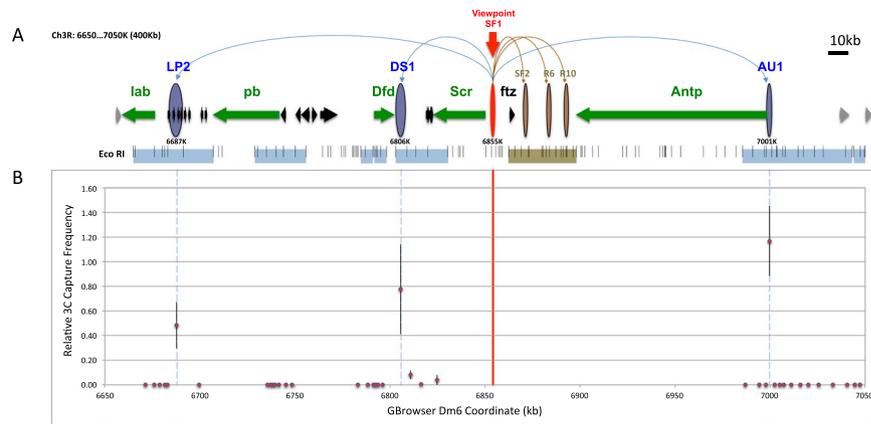


Figure 1. Chromosome Conformation Capture (3C) identification of novel SF1-Tethering-Elements (STEs) in ANT-C. (A) Diagram of ANT-C. Map coordinate is based on the *Drosophila* genome browser dm6 (BDGP R6 Plus ISO1 MT)⁹². Horizontal arrows indicate genes and their transcription direction. Green arrows represent homeotic genes; black and gray arrows represent non-homeotic genes within and outside of ANT-C, respectively. The horizontal bars below genes represent regions tested in the SF1-based 3C captures in the current study (light blue) and a previous study (brown)²². Small vertical lines represent EcoR I sites. Vertical ovals indicate SF1 (red) and STEs identified in the current study (blue) and those from previous work (brown)²². Gbrowser coordinates for the center of the captured fragment is labeled underneath the blue oval. Long curves on top represent looping interactions between SF1 and STEs. (B) Relative SF1-capture frequency of 39 EcoR I fragments in the ANT-C intergenic regions. The coordinates at the center of each fragment is plotted against the average capture frequency over control (for detail see methods)²². Red vertical line represents the position of SF1 and blue dashed lines represent the positions of STEs.

comb reduced (*Scr*) Hox gene, tethers with the SF2 insulator during early embryogenesis (Fig. 1A)²². This creates a chromatin loop that not only encloses and separates the neighbor non-Hox gene *fushi tarazu* (*ftz*) from the surrounding Hox genes, but also reconnects the interrupted *Scr* regulatory region and facilitates *Scr* distal enhancers (Fig. 1A)^{22,34,44–50}. We showed that the transient loop correlates, both in genomic extent and timing of formation, with an active chromatin domain surrounding the *ftz* transcription unit. The loop also coincides with a reduced access between the *ftz* gene and the neighboring Hox enhancers²². We further showed that SF1 also contacts several other local regions between *Scr* enhancers and the *Scr* Polycomb Response Elements (PREs) in mid and late embryogenesis, possibly regulating their access to the *Scr* promoter. Our results suggest that formation of chromatin loops can be a developmentally regulated and play important roles in gene regulation.

These findings raised several questions regarding how ANT-C is organized. For example, is SF1 only a local chromatin organizer within the *Scr-ftz* interval? How many major units, as defined by independent loops are there in ANT-C? Alternatively, can SF1 tether long-range chromatin loops beyond the *Scr-ftz* region? How many different subclasses of tether-elements can SF1 contact? What are the unique functions and *cis*- and *trans*-components of other potential STEs? In the current study we begin to address these questions by scanning for SF1-contacting points outside of the *Scr-ftz* region within ANT-C. Using chromosome conformation capture (3C) we identified three novel STEs near the *labial* (*lab*), *Deformed* (*Dfd*) and *Antennapedia* (*Antp*) Hox genes. The three STEs span the length of ANT-C, ranging from 50 to 167 kb distance from SF1. These findings suggest that SF1 is part of a larger network of chromatin loops that organize the ANT-C Hox cluster. We report that at least four different combinations of insulator proteins associate with these novel STEs, which exhibit distinct behaviors in enhancer block, enhancer bypass and boundary functions. Importantly, SF1 selectively captured with five STEs out of 17 sites that exhibit similar binding profiles for insulator proteins, suggesting that binding by insulator protein alone is insufficient to determine looping events. Further, SF1 and STEs colocalize with all but one of the major domain boundaries between repressive chromatin domains around Hox genes and the less repressive domains around non Hox domains, underscoring the functional relevance of these looping events in organizing the ANT-C Hox complex in *Drosophila* embryos.

Results

Identification of novel SF1-tethering elements (STE's) in the Antennapedia homeotic gene complex. To probe the functional range of SF1, we searched major ANT-C intergenic regions outside of the *Scr-ftz* interval for DNA regions that interact with SF1. To this end, we tested SF1 capture of 39 EcoRI fragments using high-resolution chromosomal conformation capture (3C) in 0–20 hour old embryos (Fig. 1A, see Methods)^{20,51–54}. These elements cover most of the non-transcribed regions in ANT-C (blue horizontal bars, Fig. 1A) outside the *Scr-ftz* interval (brown horizontal bar, Fig. 1A), which was described previously²². Three elements preferentially interact with SF1 (Figs 1B, S1A–C, blue ovals, Fig. 1A). Two of these, LP2 and AU1, are located near the *lab* and *Antp* genes at 184 kb and 146 kb from SF1, respectively. They capture strongly and selectively with SF1, while the surrounding regions exhibited no capture (Figs 1B, S1A,C). This is consistent with previous observations that non-specific captures due to genomic linkage is at a minimum at these distances^{51,55}.

