Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in *Drosophila*

Frédéric Bantignies,1,3,4 Virginie Roure,1,3,4 Irys Comet,1,5 Benjamin Leblanc,1 Bernd Schuettengruber,1 Jérôme Bonnet,1,6 Vanessa Tixier,1,7 André Mas,2 and Giacomo Cavalli1,*

1Institut de Génétique Humaine, CNRS UPR 1142, 141, rue de la Cardonille, 34396 Montpellier Cedex 5, France
2Institut de Modélisation Mathématique de Montpellier, CNRS UMR 5149, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France
3These authors contributed equally to this work
4Present address: Max-Planck Institute of Immunobiology, Stübeweg 51, D-79108 Freiburg, Germany
5Present address: Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
6Present address: Biotech Research & Innovation Centre, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark
7Present address: Unité de Génétique, Reproduction et Développement, INSERM-CNRS UMR 6247, Université de Clermont-Ferrand, 28 Place Henri Dunant, 63000 Clermont-Ferrand, France
*Correspondence: frederic.bantignies@igh.cnrs.fr (F.B.), giacomo.cavalli@igh.cnrs.fr (G.C.)

DOI 10.1016/j.cell.2010.12.026

**SUMMARY**

In *Drosophila melanogaster*, Hox genes are organized in an anterior and a posterior cluster, called Antennapedia complex and bithorax complex, located on the same chromosome arm and separated by 10 Mb of DNA. Both clusters are repressed by Polycomb group (PcG) proteins. Here, we show that genes of the two Hox complexes can interact within nuclear PcG bodies in tissues where they are corepressed. This colocalization increases during development and depends on PcG proteins. Hox gene contacts are conserved in the distantly related *Drosophila virilis* species and they are part of a large gene interaction network that includes other PcG target genes. Importantly, mutations on one of the loci weaken silencing of genes in the other locus, resulting in the exacerbation of homeotic phenotypes in sensitized genetic backgrounds. Thus, the three-dimensional organization of Polycomb target genes in the cell nucleus stabilizes the maintenance of epigenetic gene silencing.

**INTRODUCTION**

The organization of chromosomal domains in the cell nucleus plays an important role in the regulation of gene expression during cellular differentiation and development (Fraser and Bickmore, 2007; Williams et al., 2010). During interphase, eukaryotic chromosomes are organized as distinct domains called chromosome territories, which adopt specific structure and position in the cell nucleus. These territories are not strictly delimited in the nuclear space, allowing for some intermingling with other chromosomes (Branco and Pombo, 2006). It has been reported that chromosomal elements separated by large genomic distances are sometimes able to interact in the nuclear space. The phenomenon of long-range contacts between different chromosomal loci has been called “chromosome kissing” (Cavalli, 2007). Chromosome kissing has been reported for X chromosome inactivation, where the two copies come transiently in close proximity at the onset of the inactivation process (Okamoto and Heard, 2009). The same phenomenon has been observed in few other situations. However, it is not clear how these contacts may contribute to gene regulation in living organisms (Sexton et al., 2009; Williams et al., 2010).

Long-range gene regulation may involve epigenetic components including proteins of the Polycomb group (PcG) (Grimaud et al., 2006; Sexton et al., 2007). PcG proteins are organized into nuclear foci called PcG bodies (Alkema et al., 1997; Buchenau et al., 1998; Messmer et al., 1992; Ren et al., 2008; Saurin et al., 1998). These bodies contain silenced PcG target chromatin (Grimaud et al., 2006; Lanzuolo et al., 2007), which is made of multimeric PcG complexes bound to cis-regulatory elements named PcG response elements (PREs) (Mulder and Vernijzer, 2009; Schuettengruber et al., 2007). PcG protein binding to PREs silences genes involved in developmental patterning and cell proliferation (Merdes and Paro, 2009). In *Drosophila melanogaster*, Hox genes are the most prominent PcG targets. They are organized in two complexes: the Antennapedia complex (ANT-C) spans approximately 400 kb and comprises five Hox genes (*lab, pb, Dfd, Scr*, and *Antp*) that specify parts of the head and the anterior thorax (Kauffman et al., 1990), while the bithorax complex (BX-C) spans approximately 350 kb and contains three Hox genes (*Ubx, abd-A*, and *Abd-B*) involved in the development of the posterior thorax and the abdomen (Duncan, 1987; Lewis et al., 2003). The nuclear organization of the BX-C has been studied by Fluorescent In Situ Hybridization (FISH) and Chromosome Conformation Capture (3C) approaches (Lanzuolo et al., 2007). This study revealed the existence of long-distance interactions among the major elements bound by PcG proteins, including PREs and core promoters. Importantly, upon activation of the *Abd-B* gene, its interactions with the other genes of the complex are lost. Two particular PRE-containing elements called *Fab-7* and
Mcp, both involved in the regulation of Abd-B, have been directly implicated in chromosomal kissing events when extra copies were introduced in the fly genome as transgenes. In both cases, contacts strengthen PcG-dependent gene silencing and kissing events occur specifically at PcG bodies (Bantignies et al., 2003; Grimaud et al., 2006; Muller et al., 1999; Vazquez et al., 2006), although they are clearly tissue-specific and do not occur in several larval tissues containing polytene chromosomes (Fedorova et al., 2008). These data raise the question of whether PcG bodies might be involved in the functional compartmentalization of the Hox clusters or other Polycomb target genes.

In D. melanogaster, the ANT-C and BX-C complexes are located on the right arm of chromosome 3 (3R), separated by approximately 10 Mb of euchromatic sequences that contain more than 1,200 annotated genes and represent more than one-third of the euchromatic fraction of 3R (Figure 1A). In this work, we provide evidence that the two distant Hox clusters can be corepressed by PcG proteins via association in three-dimensional nuclear space between the Antennapedia (Antp) gene from the ANT-C and the Abd-B, Antp, and Ultrabithorax (Ubx) genes from the BX-C. We define this phenomenon as “Hox gene kissing.” Furthermore, 3C on chip (4C) confirmed Hox contacts and revealed additional Abd-B partner loci. Importantly, Hox gene kissing is conserved in D. virilis, a species evolutionarily separated from D. melanogaster by around 60 Myr. Moreover, we demonstrate that the BX-C element Fab-7 participates in Hox gene kissing, and that removal of this element weakens silencing of distant genes in the ANT-C locus, indicating that PcG-dependent chromatin contacts have a functional role in stabilizing gene silencing.

Figure 1. Kissing of Repressed Hox Genes

(A) Schematic drawing illustrating the anterior and the posterior Hox gene clusters in D. melanogaster. The colored lines represent the approximate localization of the FISH probes for Antp (red) and Abd-B (green).

(B) FISH in wild-type (WT) stage 10–11 embryos. Percentage colocalization between Antp and Abd-B. A maximum distance of 350 nm was used to define pairing between the two loci. The p values of the pairwise comparison are 2.045e-06 for Head versus PS4/5, 1.1e-04 for Head versus Posterior, 0.137 for PS4/5 versus Posterior.

(C) Characteristic examples of individual nuclei.

(D) FISH-I of Abd-B, Antp and PC. The regions chosen for image acquisition are indicated to the left. Figure images correspond to deconvolved single slices from 3D stacks. The scale bars represent 1 μm.

(E) Percentage colocalization between beat-Vc and CG17622, two non PcG-targets, which are located 10 Mb away on Chromosome 3R. The p value is 0.707.

(F) Percentage of colocalization between Antp and Abd-B in heads of stage 13–14 embryos from WT, Pcl105 and Pcl10 homozygous mutants. The p values are 1.066e-06 for WT versus Pcl105, 1.936e-09 for WT versus Pcl10, 0.259 for Pcl105 versus Pcl10.

(G) Percentage of colocalization between Antp and Abd-B in third instar larval imaginal eye discs from WT, Pcl105 and Pcl10 heterozygous mutants. The p values are 2.3e-03 for WT versus Pcl105, 2.372e-04 for WT versus Pcl10, 0.623 for Pcl105 versus Pcl10. N indicates the total number of nuclei analyzed in 3–5 embryos or tissues. Asterisks indicate that the pairwise difference between samples corresponding to the left column and the other samples is significant. See also Figure S1 and Table S1.
Figure 2. Extensive Interactions of the Fab-7 Element and Other PcG Target Genes along the 3R Chromosome Arm
(A) Whole 3R chromosome domainogram representation of WT 4C profile (top) and Pc ChIP enrichments as fold change in WT 4–12 hr embryos (bottom, black). The x axis represents chromosome 3R coordinates in Mb and the y axis of the domainogram represents domain sizes as the Log2 of the number of contiguous probes involved in the calculation of statistical scores (see Figure S2H for genomic length conversion). Purple arrowheads indicate the Fab-7 bait and the BX-C. Red arrowheads indicate strong hits within 2 Mb of the anchor region, Black arrowheads indicate the strongest long-range hits, Grey arrowheads indicate other significant long-range target regions.
RESULTS

The Nuclear Organization of the Antennapedia and Bithorax Complexes

We first analyzed the relative positioning of the two Hox complexes in the nucleus, starting with the most distal genes within each complex: Abd-B from the BX-C, and Antp from the ANT-C. We used two-color FISH in whole mount embryos and larval tissues, followed by three-dimensional image analysis (Bantignies et al., 2003). We compared embryonic and larval nuclei with different Hox gene expression profiles (Castelli-Gair, 1998; Kosman et al., 2004; Morata et al., 1994). We first focused on developmental stage 10–11, when PcG-mediated regulation of homeotic genes has already initiated (Pirrotta et al., 1995). 3D image acquisitions were carried out in ectodermal interphase diploid nuclei in three different regions along the anteroposterior axis of the embryo: the head, where both genes are repressed, a thoracic region including parasegments (PS) 4 and 5, where Antp is expressed and Abd-B repressed, and the posterior tip of the embryo (PS 13 and 14), where the reciprocal situation is observed, i.e., Abd-B is activated and Antp repressed.

