OPEN ACCESS Freely available online

PLoS GENETICS

Interactions among Polycomb Domains Are Guided by Chromosome Architecture

Bas Tolhuis[1]*, Marleen Blom[1], Ron M. Kerkhoven[2], Ludo Pagie[3], Hans Teunissen[1], Marja Nieuwland[2], Marieke Simonis[4], Wouter de Laat[4], Maarten van Lohuizen[1]*, Bas van Steensel[3]*

1 Division of Molecular Genetics and the Centre for Biomedical Genetics, Netherlands Cancer Institute, Amsterdam, The Netherlands, **2** Central Microarray Facility, Netherlands Cancer Institute, Amsterdam, The Netherlands, **3** Division of Gene Regulation, Netherlands Cancer Institute, Amsterdam, The Netherlands, **4** Hubrecht Institute–KNAW and University Medical Center Utrecht, Utrecht, The Netherlands

Abstract

Polycomb group (PcG) proteins bind and regulate hundreds of genes. Previous evidence has suggested that long-range chromatin interactions may contribute to the regulation of PcG target genes. Here, we adapted the Chromosome Conformation Capture on Chip (4C) assay to systematically map chromosomal interactions in *Drosophila melanogaster* larval brain tissue. Our results demonstrate that PcG target genes interact extensively with each other in nuclear space. These interactions are highly specific for PcG target genes, because non-target genes with either low or high expression show distinct interactions. Notably, interactions are mostly limited to genes on the same chromosome arm, and we demonstrate that a topological rather than a sequence-based mechanism is responsible for this constraint. Our results demonstrate that many interactions among PcG target genes exist and that these interactions are guided by overall chromosome architecture.

Citation: Tolhuis B, Blom M, Kerkhoven RM, Pagie L, Teunissen H, et al. (2011) Interactions among Polycomb Domains Are Guided by Chromosome Architecture. PLoS Genet 7(3): e1001343. doi:10.1371/journal.pgen.1001343

Editor: Wendy A. Bickmore, Medical Research Council Human Genetics Unit, United Kingdom

Received August 13, 2010; **Accepted** February 16, 2011; **Published** March 24, 2011

Funding: This work was supported by the NWO-ALW <http://www.nwo.nl>, the Netherlands Genomics Initiative <http://www.genomics.nl>, and the Centre for Biomedical Genetics <http://www.biomedicalgenetics.nl/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: b.v.steensel@nki.nl (BvS); m.v.lohuizen@nki.nl (MvL); b.tolhuis@nki.nl (BT)

Introduction

Polycomb group (PcG) proteins are evolutionary conserved and important epigenetic regulators that bind and modify chromatin. They repress transcription of *Homeotic* genes [1] and many other key regulators of development [2–6]. As such, PcG proteins control morphogenesis, cellular differentiation, proliferation, and cancer development [7]. PcG proteins are thought to mediate gene silencing by changing the structure of chromatin, but the exact molecular mechanisms are not yet fully understood.

In *D. melanogaster*, biochemical studies have indicated that PcG proteins exist in at least three multiprotein complexes [reviewed in: [1]]. First, the Pleiohomeotic Repressive Complex binds directly to DNA sequences. In addition, there are two Polycomb Repressive Complexes (PRC1 and PRC2) that bind indirectly to DNA via protein-protein interactions. PRC2 includes the histone methyl transferase E(Z), which deposits the characteristic lysine 27 trimethyl mark at histone H3 (H3K27me3). In turn, the chromodomain of Polycomb (PC), one of the components of PRC1, binds to this mark.

Recent genome-wide mapping studies revealed that PC and H3K27me3 tend to form large contiguous chromatin domains, which sometimes span more than 100 kb in size covering multiple genes [5,6,8,9]. These domains have been named Polycomb domains (PcDs). Other PRC1 components (PSC, and PH) and E(Z) bind as specific local enrichments within PcDs [5,9]. Some of these local enrichments coincide with well-known *cis*-regulatory sequences, called Polycomb Response Elements (PREs). PREs are essential for PcG-mediated gene silencing and thought to be initial recruitment sites for PcG proteins [1].

There are some indications that PcG proteins can be involved in long-range chromatin interactions in nuclear space. First, in *D. melanogaster*, some PREs interact over long distances with gene promoter and termination sequences, which is essential for PcG-mediated silencing [10,11]. Likewise, PcG proteins mediate the formation of a complex multi-loop chromatin structure around the silenced mammalian GATA4 locus [12]. Second, two PcDs, the two fly *Homeotic* gene clusters (Antennapedia complex – ANT-C and Bithorax complex – BX-C) were found to co-localize in nuclei of certain tissues [13,14]. This interaction is intriguing, because the *Homeotic* genes are separated by 10 Mbp. Similarly, in mammalian cells, long-range chromatin interactions were shown to be mediated by EZH2 protein (an E(z) homolog) [15]. Third, PcG proteins accumulate as microscopically visible foci in the nuclear space [14,16–19], suggesting that PcG-bound chromatin is concentrated in nuclear compartments.

These observations raise the question whether long-range chromatin interactions among PcDs are a general phenomenon. However, a systematic survey to address this question is still lacking. Therefore, we decided to map the genome-wide interactions of 4 established PcG target genes. We adapted the Chromosome Conformation Capture on Chip (4C) assay [20] to generate genome-wide maps of PcD interactions in *D. melanogaster* larval brain tissue. Our results indicate that PcDs extensively and specifically interact with each other, yet these interactions are topologically constrained by the overall chromosome architecture.

Author Summary

The folding of chromosomes inside the cell nucleus is a fascinating yet poorly understood topological problem. It is thought that certain genomic loci that are distant in the linear genome may come together in nuclear space by folding of the chromosome fiber. Previously, such a long-range interaction was found in *Drosophila* for two genomic loci that are known to be bound by the Polycomb Repressive Complex. Because hundreds of genes are known to be bound by Polycomb proteins, we asked whether such long-range contacts are more common. To address this, we optimized the Chromosome Conformation Capture on Chip (4C) technology for use in small tissue samples. Using this technique in dissected larval brains, we found that indeed Polycomb target genes interact frequently with each other, even when they are separated by megabases of sequence. However, these long-range interactions occur almost exclusively on the same chromosome arm. By using a rearranged chromosome in which segments are swapped between two arms, we demonstrate that not DNA sequence but chromosome architecture imposes this restriction. Taken together, our data demonstrate that Polycomb target genes extensively interact in nuclear space, but only when they are located on the same chromosome arm.

Results

Experimental design

We applied Chromosome Conformation Capture on Chip (4C) to map long-range chromatin interactions among PcDs on a genome-wide scale [20]. We initially focused on the *Abdominal-B* (*Abd-B*) and *Antennapedia* (*Antp*) genes, because they are part of the Bithorax (BX-C) and Antennapedia (ANT-C) complexes, which were found to interact in nuclear space by microscopy [13,14]. Importantly, these interactions occur more frequently when these *Homeotic* genes are silenced [14]. We therefore used central brain tissue from 3rd instar larvae, in which both *Antp* and *Abd-B* are repressed by PcG proteins (Figure S1) [21]. In addition, when we mapped Polycomb (PC) binding and H3K27me3 levels in central brain tissue (Text S1 and Table S1), we observed that both genes reside in chromatin domains with high levels of PC and H3K27me3 (Figure S2A, S2B). Hence, we expected *Abd-B* and *Antp* to interact in nuclear space of central brain cells.

The available 4C assay requires large numbers of cells due to a relatively low efficiency in detecting long-range interactions [22]. To accommodate this requirement, we dissected and pooled central brain hemispheres from 40 larvae. However, this still did not yield enough material for subsequent microarray analyses (data not shown). To overcome this, we implemented a modification of the 4C protocol in which the 4C PCR products are further amplified in a linear fashion using a T7 RNA amplification procedure [23]. Finally, we fluorescently labeled the amplified products with dye and hybridized them to a specially designed microarray, which covers approximately 92% of the non-repetitive fly genome. In this way, we could identify all fragments that are in close contact with a chosen locus in larval brain cells.

Detection and visualization of long-range chromatin interactions

We used recently developed [24] and novel computational tools to detect and visualize long-range chromatin interactions. Here, we briefly describe the technical aspects of these tools using 4C assay results obtained with the *Abd-B* promoter region (hereafter: *Abd-B*) as bait.

In a typical 4C experiment, reliable interactions can only be detected as genomic clusters of multiple restriction fragments with increased signals [22]. Such clusters can be identified using a sliding window filter that improves the signal/noise ratio by integrating signals from neighboring probes [20]. Indeed, after applying a running median filter (window size = 51 probes) we observed genomic clusters with increased ratios on the right arm of chromosome 3 (3R, Figure 1A). Disadvantages of this method are that it is difficult to address the statistical significance of these interactions and the choice of window size is arbitrary, limiting the identification of interacting regions to a certain size range. We used instead a multi-scale algorithm that evaluates statistical significance at all possible window sizes [24]. The results are visualized in so-called domainograms (Figure 1B) in which statistically significantly interacting regions are depicted in red, while non-interacting regions are black. This approach implements both unbiased statistics and detection of interactions at multiple scales. While domainograms are a powerful means to visualize statistically significant interactions at all scales, they do not identify the exact boundaries of interacting regions. For this purpose we used an adapted version (see Materials and Methods) of a dynamic programming algorithm [24], which identifies the statistically most probable discrete interaction domains (DIDs). DIDs represent the center of gravity of a long-range chromatin interaction (Figure 1B). We note that DIDs are an oversimplification of the 4C interaction maps, but are nevertheless useful for subsequent computational analyses.

Finally, in 4C assays the strongest signals are found centered around the chromosomal position of the bait, because interaction frequencies are inversely proportional to the distance (in base pairs) between physically linked DNA sequences [20,25]. Indeed, we observed a broad region of interacting sequences surrounding *Abd-B*, which spans several megabases and includes a large proportion of all DIDs (Figure 1B). It seems likely that many of these interactions are caused by linear proximity rather than by spatial proximity as a result of higher-order chromatin folding. Because our aim was to identify interactions caused by higher-order chromatin folding, we modeled the linear proximity effect by fitting a monotonously declining smoothing curve to the data (Figure 1A). After subtraction of this background model from the interaction data, the number of interactions close to the bait has been dramatically reduced, while long-range interactions are still detectable (Figure 1C).

Homeotic genes prefer to interact with other PcDs

To capture long-range contacts using *Abd-B* as bait, we hybridized two biological replicates of amplified PCR products to our custom microarray. The domainogram of the combined data revealed that neither chromosome 3L nor any other chromosome showed significant enrichments, whereas chromosome 3R holds all 17 DIDs (Figure 2B, and Table S2). Importantly, three centromere proximal DIDs overlapped with ANT-C, which is in line with the previously reported microscopic data [13,14].