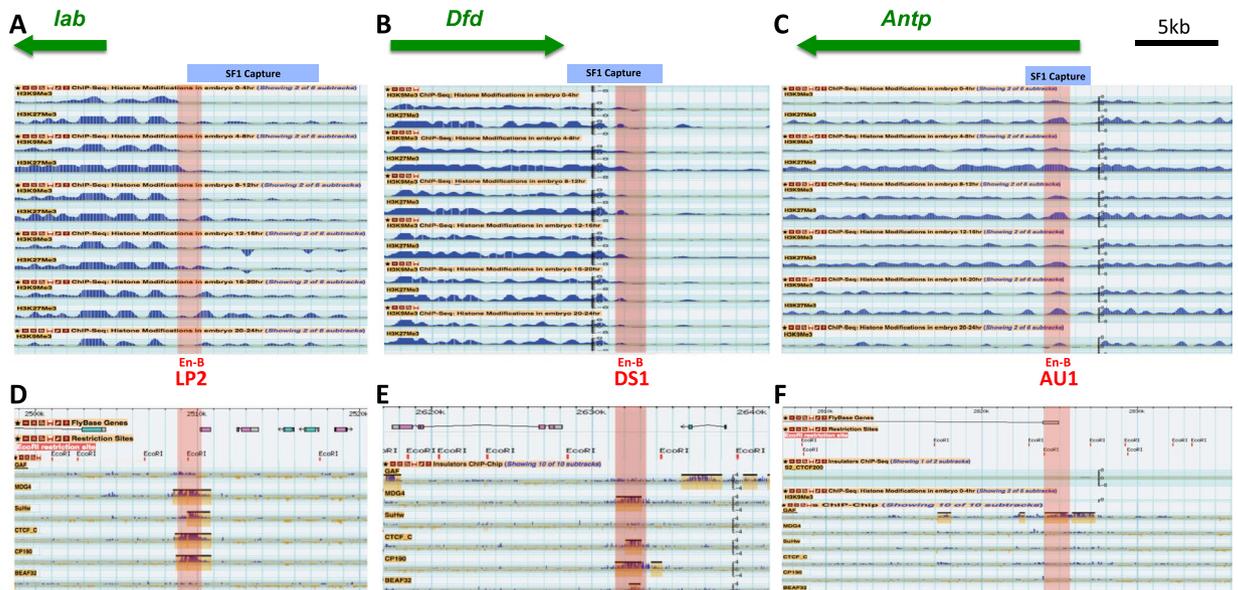


Figure 2. Novel STEs demarcate chromatin domain boundaries and bind to insulator proteins. Top, screen crops of the repressive H3K9Me3 and H3K27Me3 ChIP-seq profiles surrounding STEs in *lab* (A), *Dfd* (B), and *Antp* (C) genomic regions from different embryonic stages. The EcoRI fragments captured by SF1 are indicated by the blue-shaded horizontal bars on top. The sub-fragments containing enhancer-blocking activities (En-B, also see Fig. 3) are indicated by the red shaded vertical bars with their names indicated below. Bottom, screen captures of ChIP-Chip profiles of known insulator proteins surrounding the STEs in 0–12 h embryos⁶⁴. Yellow-shaded boxes represent called peaks for bound proteins. Map coordinate is based on ModEncode GBrowser dm3 (BDGP R5).

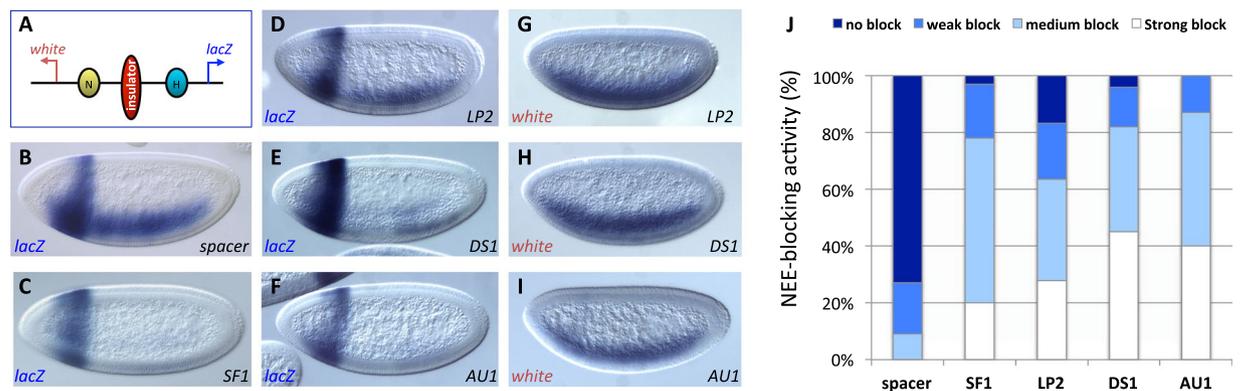


Figure 3. STEs exhibit diverse enhancer-blocking activity in transgenic *Drosophila*. (A) Diagram of the pWNHZ-STE transgene vector containing divergently transcribed *lacZ* (blue arrow) and *miniwhite* (*mw*, red arrow) reporters flanking the NEE (yellow circle) and H1 (blue circle) enhancers (see Methods). Red vertical oval: STEs or spacer control inserted between NEE and H1. (B–J) Representative images of transgenic embryos after whole mount *in situ* hybridized with the anti-*lacZ* (B–F), or the anti-*white* (G–I) RNA probes. DNA elements in these embryos are labeled at bottom right of each photo and probed used on bottom left. Embryos are shown in sagittal views with anterior to the left and dorsal up. (J) Quantitation of NEE-blocking in the whole neuroectoderm in transgenic embryos stained with the anti-*lacZ* probe (see Methods for details).