Abd-B and the Antp gene rarely colocalized in the thoracic PS and in the posterior PS, with colocalization rates of 4% and 7.6%, respectively. In contrast, the association between the loci was significantly stronger in the head where both loci are repressed (18%, Figure 1B,C). This was reflected in a global three-dimensional distance distribution skewed toward shorter distances in the head (Table S1, available online). Later during development, the association between Antp and Abd-B was re-inforced in anterior larval tissues compared to more posterior tissues such as leg and wing discs, where only Antp is active (Figure S1A and Table S2).

Antp also colocalized with the repressed Ubx gene in the head compartment of embryonic nuclei, and this interaction was lost in the trunk (PS4/5), where Antp is active. In the posterior PS, however, where both Ubx and Antp are repressed, they colocalized significantly more than in PS4/5 (Figures S1B and S1C and Table S2). This indicates that Hox gene kissing correlates with the repression status of both Hox genes.

Hox Gene Kissing Occurs at PcG Bodies and Depends on the Function of PcG Proteins

Next we combined FISH with immunostaining using a Polycomb antibody (FISH-I) to analyze the nuclear localization of Antp, Abd-B, and Ubx relative to PcG bodies. Repressed genes are associated with large PcG bodies in 80% to 85% of the cases. In embryo heads, where both Antp and Abd-B are colocalized, the two genes were found in the same PcG body (Figure 1D). In posterior segments, Abd-B is clearly outside PcG bodies in 80% of the cases (Figure 1D), although the two Hox complexes can still associate via kissing of Antp and Ubx within PcG bodies (Figure S1C). In the thoracic PS4/5, in which the Antp gene is activated, Antp is outside PcG bodies in 85% of the cases, while Abd-B and Ubx are inside (data not shown), correlating with the lack of colocalization between Antp and the other genes.

We then analyzed whether PcG proteins are required for Hox gene kissing. First, two non-PcG target genes, beat-Vc and CG17622, which are located on chromosome 3R and separated by 10 Mb like the Hox complexes, colocalize in less than 5% of the nuclei (Figure 1E and Table S1). Second, in the head of Polycomb and Polycomb-like mutant embryos, the Antp and Abd-B genes colocalize in about 5% of the nuclei, i.e., much less than in WT and close to the two control genes beat-Vc and CG17622 (Figure 1F and Table S1). This indicates that PcG proteins are required for pairing of the two loci. Even heterozygous mutations in the Pc and Pcl genes reduced the frequency of gene kissing in anterior larval tissues (Figure 1G and Table S1). These data show that Hox gene kissing occurs within PcG bodies in a PcG-dependent manner.

The BX-C Interacts Preferentially with Polycomb Enriched Regions

In order to analyze whether Hox loci only contact each other or whether they have other interacting partners in the nucleus, we developed a modified 4C protocol (see Experimental Procedures, Figure S2, and Figure S3). The Fab-7 element from the BX-C regulates the Abd-B expression and plays a role in long distance interactions inside the BX-C as well as at greater genomic distances (Bantignies et al., 2003; Lanzuolo et al., 2007). Therefore, we used Fab-7 as the 4C bait fragment in order to analyze its interaction with other loci along the 3R chromosome arm in larval brain and anterior larval discs. We found an extensive series of interactions along the 3R chromosome arm (Figure 2A). The strongest interactions occur within the BX-C. Other strong interactors are two PcG target regions, the ss and srp-pnr loci, located at a distance of 0.5 and 1 Mb, respectively (see Figure 2B). Significant interaction events are also observed at long distances. Four are particularly strong, corresponding to the NK-C, E5/ems, prospero and the ANT-C loci, and four others are weaker, corresponding to the gm, hth, pnt, and Drop loci (Figures 2A and 2C). Strikingly, all these major 4C hits are Polycomb bound regions (Schuettengruber et al., 2009). Moreover, a global analysis indicates that the 4C hits are highly enriched in Polycomb and its associated H3K27me3 repressive chromatin mark (Figure 2D).

Thus, the 4C analysis confirmed Hox gene kissing and revealed additional interactions. In order to validate the 4C results, we verified one of the strongest long-range contacts by FISH. The ltb/lbe genes from the NK homeobox gene complex (NK-C; Garcia-Fernandez, 2005) are located approximately 4.5 Mb distally from the BX-C (Figure 3A). They are organized in tandem and expressed in a few specialized cells in the embryonic para-segments and in the head (Jagla et al., 1997a, 1997b, 1998). The
BX-C and NK-C loci colocalized in approximately 20% of the nuclei in the head and in PS4/5 where both genes are repressed in most of the cells. This interaction decreased significantly in nuclei in the head and in PS4/5 where both genes are repressed. BX-C and NK-C loci colocalized in approximately 20% of the nuclei in the head and in PS4/5 where both genes are repressed.

These results confirm the observation that gene contacts between PcG target genes are more frequent when the genes are corepressed. Accordingly, as in the case of the Hox gene kissing, the contact between BX-C and NK-C occurred exclusively in a PcG body (Figure 3G).

In order to understand whether all PcG target genes associate at PcG bodies in a random manner or whether PcG-mediated gene kissing involves a subset of all target genes, we further studied by 3D-FISH the Hox gene Abd-B and the PcG targets polyhomeotic (ph), which is located on chromosome X, and hedgehog (hh), which is located 6.2 Mb distal to Abd-B on chromosome 3R (Figure S5A). The colocalization frequency was low in to the lbl/lbe locus. In all tissues, the frequencies of colocalization were similar to those observed for Abd-B and lbl/lbe in a divergently expressing tissue, while they were lower than in tissues corepressing Abd-B and lbl/lbe (Figure 3C and Table S3). This suggests that PcG-mediated silencing specifically induces gene kissing.

Gene contacts between the more remote Antp and lbl/lbe loci, which are separated by approximately 14.5 Mb, were less frequent. Nonetheless, the frequency of colocalization was significantly higher in tissues where both loci are silenced (Figure 3D and Table S3).
all embryonic tissues examined (Figures 3E and 3F and Table S3). In agreement with FISH analysis, no significant interaction between hh and Fab-7 was detected by 4C analysis (Figure 2C). Of note, hh and ph were generally localized in smaller PC bodies than those containing the Hox genes (Figure 3H), indicating that different classes of PcG bodies exist in the nucleus. Together, these data demonstrate the existence of specific long-range associations among Hox genes and other PcG target genes, which occur within PcG nuclear bodies and correlate with their transcriptional status along the anteroposterior body axis.

Hox Gene Kissing Is Evolutionarily Conserved in Drosophila Species

We reasoned that, if Hox gene kissing is functionally significant, it might be conserved through evolution. Therefore, we analyzed D. virilis, a species separated from the D. melanogaster lineage 40 to 60 Myr ago. In D. virilis, the two Hox clusters are split between Ubx and abd-A (Von Allmen et al., 1996) instead of between Antp and Ubx. They are located on chromosome 2 and are contained within two large sequence scaffolds (13047 and 12855, see http://flybase.org/cgi-bin/gbrowse/dvir), separated by more than 6 Mb (Figure 4A). A large-scale comparison of the chromosomal organization around the Hox genes in D. melanogaster and D. virilis shows that syntenic regions of limited sizes are conserved between the two species, but the global linear organization of the chromosome is highly divergent (Figure S5A). The D. melanogaster counterparts of the D. virilis genes map to independent regions that are scattered along the entire arm of chromosome 3R. In embryo head, where both genes are repressed, Antp colocalized with Abd-B in 31% of the nuclei. In posterior segments, where Abd-B is active and Antp is repressed, the colocalization frequency was strongly reduced (Figures 4B, 4C, and Figure S5B). We conclude that, despite a radical change in linear chromosome organization, silencing-dependent Hox gene kissing is evolutionarily conserved. This suggests that this phenomenon reflects specific molecular interactions rather than a three-dimensional folding coincidence of the D. melanogaster 3R chromosome arm.