Next, we wanted to know how many DIDs overlap with Polycomb domains (PcD) and H3K27me3 domains (K27D). For this purpose, we overlaid the *Abd-B* DIDs with either the PcD or K27D maps from central brain tissue (Figure 2B, and Table S1). We observed that 88 and 94% of the DIDs overlap with PcDs and K27Ds, respectively, which is significantly more than may be expected by random chance (Figure 3A, 3B, and Table S2). Thus, our data provides evidence that *Abd-B* has a preference to interact with other PcDs. In total, we identified 25 PcDs that are in contact with the *Abd-B* promoter, including ANT-C, *grain* (*grn*), *homothorax* (*hth*), *lbl/lbe*, and *Ptx1* (Table S1).

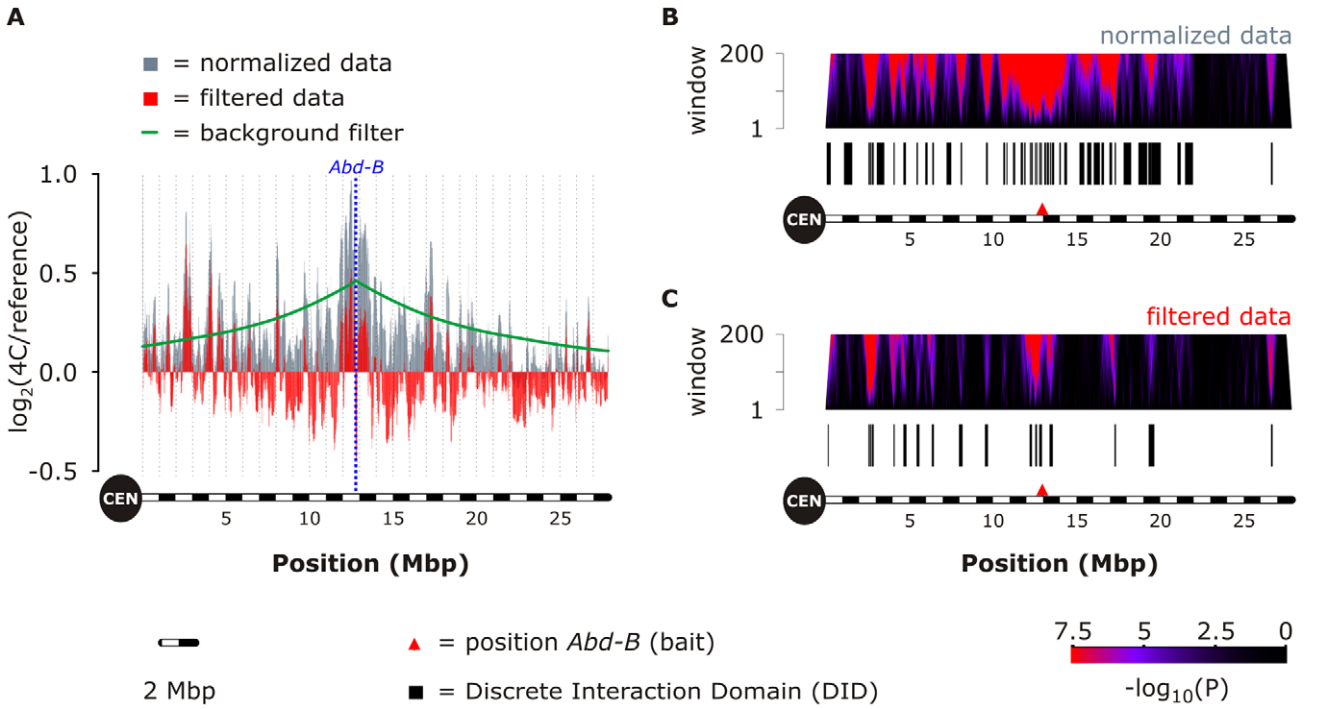

Figure 1. Detection and visualization of long-range chromatin interactions. (A) Genomic clusters of multiple restriction fragments with increased 4C signals were identified using a sliding window filter on the normalized data (running median, window = 51 probes; grey bars) along chromosome 3R. We applied a "background filter" to correct for interactions that are a consequence of short nucleotide distance to the bait sequence (green curve). After filtering the normalized data, clusters of restriction fragments that exceed the curve (running median, window = 51 probes; red bars) represent long-range chromatin interactions. X-axis indicates the chromosomal position (in Mbp), while $\log_2$-transformed 4C signals are shown on the y-axes. Position of *Abd-B* is blue dashed line.(B) Domainograms visualize long-range chromatin contacts detected using normalized data obtained with the *Abd-B* promoter as bait (red arrowhead). X-axis indicates the chromosomal position (in Mbp), while the window size of the domainogram algorithm is shown on the y-axis. Interactions are plotted using a color scale with black for non-significant and red for highly significant (scale bar). Only window size ≤200 probes (corresponding to 840 Kbp on average) are shown. We used a dynamic programming algorithm to identify non-overlapping Discrete Interaction Domains (DIDs, black rectangles bottom). (C) Domainogram of *Abd-B* data after applying the "background filter." Plotted as in B. Note, centromeric sequences (CEN) are not drawn to scale.
doi:10.1371/journal.pgen.1001343.g001

We performed a reciprocal experiment, using the *Antp* P1 promoter as bait (hereafter: *Antp*; Figure 2A, Figure S2). We aimed to: (1) confirm the interaction between BX-C and ANT-C; and (2) assess whether this bait also prefers to interact with other PcDs. Indeed, a highly significant proportion of all DIDs overlap with PcDs and K27Ds (Figure 3A, 3B, Table S2), including an interaction with the BX-C (~12.5–13.0 Mbp, Figure 2C). Finally, we identified 26 PcDs that are in close contact with *Antp*, of which 16 also interact with *Abd-B* (Table S1).

As for *Abd-B*, most interactions take place within chromosome 3R, which contains the *Antp* bait sequence. However, we found one interaction with a PcD located on chromosome 4, demonstrating that interactions between two different chromosomes can occur, albeit with low frequency.

Thus, we confirmed the previously reported long-range interaction between the *Homeotic* gene clusters [13,14], validating the specificity of our 4C assay. Importantly, we observed that the majority of other interacting regions overlap with PcDs, indicating that *Abd-B* and *Antp* both interact extensively with other PcDs.

Non-*Homeotic* PcG targets prefer to interact with other PcDs

Next, we asked whether non-*Homeotic* PcG target genes also prefer to contact other PcDs. We performed experiments using *Ptx*1 and *trachealess* (*trh*) as baits. The results are comparable to those obtained with *Abd-B* and *Antp*. First, most interactions are again mostly restricted to the chromosome arm that contains the bait (Figure 2D, 2E; Figure S2). Second, the observed overlap between DIDs and PcDs/K27Ds is highly signficant (Figure 3A, 3B). Altogether, we conclude that *Ptx1* and *trh*, like the *Homeotic* genes, preferentially interact with other PcDs.

Non-PcG target loci lack preference for PcDs

While the vast majority of all DIDs of the four PcG target genes coincide with PcDs, a small subset lacks PcDs, suggesting that PcG-target genes can sometimes interact with other chromatin types. In mammalian cells, it has been found that the genome spatially segregates into open/active and closed/inactive compartments [20,26]. Therefore, we tested the possibility that our largely inactive PcDs can interact with other inactive genome regions and together make an inactive nuclear compartment.

We first generated microarray expression profiles of brain tissue to assess the transcriptional activity of genes in DIDs. We did not observe any gene expression differences when we compared genes located inside DIDs to genes from other parts of the chromosome arm (Figure S3). This suggests that low transcriptional activity in DIDs cannot simply account for the observed long-range interactions. We note however that the resolution of 4C may be rather low, because most DIDs are large genomic regions (average size = 170 Kbp), and contain many genes (on average 22) that vary greatly in expression level.

