

1
2
3 **H3K27 modifications define segmental regulatory domains in the *Drosophila***
4 **bithorax complex.**
5

6 Sarah K. Bowman^{1,#}, Aimee M. Deaton^{1,#}, Heber Domingues², Peggy I. Wang¹, Ruslan I.
7 Sadreyev^{1,3}, Robert E. Kingston^{1,*}, and Welcome Bender^{2,*}
8

9 ¹Department of Molecular Biology, Massachusetts General Hospital, and Department of
10 Genetics, Harvard Medical School, Boston, MA 02114

11 ²Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,
12 Boston, MA 02115

13 ³Department of Molecular Biology and Department of Pathology, Massachusetts General
14 Hospital and Harvard Medical School, Boston, MA 02114

15 [#]These authors contributed equally

16 ^{*}Correspondence to REK (Kingston@molbio.mgh.harvard.edu), WB (ph: 617-432-1906, FAX:
17 617-738-0516)
18

19 **Abstract**

20 **The bithorax complex (BX-C) in *Drosophila melanogaster* is a cluster of homeotic genes**
21 **that determine body segment identity. Expression of these genes is governed by cis-**
22 **regulatory domains, one for each parasegment. Stable repression of these domains depends**
23 **on Polycomb Group (PcG) functions, which include trimethylation of lysine 27 of histone**
24 **H3 (H3K27me3). To search for parasegment-specific signatures that reflect PcG function,**
25 **chromatin from single parasegments was isolated and profiled. The H3K27me3 profiles**
26 **across the BX-C in successive parasegments showed a "stairstep" pattern that revealed**
27 **sharp boundaries of the BX-C regulatory domains. Acetylated H3K27 was broadly**
28 **enriched across active domains, in a pattern complementary to H3K27me3. The CCCTC-**

binding protein (CTCF) bound the borders between H3K27 modification domains; it was retained even in parasegments where adjacent domains lack H3K27me3. These findings provide a molecular definition of the homeotic domains, and implicate precisely positioned H3K27 modifications as a central determinant of segment identity.

Introduction

Cells in higher organisms choose among developmental pathways in response to transitory cues. The choice must be remembered; this is often described as changing the epigenetic state. One prominent mechanism to fix such choices relies on the genes of the Polycomb Group (PcG). The members of this family have been well defined by genetic and biochemical studies (Simon and Kingston, 2013). The PcG proteins are present in all cells, but they impose long-term repression on different target genes in different lineages. PcG proteins form distinct complexes with distinct functions. These include the PRC1 family of complexes, which compact chromatin and ubiquitylate histone H2A, and the PRC2 complexes, which methylate histone H3 on lysine 27. The molecular details of how PcG functions are coordinated to lead to repression are poorly understood.

The prototypic targets of PcG action are the *Drosophila* homeotic genes that control segment identity. Indeed, a mutation in *Polycomb*, the founding member of the PcG, was discovered for its segmental transformations (Lewis, 1947). In *Drosophila melanogaster*, the homeotics lie in two clusters, the Antennapedia complex (Antp-C) and the bithorax complex (BX-C). The BX-C is the better-studied cluster, due largely to the effort of E. B. Lewis. Lewis showed that mutations that transform specific segments align on the genetic map in the order of the body segments that they affect (Lewis, 1978). The evolutionary conservation of this gene order, including that in the mammalian HOX complexes, suggests that chromosomal position dictates the expression pattern. Lewis used nested deletions to show that successively more distal genetic functions of the BX-C are activated in successively more posterior segments along the body axis (Lewis, 1981). The genetic functions are now thought to encompass nine regulatory domains,

each for a different parasegment. The regulatory domains control three homeobox-containing genes of the complex, *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*) (Maeda and Karch, 2009). The regulatory domains have been roughly defined by mutant lesions (Lewis, 1978), and by enhancer traps in the BX-C, which show different anterior limits of expression (Bender and Hudson, 2000). The molecular features that define these regions, as well as the precise locations of their boundaries, have not been defined.