Hox Gene Contacts Depend on Regulatory Elements in the BX-C

In transgenic systems, Fab-7 and Mcp have been shown to be involved in intrachromosomal interactions in diploid tissues (Bantignies et al., 2003, Vazquez et al., 2006). We thus analyzed Hox gene kissing in Drosophila lines in which Fab-7 is deleted. The Fab-7(12) line corresponds to homozygous deletion of the Fab-7 region containing the PRE and the chromatin boundary portion (Mihaly et al., 1997). We performed 4C in the Fab-7(12) line and compared it to the WT profile. While most of the profile is very similar (Figure 5A), some distinct differences can be seen (Figures 5B and Figure S3). In particular, the interaction with the ANT-C is significantly reduced, while the interaction with the NK-C is increased in the Fab-7(12) line (Figure 5C). We confirmed these effects by analyzing three more deletion lines. The Fab-7(1) and the Mcp(1) lines correspond to homozygous deletion of Fab-7 and Mcp, respectively, and the Mcp(12)/Fab-7(1) stock corresponds to the homozygous deletion of both elements (Karch et al., 1994). In each of these lines, reduced levels of colocalization were observed between Antp and Abd-B (Figure S6).

We therefore conclude that the colocalization between the two Hox genes involves specific sequences within the Abd-B locus, even though these sequences do not share extensive homology with sequences from Antp. The finding that the deletion of Fab-7 and Mcp reduces but does not abolish Hox gene kissing suggests that multiple DNA regions of the BX-C contribute to this phenomenon. This notion is corroborated by the colocalization between Ubx and Antp in posterior embryonic tissues where the Fab-7/ Mcp region is located outside PcG bodies (Figures S1B and S1C).

Perturbation of Hox Gene Kissing Affects PcG-Dependent Silencing of the ANT-C Genes

What is the functional significance of the colocalization of silenced genes of the BX-C and the ANT-C? We have previously shown that long range interaction of homologous PRE/boundary...
elements reinforces PcG-mediated silencing (Bantignies et al., 2003). Therefore, we hypothesized that kissing between Hox genes might stabilize PcG-dependent gene silencing. In this case, we predicted that the perturbation of gene kissing in the Fab-7\textsuperscript{12} line would weaken silencing of ANT-C genes, while it should increase silencing at the NK-C locus. We measured, by reverse transcription followed by quantitative PCR (RT-qPCR) in eye-antennal discs, the level of Abd-B, from the BX-C, Antp, Scr, Dfd, and pb from the ANT-C, lbe from the NK-C and hh as a negative control of a gene where the frequency of contacts was not perturbed (Figures 6A and 6B). Abd-B was derepressed in the mutant, indicating that Fab-7 is involved in its repression in anterior tissues, in addition to the abdominal segments of the body plan (Galloni et al., 1993; Gyurkovics et al., 1990). Strikingly, Antp, Scr, Dfd and pb were also derepressed, showing that the partial decrease in long-distance contacts induced by the deletion of Fab-7 is sufficient to induce a decrease of silencing at the ANT-C. In contrast, the lbe gene, which contacts the BX-C with increased frequency, is significantly more repressed upon deletion of Fab-7. Finally, no effects were seen at the control hh gene.

Can the perturbation of gene kissing induce phenotypic changes in flies? Homeotic antenna to leg (A > L) transformation is not observed in flies carrying Fab-7 or Mcp deletions. However, we reasoned that the transcriptional effect may be too subtle to induce a phenotype, and that phenotypic effects might be easier to detect in a sensitized genetic background. We thus analyzed the Antp Nasobemia (Antp\textsuperscript{NS}) mutation (Talbert and Garber, 1994), in which the Antp P2 promoter is duplicated, inducing ectopic expression of the Antp protein in the antennal territory of larval imaginal discs and an incomplete antenna-to-leg (A > L) transformation (Figure 7A) that is sensitive to trxG and PcG functions (Vazquez et al., 1999 and data not shown). This increased expression is much lower than the WT expression of Antp in the wing disc, and does not correlate with altered Hox chromosome kissing (Figures S7A and S7B), suggesting that, on its own, it depends on cis-regulatory changes of the Antp locus rather than on perturbation of chromosome architecture.
We generated different lines in which the AntpNs chromosome was recombined with the Fab-712, Fab-71, Mcp1, and the McpH27Fab-71 chromosomes. As controls, the AntpNs chromosome was recombined with a WT chromosome from a w1118 line and with a mutant chromosome containing an ebony (e1) recessive marker. Three AntpNs, seven AntpNs, Fab-712, seven AntpNs, Fab-71, four AntpNs, Mcp1, three AntpNs, McpH27Fab-71 and four AntpNs, e1 recombinant lines were obtained and raised as balanced stocks at 21°C. No obvious differences were observed between these lines right after their establishment. However, clear differences emerged in subsequent generations, reaching a plateau after 3 to 5 generations: the Ns A > L phenotype was accentuated in the BX-C deletion lines (Figure 7A and Figures S7C and S7D). Transgenerational inheritance has been previously linked to PcG and trxG proteins and to the 3D organization of BX-C elements like Fab-7 (Bantignies et al., 2003; Sollars et al., 2003). Therefore, deletions in the BX-C might enhance the A > L transformations of AntpNs via loss of gene contacts and progressive decrease in chromatin silencing efficiency through subsequent generations.

In addition to A > L transformations, we also observed the emergence of outgrowth phenotypes in the eye of AntpNs combined with Fab-712, Fab-71 or McpH27Fab-71 deletions (Figure S7F). The eye phenotypes are variable, ranging from holes in the eye to large outgrowths, often emerging from ommatidia. Electron microscopy (EM) analysis of these outgrowths indicates that they resemble proximal leg or thorax-like structures, which may indicate a derepression of Antp in the eye imaginal discs. Eye phenotypes were observed in approximately 10% heterozygous AntpNs, McpH27Fab-71 and in 2% to 10% AntpNs, Fab-712 and AntpNs, Fab-71 adult flies for each line, but were totally absent in AntpNs, Mcp1 and in the recombinant control lines. Moreover, the eye outgrowths are strongly enhanced in the rare homozygous escapers, with frequencies of 60% to 80% in AntpNs, McpH27Fab-71, and 25% to 60% in AntpNs, Fab-712 and AntpNs, Fab-71, whereas they are observed in less than 10% in the other lines. A reduction in eye size and fused ommatidia were also observed in the AntpNs lines combined with BX-C deletions (Figure S7F and data not shown). Similar results were observed upon production of recombinant lines using an AntpNs stock, obtained from the Bloomington Drosophila Center. Again, these phenotypes did not appear in the two recombinant control lines obtained.

Since all these phenotypes are consistent with Antp derepression, we tested Antp expression by RT-qPCR on eye-antennal imaginal discs. Significant Antp derepression was observed in the AntpNs, Fab-712 (Figure 7B, AntpNs, Fab-71 and AntpNs, McpH27Fab-71 lines (Figure S7E) compared to the AntpNs line. In contrast, the levels of hh did not change significantly in the different lines.

Finally, we tested whether Fab-7 deletion may affect phenotypes of other Hox gene mutations. We thus analyzed the heterozygous Scr4 Scr allele combination. This chromosome carries a null mutation (Scr4) combined with a second mutation (Scr4)
in the upstream regulatory region (Southworth and Kennison, 2002). This second mutation leads to Polycomb-dependent repression of the \textit{Scr} gene in the first (T1) leg (Pattatucci et al., 1991 and data not shown). Whereas the \textit{Scr}4/\textit{+} males show an average of 6.2 sex comb teeth in the T1 leg, the additional presence of \textit{Scr}w causes a reduction to an average of 4.5, showing that the \textit{Scr}w allele represses the WT \textit{Scr} copy in trans (Figures 7C and 7E). We then tested whether the deletion of \textit{Fab}-7 may attenuate this repressive effect. Crossing the \textit{Scr}4 \textit{Scr}w chromosome with the \textit{Fab}-712 mutation induced a significant increase in the sex comb teeth (Figures 7D and 7E), showing that \textit{Scr} is partially derepressed upon loss of the \textit{Fab}-7 element. Together, our data show that the removal of BX-C regulatory elements weakens silencing at the ANT-C, demonstrating that Hox gene kissing contributes to stabilize their corepression.