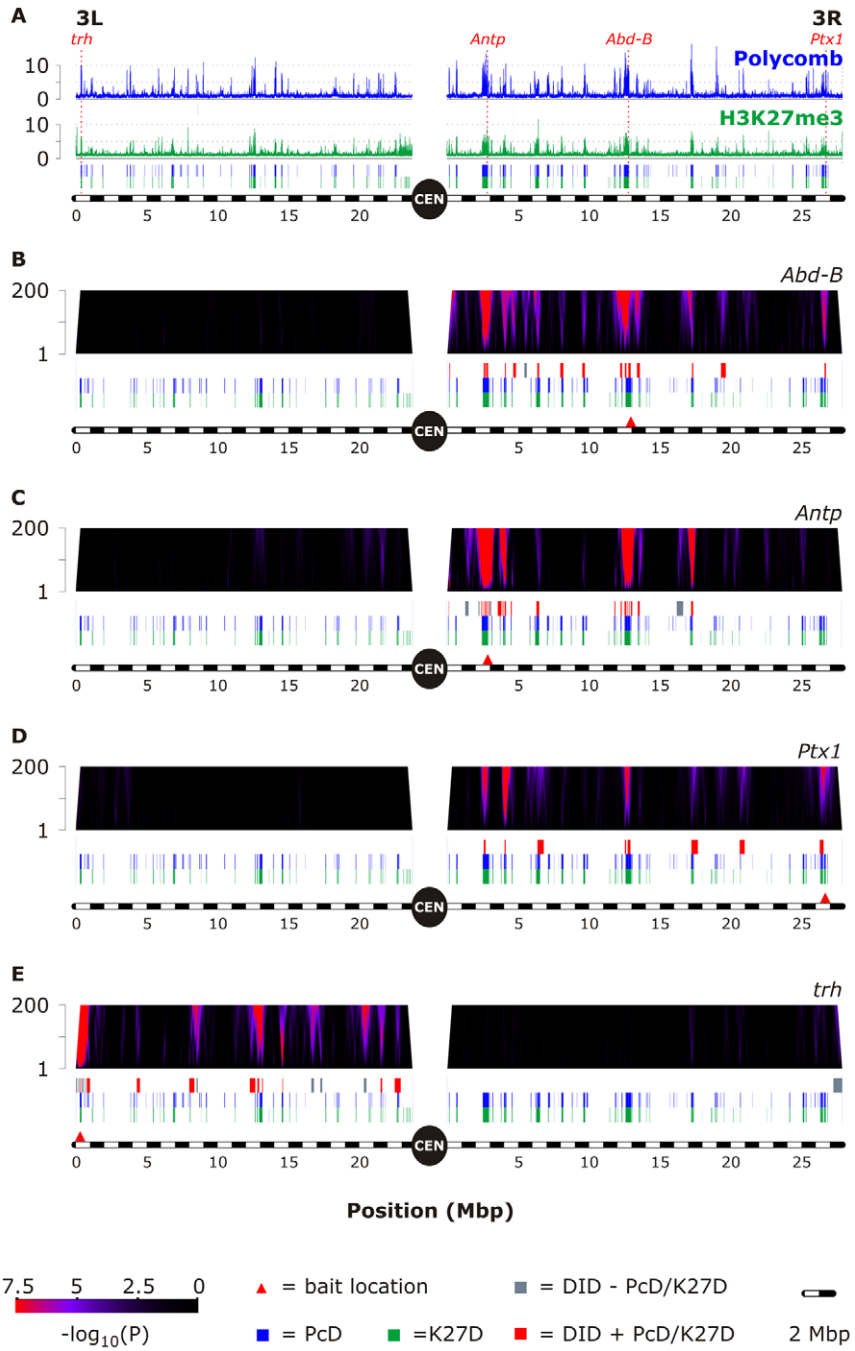

Figure 2. Long-range interactions of PcG target genes. (A) Chromatin profile of third chromosome representing Polycomb binding (blue) and H3K27me3 levels (green) as detected in 3rd instar larval central brain cells. Below graphs, the blue and green rectangles show Pc (PcD) and H3K27me3 (K27D) domains, respectively. X-axes indicate the chromosomal position (in Mbp), while the fold-change (either PC-Dam over Dam or H3K27me3 over IgG) is shown on the y-axes. Positions of the 4 PcG target baits are indicated (red). (B–E) We mapped long-range chromatin interactions using four different PcG-targets as bait, namely: *Abdominal-B* (*Abd-B*) promoter (B), *Antennapedia* (*Antp*) P1 promoter (C), *Ptx1* (D), and *trachealess* (*trh* - E). Plots are described in Figure 1B. Notably, the majority of all DIDs (red boxes) coincide with either PcDs (blue boxes) and/or K27Ds (green boxes), and a minority (grey boxes) contacts other sequences.
doi:10.1371/journal.pgen.1001343.g002

Next, we determined interactions using bait sequences containing lowly expressed genes, namely *Wnt inhibitor of Dorsal* (*wntD*) and *CG5107*. The mRNA expression levels of *wntD* and *CG5107* are equally low as the *Antp* and *Abd-B* genes (Figure 3C), but they are not part of PcDs (Figure 4A).

The interaction maps of neither the *wntD* nor the *CG5107* gene show enrichment for PcDs (Figure 4B, 4C). Less than 15% of all DIDs coincide with PcDs (Figure 3A, 3B, Table S2), which is 1.5 to 2-fold less than an overlap based on random chance. These results indicate that interactions between PcG target genes cannot be attributed to general interactions between transcriptionally inactive loci.

We also observed that actively transcribed genes do not favor interactions with PcDs. For this purpose, we used the highly expressed *Calreticulin* (*Crc*) and *RNA polymerase II 140 kD subunit* (*RpII140*) genes as bait (Figure 3C), which are not PcG-targets (Figure 4A). Again, we observed that only a small, non-significant

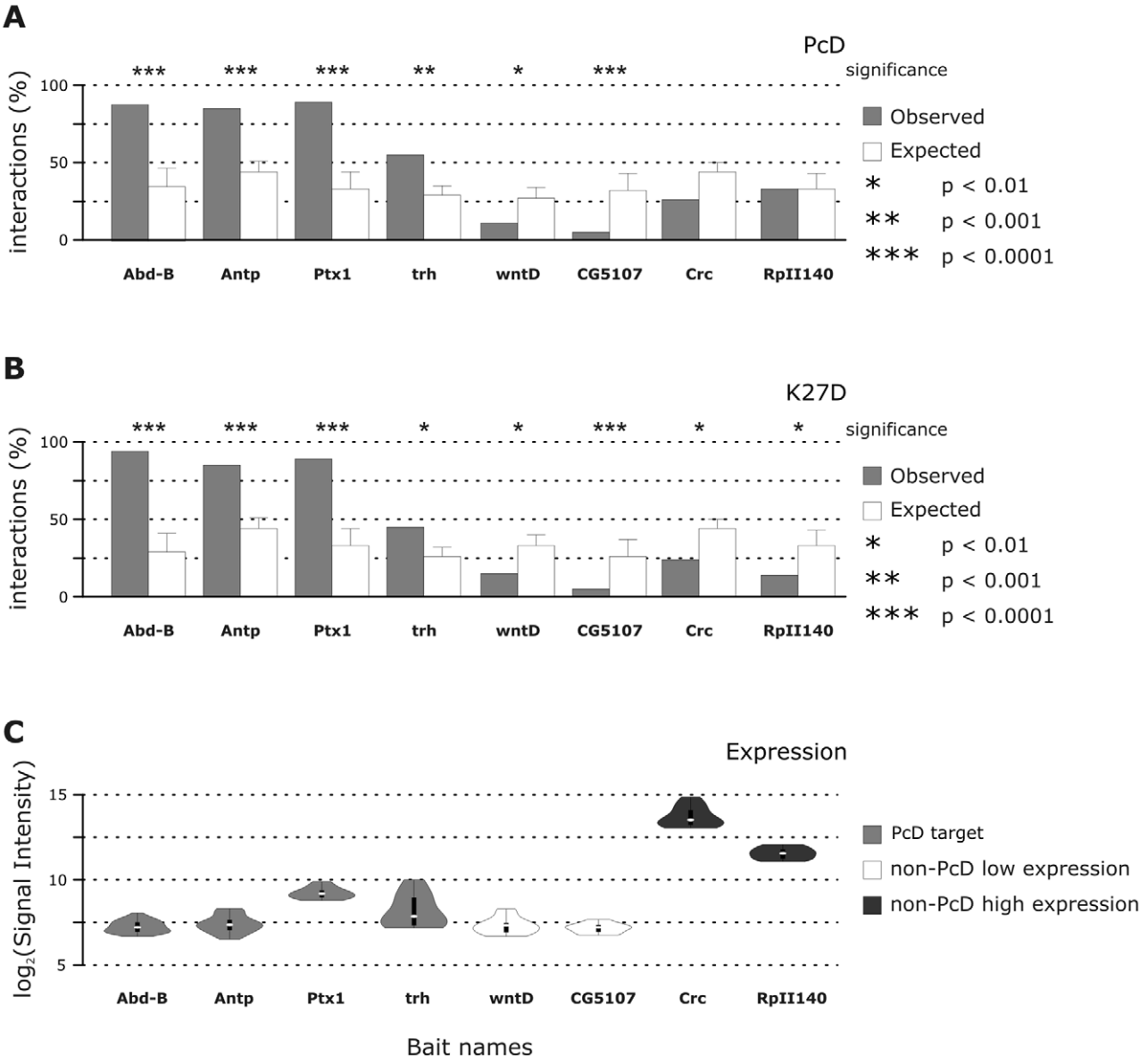

Figure 3. PcG targets interact specifically with PcDs and K27Ds. (A) Proportion of DIDs (y-axis) that overlap with PcDs is shown (observed - grey bars) and compared to an overlap based after random permutation (expected – white bars). Random sampling was repeated 1000-times and the bars show the average random overlap, while the error-bars show the standard deviation. Significance levels were determined using Pearson's chi-square (χ^2) goodness of fit test and are indicated. Bait names are given (x-axis). (B) Proportion of DIDs that overlap with K27Ds. Plot details are described in A. (C) mRNA expression levels of all baits as detected in 3rd instar larval central brain cells are shown in violin plots. Violin plots are a combination of box plots and smoothed histograms. Median values are indicated by white dot, lower and upper quartiles by black rectangle, and sample minima/maxima by black line. Smoothed histograms of the four PcG-target baits are depicted in light grey, lowly expressed non-target genes in white and highly expressed genes in dark grey. Data represents four replicate microarray experiments. Genes were probed multiple times in each experiment, giving 48 (*Antp*), 32 (*Abd-B*), and 16 (all other genes) data points. Notably, *Ptx1* is higher expressed than the three other PcG targets, because it is expressed in a subset of larval brain cells (Figure S6, Text S1).
doi:10.1371/journal.pgen.1001343.g003

fraction of DIDs match with PcDs (Figure 3A and 3B, Figure 5D and 5E, and Table S2).

Taken together, our data strongly indicates that long-range interactions of the four PcG-target genes are highly specific with a strong preference for other PcG target loci.

Large PcDs interact more often

When we further characterized interacting PcDs we observed two distinctive features. First, interacting PcDs encompass significantly larger genomic regions than non-interacting PcDs (Figure 5; Table S3). Second, we overlaid our PcD map with a computational prediction of PRE sequences [27] and observed that interacting PcDs contain significantly more predicted PRE sequences than non-interacting PcDs (Table S3). In contrast, we observed no differences for gene density, repetitive DNA sequences, and non-coding evolutionary conserved sequences (Table S3). These results suggest that, at least for our 4 PcG baits in larval brain, increased size and presence of PREs may contribute to the frequency of interaction between two PcDs.