Both PRC1 and PRC2 are required to maintain repression of homeotic genes in appropriate parasegments, so these complexes may generate segment-specific chromatin features in the regulatory domains of the BX-C. We initially evaluated H3K27me3 patterns because they reflect the central function of PRC2. The entire BX-C is heavily marked with H3K27me3 in whole embryos (Schuettengruber et al., 2009; Nègre et al., 2011). This histone modification is necessary for maintaining repression of BX-C genes (Pengelly et al., 2013), yet these genes are not repressed throughout the body. It seemed likely that parasegment-specific patterns of histone modification might be obscured when all the parasegments are pooled for analysis. Indeed, Papp and Müller (2006) observed loss of the K27me3 mark across the *Ubx* transcription unit in haltere and third leg imaginal discs, tissues where *Ubx* is transcribed. These discs include cells of two parasegments (PS5&6), and so it was not clear how the K27me3-free regions might correlate with transcription units or with the genetic regulatory domains. The issue could be resolved if H3K27me3 patterns could be studied in single parasegments.

Results

The chromatin features of the BX-C have not been studied in individual parasegments because of the technical challenges of cell isolation and molecular analysis on small samples. To address this, we marked single parasegments with a combination of Gal4 and Gal80 drivers, and developed a nuclear sorting-chromatin immunoprecipitation-sequencing protocol. Gal4, a transcriptional activator, was expressed in a series of parasegments with a defined anterior boundary, and Gal80 was expressed in a pattern shifted one parasegment more posterior. Gal80

binds to and inactivates Gal4, leaving Gal4 activity in a single parasegment (Figure 1A). Gal4 and Gal80 expression domains were established either by enhancer trapping or by selecting an enhancer to achieve the desired expression pattern (Figure1-figure supplement 1). Various combinations of Gal4 and Gal80 drivers were combined genetically to limit Gal4 activity to each of parasegments 4, 5, 6, and 7, which approximately correspond to the 2nd thoracic through the 2nd abdominal segments (Figure 1B, and Figure1-figure supplements 1 and 2; see Methods for details).

For isolation of marked nuclei, we used Gal4 to drive an mCherry-RanGAP fusion protein, from the INTACT system (Deal and Henikoff, 2011; Steiner et al., 2012). Staged embryos with single parasegment stripes of mCherry were formaldehyde-fixed and disrupted, and their nuclei were sorted on a FACS instrument (Figure1-figure supplement 1C). DNA enriched by chromatin immunoprecipitation (ChIP) was prepared for paired-end sequencing, using a picogram-scale library protocol (Figure 1-figure supplement 1D, (Bowman et al., 2013)).

The largest differences in H3K27me3 enrichment between parasegments occurred at the BX-C (Figure 2A; Figure 2-figure supplement 1). While the H3K27me3 profiles in our four adjacent parasegments were very similar elsewhere in the genome (Figure 2-figure supplement 1), the BX-C exhibited a striking "stairstep" pattern. Large expanses of H3K27me3 enrichment were lost as we moved from anterior to posterior along the body axis (Figure 2C to F), revealing the locations of the regulatory domains. This is reminiscent of the temporal activation pattern of the Hox-D cluster in mouse, which correlates with loss of H3K27me3 (Soshnikova and Duboule, 2009). Our analysis of single parasegments locates distinct domain borders where K27me3 status changes abruptly. The contrast between the whole embryo profile (Figure 2B) and the single segment profiles (Figure 2C to F) shows how isolated cell types can identify chromatin changes linked to cell identity.

In each isolated parasegment, the extent of K27me3 coverage was largely consistent with genetic studies indicating the location of BX-C active and repressed regulatory domains. The PS4 H3K27me3 pattern is mesa-like across the complex (Figure 2C), with only a few narrow