**DISCUSSION**

The Nature of Hox Gene Contacts

4C and FISH measurements reveal clear differences in the degree of interactions within the BX-C as opposed to those between BX-C and ANT-C genes. Earlier FISH experiments within the BX-C indicate contacts (as defined by FISH distances \( \leq 350 \, \text{nm} \)) between the \textit{Fab}-7/\textit{Abd-B} and the \textit{bxd/Ubx} regions in the range of 85% (Lanzuolo et al., 2007). This is much higher than the 15 to 20% observed here between \textit{Fab}-7/\textit{Abd-B} and \textit{Antp} in the same tissues, i.e., head of the embryo (Figure 1B). These FISH data are concordant with 4C, where the signals within the BX-C are far stronger than those between \textit{Fab}-7 and the ANT-C. Therefore, the higher-order interactions between elements of the BX-C are tighter than those between heterologous loci such as the BX-C and the ANT-C. We propose that this reflects a hierarchy of higher-order structures in the nucleus, where chromatin elements are much more likely to interact with neighboring partners in cis before engaging in interactions with remote partners. A 3D loop-structure model has been proposed for the BX-C (Lanzuolo et al., 2007). These structures may be highly dynamic entities enabling contacts...
between epigenetic elements from distant loci when they come into three-dimensional proximity in the same nuclear compartment.

**The Discovery of a 3D Polycomb Interactome Gene Network**

Our 4C analysis clearly reveals that the gene kissing is not limited to Hox genes, since we found that the BX-C is engaged in additional contacts with other partners on the same chromosome arm. The fact that all the main interaction partners are PcG-bound chromatin domains, suggests that PcG proteins contribute to the establishment of long-range contacts among their target genes in the 3D nuclear space. However, most PcG target genes are cobound by many other chromatin factors, including insulator proteins like CTCF and Su(Hw) (Bushey et al., 2009; Negre et al., 2010), and these proteins may contribute to drive long-distance contacts via chromatin boundary or insulator elements present at PcG target loci.

The 4C analysis measured chromosome contact frequencies from a large population of fixed cells, however it cannot distinguish whether these interactions are simultaneous. Two-color FISH experiments validated 4C contacts between BX-C and ANT-C and between BX-C and NK-C. Three-color FISH experiments, where all three complexes are silenced, revealed that the contacts are not simultaneous, at least during embryogenesis (Figure S4). We thus propose that each of the Hox loci is in dynamic contact with the other one, as well as with other PcG target loci. The existence of multiple gene contacts as we discovered in 4C and FISH can thus explain why the two Hox loci do not interact in all nuclei, but only in a significant minority of them. Interestingly, the deletion of Fab-7 reduced Antp-Abd-B contacts while it increased Abd-B-lbl contacts. This might suggest that, once free from one interacting partner, Abd-B chromatin might be available for increased interactions with some of the other partners. In the future, further 4C studies and high-throughput FISH analysis with probes directed against the interacting 4C loci and the intervening regions should extend these observations to gain a systematic understanding of Polycomb dependent gene kissing. In this context, a study which these observations to gain a systematic understanding of the interacting 4C loci and the intervening regions should extend and high-throughput FISH analysis with probes directed against PcG target loci.

Evolutionary Conservation and Functional Significance of Hox Gene Kissing

The evolutionary conservation of Hox gene kissing between D. melanogaster and D. virilis strongly suggests that selective pressure maintains the spatial proximity of Hox complexes. This fact reinforces the idea that gene kissing might stabilize PcG-mediated silencing of Hox genes, contributing to the specification of body structures along the anteroposterior axis. Fab-7 and Mcp might represent examples of bifunctional regulatory elements, as they appear to have two distinct roles: on the one hand, they regulate the expression of their flanking genes in cis and, on the other, they mediate long-distance regulatory interactions with Hox genes in the ANT-C. We propose that one of the roles of Fab-7 and Mcp was to promote locus-wide chromatin condensation of the ancestral Hox gene cluster by interacting with more anterior genes. The maintenance of these interactions after the physical split of the cluster may have originated the long-distance pairing that is still observed today.

The analysis of deletions of BX-C elements reveals important features of long-range regulation. Within the BX-C, we noticed that the removal of Fab-7 reduces PcG-dependent silencing of the abd-A and Ubx genes (2- to 3-fold derepression, data not shown). This result indicates that multiple regulatory interactions take place for the strong and faithful repression of the BX-C in anterior segments. In this system, we exclude secondary effects in trans of Abd-B on Ubx and abd-A because the products of posterior Hox genes always repress, not activate, more anterior genes (a phenomenon called posterior dominance, Kuziora, 1993; Struhl and White, 1985). Indeed, this feature might attenuate the transcriptional effects of loss of long-distance Hox contacts.

A second level of higher-order chromatin association involves larger scales, such as between the BX-C and ANT-C loci. These contacts are probably not indispensable for repression. For instance, Antp is silenced in every anterior cell, even if it contacts the BX-C in a minority of the nuclei. Cis regulatory elements are probably able to maintain silencing to a sufficient degree. However, small but clear transcriptional derepression of multiple ANT-C genes was observed upon mutations in the BX-C. In this context, one should note that the BX-C contacts many other loci in addition to the ANT-C. Likewise, the ANT-C may contact several other PcG target regions. These other contacts could functionally complement a loss of pairing upon BX-C mutations. ANT-C may maintain spatial association with other regions within the BX-C or with other domains that participate in the same spatial network of PcG target genes, thus remaining in an appropriate regulatory environment. However, this compensation cannot be complete, as illustrated by the increase in homoeotic phenotypes seen in the sensitized Antp and Scr backgrounds. To our knowledge, this is the first report that links the function of a chromatin element involved in 3D chromosome contacts to specific phenotypes.

**Hox Contacts and Chromosomal Rearrangements**

Three different splitting events of the ancestral Hox gene complex have been detected to date in the Drosophila genus. Two of them, mentioned in this study, are represented by the separation of Antp and Ubx in D. melanogaster and of Ubx and abd-A in D. virilis. A third split has been described in D. buzzatii (Negre et al., 2003). This species, a member of the D. repleta group, shows a split between Ubx and abd-A as with D. virilis (Ranz et al., 1997), and an additional split, whereby the most anterior gene, lab, has been relocated flanking the two posterior Hox genes abd-A and Abd-B. In light of our data, the most likely scenario that might explain this third split involves translocation of the entire lab locus to the border of the posterior cluster due to PcG-dependent spatial proximity between the anterior and posterior Hox clusters. Related to this point, in mammalian cells the proximity of chromosome or chromosomal loci has been suggested to induce chromosomal rearrangements between

Cell 144, 214–226, January 21, 2011 ©2011 Elsevier Inc. 223
them (Branco and Pombo, 2006; Lin et al., 2009; Nikiforova et al., 2000; Roix et al., 2003). Moreover, chromosome kissing events dependent on colocalization in transcription factories were also shown to be correlated with a high rate of translocation (Osborne et al., 2007). It is interesting to note that Hox gene clusters have been submitted to considerable rearrangements during evolution of the animal kingdom (Garcia-Fernandez, 2005). We propose that split Hox clusters might have contributed to evolution of chromosomes bearing them.

In conclusion, the data described here show that the specific nuclear organization imposed by the PcG proteins in Drosophila diploid tissues influences the maintenance of epigenetic states and might contribute to genome evolution.

**EXPERIMENTAL PROCEDURES**

**Fly Stocks and Handling**

Fly stocks were maintained over the Pcl10/ KrGFP-TM3,Sb line was used as wild-type (WT) D. melanogaster. A w stock (#15100-1051, from the Tucson species stock center) was used as WT D. virilis. The Pcl10/ KrGFP-CyO stock was used for selection of homozygous Pcl10 mutants (Bantignies et al., 2003). A Pcl10/ KrGFP-TM3,Sb stock was used for the selection of homozygous Pcl10 mutants. The Fab-7/2, Fab-7/1, Mcp1, and Mcc2/Fab-7 deletion lines were described in (Karch et al., 1994; Mihaly et al., 1997). The Antp(C) stock used in this study was provided by W. Gehring. Antp(C) stock from Bloomington was also used, although this stock shows lower penetrance of the A > L phenotype. All the Antp(C) recombinant lines were maintained over the KrGFP-TM3,Sb balancer (from stocks BL#1935 of the Bloomington Drosophila Stock Center). The ebony (e), Scr(C) and Sco(C) stocks were from Bloomington (BL#2558, BL#118B and BL#809, respectively). Staged eggs were collected on agar plates with standard yeast/fresh yeast medium. Antp and Scr mutant flies were grown at 21°C, and sex comb teeth were counted under a Nikon SMZ1000 binocular at 80× magnification.

**RT-qPCR**

Third-instar larval imaginal eye-antennal discs were dissected in Schneider’s Drosophila Medium (GIBCO) and 30–40 discs were taken for RNA isolation using TRizol reagent (Invitrogen). 300–400 ngs of total RNA were used for the RT reaction. RT was performed using the SuperScript II First Strand Synthesis Kit from Invitrogen following the manufacturer’s instructions and using hexamer primers. cDNA quantifications were performed by real-time PCR, using a Roche Light Cycler and the Light Cycler FastStart DNA Master SYBR green I kit. NdeI digested genomic DNA served for the standard curve. Expression levels were normalized to Rp49 and multiplied by 1.109. Primer sequences are listed in Table S4.

**Two-Color 3D-FISH and FISH-I**

These procedures are described in the Extended Experimental Procedures.