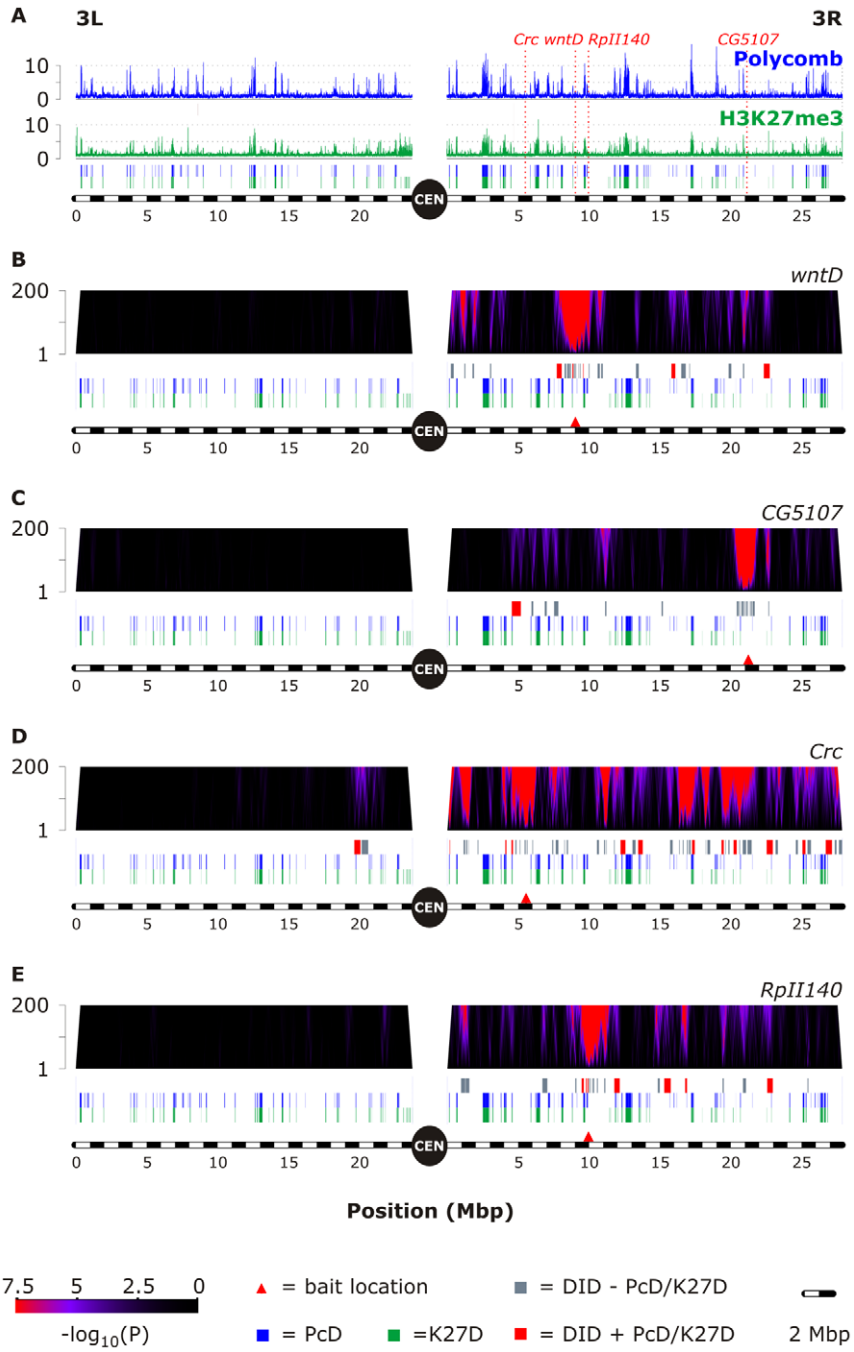

Figure 4. Non-target genes interact with other sequences than PcG target genes. (A) Position of four 4C baits that are not PcG target genes (red), plotted as in Figure 2A. (B–E) Long-range chromatin interactions for *Wnt inhibitor of Dorsal* (*wntD* - B), *CG5107* (C), *Calreticulin* (*Crc* - D), and *RNA polymerase II 140 kD subunit* (*RpII140* - E). Plots are as in Figure 1B and Figure 2.
doi:10.1371/journal.pgen.1001343.g004

Long-range interactions are predominantly limited to a single chromosome arm

We noticed that long-range interactions of all our baits predominantly occur within the same adult chromosome arm (Table S2). For example, *Abd-B* only interacts with sequences on chromosome 3R (Figure 6A). However, we also identified rare cases of interactions between loci on different chromosome arms. For instance, the *Ptx1* locus (chromosome 3R) interacts with a discrete region on the X, and *Antp* with a PcD on chromosome 4 (Table S2). The highly expressed *Crc* (chromosome 3R) gene interacts with two regions on chromosome 3L and one on chromosome 4, indicating that not only PcG targets can interact with other chromosome arms. Finally, the *trh* gene (chromosome 3L) shows the highest frequency of interactions with other chromosome arms (Figure 6A), although this only represents 19% of all DIDs (i.e. 6 out of 31).

We observed that between 81 and 100% of all interactions for all eight bait sequences occur within a single chromosome arm (Figure 6B). When we combined all data, we observed that on average 95% of all interactions (n = 201) are limited to a single

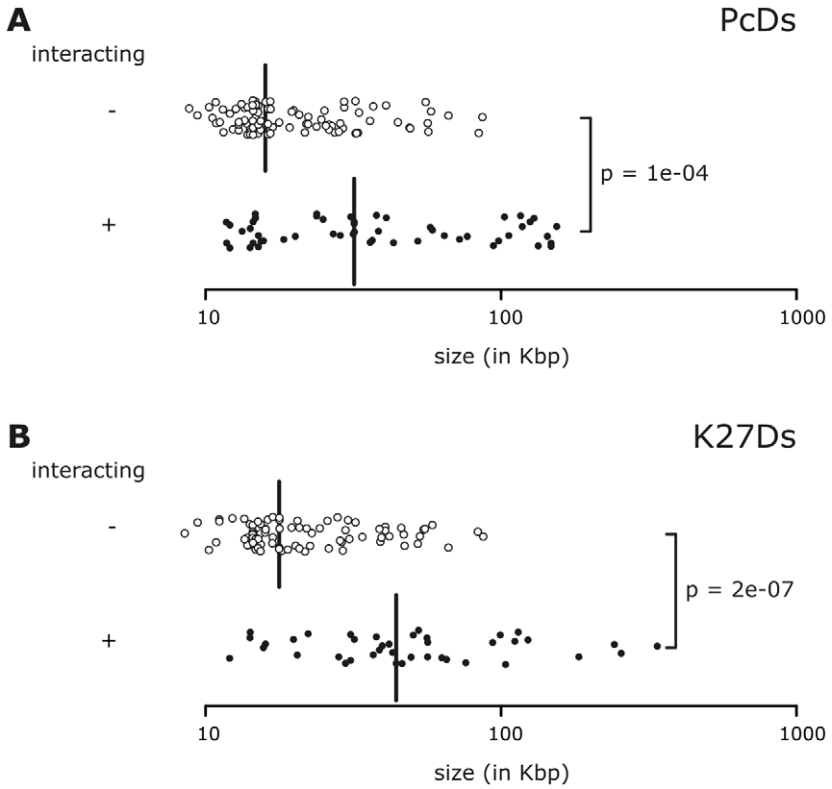

Figure 5. Large PcDs interact more often than small PcDs. (A) Strip charts show size differences between PcDs that interact with at least one of the 4 PcG baits (closed circles) and PcDs that do not interact (open circles). (B) Similar to a, but now for K27Ds. The x-axes indicate domain size in Kbp ($\log_{10}$ scale), and each dot represents a single domain on chromosome 3. Perpendicular bars are median values and p-values were calculated using a two-sided non-parametric Mann-Whitney U test.
doi:10.1371/journal.pgen.1001343.g005

chromosome arm (Figure 6B). Thus, interactions within a chromosome arm are prevalent and contacts between sequences on different chromosome arms are rare.

Interactions are constrained by chromosome architecture

What mechanism limits interactions to a single chromosome arm? To date, it is well established that chromatin contacts can occur between *cis*-acting DNA elements [28]. It is possible that long-range chromatin interactions are driven by high-affinity interacting DNA elements, which are confined to a single chromosome arm. Alternatively, it is possible that interactions are topologically constrained by overall chromosome architecture, for example because each chromosome arm is a spatially confined unit.

We tested these two models by analyzing PcD interactions using fly strain *In(3LR)sep* that carries a chromosomal inversion [29]. This inversion (Figure 7A) rearranges the centromere-proximal ~5.8 Mbp of chromosome 3R to the left arm [29]. This region contains 17 PcDs, including the ANT-C and *grn*. The other PcDs on 3R, including those annotated as *hth*, BX-C, *lbl/lbe*, and *Ptx1*, remain on the right arm. In addition, approximately 17 Mbp of chromosome 3L, including 43 PcDs, is moved to the right arm. For clarity, we will refer to the wild-type left and right arms as $3L^{WT}$ and $3R^{WT}$, and to the inverted arms as $3L^{IN}$ and $3R^{IN}$, respectively. We reasoned that if long-range contacts are only possible within the same arm, then on the inverted chromosome interactions across the inversion breakpoints, such as between BX-C and ANT-C, will be abolished and new interactions may be formed. On the other hand, if interactions depend solely on local *cis*-regulatory DNA sequences then they will not be affected by the inversion.

To discriminate between these possibilities, we first mapped the long-range chromatin contacts in central brain nuclei of homozygous *In(3LR)sep* larvae using *Abd-B* as bait. For the domainograms (Figure 7B), we plotted the chromosomal position of probes identical to wild type, which facilitates direct comparison to wild type data (Figure 2B). Remarkably, all significantly enriched regions are located on $3R^{IN}$ (Figure 7B, Figure 8A, and Table S4), demonstrating that long-range contacts are again limited to a single chromosome arm. Specifically, among PcDs that remained to the right of the centromere we observed contacts that we previously identified in wild type cells, including *hth*, *lbl/lbe*, and *Ptx1* (Table S4, and Figure 7B). In contrast, all interactions on the inverted 5.8 Mbp, including ANT-C, were abolished. Instead, new contacts occur on the 17 Mbp that is inverted from chromosome $3L^{WT}$ to $3R^{IN}$ (Figure 7B), including PcG targets *dorsocross* (*doc*) genes and Iroquois-complex (IRO-C; *araucan*, *caupolican*, and *mirror*).

We observed similar results when we mapped interactions using *Antp* as bait. Several interactions that occur on $3R^{WT}$, such as with BX-C, *hth*, and *lbl/lbe* are lost on $3R^{IN}$ and many new contacts are found on $3L^{IN}$ (Figure 7C, Table S4). Interactions are mostly limited to $3L^{IN}$, and only 2 out of 18 DIDs occur with sequences on chromosome 4 (Figure 8A, Table S4). Similar to wild type, we observed that ~95% of all *Abd-B* and *Antp* interactions (n = 37) occur within a single chromosome arm (Figure 8A).