DS1, the third region captured by SF1, is located immediately downstream of the *Dfd* and *Scr* transcription units, approximately 50 kb from SF1 (Figs 1B, S1B). Since genomic elements within such distances are known to capture at relatively high frequencies, we applied a distance-frequency reference curve for the *Drosophila* genome to evaluate the significance of capture (dotted curve, Fig. S1B)²². Of the four fragments in the *Dfd-Scr* intergenic region, the EcoRI element at 6806K captured with SF1 at a frequency significantly above the expected value ($P < 0.0016$), whereas the other three fragments captured SF1 did not. Negative controls without cross-linking or without ligase yielded no capture products as expected (not shown). The identification of these novel STEs in the distal regions of ANT-C extends the SF1 contact range beyond the local *Scr-ftz* region, revealing a long-range chromatin loop network that could organize the entire Hox cluster.

Novel STEs colocalize with chromatin boundaries that separate distinct histone modification domains. Insulator-tethered chromatin loops are known to impede the spread of silent or active chromatin and therefore often correspond to boundaries between domains of distinct histone modification marks^{22,23,56–60}. For example, SF1 and its local partner SF2 flank an active chromatin domain that encloses the *ftz* gene and separate it from the surrounding silent chromatin in the early *Drosophila* embryo²². To probe whether the newly identified STEs correspond to such chromatin domain boundaries, we examined the histone modification profiles surrounding these elements in the modENCODE database^{61–63}. We found that the STEs in the *lab* and *Dfd* regions indeed colocalize with transition points between active and repressive chromatin in the *Drosophila* embryos (Fig. 2A,B). The ~8-kb *lab* STE is located ~5 kb upstream of the *lab* promoter, separating it from multiple small non-Hox genes (Fig. 2A). The ~5.5-kb *Dfd* STE is located ~3 kb downstream of the *Dfd* gene, separating it from a series of small tRNA and non-coding RNA genes (Fig. 2B). In both cases, the STEs (blue shaded bar on top) contain the border between the Hox gene domains with higher levels of repressive H3K9me3 and H3K27me3 marks, and the non-Hox gene regions that are depleted of these marks (Fig. 2A,B). In contrast to these two STEs, the *Antp* STE does not clearly demarcate a boundary between different histone modification domains. The repressive histone marks appear to be enriched to a comparable level on both sides of the STE (Fig. 2C). It is possible that the *Antp* STE separate distinct chromatin domains only in selected tissues. It is also possible that loops formed at the *Antp* STE provide a different genomic or regulatory function.

To understand the different boundary behaviors exhibited by these new STEs, we further examined their binding profiles by known insulator/boundary proteins in the *Drosophila* embryo using the ModENCODE chromatin immunoprecipitation (ChIP) database^{61,63,64}. All three new STEs are bound by insulator/boundary proteins, although each with a different combination. For example, both *lab* and *Dfd* STEs are bound by dCTCF, CP190, and Mod (mdg4) (Fig. 2D,E). However, the *lab* STE is also bound by SuHw, whereas the *Dfd* STE is not. In comparison, the *Antp* STE is bound only by GAF, a zinc finger/BTB domain insulator protein that also plays a role in transcription activation and PcG and TrxG mediated functions (Fig. 2F)^{65–71}. The difference in the insulator complexes assembled on these STEs may determine their boundary/barrier behavior. The binding profiles of the new distal STEs are reminiscent of those of SF1 and its local STEs (Fig. S1)²². For example, SF1 and its local partner SF2 are bound by dCTCF, CP190 and Mod (mdg4). Although SF1 does not directly bind to GAF, R10, an STE located at the end of the *Antp* transcription unit, binds strongly to GAF. These findings suggest that SF1 can contact genomic loci bound by different insulator complexes and possibly serve different functions.

Novel STEs contain enhancer-blocking activities in transgenic *Drosophila*. An important function of chromatin insulators is to modulate enhancer-promoter communications^{16,47,72–77}. The new STEs are located near Hox regulatory elements and may be involved in organizing enhancer traffic. We tested new STEs for their insulator function using an enhancer-blocking assay in transgenic *Drosophila* embryos^{22,34,78–80}. The assay transgene contains divergently transcribed *white* and *lacZ* reporters driven by two tissue-specific enhancers, the Neuroectoderm Enhancer (NEE) and the Hairy Stripe 1 enhancer (H1, Fig. 3A). Known insulators such as SF1 but not a neutral spacer, when inserted between the two enhancers, can block the distal NEE from the *lacZ* reporter, reducing its expression in the horizontal neuroectoderm, without affecting the H1 driven *lacZ* expression in the head (Fig. 3B,C, see quantitation in Fig. 3J)^{22,34,78–80}. We first dissected the ~9-kb *lab* STE in the transgenic insulator assay. The proximal end of the EcoRI fragment overlaps with ChIP-seq peaks for multiple insulator proteins (Fig. 2D). We tested a ~1.4-kb fragment, named LP2, that includes the major peaks of insulator binding and found it to contain significant enhancer-blocking activity (Fig. 3D, also see quantitation in Fig. 3J). To confirm that the loss of the *lacZ* expression is not due to a silencing effect from the insulator, we also examined the divergently transcribed mini-*white* reporter. We found a reduction of H1-driven *white* expression in the head but not NEE-driven *white* expression in the ventral lateral region (Fig. 3G). This result indicates that LP2 contains enhancer-blocking/insulator activity. We then examined the STE fragments from the *Dfd* and *Antp* regions. Each of these regions contains a single enhancer-blocking sub-fragment, named DS1 and AU1, respectively (Figs 2B and 3E–J). In both cases, the enhancer-blocking fragment colocalizes with the DNA region bound by insulator proteins (Fig. 2E,F). Our enhancer-blocking results indicate that all three distal STEs function as autonomous insulators, independent of their native genomic context and away from SF1. It is possible that the STEs tether chromatin loops with insulator complexes near the insertion sites to block transgenic enhancers. Previous studies indicate that enhancer-blocking and chromatin-blocking function of a boundary element may be mediated by distinct mechanisms and separable^{29,76,77}. The differential boundary behaviors exhibited by the STEs further suggest that these activities could be provided by different insulator complexes.