gaps at nucleosome-free regions (Mito et al., 2005). The BX-C is almost completely repressed in PS4, with only a few neuronal cells on the midline showing UBX expression in embryos. The PS5 regulatory domain, associated with the *bithorax* (*bx*) mutant lesions, is defined by the low level of H3K27me3 modification across the leftmost (proximal) 92 kb of the complex (Figure 2D). There is a sharp transition to full H3K27me3 coverage at a position upstream of the *Ubx* transcription start site. The low residual level of H3K27me3 marks in the 92 kb PS5 domain is likely due to partial contamination from more anterior parasegments in our preparation of PS5 nuclei (Figure 1-figure supplement 2). In PS6, the leftmost 137 kb of the BX-C is clear of the K27me3 mark (Figure 2E); the additional K27me3-clear region reveals the PS6 domain. The PS6 regulatory domain is the one best defined by genetic analysis, with a series of rearrangement breakpoints giving *bithoraxoid* (*bxd*) phenotypes (Bender and Lucas, 2013). The transition point to full K27 methylation coincides with the genetically defined *Fub* border between the PS6 and the PS7 regulatory domains (Bender and Lucas, 2013). In PS7, an additional 36 kb was depleted of H3K27me3 (Figure 2F). This region includes the *abd-A* transcription unit, and is associated with the *infraabdominal-2* (*iab-2*) class of genetic mutations. We observed a reproducible mound of H3K27me3 in PS7 nuclei close to the *Ubx* promoter (arrow in Figure 2F). This may reflect repression of *Ubx* by *abd-A* in this parasegment (Karch et al., 1990), mediated by the PcG.

The regulatory domains defined by the sharp transitions in K27me3 enrichment do not correspond in every detail to expectations from genetic studies. Prior mapping of domains has relied upon random insertion of transgenic reporter elements into the BX-C to interrogate local enhancer activity. These enhancer traps show sharp anterior edges to their expression patterns, with the anterior-most labeled parasegment corresponding to the domain in which the insertion resides (Bender and Hudson, 2000). In the current study, the border between the PS5 and PS6 domains, defined by the discontinuity in K27me3 enrichment in PS5 nuclei, lies 14 kb upstream of the *Ubx* promoter. However, transgenic reporter elements inserted between the *Ubx* promoter and the K27me3-defined border (presumably part of the PS5 regulatory domain) show patterns

with a PS6 anterior limit. Perhaps the promoters of these reporters interact with the *Ubx* promoter in distinctive ways. There was also prior evidence of an active BX-C domain in PS4. An enhancer trap 3 kb downstream of the *Ubx* start site shows a PS4 anterior border (Casares et al., 1997), and a ncRNA with a PS4 anterior expression limit is transcribed from the region 5-15 kb downstream of the *Ubx* start (Pease et al., 2013). Despite the activity of these regulatory regions in PS4 cells, we did not detect a decrease in K27me3 at these locations. The PS4 active state may be early and transient, limited to a subset of PS4 cells, or resistant to H3K27 methylation.

We have profiled additional histone modifications and several chromosomal proteins in single parasegments to understand better how domains are established and bounded. The sharp discontinuities in K27 methylation coincide with binding sites of the CCCTC-binding factor (CTCF), which is thought to impose boundaries on regulatory domains in many systems (Ohlsson et al., 2010). Figure 2G shows the CTCF ChIP pattern in unsorted nuclei, similar to that previously reported (Nègre et al., 2010). CTCF in PS7 nuclei gives a pattern virtually identical to the whole embryo pattern (Figure 2H). Apparently, the activation state of a domain is not coupled to the presence of CTCF at its borders. We note that our experimental conditions do not detect robust CTCF enrichment in some BX-C locations that have been reported by others (Nègre et al., 2010), notably at the border between regulatory regions for PS7 and PS8, and proximal to the 3' end of the *Ubx* transcription unit.

CTCF is present in a protein complex with Centrosomal Protein 190 (CP190) (Mohan 2007), and the binding sites of these proteins often overlap (Nègre 2010). The CP190 sites largely coincide with CTCF sites in the BX-C, and, as with CTCF, there is little difference between the CP190 pattern in unsorted nuclei (Figure 2I and Nègre et al., 2010) and in PS7 nuclei (Figure 2J). There is a noteworthy CP190 site immediately proximal to the 3' end of the *Ubx* transcription unit, separating *Ubx* from the adjacent gene (*modular serine protease*). In the Antennapedia complex, the other homeotic gene complex in flies, CP190 peaks often appear without coincident CTCF peaks (Figure 2-figure supplement 2).