Next, we tested whether the dramatically changed interactions affect expression of genes on the *In(3LR)sep* chromosome. In a series of microarray expression profiling experiments we could not

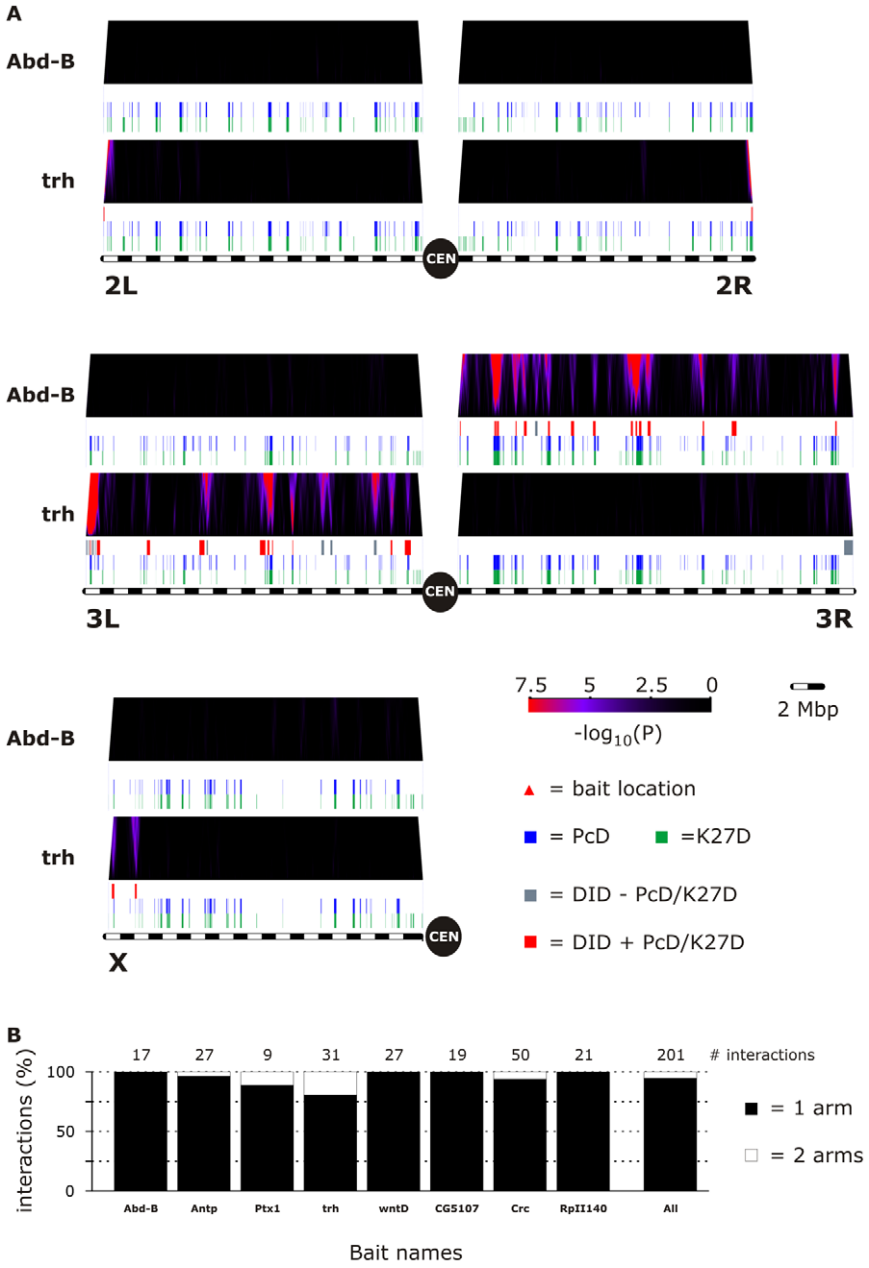

Figure 6. Long-range interactions predominantly occur within a single chromosome arm. (A) Domainograms show two illustrative examples of genome-wide interaction maps using either *Abd-B* (top) or *trh* (bottom) as bait. Plots are described in Figure 1B and Figure 2. Note, the small repeat-rich chromosome 4 is not shown. (B) Bar plots depict fraction of interactions within bait containing chromosome arm (black) and interactions with other chromosome arms (white). Bait names are indicated (x-axis) as well as absolute number of interactions (top). Across eight data sets we identified 201 interactions (all), of which ~95% are limited to a single chromosome arm, which also contains the bait sequence, and only 5% occurs between sequences on two chromosome arms.
doi:10.1371/journal.pgen.1001343.g006

detect significant changes relative to wild type gene expression, when we compared genes residing on either $3L^{IN}$, $3R^{IN}$, or other chromosome arms (Figure S4). We additionally performed genetic experiments to test whether gene expression is altered on the *In(3LR)sep* chromosome. *Sex combs reduced* (*Scr*) expression levels in larval imaginal discs correlate strongly with the number of sex comb teeth on adult male 1st thoracic legs [30,31]. We further sensitized *Scr* expression by crossing either the wild type or *In(3LR)sep* chromosome to several *Scr*-mutant alleles. As expected, we observed approximately 6 to 7 sex comb teeth per leg on average for the *Scr*/$3R^{WT}$ (functional *Scr* allele on wild type chromosome 3R) heterozygous animals (Figure S4) [31]. We observed similar numbers on *Scr*/$3L^{IN}$ flies, which carry the functional *Scr* allele on inverted chromosome (Figure S4). Taken together, our experiments failed to detect altered gene expression on the *In(3LR)sep* chromosome.

We further examined the interactions on the *In(3LR)sep* chromosome, and observed that, like on the wild type chromosome, *Abd-B* and *Antp* prefer to contact other PcDs. We observed PcD-specific interactions for 4 out of 6 new *Abd-B* contacts on the 17 Mbp that rearranged from $3L^{WT}$ to $3R^{IN}$ (Figure 7B, left). In addition, *Antp* showed 7 PcD-specific interactions out of 10 new

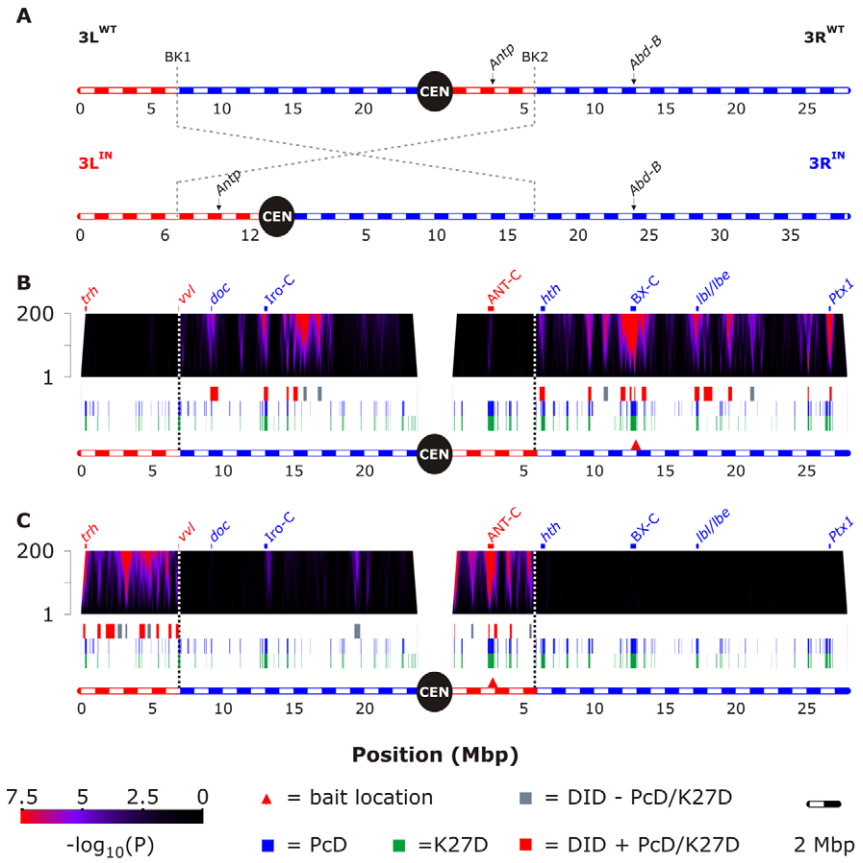

Figure 7. Interactions are constrained by chromosome architecture. (A) A schematic overview of the wild type and inverted chromosome 3. The wild type strain used in this study (w^{1118}) is shown on top. Dashed lines indicate the annotated positions of the break points (BK1 and BK2) present in the inverted chromosome of *In(3LR)sep* (bottom). Sequences on the inverted left arm $3L^{IN}$ are red and $3R^{IN}$ blue. Positions of the baits are indicated. (B–C) Domainograms show interactions of *Abd-B* (B) and *Antp* (C) on the inverted third chromosome. Note that all domainograms are plotted in the wild type orientation, which allows direct comparison with wild type data (Figure 2B, 2C). The colors on the x-axes indicate location on inverted chromosome (red = $3L^{IN}$; blue = $3R^{IN}$). Exact positions of the break points are highlighted by dashed vertical lines. Rest of plot details are described in Figure 1B and Figure 2.
doi:10.1371/journal.pgen.1001343.g007

contacts with sequences that originate from $3L^{WT}$ and are now on $3L^{IN}$ (Figure 7C, left). All together, a highly significant fraction (70% for *Antp* and 74% for *Abd-B*) of all DIDs coincide with PcDs (Figure 8B) and K27Ds (data not shown).

In summary, we conclude that on the *In(3LR)sep* chromosome *Abd-B* and *Antp* preferentially contact PcDs that are located on the same chromosome arm, demonstrating that interactions are constrained by the overall chromosome architecture.

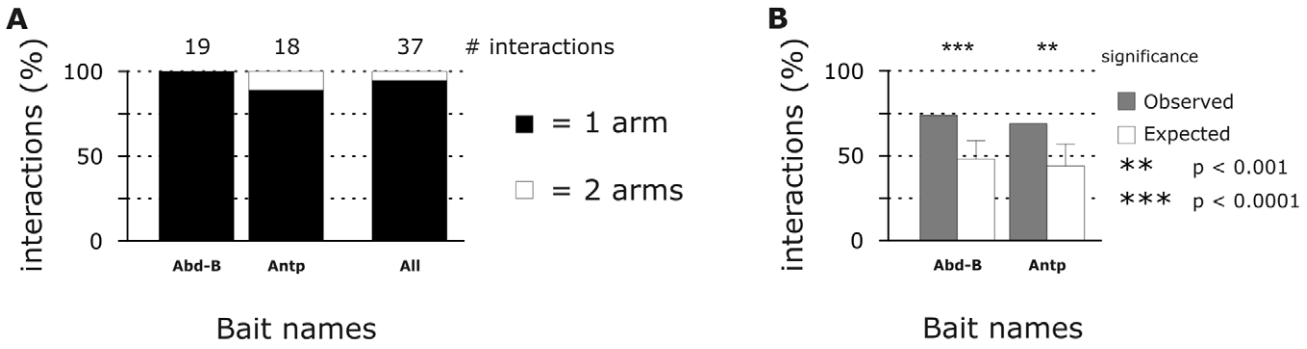

Figure 8. On the inverted third chromosome *Homeotic* genes preferentially contact PcDs located on the same chromosome arm. (A) Bar plots show quantification of interactions within single chromosome arm and between two chromosome arms in brain tissue from *In(3LR)sep* larvae. Plot is described in Figure 6B. Like on the wild type third chromosome (Figure 6B), on the inverted third chromosome 95% of all interactions are intrachromosomal and only 5% is interchromosomal. (B) Proportion of DIDs that overlap with PcDs. Plot is described in Figure 3A.
doi:10.1371/journal.pgen.1001343.g008

Discussion

We successfully adapted the 4C method to systematically map long-range chromatin contacts with limited material from a single fly tissue. With this method, we detect interactions between the ANT-C and BX-C in central brain. This observation is in good agreement with earlier microscopic reports [13,14], underscoring the strength of our method. Importantly, we further demonstrate that not only the two *Homeotic* gene clusters, but also many other PcG target genes interact, suggesting that long-range chromatin contacts between PcDs are common in central brain tissue. The control fragments (*wntD*, *CG5107*, *Crc*, and *RpII140*), which do not reside in PcDs, have interactions that are distinct from PcDs, emphasizing the specificity of our findings.