SF1-STE pairing can facilitate enhancer-bypass. We further tested whether the enhancer-blocking activity of the new STEs is cancelled by their tandem arrangement with SF1. Tandem pairing between certain insulators are known to neutralize their enhancer-blocking function^{16,47}. This “enhancer bypass” phenomenon suggests that the formation of chromatin loops is necessary for enhancer blocking^{16,47,71,81–86}. It is also an indication that the two insulators interact with each other directly. To this end, each of the new STEs was placed in tandem with SF1 and inserted between the NEE and the H1 enhancers (Fig. 4A). Transgenic embryos were examined for blocking of NEE enhancer on the *lacZ* reporter gene. For AU1 and DS1, we observed a strong recovery of NEE-driven *lacZ* expression in transgenic embryos, suggesting that pairing of SF1 and these two STEs neutralized their enhancer-blocking function (Fig. 4B,C,H, compared to Fig. 3E,F,J). However, we observed no enhancer-bypass with the LP2-SF1 pairing (Fig. 4D,H). We have recently found that for certain SF1-STE pairings, enhancer-bypass occurs in an orientation-dependent fashion. For example, SF1-SF2 pairing in a forward-forward (SF1-SF2-ff or SF2-SF1-ff) arrangement was shown to mediate enhancer-bypass (Fig. 4E,H)²². However, when SF2 is placed in an inverted orientation relative to SF1 (SF1-SF2-fr), bypass was not observed (Fig. 4F,H). In fact, an augmentation of insulator activity beyond that of SF1 or SF2 alone was observed (Figs 3J and 4H)²². This result

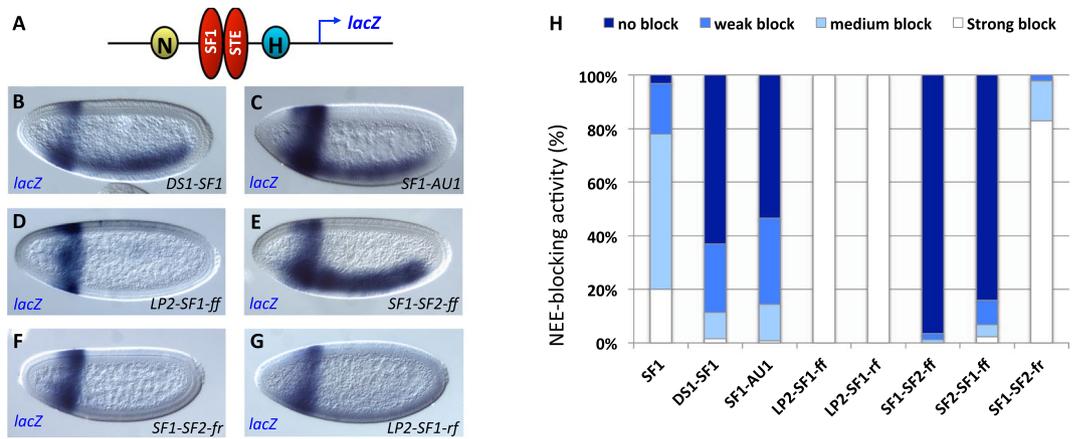


Figure 4. SF1-STE pairing mediates enhancer bypass. (A) Diagram of the enhancer-bypass transgene containing SF1 and STE placed in tandem between NEE and H1 in the pWNHZ vector (red ovals, also see Methods). (B–G) Representative images of transgenic embryos after whole mount *in situ* hybridization with anti-*lacZ* probes. DNA elements in these embryos are labeled at bottom right of each photo and probes used on bottom left. Orientation of the elements relative to the endogenous arrangement is indicated by “f” (forward) or “r” (reverse). (H) Quantitation of NEE-blocking in by tandem insulators in bypass transgenic embryos (see Methods for details).

suggests that formation of chromatin loops by insulator pairing may be affected by the relative orientation of the DNA sequences and/or protein complexes. Based on these findings, we also tested LP2-SF1 pairing in opposite orientation (LP2-SF1-rf). As shown in Fig. 4, inverting LP2 did not lead to enhancer-bypass, but rather an augmentation of NEE block (Fig. 4F,J, compare to Fig. 3D,J). The inability to cancel with SF1 might be due to the unique insulator complexes bound at LP2. Although LP2, as SF2 and DS1, is bound by dCTCF, CP190 and Mod (mdg4), it is also bound by SuHw whereas SF2 and DS1 are not. This difference might contribute to its inability to cancel with SF1. Certain insulators are known to not bypass when placed in tandem and this was attributed to their not interacting with each other *in vivo*⁷⁷. Since LP2 was identified through its capture with SF1, it is possible that SF1 and LP2 do not directly interact with each other, and that their capture in 3C could be mediated by mutual partners.

Discussion

Chromatin looping plays essential roles in gene regulation as it modulates both enhancer-promoter communications and the extent of active or repressed chromatin^{16–25,87}. However, high-resolution characterization of chromatin loops, especially in the context of animal development, has been rare^{22,88,89}. We have dissected the genomic loops tethered by the SF1 insulator at high resolution to probe their function in the *Drosophila* Hox cluster ANT-C. Our previous studies indicate that SF1 organizes a series of local chromatin loops in the *Scr-ftz* interval to maintain independent regulation of *Scr* and *ftz*²². These findings raised further questions on SF1’s functional range of and its level of connectivity. To address these questions, we probed for additional STEs in ANT-C. We identified three distal STEs, LP2, DS1 and AU1 in the *labial*, *deformed* and *Antp* gene regions, extending the SF1’s contact range beyond the local *Scr-ftz* region to include the full extent of ANT-C (red vertical ovals, Fig. 5). Although the new distal STEs bind to diverse insulator proteins they all exhibit enhancer-blocking activity in transgenic assays, an indication that they tether chromatin loops autonomously and independent of their native genomic context. This is in contrast to some proximity STEs located in the *Scr-ftz* region, such as R2 and R6, which do not all contain constitutive enhancer-blocking activity²².