Acetylation of H3K27 (H3K27ac) is correlated with active enhancer regions (Heintzman et al., 2009; Kharchenko et al., 2011), and is anti-correlated with PcG function (Tie et al., 2009), prompting us to examine its pattern in the BX-C. H3K27ac showed the opposite pattern to H3K27me3. There is relatively little H3K27ac across the BX-C in unsorted nuclei, but in PS7, the active domains show broad enrichment for H3K27ac (Figure 3 B and C). The H3K27ac levels differ in the three active domains, for reasons we do not yet understand. Within each domain, however, the coverage is boundary-to-boundary, although many enhancers in the BX-C are not active in embryos. Overall, H3K27ac is a mirror to H3K27me3, in extent, if not in density.

ChIP of serine-5-phosphorylated RNA polymerase II (Pol-II) (transcriptionally engaged polymerase) in whole embryos shows the transcription units for *Ubx*, *abd-A*, and *Abd-B* (Figure 3D). Embryonic non-coding transcripts (Bae et al., 2002; Pease et al., 2013) are not apparent from the Pol-II density, either because they are limited to very early stages of embryogenesis, or because the levels of transcription are too low. There is an unexpected Pol-II peak ~14.5 kb upstream of the *Ubx* promoter, coincident with the CTCF binding site at the border between the PS5 and PS6 domains (arrow in Figure 3D). There is no reported transcript initiating at this position (Pease et al., 2013), although a short transcript could have been missed. Pol-II in PS7 (Figure 3E) marks the *Ubx* and *abd-A* transcription units, as expected from the UBX and ABD-A protein patterns (Figure 1-figure supplement 2). There is also a Pol-II peak at the major *Abd-B* promoter (arrow in Figure 3E), although *Abd-B* RNA and protein are absent from PS7 (Celniker et al., 1990; Boulet et al., 1991) and this region is H3K27 trimethylated. Analogous localization of Pol-II at the *Abd-B* promoter in wing and haltere discs was reported by Chopra et al. (2009), who argued that these polymerase molecules were "paused". The H3K4me3 mark aligns with the *Ubx* and *abd-A* promoters in PS7 nuclei (but not *Abd-B*; Figure 3G) and is reduced in mixed cells (Figure 3F); this histone mark has been seen at active promoters in other systems (Ruthenburg et al., 2007).

The POLYCOMB subunit of PRC1 is known to bind to the H3K27me3 modification through

its chromodomain (Fischle et al., 2003), and so PRC1 localization might be expected to correspond to the stair-step H3K27me3 pattern. Published ChIP-CHIP and ChIP-seq profiles for PC across the BX-C vary from broad coverage matching the H3K27me3 profile (Kwong et al., 2008) to sharp peaks at the known PREs (Gonzalez et al., 2014). The difference may be due to different procedures for chromatin fragmentation prior to immunoprecipitation (Straub et al., 2013). Our PC profiles show sharp peaks at PREs in nuclei from whole embryos (Figure 3H), matching the profile for POLYHOMEOTIC, another PRC1 subunit (Figure 3J). In PS7 nuclei, we find both the POLYCOMB and POLYHOMEOTIC subunits of PRC1 are bound to the Polycomb Response Elements of the active PS6 and PS7 domains at levels comparable to those in the whole embryo profile (Figure 3H-K). In the active PS5 domain, the levels of PC and PHO at the PS5 PRE are reduced, relative to the whole embryo profile. Papp and Müller (2006) have reported reduced retention of PRC1 and 2 components at the PRE in the PS5 domain in haltere and third leg imaginal discs, where the PS5 domain is in the active state and K27me3 is absent. Kwong et al (2008) have made analogous observations regarding PRC retention in imaginal discs of the third thoracic segment. PRC1 presence is apparently not a determinant of repressed domains, although its subunits are clearly needed for PcG repression (Simon and Kingston, 2013). It will be important to map H2A ubiquitylation (a function of the dRING subunit of PRC1) across the domains, especially given recent findings that H2A ubiquitylation promotes PRC2 recruitment and H3K27 trimethylation (Blackledge et al, 2014; Cooper et al, 2014; Kalb et al, 2014).