PcG targets show a strong preference for interaction with other PcG targets, suggesting that PcG proteins help to establish these interactions. This is in line with earlier *in vivo* and *in vitro* results that indicated that PcG proteins can keep certain DNA sequences together [10,11,14,15,32]. However, interactions among PcG target genes are constrained by overall chromosome architecture, because our data demonstrate that loci need to be on the same chromosome arm for efficient interaction

Interpretation of the 4C data

DIDs range in size from 6 to 600 kb, with an average of ~170 kb. Thus, we find highly local strong enrichments as well as moderate enrichments over larger regions, which may reflect different types of long-range interactions. We emphasize that interactions as detected by 4C technically represent events of molecular proximity of DNA sequences, and not necessarily physical binding. Based on the current data it is therefore not possible to identify within the DIDs sequence elements that may mediate direct contact with other DIDs. It is conceivable that contacts between PcDs may occur at any position within the PcDs; if PcG protein complexes have an intrinsic propensity to aggregate, as has been observed in vitro [33], then large PcDs may have a higher chance of interacting with each other due to their larger 'sticky' surface area.

The 4C method is a cell population based assay that only detects the most frequent interactions in the population. Previous 4C studies in mammalian cells suggested an extensive network of long-range interactions [20]. Our data also suggests an extensive network among interacting PcDs. However, microscopic studies in mammalian cells revealed that specific long-range interactions occur only in a small proportion of the cells [20]. Likewise, only a proportion of *D. melanogaster* cells show contacts between the two *Homeotic* clusters [13,14]. Therefore, 4C data have to be carefully interpreted, because the identified interactions are in part stochastic and do not all occur simultaneously [20]. As a consequence, we do not know how many PcDs interact in a single cell, but we do know the most common interactions in the population of larval brain cells.

Chromosome arm territories

Interphase chromosomes in most eukaryotes occupy distinct "territories" inside the nucleus, with only a limited degree of intermingling [34,35]. Although some studies have reported interactions between some loci that are on two different chromosomes (interchromosomal) [reviewed in: [28]], unbiased 4C mapping in mouse tissues has indicated that interactions within the same chromosome (intrachromosomal) occur much more frequently than between different chromosomes [20]. In addition, a recent genome-wide map of chromatin interactions in human cells showed that intrachromosomal interactions occur with higher frequency than interchromosomal contacts [26]. Our data are in agreement with these observations, and show that most interactions are even limited to single chromosome arms, at least in *D. melanogaster* larval brain. Since our experiments indicate that a topological mechanism prevents interactions between the two arms of a chromosome, we propose that each arm (rather than the chromosome as a whole) forms a distinct territory. This is consistent with early microscopy studies of chromosome architecture in *D. melanogaster*, which suggested that chromosome arms are units of spatial organization [36–38].

What topological mechanism may limit contacts between the two chromosome arms? About ~16 Mbp of pericentric heterochromatin are located in between the two euchromatic arms of chromosome 3 [39]. This heterochromatin region could act as a long spacer and prevent efficient interactions between DNA fragments that are located on either chromosome arm. However, our data show that interactions within one arm can span even longer distances, such as between *Ptx1* and *grn* (~22.7 Mbp). Another explanation may be that the pericentric regions of all chromosomes assemble into a nuclear compartment, called chromocenter [40]. This large structure could physically obstruct interactions between chromosome arms.

Previous reports have demonstrated that certain PcG-bound PREs can pair *in trans* (i.e. when they are located on different chromosomes). First, a *Fab7* PRE-element integrated on the X chromosome (*Fab-X*) was often found in close spatial proximity to the endogenous *Fab-7* in the BX-C (chromosome 3R) [41], although this phenomenon appears to be tissue-specific and dependent on the transgene integration site [13]. Second, a microscopy assay based on Lac repressor/operator recognition showed that the *Mcp* PRE-element is able to pair with copies of that same element inserted at remote sites in the genome either *in cis* (i.e. when they are located on the same chromosome) or *in trans* [42]. Our 4C experiments also identified cases of *trans* interactions between endogenous loci, although they occur with low frequency (approximately 5% occurs *in trans;* Figure 6B).

Although rare, such interactions between loci on different chromosome arms are of interest, because they indicate that the topological constraints imposed by chromosome architecture can in principle be overcome. In mammalian cells, there is evidence that the relative position of a gene locus within its chromosome territory (CT) influences its ability to form either *cis*- or *trans*-interactions. Peripheral regions of mammalian CTs intermingle their chromatin [43], which may allow for interactions between chromosomes. Indeed, more *trans* contacts are identified by 4C using a bait that often resides in the CT periphery compared to a bait located in the interior of a CT [20]. Likewise, activation of the *HoxB* gene cluster during differentiation coincides with relocation away from its CT interior [34], and the active *HoxB1* gene more frequently contacts sequences on other chromosomes compared to the inactive gene [44]. In line with this, we observed a varying degree of *trans* interactions among eight bait sequences, suggesting distinct capacities to contact chromatin on other chromosomes. The *trh* gene has the strongest capacity to contact other chromosome arms. Interestingly, *trh* is located within 500 Kbp of the telomere of chromosome 3L, and 5 out of 6 contacts *in trans* occur within 500 Kbp of other telomeres (Figure 6A, and Table S2). Thus, interactions between chromosome arms may be possible if loci are favorably positioned on the edge of chromosome (arm) territories, which could be the case for telomeric sequences in larval brain cells.

Relevance for gene regulation

The experiments with strain *In(3LR)sep* showed dramatic changes in interactions, such as loss of contacts between the

Homeotic gene clusters and gain of contacts with other PcDs. Despite these changed interactions, we did not find convincing evidence for global gene expression alterations on the *In(3LR)sep* chromosome. This raises the question: how relevant are long-range chromatin contacts between PcG-target genes for regulation of expression?

The lack of detectable expression changes may indicate that long-range interactions have only quantitatively subtle effect on the regulation of gene expression. Nevertheless, such subtle effects on gene expression could be very important for long-term viability and species survival. *In(3LR)sep* animals do suffer from an overall reduced viability during several stages of development (data not shown), which may indicate a generally reduced fitness, possibly due to an dimly altered regulation of gene expression.

Alternatively, PcG gene regulation may not be affected in strain *In(3LR)sep*, because *Abd-B* and *Antp*, although they no longer interact with each other, still prefer to interact with other PcDs (Figure 8B), suggesting that it is not relevant which PcDs interact. In such a model, the complement of all interactions contributes to PcG-mediated gene silencing in a population of cells.

Finally, it is interesting to note that over ~100 million years of evolution of the *Drosophila* genus, exchange of genes between chromosome arms has been rare despite extensive rearrangements within each arm [45,46]. Chromosome arm territories ensure that genes within a single arm are relatively close compared to genes on other arms, which may have resulted in an increased chance of rearrangements within one arm. Alternatively, the importance of long-range interactions among sets of genes, which are topologically limited to the same arm, may have contributed to the selective pressure that has led to this remarkable conservation of the gene complement of each chromosome arm.

Materials and Methods

Fly stocks and handling

Flies were raised at 25°C on standard cornmeal/molasses/agar medium. The wild type (w^{1118}), *In(3LR)sep* (*In(3LR)sep, vvl*^{sep}, *kni*$^{ri-1}$, *p*^{p}), *Scr*^{1} (*Scr*^{1} *p*^{p} *cu*^{1}*/TM3, Sb*^{1} *Ser*^{1}), and *Scr*^{4} (*pb*^{1} *Scr*^{4} *red*^{1} *e*^{1}*/TM3, Sb*^{1}) stocks were obtained from the Bloomington stock center (Bloomington, USA). Transgenic DamID stocks, *M{UAS-Pc::Dam}ZH51C/CyO, P{Dfd-EYFP}* and *M{UAS-Dam}ZH51C/CyO*, were made by the BestGene corporation, using the phiC31-integrase system (see below). For detecting *Scr* gene expression we crossed wild type and *In(3LR)sep* to *Scr*^{1} and *Scr*^{4} and selected male adult offspring that lacked the *Sb*^{1} phenotypic marker from the balancer chromosome.

4C in larval brain

We adapted a previously described 4C protocol [20] for *D. melanogaster* larval central brain tissue. For each experiment we dissected and pooled 40 pairs of central brain hemispheres. With gentle forceps manipulation we quickly removed the central nervous system from each larva, while leaving the mouth hook and imaginal discs attached, and placed it in PBS on ice. After 40 dissections, we replaced the PBS with 2% formaldehyde in PBS and fixed the samples for 15 minutes at room temperature while tumbling. The reaction was quenched for 5 minutes by adding 125 mM glycin final concentration, followed by three quick washes with cold PBS. Next, we removed the mouth hook, imaginal discs, and ventral nerve cord with gentle forceps manipulation. The remaining central brain hemispheres were placed into ice cold cell lysis buffer (10 mM Tris – pH 8.0; 10 mM NaCl; 0.2% Nonidet P40; 1x Complete protease inhibitors – Roche) and the tissue and cells were disrupted by repetitively pipeting up and down several times to isolate nuclei. Finally, the nuclei were centrifuged for 15 minutes with 2700 g at 4°C and supernatant was removed afterwards.

On this nuclei pellet, we applied the 3C procedure as described [47], yielding ligation products between *Eco*RI fragments. We used the original 4C protocol to obtain circularized DNA templates for the 4C PCR [20], and either *Dpn*II or *Nla*III as frequently cutting restriction enzymes.

We supplemented our reverse 4C PCR primers with a T7 promoter sequence overhang (Table S5). 4C PCR reactions were performed directly on the circularized 4C templates using *Taq* polymerase (Invitrogen) and conditions were as follows: 94°C for 2 min; 30 cycles of 94°C for 15 s, 60°C for 1 min and 72°C for 1 min; and a final elongation step of 72°C for 5 min. We purified the PCR products using the Qiaquick PCR purification kit (Qiagen).

We further amplified the 4C PCR products in a strictly linear fashion using a T7 RNA amplification step. For this we used the Superscript RNA amplification system (Invitrogen) according to the manufacturer's guidelines. As a reference, we digested genomic DNA from w^{1118} third instar larvae with *Dpn*II. To each restriction end, we added adaptors containing the T7 promoter sequence by means of ligation. We amplified 1 µg of this DNA using the Superscript RNA amplification system.

The 4C and reference samples were differentially labeled with the ULS system (Kreatech), using Cy3 and Cy5 dyes (Amersham). For each biological replicate we used opposite dye orientations to correct for dye bias, we used the average signal of the replicates as output. For all microarray experiments we used the same reference sample. Microarray probes were designed in close proximity to *Eco*RI restriction sites across the fly genome. These probes were used on custom designed tiling microarrays (Agilent Technologies).