Of the three new STEs, both LP2 and DS1 are located near the end of a Hox transcriptional unit, separating the Hox region from a neighboring non-Hox gene region. Both are bound by dCTCF, CP190 and Mod (mdg4), and both insulators colocalize with boundaries between domains of repressive histone marks associated with Hox genes and the less repressive non-Hox domains (red vertical lines, Fig. 5C). In these respects, LP2 and DS1, as well as previously described SF1, SF2 and R10, belong to the “Class I insulators” as defined by the ChIP-seq profile of insulator proteins (orange triangles, Fig. 5)⁶⁴. Intriguingly, out of 13 Class-I insulator protein ChIP sites in the ANT-C intergenic regions, SF1 only captures with four (Fig. 5A,B). This suggests that besides binding of known insulator complexes, other unknown factors may help determining the selectivity of looping events. Further, we showed that DS1 or AU1, when placed in tandem arrangement with SF1, both can mediate enhancer-bypass, whereas LP2 does not. Enhancer-bypass provides evidence that the pairing insulators can interact directly with each other, precluding both partners from looping with other sites. Therefore, inability of the LP2-SF1 pairing to bypass enhancers suggests that the two insulators might not directly contact each other. In this light, their capture in our 3C experiment could result from both elements interacting with a mutual partner simultaneously. Alternatively, since LP2 is the only STE that also interacts with the SuHw insulator protein (Class II), it is possible that the unique combination of insulator complexes interferes with enhancer-bypass (Fig. 5).

The remaining STE, AU1, is bound by GAF, a multi-faceted insulator factor also involved in transcription activation and PcG and Trx-G mediated chromatin organization^{65–69,71,90}. As observed above with the Class I insulators, SF1 selectively interacts with R10 and AU1 among multiple other GAF sites present in the ANT-C

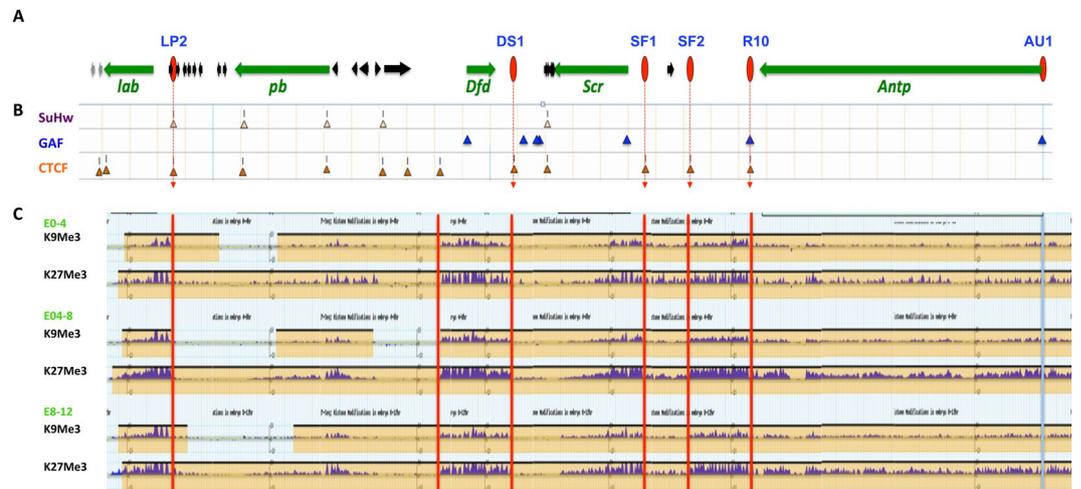


Figure 5. Selective interactions between diverse STEs organize the ANT-C Hox cluster. **(A)** ANT-C genomic map. Horizontal arrows represent Hox (green) and non-Hox (black and grey) genes with gene names below. Red ovals and red dash lines indicate SF1 and STE insulators. **(B)** Distribution of major insulator proteins across ANT-C. ChIP-seq peaks of SuHw (class II), GAF, and dCTCF (class I) insulator proteins are indicated by yellow, blue and orange triangles, respectively^{64,92}. **(C)** Repressive histone domains in ANT-C during *Drosophila* development. Screen crop of ChIP-seq profiles of H3K9Me3 and H3K27Me3 in 0–4, 4–8 and 8–12 hour old *Drosophila* embryos (<http://www.genome.gov/modENCODE>). Red vertical lines indicate the locations of boundaries between distinct chromatin domains. Light blue vertical line indicates the absence of a clear domain boundary at AU1.

intergenic regions. AU1 interacts exclusively with GAF and does not clearly demarcate a boundary between distinct histone modification domains (Fig. 5). These suggest that although GAF insulators can tether chromatin loops, mediate enhancer block and enhancer bypass, they do not possess an intrinsic domain boundary or barrier activity²⁹. One other major GAF sites in the ANT-C intergenic regions is R10, an STE located at the other end of the ~110-kb *Antp* transcription unit. Together with AU1, the two GAF insulators may play unique roles in modulating *Antp* enhancer traffic, rather than separating silent and active chromatin domains. Previous studies indicate that enhancer-blocking and the barrier function of a boundary element may be mediated by distinct mechanisms^{29,76,77}. Our results support the notion that distinct insulator complexes, or combination thereof, dictate unique insulator and boundary behavior and functions.