The profiles for SU(Z)12, a PRC2 component (Figure 3L, M), look similar to those of the PRC1 proteins, POLYCOMB and POLYHOMEOTIC. There are major peaks at the known PREs in both whole embryo and PS7 nuclei, although in PS7, the PREs in active domains have reduced levels of SU(Z)12. This result is again consistent with the prior work of Papp and Müller (2006) on the PS5 PRE in imaginal discs. SU(Z)12 is also present at the transcription start sites of *Ubx*, *abd-A*, and *Abd-B*; binding at these promoters is more apparent than it is for PRC1 proteins.

Discussion

The Polycomb Group repression system is often described as a cellular memory mechanism, which can impose lifelong silencing of a gene in response to a transitory signal. That view seems valid, but the concept of a PcG regulatory domain is much richer. In the PS6 domain of the BX-C, for example, there are many enhancers to drive *Ubx* expression in specific cells at specific developmental times, all of which are blocked in parasegments one through five, but active in parasegments 6 through 12 (Bender and Lucas, 2013). Individual enhancers need not include a segmental address that is specified, for example, by gap and pair-rule DNA-binding factors; their function is segmentally restricted by the domain architecture. Indeed, these enhancers will drive expression in a different parasegment when inserted into a different domain (as in the *Cbx* transposition, (Peifer et al., 1987)). Each domain has a distinctive collection of enhancers; the UBX pattern in PS5 is quite different from that in PS6. Thus, there are two developmental programs for *Ubx*, one in each of these parasegments, without the need for a duplication of the *Ubx* gene. Other loci with broad regions of H3K27 methylation may likewise be parsed into multiple domains, once we examine histone marks in specific cell types.

The all-or-nothing H3K27me3 coverage of the BX-C parasegmental domains validates and refines the domain model. In particular, K27me3 is uniformly removed across the PS5 and PS7 domains in PS5 and PS7, even though the activated genes in those parasegments (*Ubx* and *abd-A*, respectively) are only transcribed in a subset of cells. It is interesting that both PRC1 and PRC2 components have binding patterns that do not fully reflect function (repression and K27 methylation, respectively), indicating the possibility that function of these complexes is regulated separately from binding (Papp and Mueller, 2006.) The challenges now are to understand how PcG regulated domains are established, differently in different parasegments, and to describe the molecular mechanisms, including changes in chromosome structure, that block gene activity in H3K27 trimethylated domains.

Acknowledgements: The datasets supporting the results of this report are available in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE55257. We are grateful to Barret Pfeiffer for the pBPG80Uw plasmid containing a codon-optimized version of Gal80. Paul Talbert and Steve Henikoff supplied us with a UAS/INTACT P element. Stephanie Terrizzi and Jodi Moore assisted with nuclear sorting. PH and SU(Z)12 antiserum were a gift from Jürg Müller, and CTCF antiserum was a gift from Victor Lubanenkova. Matt Simon, Michele Markstein, and Jesse Cochrane gave helpful comments on the manuscript. S.K.B. was supported by the Damon Runyon Cancer Research Foundation. This work was supported by grants from the NIH, GM43901 to R.E.K. and GM28630 to W. B.

Competing Interests: None.

Materials and Methods:

Gal4 and Gal80 sources

A P element was built containing both Gal4 and Gal80 (Figure 1-figure supplement 1A). The P promoter initially drives Gal4 transcription. The P-element used the Gal4 gene with an HSP70 terminator derived from pGawB (Brand and Perrimon, 1993), fused in frame to the P transposase gene after the 130th amino acid. The Gal4 gene was followed by a synthetic FRT site plus a polylinker. The Gal80 coding sequence with an HSP70 terminator, derived from pBPG80Uw (a gift of B. Pfeiffer and G. Rubin) was inserted into this polylinker. Another synthetic FRT site was inserted near the start of the Gal4 coding region at a unique *SphI* site. Thus, the Gal4 coding sequence is flanked by FRT sites, so that flipase-induced recombination can turn the P element into a Gal80 producer, with the same expression pattern. The P element was transformed into a random chromosomal position, and then used to swap into the position of an existing P enhancer trap in the BX-C (Bender and Hudson, 2000) or at the *Antennapedia* promoter. P element swapping (Sepp and Auld, 1999) is a low frequency event, but the lines