**Microscopy and Image Analysis**

Microscopy and 3D image analysis were as previously described (Bantignies et al., 2003). Minor modifications and EM are described in the Extended Experimental Procedures.

**Chromosome Conformation Capture on Chip (4C)**

Flies were grown at 25°C. Third-instar larval brain and anterior discs from 200 larvae were dissected in Schneider’s Drosophila Medium and used for the 4C. The 4C was performed as previously described (Hagege et al., 2007; Miele and Dekker, 2009) with the main differences being the use of DpnII (New England Biolabs), a 4 bp cutter restriction enzyme, and a fixation in 3% para-formaldehyde for 30 min, maximizing sensitivity and resolution of contact detection. The 4C method includes an “anchor biotinylated primer extension” procedure that is described in detail in the Extended Experimental Procedures. Microarray analyses are described in the same section and Figure S2 and Figure S3.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures, seven figures, and six tables and can be found with this article online at doi:10.1016/j.cell.2010.12.026.

**ACKNOWLEDGMENTS**

We would like to thank W. Gehring, F. Karch, H. Gyrkovic, J.M. Dura, and Alain Pelisson for fly lines and plasmids. We also thank T. Sexton and D. Cribbs for stimulating discussions; B. Pfeffer and S. Celinker for sharing the D. virilis BX-C sequence before publication; Chantal Cazevieille for electron microscopy at the Centre de ReSSources en Imagene Cellulaire of Montpellier. F.B. is supported by the CNRS; V.R. and I.C. were supported by the Ministère de l’Enseignement Supérieur, the Association pour la Recherche sur le Cancer, and the Agence Nationale de la Recherche. B.L. was supported by the Ministère de l’Enseignement Supérieur, B.S. was supported by the Fondation de la Recherche Médicale. G.C. was supported by grants of the CNRS, the Human Frontier Science Program Organization, the European Union FP6 (Network of Excellence, The Epigenome, and STREP 3D Genome), the European Research Council (ERC-2006-AdG No 232947), by the Indo-French Centre for Promotion of Advanced Research, by the Agence Nationale de la Recherche, and by the Ministère de l’Enseignement Supérieur, ACI BCMS.

Received: June 22, 2009
Revised: September 22, 2010
Accepted: December 17, 2010
Published: January 20, 2011

**REFERENCES**


EXTENDED EXPERIMENTAL PROCEDURES

FISH

Two-Color 3D-FISH and FISH-I Procedures

Two-color FISH was performed as previously described (Bantignies et al., 2003) and a detailed protocol is available at http://www.epigenome-noe.net/WWW/researchtools/protocols.php. Briefly, embryos were dechorionated with bleach and fixed in buffer A (60 mM KCl; 15 mM NaCl; 0.5 mM spermidine; 0.15 mM spermine; 2 mM EDTA; 0.5 mM EGTA; 15 mM PIPES, pH 7.4) with 4% paraformaldehyde for 25 min in the presence of heptane. Embryos were then devitellinized by adding methanol to the heptane phase. They were sequentially re-hydrated in PBT (PBS, 0.1% Tween 20). Larval imaginal discs were dissected in PBS and fixed in PBT with 4% paraformaldehyde for 20 min. Tissues were treated with 100-200 µg/ml RNaseA in PBT for 2 hr at Room Temperature (RT), then incubated in PBStr (PBS, 0.3% Triton) for 1 hr. Tissues were then sequentially transferred into a pre-Hybridization Mixture (pH:M: 50% formamide; 4XSSC; 100 mM NaH2PO4, pH 7.0; 0.1% Tween 20). Tissue DNA was denatured in pHM at 80°C for 15 min. The pHM was removed, and denatured probes diluted in the FISH Hybridization Buffer (FHB: 10% dextran sulfate; 50% deionized formamide; 2XSSC; 0.5 mg/ml Salmon Sperm DNA) were added to the tissues without prior cooling. Hybridization was performed at 37°C overnight with a gentle agitation. Posthybridization washes were performed, starting with 50% formamide, 2XSSC, 0.3% CHAPS and sequentially returning to PBStr.

For FISH-I, after posthybridization washes, embryos were blocked in PBStr-10% Normal Goat Serum (NGS) for 2hr at RT, and incubated overnight at 4°C with a PC rabbit polyclonal antibody (described in Martinez et al., 2006) at a dilution of 1:250 in PBStr-10% NGS. Embryos were washed several times in PBStr, blocked again in PBStr-10% NGS for 1h at RT, and incubated for 1h at RT with an anti-rabbit-Cy5 (Jackson Laboratories) at a dilution of 1:200 in PBStr-10% NGS. DNA was counterstained with DAPI and embryos were mounted in Prolong Antifade medium (Molecular Probes).

Probe Labeling

FISH Probes were initially labeled by nick translation with biotin-14-dATP (Invitrogen life technologies) or digoxigenin-11-dUTP (Roche Diagnosis) according to manufacturers’ instructions. During the course of this work, we switched to a more sensitive direct labeling method, using the FISH Tag DNA Green Kit (Alexa Fluor 488 dye) and the FISH Tag DNA Orange Kit (Alexa Fluor 555 dye) (Roche Diagnosis) according to manufacturers’ instructions. For three-color FISH, we used in addition the FISH Tag DNA Far Red Kit (Alexa Fluor 647 dye). Approximately 100 ngs of each labeled probe diluted in 30 µl of FHB were used for hybridization.

Description of the Probes

FISH Probes are made of genomic PCR fragments or of plasmid fragments (see Table S5), Each probe covered 12 to 25 kb of the relevant genomic region: the Antp probe contained the P1-PRE region (Zink et al., 1991) (or its equivalent for the Antp D. virilis probe), the Abd-B probe included the Fab-7 PRE and the distal DNA region (or its equivalent for the Abd-B D. virilis probe), the Ubx probe included the bxd PRE and the proximal region (the Abd-B and Ubx probes are separated by approximately 130 kb). The lbl probe spanned the lbl promoter region, covering the intergenic region between lbl and lbe. The hh and ph probes spanned their upstream promoter regions, in which PREs have been identified (Bloyer et al., 2003; Maurange and Paro, 2002). The specificity of each probe was systematically verified by FISH on polytene chromosomes.

Microscopy and Image Analysis

Due to somatic pairing of homologous chromosomes, the majority of the embryonic stage 10-11 nuclei show a single FISH spot for each probe. The two probes were considered colococalized when the two FISH signals overlapped or were juxtaposed (distance between the two centers of the signals ≤ 350 nm). For each case, 3D stacks were collected from 3-5 different tissues (optical sections were collected at 0.5 µm intervals along the z-axis), and relative distances between FISH signals were analyzed in approximately 70 to 130 nuclei per 3D stack using Metamorph software (Universal Imaging Corp.). The mean percentage of colocalization was calculated for each stack. For graphic display, the mean values were pooled and displayed in the y axis. Details of the statistical analysis are provided below. Figure images correspond to deconvolved single slices from 3D stacks, using the Huygens MLE single tif procedure (Scientific Volume Imaging). For electron microscopy (EM), the samples were coated with a 10 nm thick gold film and then examined under a scanning electron microscope (Hitachi S4000) using a lense detector with an acceleration voltage of 10kV at calibrated magnifications.

Statistical Analysis of FISH Data

For statistical analysis of the percentage of pairing, we performed pairwise comparisons. For each sample, the raw distance data from all stacks were pooled. These distances were then coded: for distances up to 350 nm, pairing was assumed and the distance was given the value 0. The value 1 was assigned to all other distances. Each distance was consequently replaced by a 0 or 1, giving rise to Bernoulli random variables. These Bernoulli variables are independent, since the presence of pairing or non-pairing in a given nucleus is not assumed to depend on pairing events in other nuclei of the sample under consideration. Consequently, the total number of pairings follows a binomial distribution as a sum of independent identically distributed Bernoulli variables. We then performed two-sample tests on proportions. Gaussian approximation of the binomial distribution is always valid in our framework and was applied to obtain p-values (R language routines were implemented by a semi-automatic procedure). For statistical analysis of the 3D distance distribution differences, we performed Wilcoxon tests on pooled distances (Table S1, Table S2, and Table S3).
4C

Details of the 4C Procedure

Excepted when mentioned, all primer extension and PCR reactions were performed with the FastStart DNA MasterPLUS SYBR Green I Kit (Roche Applied Science) using the LightCycler 2.0 Real-Time PCR System (Roche Applied Science). All incubations were performed in a Thermomixer (Comfort, Eppendorf).

The DNA obtained after the 3C procedure was resuspended in 44 μl of 10mM Tris-HCl, pH 8.0 and used for the 4C. The copy number of the anchor fragment was determined by real-time PCR using the Forward and Reverse Nested Anchor Primers.