Using SF1 as the 3C viewpoint, we have now identified a total of six major insulator/boundary elements across the ANT-C Hox complex (Fig. 5A). SF1 selectively captured with these five STEs out of ~20 sites that exhibit similar binding profiles for insulator complexes, suggesting that binding by insulator proteins alone is insufficient to determine looping events. Further, SF1 and STEs occupy all but one of the major domain boundaries between repressive chromatin domains around Hox genes and the less repressive domains around non Hox domains in ANT-C, underscoring the functional relevance of these looping events in organizing the Hox complex in *Drosophila* embryos (Fig. 5B). Use of complementary approaches including transgenic enhancer-block and enhancer-bypass assays provided unique and novel insights on the properties of the STEs and the mechanisms that govern their long-range interactions. Our results highlight the importance of defining chromatin loops in high resolution and with higher specificity. Future work should identify tissue and developmentally regulated factors that dictate the formation of these loops during animal development.

Materials and Methods

Chromosome Conformation Capture (3C). The 3C experiments were performed according to published protocols and as in our previous report^{22,51}. All captures were repeated minimally three times (biological replicates). Approximately 3×10^7 nuclei were collected from 0–20 hour old embryos and used in the chromatin preparation according to existing protocols^{17,22}. The optimal quantity of template DNA used in PCR reactions was determined empirically by serial dilutions. Briefly, 3C samples were amplified for 20–22 cycles with the outer primer pair (Table S1). Five to ten percent of the outer PCR reaction was amplified with nested inner primers. Capture products were then fractionated on agarose gels and digitally imaged. Quantitation and analysis were done using the Image-J software. All primers were designed to be ~100–150 bp from the restriction site so that all capture products are comparable in size. To generate the control template, purified fly genomic DNA was digested with Eco RI or other restriction enzymes and ligated at a concentration of ~500 ng/ μ l. The frequencies of capture expressed as Relative Crosslinking, $\text{PCR}_E/\text{PCR}_C$, is generally plotted over distance⁹¹. The distance-capture frequency curves were generated using data from ~500 captures between sites with linear distance up to ~200 kb apart, of which the relevant distance range were shown in Fig. 1C. We also generated separate curves for conventional and quantitative PCRs to control for the data range. Statistical analysis and charts were made using the Microsoft Excel and JMP programs. The p-values were calculated using LSMeans Differences Student's t-test.

Enhancer-blocking assay in transgenic *Drosophila* embryos. Enhancer-blocking assay, including spacer- and SF1-containing transgenic *Drosophila* lines were described previously^{22,34,77,78}. Individual STE sub-fragments were cloned by PCR (see Table S2 for primers), purified after further digestions, and inserted into the Not I site between the NEE and H1 enhancers in pWNHZ vector (Fig. 3A). For transgene containing tandem SF1 and STE, the two insulators were inserted in the same relative position and orientation as in their native genomic loci (Fig. 4A). For transgenes containing LP2, LP2-SF1-ff, LP2-SF1-rf, SF1-SF2-ff, SF2-SF1-ff, and SF1-SF2-fr, genomic integration at the VK33 attP site were mediated by phiC-31 site-specific insertion. For these transgenes, a phiC31 attB site was inserted at the Nsi I site downstream of the *miniwhite* gene (Fig. 3D,G,J). Microinjections were performed in the Cai lab or by Rainbow Transgene (Camarillo, CA). *In situ* hybridization with *lacZ*, *white* anti-sense RNA probes were performed as previously described⁴⁷. Whole mount *in situ* hybridization and visual assessments of reporter expression were performed according to existing procedures^{34,78}. For each pWNHZ transgene, 50–100 embryos were scored double-blindly from at least three independent insertion lines. Briefly, blastoderm stage embryos were scored for *lacZ* level in the NEE domain against the H1 domain (Figs 3B and 4B). Based on the ratio of the *lacZ* in H1/NEE domains, embryos were ranked into four categories from no-block: $H/N \leq 1$, weak block: $H/N = 2$, medium block: $H/N = 4$, and strong block: $H/N \geq 5$.

Data Availability Statement

The authors will make all data available upon publication of the manuscript.

References

- Lancôt, C., Cheutin, T., Cremer, M., Cavalli, G. & Cremer, T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nature Reviews Genetics* **8**, 104–115 (2007).
- Misteli, T. Beyond the sequence: cellular organization of genome function. *Beyond the sequence: cellular organization of genome function* (2007).
- Pombo, A. & Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nature Reviews Molecular Cell Biology* **16** (2015).
- Boney, B. & Cavalli, G. Organization and function of the 3D genome. *Nature reviews. Genetics* **17**, 661–678 (2016).
- Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).
- Krijger, P. & de Laat, W. Regulation of disease-associated gene expression in the 3D genome. *Nature Reviews Molecular Cell Biology* **17** (2016).
- Uhler, C. & Shivashankar, G. V. Regulation of genome organization and gene expression by nuclear mechanotransduction. *Nature Reviews Molecular Cell Biology* **18** (2017).
- Rennie, S., Dalby, M., van Duin, L. & Andersson, R. Transcriptional decomposition reveals active chromatin architectures and cell specific regulatory interactions. *Nature communications* **9**, 487 (2018).
- Dean, A. In the loop: long range chromatin interactions and gene regulation. *Brief Funct Genomics* **10**, 3–10 (2011).
- Murrell, A. Setting up and maintaining differential insulators and boundaries for genomic imprinting. *Biochem Cell Biol* **89**, 469–478 (2011).
- Yang, J. & Corces, V. G. Chromatin insulators: a role in nuclear organization and gene expression. *Adv Cancer Res* **110**, 43–76 (2011).
- Holwerda, S. J. & de Laat, W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120369 (2013).
- Maksimenko, O. & Georgiev, P. Mechanisms and proteins involved in long-distance interactions. *Front Genet* **5**, 28 (2014).
- Mizuguchi, T. *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* **516**, 432–435 (2014).
- Murrell, A., Heeson, S. & Reik, W. Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nature Genetics* **36** (2004).
- Muravyova, E. *et al.* Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science* **291**, 495–498 (2001).
- Blanton, J., Gaszner, M. & Schedl, P. Protein:protein interactions and the pairing of boundary elements *in vivo*. *Genes Dev* **17**, 664–675 (2003).
- Horike, S., -I., Cai, S., Miyano, M., Cheng, J. -F. & Kohwi-Shigematsu, T. Loss of silent-chromatin looping and impaired imprinting of *DLX5* in Rett syndrome. *Nature Genetics* **37** (2004).
- Yusufzai, T. M., Tagami, H., Nakatani, Y. & Felsenfeld, G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* **13**, 291–298 (2004).
- Simonis, M. *et al.* Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* **38**, 1348–1354 (2006).
- Fujioka, M., Sun, G. & Jaynes, J. B. The *Drosophila* eve insulator Homie promotes eve expression and protects the adjacent gene from repression by polycomb spreading. *PLoS Genet* **9**, e1003883 (2013).
- Li, M. *et al.* An Organizational Hub of Developmentally Regulated Chromatin Loops in the *Drosophila* Antennapedia Complex. *Mol Cell Biol* **35**, 4018–4029 (2015).
- Zhao, K., Hart, C. M. & Laemmli, U. K. Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* **81**, 879–889 (1995).
- Donze, D., Adams, C. R., Rine, J. & Kamakaka, R. T. The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. *Genes Dev* **13**, 698–708 (1999).
- Byrd, K. & Corces, V. G. Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *J Cell Biol* **162**, 565–574 (2003).
- Moon, H. *et al.* CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep* **6**, 165–170 (2005).
- Mohan, M. *et al.* The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *Embo J* **26**, 4203–4214 (2007).
- Bushey, A. M., Ramos, E. & Corces, V. G. Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev* **23**, 1338–1350 (2009).
- Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
- Bell, A. C. & Felsenfeld, G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482–485 (2000).
- Hark, A. T. *et al.* CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* **405**, 486–489 (2000).
- Splinter, E. *et al.* CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* **20**, 2349–2354 (2006).

33. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
34. Belozerov, V. E., Majumder, P., Shen, P. & Cai, H. N. A novel boundary element may facilitate independent gene regulation in the Antennapedia Complex of *Drosophila*. *EMBO J.* **22**, 3113–3121 (2003).
35. de Almeida, C. *et al.* Critical Role for the Transcription Regulator CCCTC-Binding Factor in the Control of Th2 Cytokine Expression. *The Journal of Immunology* **182**, 999–1010 (2009).
36. Kehayova, P., Monahan, K., Chen, W. & Maniatis, T. Regulatory elements required for the activation and repression of the protocadherin- α gene cluster. *Proceedings of the National Academy of Sciences* **108**, 17195–17200 (2011).
37. Aoki, T., Sarkeshik, A., Yates, J. & Schedl, P. Elba, a novel developmentally regulated chromatin boundary factor is a hetero-tripartite DNA binding complex. *eLife* **1** (2012).
38. Ikeda, H., Sone, M., Yamanaka, S. & Yamamoto, T. Structural and spatial chromatin features at developmental gene loci in human pluripotent stem cells. *Nature Communications* **8** (2017).
39. Beck, S., Faradji, F., Brock, H. & Peronnet, F. In *Hox Genes: Studies from the 20th to the 21st Century* (ed Jean S. Deutsch) 41–62 (Springer New York, 2010).
40. Lanzuolo, C. & Orlando, V. Memories from the polycomb group proteins. *Annu Rev Genet* **46**, 561–589 (2012).
41. Pirrotta, V. Polycomb: the genome. PcG, trxG, and chromatin silencing. *Cell* **93**, 333–336 (1998).
42. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38**, 413–443 (2004).
43. Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
44. Gindhart, J. G. Jr. & Kaufman, T. C. Identification of Polycomb and trithorax group responsive elements in the regulatory region of the *Drosophila* homeotic gene *Sex combs reduced*. *Genetics* **139**, 797–814 (1995).
45. Gindhart, J. G. Jr, King, A. N. & Kaufman, T. C. Characterization of the cis-regulatory region of the *Drosophila* homeotic gene *Sex combs reduced*. *Genetics* **139**, 781–795 (1995).
46. Gorman, M. J. & Kaufman, T. C. Genetic analysis of embryonic cis-acting regulatory elements of the *Drosophila* homeotic gene *sex combs reduced*. *Genetics* **140**, 557–572 (1995).
47. Cai, H. N. & Shen, P. Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science* **291**, 493–495 (2001).
48. Southworth, J. W. & Kennison, J. A. Transvection and Silencing of the *Scr* Homeotic Gene of *Drosophila melanogaster*. *Genetics* **161**, 733–746 (2002).
49. Calhoun, V. C., Stathopoulos, A. & Levine, M. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila* Antennapedia complex. *Proc Natl Acad Sci USA* **99**, 9243–9247 (2002).
50. Calhoun, V. C. & Levine, M. Long-range enhancer-promoter interactions in the *Scr*-*Antp* interval of the *Drosophila* Antennapedia complex. *Proc Natl Acad Sci USA* **100**, 9878–9883 (2003).
51. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
52. Mito, Y., Henikoff, J. G. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**, 1090–1097 (2005).
53. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* **16**, 1299–1309 (2006).
54. Belton, J. M. *et al.* Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods* (2012).
55. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
56. Kellum, R. & Schedl, P. A group of *scs* elements function as domain boundaries in an enhancer-blocking assay. *Mol Cell Biol* **12**, 2424–2431 (1992).
57. Van Bortle, K. *et al.* *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res* **22**, 2176–2187 (2012).
58. Li, G. & Zhou, L. Genome-Wide Identification of Chromatin Transitional Regions Reveals Diverse Mechanisms Defining the Boundary of Facultative Heterochromatin. *PLoS ONE* **8** (2013).
59. Bowman, S. K. *et al.* H3K27 modifications define segmental regulatory domains in the *Drosophila* bithorax complex. *Elife* **3**, e02833 (2014).
60. Stadler, M. R., Haines, J. E. & Eisen, M. B. Convergence of topological domain boundaries, insulators, and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early *Drosophila melanogaster* embryo. *eLife* **6** (2017).
61. Contrino, S. *et al.* modMine: flexible access to modENCODE data. *Nucleic acids research* **40**, 8 (2012).
62. Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**, 1813–1831 (2012).
63. Celniker, S. E. *et al.* Unlocking the secrets of the genome. *Nature* **459**, 927–930 (2009).
64. Negre, N. *et al.* A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* **6**, e1000814 (2010).
65. Lu, Q., Wallrath, L. L., Granok, H. & Elgin, S. C. (CT) n (GA) n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the *Drosophila* *hsp26* gene. *Mol Cell Biol* **13**, 2802–2814 (1993).
66. Becker, P. B. *Drosophila* chromatin and transcription. *Semin Cell Biol* **6**, 185–190 (1995).
67. Chopra, V. *et al.* Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3. *Developmental Biology* **317**, 660–670 (2008).
68. Lomaev, D. *et al.* The GAGA factor regulatory network: Identification of GAGA factor associated proteins. *PLOS ONE* **12** (2017).
69. Bejarano, F. & Busturia, A. Function of the Trithorax-like gene during *Drosophila* development. *Dev Biol* **268**, 327–341 (2004).
70. Ohtsuki, S., Levine, M. & Cai, H. N. Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev* **12**, 547–556 (1998).
71. Melnikova, L. *et al.* Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proc Natl Acad Sci USA* **101**, 14806–14811 (2004).
72. Ameres, S. L. *et al.* Inducible DNA-loop formation blocks transcriptional activation by an SV40 enhancer. *Embo J* **24**, 358–367 (2005).
73. Bondarenko, V. A., Jiang, Y. I. & Studitsky, V. M. Rationally designed insulator-like elements can block enhancer action *in vitro*. *Embo J* **22**, 4728–4737 (2003).
74. Scott, K. C., Taubman, A. D. & Geyer, P. K. Enhancer blocking by the *Drosophila* gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* **153**, 787–798 (1999).
75. Singh, V. & Srivastava, M. Enhancer Blocking Activity of the Insulator at H19-ICR Is Independent of Chromatin Barrier Establishment. *Molecular and Cellular Biology* **28**, 3767–3775 (2008).
76. Recillas-Targa, F. *et al.* Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci USA* **99**, 6883–6888 (2002).
77. Majumder, P. & Cai, H. N. The functional analysis of insulator interactions in the *Drosophila* embryo. *Proc Natl Acad Sci USA* **100**, 5223–5228 (2003).
78. Cai, H. N. & Levine, M. The gypsy insulator can function as a promoter-specific silencer in the *Drosophila* embryo. *EMBO J* **16**, 1732–1741 (1997).