were constructed with UAS/GFP as a marker for Gal4 activity, and thousands of first instar larvae could be quickly screened for one with the expected segmentally-restricted GFP pattern. A derivative of this vector was made by replacing the HSP70 terminator downstream of the Gal80 coding sequence with the poly(A) addition site from the *Drosophila Alcohol Dehydrogenase (Adh)* gene, recovered as a 317 bp PCR fragment from genomic DNA.

We have also taken known embryonic enhancers from the BX-C and inserted them into a second vector (Figure 1-figure supplement 1B), which was derived from the initial vector with the *Adh* terminator. At a *Bam*H1 site immediately upstream of the 5' FRT site, a fragment was inserted containing: 1) an 841 bp fragment from the 5' end of the *gypsy* mobile element (bases 189-1029), 2) an *Asc*I site, and 3) a fragment from the *engrailed* gene (-595 bp to +243 bp relative to the major transcription start site). The *gypsy* fragment includes twelve tandem binding sites for SUPPRESSOR OF HAIRY WING (SU(HW)). The *engrailed* fragment includes the proximal PRE ("PSE2" (DeVido et al.)), and is designed to fuse the first nine amino acids of ENGRAILED to the amino terminal end of Gal4. Enhancer fragments from the BX-C were cloned by PCR with *Asc*I extensions on the primers, and inserted into the *Asc*I site between the *gypsy* and *engrailed* fragments. The *engrailed* PRE was included to maintain the segmentally-restricted pattern through most of embryonic development, and the SU(HW) binding sequences were intended to insulate the promoter from endogenous enhancers near the insertion site. Transformants were recognized by Gal4-driven GFP patterns in young larvae, and, again, any Gal4 producer could be converted to a Gal80 producer. Sequences of these P elements are available on request.

Swap positions

P element swapping was used to replace a P element at the *Antennapedia* P1 promoter (P[XP] Antp^{d02480}, target site duplication: 3R: 2,825,198 -2,825,205 (Thibault et al., 2004)), and a P element in the PS8 (*iab-3*) domain of the BX-C (HCJ192, target site duplication 3R: 12,673,279-12,673,286 (Bender and Hudson, 2000)). Other swaps were made with P elements

in the PS5 (bx) and PS6 (bxd) domains, but these did not give sufficiently strong or uniform expression patterns. The *Antp* swap was subjected to P transposase to induce "local hopping" and thereby increase the Gal4 expression level.

Enhancer fragments

Fragments used as enhancers inserted into the vector of Figure 1, figure supplement 1B included:

abx enhancer (for PS5 anterior limit) 3R: 12,508,951-12,513,096

pbx enhancer (for PS6 anterior limit) 3R: 12,598,546-12,600,177

iab-2 enhancer (for PS7 anterior limit) 3R: 12,636,236-12,639,140

iab-3 enhancer (for PS8 anterior limit) 3R: 12,664,301-12,666,886

Other fragments tested covered the "bx" and "bxd" enhancers, but these did not give sufficiently strong or uniform expression patterns. One insertion of a P element with the iab-3 enhancer fortuitously inserted 1.4 kb upstream of the *Ubx* transcription start site (target site duplication, 3R: 12,561,577-12,561,584). This gave a strong pattern with a sharp PS6 anterior limit, and was used as the Gal4 source for PS6 nuclei. Insertions with the abx and pbx enhancers were subjected to P transposase to adjust the levels of Gal4 and Gal80. *D. melanogaster* genome coordinates are from release 5.57.