Step 1: Biotinylated Primer Extension from the Anchor Fragment

The reaction volume was 20 μl and contains 10^3 anchor copies of the 3C sample (input) and 0.5 μM of the Biotinylated Primer. Two reactions were performed in parallel. The primer extension program was: 1 cycle of (95°C, 10 min); 30 cycle of (95°C, 5 s; 65°C, 5 s; 72°C, 12 s). The two reactions were pooled, the volume adjusted to 400 μl with water and separated into two samples (2 x 200 μl) for step 2.

Step 2: Immobilization on Streptavidin Beads and Denaturation

Two batches of 1 mg of Magnetic Streptavidin beads (Dynabeads MyOne Streptavidin C1, Invitrogen) were washed twice with 2 x Binding Buffer (10 mM Tris-HCl pH 8.0; 2 M NaCl) and resuspended in 200 μl of this buffer. The two samples were added to the two bead batches. Incubation was performed at 25°C under agitation at 1400 rpm for 1 hr. The supernatants were discarded. Beads were then resuspended in 500 μl of Denaturing Buffer (0.1 N NaOH), and incubated at 25°C under 1400 rpm for 10 min. Beads were washed once with 900 μl of Denaturing Buffer and three times with 900 μl of 10 mM Tris-HCl pH 8.0.

Step 3: Poly-G Tail Synthesis

Beads were resuspended in 47.5 μl of Poly-G tail synthesis Buffer (1.05 x TdT Buffer (New England Biolabs (NEB)), 1.05 x CoCl2 Solution (NEB), 5.26 μM dGTP). Suspensions were incubated at 95°C under 1400 rpm for 1 min, and then incubated on ice for 1 min. 2.5 μl of 20 units/μl TdT (NEB) were then added to each sample, and incubated at 37°C under 1400 rpm for 20 min. The TdT was inactivated at 70°C under 1400 rpm for 10 min. Beads were then washed 3 times with 900 μl of 10 mM Tris-HCl pH 8.0.

Step 4: Specific Elution of Immobilized Products by Directed Digestion of the Anchored Fragment, Purification, and Concentration

The reaction volume was 20 μl and contained 10 μl of the specific elution product and 0.5 μM of the Universal Linker Synthesis Template. The linker synthesis program was: 1 cycle of (95°C, 1 min); (80°C to 50°C with 0.01°C/s); (4°C, pause). Beads were then transferred to 1.5 ml eppendorf tubes. 50 μl of cold Specific Elution Solution (1X Buffer 2; 250 units of 50 units/μl HinfI, NEB) were added to the beads, and incubated at 37°C under 1400 rpm for 2 hr. The supernatants of the two samples were then collected and pooled into a new tube. Beads were then washed twice with 90 μl of 10 mM Tris-HCl pH 8.0, and all supernatants collected during these washes were pooled (total volume of 480 μl). The sample was extracted with Phenol/Chloroform/IsoAmyl Alcohol pH 8.0 (PCI). The aqueous phase was EtOH precipitated in the presence of 500 mM NaAc, pH 5.2 and 35 μg of glycogen, at 80°C for 10 min. After centrifugation at 4°C and 20000 rcf for 90 min, the DNA pellet was washed with 70% EtOH, air-dried, and resuspended in 11 μl of 10 mM Tris-HCl pH 8.0. At this step, the anchor copy number was determined, and the yield was approximately 50% of the input.

Step 5: Linker Synthesis Using a Biotinylated Universal Linker Synthesis Template

The reaction volume was 20 μl and contained 10 μl of the specific elution product and 0.5 μM of the Universal Linker Synthesis Template. The linker synthesis program was: 1 cycle of (95°C, 10 min); 50 cycle of (95°C, 5 s; 65°C, 10 s; 72°C, 5 μl). Linker synthesis product volume was adjusted to 200 μl with water for step 7.

Step 6: Second Immobilization on Streptavidin Beads

1 mg of Magnetic Streptavidin beads was prepared as described in step 2. The Linker products were added to the beads, and incubated at 25°C under 1400 rpm for 1 hr. The supernatant was collected into a new tube.

Step 7: Eution by Denaturation, Purification, and Concentration

Beads were mixed with 260 μl Denaturing Buffer and incubated at 25°C under 1400 rpm for 10 min. Supernatant was collected and mixed with 66 μl of 5X Equilibration Buffer (1:1 of 1M Tris-HCl pH 8.0 and 1 N HCl). Beads were washed with 260 μl Denaturing Buffer and the supernatant mixed with 66 μl of 5X Equilibration Buffer. The three supernatants (collected in steps 6 and 7) were pooled and split into two samples (2 x 520 μl).

The samples were extracted with PCI, and the aqueous phase was EtOH precipitated as described in step 4 (but without adding NaAc). The DNA pellets were resuspended in 21 μl of 10 mM Tris-HCl pH 8.0, and pooled. At this step, the anchor copy number was determined, and the yield was approximately 20% of the input.

Step 8: DNA Amplification, Purification, and Concentration

8 reactions were performed in parallel using the FastStart High Fidelity PCR System, dNTPack (Roche Applied Science). Each reaction of 20 μl contained 5 μl (10^6 anchor copies) of the sample in 1 x FastStart High Fidelity Reaction Buffer without MgCl2, 3.5 mM MgCl2, 5% (v/v) DMSO, 2 μM Nested Anchor Primer (Forward), 2 μM Universal Primer, 0.25 mM PCR Grade Nucleotide Mix, 3.5 units of the FastStart High Fidelity Enzyme Blend. The amplification program was: 1 cycle of (95°C, 10 min); 35 cycle of (95°C, 10 s; 68°C, 20 s; 72°C, 60 s); 1 cycle of (72°C, 5 min).

All sample volumes were adjusted to 100 μl with 10 mM Tris-HCl pH 8.0 and purified using the NucleoSpin Extract II kit (Macherey-Nagel). Elution was performed with 50 μl Buffer NT (Macherey-Nagel), and samples were pooled and supplemented with 70 μl 10 mM MgCl2.
Tris-HCl pH 8.0. The sample was extracted with PCI, and the aqueous phase was EtOH precipitated as described in step 4. The DNA pellet was resuspended in 22 μl of 10 mM Tris-HCl pH 8.0. More than 4 μg of DNA were usually obtained after this amplification step, and 500 ng were used to check the fragment size by electrophoresis on a 4% agarose denaturing gel. The fragment size is expected to range from 100 bp (anchor fragment alone) to 500 bp.

**DNA Sequences for the 4C Procedure**

**Anchor Fragment**

Anchor fragment is a 2179 bp DpnII fragment located in the Fab-7 regulatory element of the BX-C. Its genomic coordinates in genome release 5.17 are: 12,724,247–12,726,425.

**Primers for Copy Number Quantification of the Anchor Fragment (Steps 1, 4, and 7)**

Nested Anchor Primer (Forward) starts 87 bp upstream the 3' DpnII restriction site of the anchor fragment and its sequence is: CACACAATTCCTTGCCAAGT

Nested Anchor Primer (Reverse) starts 43 bp upstream the 3' DpnII restriction site of the anchor fragment and its sequence is: GAGCGTGCCAAAAATCTTAGAA

Biotinylated Anchor Primer (step 1) starts 179 bp upstream of the 3' DpnII restriction site of the anchor fragment and its sequence is: BiotinTEG-ACGCCTGCTGTGACATTTGCCAA

Specific Elution Probe (step 4) starts 126 bp upstream the 3' DpnII restriction site of the anchor fragment and its sequence is: GGAATTCGTGACAGTGCAGATCTTGAACAAATGGTATTTTATTTAT

Universal Linker Synthesis Template (step 5) sequence is BiotinTEG-GTCGGATGAAACGAGTACTACCGACCCCCCCCCCCCC2'-3'-DideoxyC

**Primers for the Final DNA Amplification (Step 8)**

Nested Anchor Primer (Forward), described above: CACACAATTCCTTGCCAAGT

Universal Primer (Reverse): GATGAAACGAGTACTACCGACC

**Microarray Signal Improvement (Illustrated in Figures S2A–S2F)**

The 4C and DpnII digested genomic DNA (Control) samples were hybridized by Roche Nimblegen to a 380K microarray design (GEO: GPL10867) tiling euchromatic regions of the *Drosophila melanogaster* genome. This tiling array has 219,005 oligonucleotide probes mapping to chromosome arm 3R with an average oligonucleotide length of 50 bp and an average inter-oligo spacing of 120 bp.

Microarray data have been deposited in the Gene Expression Omnibus (GEO) database with accession number GSE23887 at http://www.ncbi.nlm.nih.gov/geo. We analyzed microarray data from two biological 4C replicates from WT (WT-r1, WT-r2, GEO: GSM589073 and GSM589074) and Fab-7^YZ^ (F12-r1, F12-r2, GEO: GSM589075 and GSM589076) larval discs. *I_{dc}* and *I_{ctr}* represent hybridization levels of 4C and Control samples respectively, and the average signal intensity (*A*) and log-ratio (*M*) values are thus classically defined as:

\[
A = (\log_2(I_{dc}) + \log_2(I_{ctr}))/2
\]

\[
M = \log_2(I_{dc}) - \log_2(I_{ctr})
\]

Raw microarray signal qualities were improved using low-distortion normalization and probe filtering steps described below. 4C profiles were then analyzed using non-parametric statistics and random permutation tests (Figure S2A).