DNA Adenine methyltransferase IDentification (DamID) and Chromatin Immunoprecipitation (ChIP) on chip

We cloned either Pc::Dam or Dam open reading frames from our previous transfection plasmids (Tolhuis et al. 2006) into the pUAST-attB vector (gift from Konrad Basler, Zurich, Switzerland). Plasmid integrity was verified using sequencing. The BestGene corporation used these vectors to create transgenic flies with the PhiC31-integrase system [48]. Vectors were injected into embryos of strain M{3xP3-RFP.attP'}ZH-51C (with M{vas-int.Dm}ZH-2A), which contains the attP-landing site at cytogenetic location 51C. Transgenic strains were balanced over either CyO, P{Dfd-EYFP}2 or CyO. Importantly, the DamID assay depends on very low expression levels to avoid mistargeting [49,50]. We achieved this by depending solely on leaky expression of the uninduced pUAST promoter in the absence of a Gal4 driver. For a single DamID experiment, we used 10 *M{UAS-Pc::Dam}ZH51C/CyO, P{Dfd-EYFP}2* central brains from 3rd instar larvae, and as reference sample for the microarray we used 10 brain from *M{UAS-Dam}ZH51C/CyO*.

For the ChIP-chip we used the Red ChIP kit (Diagenode) according the manufacturers guidelines. We used the following antibodies, anti-H3K27me3 (Upstate 07-449), and anti-IgG (non-specific control; Abcam ab17890). We used central brains from 50 3rd instar larvae (w^{1118}) in a single ChIP experiment.

Amplification of DamID and ChIP isolated DNA fragments and microarray analysis were performed as described [6,51] using custom oligonucleotide tiling arrays (NimbleGen Systems).

Discrete domains of either Pc or H3K27me3 enrichment were calculated using the same dynamic algorithm to extract DIDs from the 4C data (see below). A validation of the data is described in the Text S1 and Figure S2.

Gene expression profiling

For a single microarray experiment, we isolated 20 central brains from 3rd instar larvae for both w^{1118} (i.e. wild type) and *In(3LR)sep* fly strains. Tissue samples from each fly strain were collected in 1 ml TriZol (Invitrogen) reagent, and total RNA was isolated according to the manufacturer's protocol. For each fly strain we collected 4 independent RNA isolates. We amplified 1 µg of total RNA of each isolate using the Superscript RNA amplification system (Invitrogen). We labeled 1 µg of amplified RNA with the ULS system (Kreatech), using either Cy3 or Cy5 dyes (Amersham). We paired independent aRNA samples from w^{1118} (i.e. wild type) and *In(3LR)sep* fly strains and performed dual channel microarray hybridizations, including technical replicates (i.e. dye swaps). Our microarrays contain a 15k INDAC oligomer set that probes 93% of all protein coding genes in the *D.melanogaster* genome.

We used the LIMMA package (R package for statistical computing; http://www.r-project.org) to obtain loess normalized $\log_2$-transformed *In(3LR)sep* over w^{1118} ratios. In addition, we determined $\log_2$-transformed signal intensities from all w^{1118} channels to obtain wild type expression data.

Data processing

Each independent replicate 4C experiment was normalized using a loess algorithm (R package for statistical computing; http://www.r-project.org). For each bait two replicates were combined and analyzed.

We designed a "background filter" to correct for interactions that are a consequence of short nucleotide distance to the bait sequence. We only applied this filter to data that comes from the bait containing chromosome (e.g. chromosome 3R for *Abd-B*), because we consider other chromosome arms to be independent entities. First, we ranked data based on distance of the probe to the bait sequence (Figure S5). Next, we applied a loess algorithm to this ranked data (for each probe in the data a quadratic polynomial is fitted using a subset of the data), yielding smoothened representation of the data as a loess curve. The curve shows high values for probes located at short physical distance from the bait and gradually decreases as physical distance increases. As such, the loess curve represents a hypothetical linear conformation of the chromosome. Any long-range interaction will be detectable as a region with 4C signals that exceed the loess curve in that same position. To obtain these regions we subtracted the loess curve from the normalized 4C signals. Importantly, the degree of smoothening in a loess algorithm is determined by the smoothening parameter alpha (α), which determines the size of the subset of the data that is used to fit each local polynomial. We arbitrarily chose an α of 2, because it yields the most gradual degree of smoothening (Figure S5). However, we also tested smaller α values, which all gave similar results (Table S6), indicating that our results are independent of the parameters that we used.

We used this corrected data as input for the domainogram and DID detection algorithms [24]. Originally, the DID algorithm discretizes the domainograms into a set of multi-layered and partially overlapping discrete domains of enrichment, referred to as BRICKs [24]. However, we extracted only the DID with the smallest window size of each set of partially overlapping discrete domains, because this will represent the center of gravity of a long-range chromatin interaction. Stringency of the DID algorithm is provided by the user defined bias factor γ [24]. We empirically determined determined γ, such that no DIDs were identified in randomly permutated data.

To test the significance of overlap between DIDs and PcDs, we randomly permuted the DIDs' coordinates across the bait containing chromosome a thousand times, while maintaining their size aspects. We used Pearson's chi-square (χ^2) goodness of fit test to evaluate the difference between the observed coincidence of DIDs and PcDs/K27Ds and the expected/random coincidence (average of 1000 random shufflings).

To test whether interacting PcDs have distinctive genomic features, we divided all PcDs from the third chromosome into two categories: 1) interacting with at least one of our 4 PcG baits (*Abd-B*, *Antp*, *Ptx1*, *trh*); and 2) non-interacting with our baits. We used the non-parametric Mann-Whitney U test to test for significant differences between the two groups.

We analyzed the following genomic features (based on the April 2004 release of the UCSC genome browser): size of domains, PRE sequences frequency, gene density, repetitive DNA sequences frequency and non-coding conserved sequences frequency. Our data demonstrated that interacting PcDs are significantly larger. In order to avoid size bias, we corrected the frequencies of PREs, genes, and non-coding conserved sequences by calculating the feature frequencies per 1 Kbp domain.

All further statistical analyses were performed using the R package for statistical computing (http://www.r-project.org); scripts are available upon request.

We deposited all our microarray data sets (4C, DamID, ChIP-chip, and gene expression profiles) to the GEO website under accession number: GSE26694 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE26694).

Supporting Information

Figure S1 PcG proteins repress expression of *Homeotic* genes in central brain tissue. (A–B) In wild type larvae, Antp antibody stains nuclei in the ventral nerve cord (VNC) cells, while it is absent from the central brain cells (CB). (C–D) Reduction of the PcG protein PH by RNA interference results in detection of Antp in subsets of central brain cells. (E–F) In wild type larvae, Abd-B antibody only stains nuclei of cells in the most posterior compartment of ventral nerve cord (arrow). (G–H) Reduction of the PcG protein PH by RNA interference results in detection of Abd-B in a subset of in central brain cells. Left panels show projections of consecutive confocal sections through the CNS recorded at 4 µm intervals. Right panels depict a single confocal section through the right central brain. Scale bar is 50 µm. Anterior is always up. Images are false colored for better contrast.
Found at: doi:10.1371/journal.pgen.1001343.s001 (3.80 MB TIF)

Figure S2 Validation of PC binding and H3K27me3 chromatin maps in central brain tissue. (A–D) Chromatin maps of specific gene loci: Bithorax complex (A), Antennapedia complex (B), *Ptx1* locus (C), and *trachealess* (*trh*) locus (D). Polycomb binding (top) and H3K27me3 levels (middle) are highly similar in these loci, yielding approximately the same boundaries for PcDs (blue) and K27Ds (green). Gene positions are given at the bottom, with target genes highlighted in cyan, while nontargets are grey. Some gene symbols are indicated, and the gene that contains the bait sequence is highlighted in bold. Scale bar is 50 Kbp. (E) On a genome-wide scale, a large proportion of all PcDs and K27Ds co-occur/(partially) overlap in central brain tissue (observed), such overlap is not seen in a random overlap (expected). Random overlap was based on 1000 random permutations of domain positions, while maintaining their size aspects. Expected is the average random overlap after 1000 samplings, error bars indicate standard deviation. (F–G). The borders of all domains were aligned (relative position = 0) and the average H3K27me3 levels (F, green line) or PC binding (G, blue line) is plotted. Positions outside domains are indicated by negative values, and positive values are inside domains (x-axes). The y-axes represent average fold-change in signal. (F) Inside PcDs (blue area) the H3K27me3 levels are ~2 fold higher than outside (white area). (G) Likewise, Pc binding is

~2-fold increased inside K27Ds (green area). (H). Violin plots show distribution of PcD (blue, top) and K27D (green, bottom) sizes (x-axis). Median sizes are indicated. (I). Distribution of gene content in PcDs. Number of genes inside a PcD are given on the x-axis and the total fraction of PcDs with a specific number of genes is on the y-axis. More than 60% of all PcDs contains more than 2 genes. (J). Violin plots show distributions of mRNA signal intensities (y-axis) of PC (blue), and H3K27me3 (green) targets, and nontarget genes (grey). mRNA signal intensities of PC and H3K27me3 targets are slightly reduced. P-values were calculated using a two-sided non-parametric Mann-Whitney U test.
Found at: doi:10.1371/journal.pgen.1001343.s002 (7.73 MB TIF)

Figure S3 mRNA expression levels of interacting genes on chromosome 3. Violin plots show distributions of mRNA signal intensities (y-axis) of genes that interact with PcG baits (*Abd-B*, *Antp*, *Ptx1*, and/or *trh*; blue), low expressed genes (*wntD* and/or CG5107; green), and high expressed genes (*Crc* and/or *RpII140*; red). The remainder genes that do not interact are shown in grey. Median values and population sizes (N) are indicated.
Found at: doi:10.1371/journal.pgen.1001343.s003 (0.28 MB JPG)

Figure S4 Gene expression is not detectably changed on *In(3LR)sep* chromosome. (A) Violin plots show distribution of gene expression changes (log_2-transformed ratios of *In(3LR)sep* over wild type) of genes annotated to $3L^{IN}$ (red), $3R^{IN}$ (blue), or other chromosome arms (grey). Median values and population sizes (N) are indicated. (B) Histograms show frequency distributions of number of sex combs per 1st thoracic leg (x-axes) for crosses between wild type and mutant *Sex combs reduced* alleles (*Scr*; green) and crosses between *In(3LR)sep* and mutant *Scr* alleles (red). We examined Scr^1 (left panel) and Scr^4 (right panel) alleles. Total number of counted legs per group is indicated (n). P-values were calculated using a two-sided non-parametric Mann-Whitney U test.
Found at: doi:10.1371/journal.pgen.1001343.s004 (0.43 MB JPG)