79. Majumder, P. *et al.* Diverse transcription influences can be insulated by the *Drosophila* SF1 chromatin boundary. *Nucleic Acids Res* (2009).
80. Li, M., Belozherov, V. E. & Cai, H. N. Modulation of chromatin boundary activities by nucleosome-remodeling activities in *Drosophila melanogaster*. *Mol Cell Biol* **30**, 1067–1076 (2010).
81. Morris, J. R., Chen, J. L., Geyer, P. K. & Wu, C. T. Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. *Proc Natl Acad Sci USA* **95**, 10740–10745 (1998).
82. Chen, J. L. *et al.* Enhancer action in trans is permitted throughout the *Drosophila* genome. *Proc Natl Acad Sci USA* **99**, 3723–3728 (2002).
83. Kuhn, E. J., Viering, M. M., Rhodes, K. M. & Geyer, P. K. A test of insulator interactions in *Drosophila*. *Embo J* **22**, 2463–2471 (2003).
84. Golovnin, A. *et al.* Integrity of the Mod(mdg4)-67.2 BTB domain is critical to insulator function in *Drosophila melanogaster*. *Mol Cell Biol* **27**, 963–974 (2007).
85. Kyrchanova, O., Toshchakov, S., Parshikov, A. & Georgiev, P. Study of the functional interaction between Mcp insulators from the *Drosophila* bithorax complex: effects of insulator pairing on enhancer-promoter communication. *Mol Cell Biol* **27**, 3035–3043 (2007).
86. Kyrchanova, O., Chetverina, D., Maksimenko, O., Kullyev, A. & Georgiev, P. Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements. *Nucleic Acids Res* **36**, 7019–7028 (2008).
87. Lewis, A. & Murrell, A. Genomic imprinting: CTCF protects the boundaries. *Curr Biol* **14**, R284–286 (2004).
88. Kang, Y., Kim, Y., Kang, J., Yun, W. & Kim, A. Erythroid specific activator GATA-1-dependent interactions between CTCF sites around the β -globin locus. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1860**, 416–426 (2017).
89. Tchurikov, N. A. *et al.* Genome-wide mapping of hot spots of DNA double-strand breaks in human cells as a tool for epigenetic studies and cancer genomics. *Genomics Data* **5**, 89–93 (2015).
90. Ohtsuki, S. & Levine, M. GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. *Genes Dev* **12**, 3325–3330 (1998).
91. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**, 1453–1465 (2002).
92. dos Santos, G. *et al.* FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res* **43**, D690–697 (2015).

Author Contributions

Mo Li: conception of intellectual ideas, experimental design, performing 3C capture, performing enhancer block assay, manuscript writing Zhibo Ma: conception of intellectual ideas, experimental design, performing enhancer block assay, critical reading of manuscript Sharmila Roy: performing enhancer block assay Sapna K Patel: performing enhancer block assay Derrick C. Lane: performing enhancer block assay Carly R. Duffy: performing enhancer block assay Haini N. Cai: conception of intellectual ideas, experimental design, manuscript writing.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-33588-4>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018