**Step 1: Low-Distortion Normalization (Figures S2B and S2C)**

To reduce biases associated with variations of the average raw signal intensities, we first directly applied lowess normalization (Cleveland, 1979). However, this leads to a large distortion of the data as illustrated on the MA plots (Figure S2B, top right panel). A similar problem of normalization was previously addressed to optimize correction for the dye bias that may exist in dual channel hybridizations for ChIP-on-chip data (Peng et al., 2007). Their strategy consists in a rotation in the \(\{A, M\}\) plane which is applied prior to lowess normalization. We therefore tested different rotation angles and selected values leading to reduced distortions during array hybridizations for ChIP-on-chip data (Peng et al., 2007). Their strategy consists in a rotation in the \(\{A, M\}\) plane which is applied prior to lowess normalization, could have impacted results of our 4C data analysis, we also tested alternative domainogram analysis using slightly modified rotation angles or even raw 4C signals \((I_{dc})\) as enrichment signal. This analysis resulted in slight changes in the statistical scores, but the final conclusions were unchanged (data not shown).

**Step 2: Probe Filtering (Figures S2D–S2F)**

Since microarray probes were originally designed using the FlyBase release 4 genome sequence, we first realigned all probes to the FlyBase release 5 genome. There were no differences in chromosome 3R sequence between the releases, although we were able to remove probes matching to multiple genomic locations (Table S6 sheet1). We noticed large differences in raw hybridization levels of Control samples, with weaker signals for probes overlapping or very close to DpnII sites compared to probes located further away (Figure S2D). This is not the case of 4C DNA (this depends on the elongation step 1 described in the 4C protocol above, which produces fragments of up to 200-300 bp). The result of this difference is that only the signals from probes located close to DpnII sites
probes in the BX-C and probes in the non-BX-C regions were permutated independently (Figure S3D). The genomic coordinates (Simonis et al., 2006). To avoid contamination of long-distance measurement by highly enriched anchor-biased measurements, We used a random permutation test to estimate the significance of statistical scores associated with long-distance interactions probability test.

To extract features from the 4C profiles, we applied a non-parametric transformation (de Wit et al., 2008). This method generates combined profiles for WT and experiments. We have checked recently published e4C raw data (Schoenfelder et al., 2010), and observed similar scatter plots between replicate cis molecules are present in only one experiment but not in the other. This is not specific to our dataset since we this does not saturate the whole spectrum of the possible long-range interactions and explains why 4C data are of lesser quality start from approximately 20,000,000 cells for each 4C experiment. At best,

\[ W_2 = \{ \text{second closest to DpnII site} \} \]

\[ W_1 = \{ \text{closest overlapping DpnII site} \} \]

with \( Q(x) = (\text{rank}(x) - 0.5)/N \)

where \( N \) is the number of microarray probes realigned on the Flybase 5 genome release

At the end of this filtering procedure, 99% of the selected probes were less than 260 base pairs from a DpnII site (Figure S2F), 70% did not overlap with a DpnII site, and only 2% of the probes including a DpnII site had less than 25 nucleotides specifically hybridizing to the matching DpnII fragment (data not shown). Detailed numbers of probes and DpnII fragments considered during our analysis are summarized in Table S6 sheet 1.

**Statistical Analysis of 4C Data (Figures S2G, S2H, and S3)**

**Domainograms**

To extract features from the 4C profiles, we applied a non-parametric transformation (de Wit et al., 2008). This method generates statistical scores for all possible windows of contiguous probes and allows their visualization with domainograms. Therefore, this is a multiscale analysis that can be thought as the application of sliding windows of the type used in previous analyses (Simonis et al., 2006), but without pre-imposing a fixed window size, and using enrichment ranks instead of levels. One can thus view results with sliding windows of increasing size, starting from a single data point (hybridization to a single oligonucleotide in the chip) up to the whole chromosome size. The domainograms are representations of the 4C profiles as heat maps in a \((x, y)\) graph, where \( x = \text{chromosome coordinate in Mb}, y = \text{domain size expressed in number of contiguous probes involved in the calculation of statistical scores} \) (Figures S2G and S2H). We performed two independent experiments with WT-r1 paired with F12-r1 and WT-r2 paired with F12-r2. To illustrate the overall reproducibility of our 4C experiments (Figure S3A), we generated scatter plots of replicates WT-r1 versus WT-r2, and F12-r1 versus F12-r2, as well as of paired experiments, WT-r1 versus F12-r1 and WT-r2 versus F12-r2 (Figure S3B). One can see the good degree of reproducibility in pairwise comparisons. As expected, the majority of the data points are not more different between two WT experiments than between WT and mutant experiments, since the significant differences concern a small subset of the data points. The most likely reason for having signals close to zero in one but not the other replicate depends on the fact that we start from approximately 20,000,000 cells for each 4C experiment. At best, Fab-7 can interact with 2 molecules per cell, and 3C efficiency being lower than 1% one can expect a maximum of 200,000 different DNA molecules being present in the original sample. Many of these are coming from cis-loci located close to the bait, so that one is left with few thousand different molecules. Most likely, this does not saturate the whole chromosome of possible long-range interactions and explains why 4C data are of lesser quality compared to ChIP (where in each cell one may have thousands of protein-DNA interactions as opposed to two). The data are therefore more noisy as some molecules are present in only one experiment but not in the other. This is not specific to our dataset since we have checked recently published e4C raw data (Schoenfelder et al., 2010), and observed similar scatter plots between replicate experiments.

We also generated domainograms for each replicate (Figure S3C). We then computed and analyzed statistical scores from combined profiles for WT and Fab-7 (Figures 2A and 5A). These combined profiles were obtained by combining each pair of probe scores (defined as \( Q(x) = (\text{rank}(x) - 0.5)/N \), see Probe filtering section) from the two replicate experiments and using Fisher’s combined probability test.

**Determination of Significant Long-Range Interactions outside the BX-C**

We used a random permutation test to estimate the significance of statistical scores associated with long-distance interactions (Simonis et al., 2006). To avoid contamination of long-distance measurement by highly enriched anchor-biased measurements, probes in the BX-C and probes in the non-BX-C regions were permutated independently (Figure S3D). The genomic coordinates of BX-C and non-BX-C regions, according to FlyBase release 5 genome annotations, are described in Table S6 sheet 2. For both WT and Fab-7, we used 2000 random permutations to generate a distribution of maximum scores found at each domain size in the non-BX-C regions. We then used this distribution to calculate the threshold values corresponding to 99% confidence according to our random permutation test (Figure S3E). We identified regions of significant interaction by comparing scores derived from 4C profiles to the thresholds and generated black and white (binarized) domainograms with significant domains as black areas (Figure 5B). Note that repeating the analysis by permuting signals coming from all probes without separation of the BX-C from the rest of the chromosome showed contacts with the same regions, suggesting that the background modeling procedure is not a main determinant of the result under our experimental setting. Furthermore, a simpler analysis following previously published procedures (Simonis et al., 2006) with a fixed sliding window of either 9 or 298 Kb, confirms the strongest interactions found with the domainogram. Finally, the changes in interactions between Fab-7 and the ANT-C or the NK-C loci in the WT versus the
Correlation of Non-BX-C Interactions with PcG Chromatin Marks

In each significantly interacting non-BX-C region, all 4C hits locally presenting maximal scores at minimum domain size were extracted ($N = 77$). We then used previously published chromatin profiles from 4-12 hr Drosophila embryos (Schuettengruber et al., 2009) to generate distributions of Pc, H3K27me3 and H3K4me3 enrichments, centered on the location of each selected 4C hit. As an indicator of significance of the observed patterns, we generated parallel distributions of Pc, H3K27me3 and H3K4me3 enrichment, centered on the location of 200 random selections of the same number of sites. To be consistent with the DpnII restriction map, these random sites were chosen among the set of filtered probes in the non BX-C regions (Figure 2D).

Comparison of the 4C Interactions at ANT-C, NK-C, and hh Regions in WT versus Fab-712

We defined ANT-C, NK-C, and hh regions according to FlyBase release 5 genome annotations and including neighboring intergenic sequences. The genomic coordinates of these 3 regions are detailed in Table S6 sheet 3. We calculated a maximal score of interaction for these three regions in WT and Fab-712 experiments by scanning the entire domains, and represented these scores as a graph (Figure 5C). To analyze the difference in WT versus Fab-712, we computed a null distribution of differences between maximal scores for the three regions using 5000 independent random permutations of the WT and Fab-712 4C profiles (including the BX-C probes). This led to 5000 × 5000 random difference values per region. Then, for each region, we associated a p value to the actual difference between WT and Fab-712 according to its rank in the corresponding null distribution (Figure 5C). Interestingly, the maximal 4C interaction score within the ANT-C region in WT corresponds to a 2.7 kb sequence right at the promoter region of the Antp gene, corresponding to 14 consecutive probes mapping to seven DpnII fragments.