Figure S5 Background filter. (A) Background filter applied to chromosome 3R using illustrative data, *Abd-B* (left) and *Ptx1* (right). Bait sequence is positioned to the left of each graph (blue dashed line). All other sequences are plotted according to their physical distance to the bait sequence (x-axis). Data is plotted as log2-transformed 4C signals (y-axis) after applying a running median with a window of 50 probes (grey bars). The background filters (green lines) are highest closest to the bait sequence and gradually decrease as linear distance increases. As such, the filters represent a hypothetical linear conformation of the chromosome without any long-range chromatin interactions. For this filters a smoothening parameter (α) of 2 was used. (B) The smoothening parameter (α) affects the shape of the background filter. *Abd-B* 4C signal intensities (y-axis) after applying a running median with a window of 50 probes (grey bars) are plotted against chromosomal probe positions (x-axis). The position of the bait sequence is indicated (blue dashed line). Colored lines are background filters with different α values (red: $\alpha = 0.25$; blue: $\alpha = 0.5$; yellow: $\alpha = 1$; and green: $\alpha = 1$). The lower the α value the more tightly the line follows the original data. Higher α values show a more gradual slope.
Found at: doi:10.1371/journal.pgen.1001343.s005 (1.11 MB JPG)

Figure S6 Expression of Ptx1 in larval central nervous system. Ptx1 positive cells are found both in the central brain and ventral nerve cord (VNC). In the VNC Ptx1 is expressed in a segmental manner (right panel) that was also observed at embryonic stages [52]. In central brain (middle panel), a subset of all cells express Ptx1, while the majority does not.
Found at: doi:10.1371/journal.pgen.1001343.s006 (1.88 MB TIF)

Table S1 Genomic coordinates of PcDs and K27Ds in larval brain tissue, including gene annotation and long-range chromatin interactions.
Found at: doi:10.1371/journal.pgen.1001343.s007 (0.19 MB XLS)

Table S2 Genomic coordinates of DIDs in wild type larval brain tissue, including gene annotation and PcD/K27D specific interactions.
Found at: doi:10.1371/journal.pgen.1001343.s008 (0.07 MB XLS)

Table S3 Comparison of genomic features between interacting and non-interacting PcDs.
Found at: doi:10.1371/journal.pgen.1001343.s009 (0.02 MB XLS)

Table S4 Genomic coordinates of DIDs in *In(3LR)sep* larval brain tissue, including gene annotation and PcD/K27D specific interactions.
Found at: doi:10.1371/journal.pgen.1001343.s010 (0.03 MB XLS)

Table S5 Description of bait sequences used in this study, including 4C PCR primer sequences.
Found at: doi:10.1371/journal.pgen.1001343.s011 (0.01 MB XLS)

Table S6 Comparison of background filters with various levels of smoothening parameter (α) to test effects on overlap between DIDs and PcDs/K27Ds.
Found at: doi:10.1371/journal.pgen.1001343.s012 (0.03 MB XLS)

Text S1 Additional information on PC and H3K27me3 maps in larval brain tissue.
Found at: doi:10.1371/journal.pgen.1001343.s013 (0.08 MB DOC)

Acknowledgments

We thank K. Basler, W. Brugman, U. Braunschweig, G. Filion, J. Kind, J. Vissers, A. Sparmann, G. Gargiulo, M. Vogel, B. van Veen, J. Noordermeer, and L. Fradkin for technical assistance and/or critically reading the manuscript.

Author Contributions

Conceived and designed the experiments: BT MvL BvS. Performed the experiments: BT MB HT. Analyzed the data: BT. Contributed reagents/materials/analysis tools: RMK LP HT MN MS WdL. Wrote the paper: BT MvL BvS.

References

1. Schwartz YB, Pirrotta V (2007) Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 8: 9–22.
2. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, et al. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441: 349–353.
3. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 20: 1123–1136.
4. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, et al. (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 125: 301–313.
5. Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, et al. (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet 38: 700–705.
6. Tolhuis B, de Wit E, Muijrers I, Teunissen H, Talhout W, et al. (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. Nat Genet 38: 694–699.

7. Sparmann A, van Lohuizen M (2006) Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer 6: 846–856.
8. Negre N, Hennetin J, Sun LV, Lavrov S, Bellis M, et al. (2006) Chromosomal distribution of PcG proteins during Drosophila development. PLoS Biol 4: e170. doi:10.1371/journal.pbio.0040170.
9. Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, et al. (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol 7: e13. doi:10.1371/journal.pbio.1000013.
10. Cleard F, Moshkin Y, Karch F, Maeda RK (2006) Probing long-distance regulatory interactions in the Drosophila melanogaster bithorax complex using Dam identification. Nat Genet 38: 931–935.
11. Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V (2007) Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. Nat Cell Biol 9: 1167–1174.
12. Tiwari VK, McGarvey KM, Licchesi JD, Ohm JE, Herman JG, et al. (2008) PcG proteins, DNA methylation, and gene repression by chromatin looping. PLoS Biol 6: e306. doi:10.1371/journal.pbio.0060306.
13. Fedorova E, Sadoni N, Dahlsveen IK, Koch J, Kremmer E, et al. (2008) The nuclear organization of Polycomb/Trithorax group response elements in larval tissues of Drosophila melanogaster. Chromosome Res 16: 649–673.
14. Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, et al. (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. Cell 124: 957–971.
15. Tiwari VK, Cope L, McGarvey KM, Ohm JE, Baylin SB (2008) A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. Genome Res 18: 1171–1179.
16. Alkema MJ, Jacobs J, Voncken JW, Jenkins NA, Copeland NG, et al. (1997) MPc2, a new murine homolog of the Drosophila polycomb protein is a member of the mouse polycomb transcriptional repressor complex. J Mol Biol 273: 993–1003.
17. Buchenau P, Hodgson J, Strutt H, Arndt-Jovin DJ (1998) The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing. J Cell Biol 141: 469–481.
18. Ficz G, Heintzmann R, Arndt-Jovin DJ (2005) Polycomb group protein complexes exchange rapidly in living Drosophila. Development 132: 3963–3976.
19. Hernandez-Munoz I, Taghavi P, Kuijl C, Neefjes J, van Lohuizen M (2005) Association of BMI1 with polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. Mol Cell Biol 25: 11047–11058.
20. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, et al. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38: 1348–1354.
21. Bello B, Holbro N, Reichert H (2007) Polycomb group genes are required for neural stem cell survival in postembryonic neurogenesis of Drosophila. Development 134: 1091–1099.
22. Simonis M, Kooren J, de Laat W (2007) An evaluation of 3C-based methods to capture DNA interactions. Nat Methods 4: 895–901.
23. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, et al. (1990) Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A 87: 1663–1667.
24. de Wit E, Braunschweig U, Greil F, Bussemaker HJ, van Steensel B (2008) Global chromatin domain organization of the Drosophila genome. PLoS Genet 4: e1000045. doi:10.1371/journal.pgen.1000045.
25. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295: 1306–1311.
26. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289–293.
27. Ringrose L, Rehmsmeier M, Dura JM, Paro R (2003) Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. Dev Cell 5: 759–771.
28. de Laat W, Grosveld F (2007) Inter-chromosomal gene regulation in the mammalian cell nucleus. Curr Opin Genet Dev 17: 456–464.
29. Craymer L (1981) Techniques for manipulating chromosomal rearrangements and their application to Drosophila melanogaster. I. Pericentric inversions. Genetics 99: 75–97.
30. Pattatucci AM, Kaufman TC (1991) The homeotic gene Sex combs reduced of Drosophila melanogaster is differentially regulated in the embryonic and imaginal stages of development. Genetics 129: 443–461.
31. Southworth JW, Kennison JA (2002) Transvection and silencing of the Scr homeotic gene of Drosophila melanogaster. Genetics 161: 733–746.
32. Lavigne M, Francis NJ, King IF, Kingston RE (2004) Propagation of silencing; recruitment and repression of naive chromatin in trans by polycomb repressed chromatin. Mol Cell 13: 415–425.
33. Francis NJ, Kingston RE, Woodcock CL (2004) Chromatin compaction by a polycomb group protein complex. Science 306: 1574–1577.
34. Heard E, Bickmore W (2007) The ins and outs of gene regulation and chromosome territory organisation. Curr Opin Cell Biol 19: 311–316.
35. Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet 8: 104–115.
36. Hochstrasser M, Mathog D, Gruenbaum Y, Saumweber H, Sedat JW (1986) Spatial organization of chromosomes in the salivary gland nuclei of Drosophila melanogaster. J Cell Biol 102: 112–123.
37. Hochstrasser M, Sedat JW (1987) Three-dimensional organization of Drosophila melanogaster interphase nuclei. I. Tissue-specific aspects of polytene nuclear architecture. J Cell Biol 104: 1455–1470.
38. Marshall WF, Dernburg AF, Harmon B, Agard DA, Sedat JW (1996) Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in Drosophila melanogaster. Mol Biol Cell 7: 825–842.
39. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, et al. (2000) The genome sequence of Drosophila melanogaster. Science 287: 2185–2195.
40. Mathog D, Hochstrasser M, Gruenbaum Y, Saumweber H, Sedat J (1984) Characteristic folding pattern of polytene chromosomes in Drosophila salivary gland nuclei. Nature 308: 414–421.
41. Bantignies F, Grimaud C, Lavrov S, Gabut M, Cavalli G (2003) Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. Genes Dev 17: 2406–2420.
42. Vazquez J, Muller M, Pirrotta V, Sedat JW (2006) The Mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. Mol Biol Cell 17: 2158–2165.
43. Branco MR, Pombo A (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol 4: e138. doi:10.1371/journal.pbio.0040138.
44. Wurtele H, Chartrand P (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome Res 14: 477–495.
45. Bhutkar A, Schaeffer SW, Russo SM, Xu M, Smith TF, et al. (2008) Chromosomal rearrangement inferred from comparisons of 12 Drosophila genomes. Genetics 179: 1657–1680.
46. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. (2007) Evolution of genes and genomes on the Drosophila phylogeny. Nature 450: 203–218.
47. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10: 1453–1465.
48. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A 104: 3312–3317.
49. van Steensel B, Henikoff S (2000) Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol 18: 424–428.
50. van Steensel B, Delrow J, Henikoff S (2001) Chromatin profiling using targeted DNA adenine methyltransferase. Nat Genet 27: 304–308.
51. Greil F, van der Kraan I, Delrow J, Smothers JF, de Wit E, et al. (2003) Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. Genes Dev 17: 2825–2838.
52. Vorbruggen G, Constien R, Zilian O, Wimmer EA, Dowe G, et al. (1997) Embryonic expression and characterization of a Ptx1 homolog in Drosophila. Mech Dev 68: 139–147.