SUPPLEMENTAL REFERENCES


Figure S1. Hox Gene Kissing in Embryos and Imaginal Discs, Related to Figure 1

(A) FISH comparing the percentage of colocalization between Antp and Abd-B from different third instar larval imaginal discs. The P-values of the pairwise comparison are 3.137e-03 for Eye versus Antenna, 1.083e-09 for Eye versus Wing, 6.451e-11 for Eye versus Leg, 2.003e-03 for Antenna versus Wing, 3.483e-04 for Antenna versus Leg, 0.592 for Wing versus Leg. N indicates the total number of nuclei analyzed in 3-5 tissues. Asterisks indicate that the pairwise difference between samples corresponding to the left column and the other samples is significant. On the right, characteristic examples of individual nuclei. Figure images correspond to deconvolved single slices from 3D stacks.

(B) FISH in wild-type stage 10-11 embryos. Percentage colocalization between Antp and Ubx loci. The P-values are 2.093e-04 for Head versus PS4/5, 0.553 for Head versus Posterior, 1.511e-03 for PS4/5 versus Posterior. N indicates the total number of nuclei analyzed in 3-4 embryos. Asterisks indicate that the pairwise difference between samples corresponding to the left column and the other samples is significant.

(C) A characteristic example of an individual nucleus from the posterior segment of the embryo. Dapi staining, the Ubx locus, the Antp locus, and the merge of the three channels are shown. The scale bars represent 1 μm.

See also Table S2.
Figure S2. 4C Analysis, Signal Normalization, Probe Filtering, and Schematic Representation of the Generation of Domainograms from 4C Profiles, Related to Figure 2
(A) Schematic overview of our 4C data analysis. A detailed explanation is provided in the supplemental experimental procedures.
(B) Control MA plots for rotation and normalization of tiling array hybridization signals from F12-r1 (see Figure S3). The purple line (top left panel) indicates the
slope of a bias leading to strong distortions when applying lowess normalization (top right panel). These distortions were reduced by rotation of the raw signals (bottom left panel) before lowess normalization (bottom right panel).

(C) Scatter plot of 4C DNA hybridization levels versus normalized log-ratios from F12-r1 experiment.

(D) Distinction between two classes of the tiling array probes according to DpnII restriction site proximity, with hybridization levels from probes close to DpnII sites (orange) and “far” from DpnII sites (blue), compared to hybridization levels from the whole set of probes (gray) taken from experiment F12-r1.

(E) Control MA plots for probes filtering showing probes next to- (chocolate in left panel; right panel) and “far” from- (blue in left panel; center panel) DpnII sites. Results of the refined probe filtering are indicated in red for best ($B_2$) and green for worst ($W_1$) probes.

(F) Log-ratio distribution centered on DpnII sites from normalized and filtered probes of the F12-r1 experiment.

(G) Sketch of a short 4C profile and its domainogram representation to the right.

(H) Conversion on the y axis domainogram of the log2 number of probe scale to a genomic length scale.
Figure S3. Reproducibility of 4C Experiments and Statistical Analysis of Long-Range Interactions, Related to Figure 2 and Figure 5
(A) Processed 4C profiles as log-ratios on chromosome 3R, showing from top to bottom: WT-r1, WT-r2 (replicates 1 and 2 of 4C experiments in WT), F12-r1 and F12-r2 (replicates 1 and 2 of 4C experiments in the Fab-712 line), with WT-r1 paired with F12-r1, and WT-r2 paired with F12-r2.
(B) Control scatter plots showing per probe reproducibility of processed 4C log-ratios between replicates as well as between paired experiments.
(C) Per replicate 4C profile domainograms on chromosome 3R corresponding to, from top to bottom: WT-r1, WT-r2, F12-r1 and F12-r2 experiments.
(D) Example of a domainogram corresponding to one of the random permutations of the 4C profiles that were used to estimate statistical score thresholds. The random permutation was applied independently for BX-C and non BX-C signals, as indicated below with the black box and the green thick line.
(E) Distribution of maximum scores per probe number from WT (dark blue) and Fab-712 (purple) 4C experiments and randomly permuted signals (black) in non BX-C regions. The 99% confidence threshold is indicated by red (WT) and green (Fab-712) lines (they are at very similar levels and thus superimposed in the graph).
### Figure S4. Rare “Ménage à Trois” between Antp, Abd-B, and lbl/lbe, Related to Figure 3

(A) Three-color FISH with the Abd-B, Antp, lbl/lbe loci. Characteristic examples of individual nuclei. Interactions between Abd-B and Antp and between Abd-B and lbl/lbe are seen in 15%–20% of the nuclei in the head of stage 10–11 embryos. Two examples of “ménage à trois” associations seen in approximately 1% of the nuclei. Interestingly, when the “ménage à trois” associations occur, one of the three loci was systematically unpaired (here, the Abd-B locus was clearly unpaired in the two examples).

(B) Schematic drawing illustrating the different examples seen with three-color FISH. The color circles represent the FISH signals (red for Antp, green for Abd-B, blue for lbl), and the 3R chromosome arm is represented by a thin dashed line.
Figure S5. Chromosomal and Nuclear Positions of Antp and Abd-B in D. virilis, Related to Figure 4

(A) Large-scale comparison of the gene organization around the Hox regions between D. virilis and D. melanogaster. Connecting lines match syntenic regions. Numbers in the boxes represent the approximate size of syntenic regions in kb. Solid lines are used for the Hox syntenic regions, dashed lines for other chromosomal regions. Numbers in D. mel. Chromosome 3R represent the cytological position of each region.

(B) Quantitative analysis of the distribution of FISH 3D distances in different regions of stage 11-12 D. virilis embryos, corresponding to Figure 4B. The mean value of the distances (in µm) and the N, which indicates the total number of nuclei analyzed, are indicated for each distribution. The statistical significance of differences between the distributions was examined by a Wilcoxon test and the p value is indicated under the graph.

Head (mean=0.52, N=192)
Posterior (mean=0.85, N=211)
P = 1.3e-15
Figure S6. Hox Gene Kissing Is Significantly Reduced upon Other Deletions of the Fab-7 or Mcp Elements, Related to Figure 5

(A) FISH in heads of stage 10-11 embryos comparing the percentage of colocalization between Antp and Abd-B loci in the different genotypes. The p values of the pairwise comparison are 5.733e-03 for WT versus Fab-71, 0.071 for WT versus Mcp1, 0.011 for WT versus McpH27Fab-71. N indicates the total number of nuclei analyzed in 3-5 embryos. Blind counting was performed to avoid biases.

(B) Same blind quantification as (A) in third instar larval eye-antennal discs. Approximately the same number of nuclei was counted in the antennal and the eye regions of the discs in 4–5 antennal discs and 3–4 eye discs. The data were pooled for homogeneity and statistical analysis. The p values are 6.103e-05 for WT versus Fab-71, 0.148 for WT versus Mcp1, 0.015 for WT versus McpH27Fab-71. Asterisks indicate that the pairwise difference between samples corresponding to the left column and the other samples is significant.
Figure S7. Loss of Hox Gene Kissing Decreases Silencing and Induces Phenotypic Transformation in the Sensitized Antp<sup>Ns</sup> Background, Related to Figure 7

(A) RT-qPCR analysis of Antp, hh and Abd-B in WT and Antp<sup>Ns</sup> eye-antennal discs (EAD) and WT wing discs (WD). Expression levels were normalized to Rp49 and multiplied by 1.10<sup>4</sup>. Error bars represent the standard deviation of three independent experiments.

(B) FISH in antennal and eye larval discs comparing the percentage of colocalization between Antp and Abd-B loci in WT and Antp<sup>Ns</sup>. N indicates the total number of nuclei analyzed in 3–4 tissues. Blind counting was performed to avoid biases. No significant difference was observed between the two genotypes.

(C) Electron microscopy images showing examples of the antenna-to-leg (A > L) transformations in Antp<sup>Ns</sup> heterozygous flies and in Antp<sup>Ns</sup> heterozygous flies containing deletions of the BX-C or e<sup>1</sup> as a negative control. White arrows indicate the A > L transformations that are stronger in Antp<sup>Ns</sup> heterozygous flies containing BX-C deletions.
(D) Quantification of the A > L phenotypes in heterozygous lines. N indicates the total number of flies counted, which were collected from 2 to 4 independent experiments.

(E) RT-qPCR analysis of Antp and hh. Expression levels were normalized to Rp49 and multiplied by $1.10^3$. Error bars represent the standard deviation of three independent experiments.

(F) Eye ectopic outgrowth phenotypes seen at a low frequency in the different genotypes.