INTACT marker

Flies with the INTACT reporter in a P element on the third chromosome (w^{1118} ; $p[w^+;$ UASRG]6) were generously provided by Paul Talbert and Steve Henikoff. This P element was crossed to a P transposase source (P{ry[+t7.2]=Delta2-3}99B), and offspring were screened initially for darker eye color, most likely due to multiple copies of the INTACT element. Such stocks with INTACT insertions on the second or third chromosome were then crossed to a Gal4 source, and those with brightest mCherry expression were used for subsequent efforts to mark single parasegments. The higher copy INTACT chromosomes are designated "INTACT_n on II"

and "INTACT_n on III".

Genotypes:

To mark PS4, males homozygous for [w; abx en.>Gal4, pbx en.>Gal80; Antp swap>Gal4*, INTACT on III] were crossed to females homozygous for [w; INTACT_n on II; INTACT_n on III].

To mark PS5, males homozygous for [w; abx en.>Gal4*, pbx en.>Gal80*, UAS-GFP*] were crossed to females homozygous for [w; INTACT_n on II; INTACT_n on III].

To mark PS6, males homozygous for [w; INTACT_n on II; iab-2 en.>Gal80, iab-3 swap>Gal80, INTACT on III] were crossed to females [w; bxd insert>Gal4, iab-2 en.>Gal80, iab-3 en.>Gal80, INTACT on III / TM6]. Only half of the resulting embryos had a Gal4-driven stripe.

To mark PS7, males homozygous for [w; iab-2 en.>Gal4, iab-3 swap>Gal80, INTACT on III] were crossed to females homozygous for [w; INTACT_n on II; INTACT_n on III].

Insertions treated with P transposase are marked above with an asterisk (*).

Embryo collection and nuclear preps

Agar plates with yeast paste were put on the PS4, PS5 or PS7 laying cages for 8 h at 25°C (or 16 h at 18°C), and then the eggs were aged for another 5h at 25°C (or 10 h at 18°C). For PS6, a 6 h collection was aged for a further 4 h. The eggs were dechorionated (2 min in 50% bleach) and fixed in fixation buffer (5% formaldehyde, 100 mM NaCl, 50 mM Hepes, 1 mM EDTA, 0.5 mM EGTA, pH 8.0) under heptane, on a rotator for 15 min at 25°C. The fixed embryos were washed with stop buffer (125 mM glycine, 130 mM NaCl, 7 mM Na₂HPO₄, 3 mM KH₂PO₄, 0.1% Triton X100, pH8.0) for 2 min at 25°C, and then briefly with PBS (130 mM NaCl, 7 mM Na₂HPO₄, 3 mM KH₂PO₄) plus 0.1% Triton X100. The washed embryos were collected on a nytex filter, and examined under an epifluorescence stereoscope. Embryos that appeared too old, with fluorescent nuclei anterior to the desired parasegment, were manually removed. The

remaining eggs were collected and weighed in a 1.5 ml microfuge tube, flash frozen in liquid nitrogen, and stored at -75°C.

For the isolation of nuclei, ~100 mg of embryos were suspended in 5 ml of BBT buffer (55mM NaCl, 40mM KCl, 15mM MgSO₄, 5mM CaCl₂, 10mM Tricine, 20mM dextrose, 50mM sucrose, 0.1% bovine serum albumin (Miles PENTEX), 0.01% Triton X-100, pH7.0) plus 100 µl protease inhibitor cocktail (Roche Complete, EDTA-free), and disrupted with 10 strokes in a Dounce homogenizer with a loose pestle, then 10 strokes with a tight pestle. The homogenate was spun at 275 x g for 1 min to remove large debris. The resulting supernatant was spun at 1000 x g for 10 min to pellet the nuclei. The pellet was resuspended in 2 ml BBT buffer plus 40 µl protease inhibitor, and homogenized again, 20 strokes with the tight pestle. The homogenate was passed through a cell strainer cap with a 40 µm nylon mesh and stored at 0°C. Hoechst 33342 dye was added (2 µl of 10 mg / ml solution) before sorting.

Sorting

Nuclei were sorted on a BD Biosciences FACS Aria IIu instrument. The fluorescence from Hoechst 33342 (405 nM excitation, 425-475 nm emission) was used to select for a 2N DNA content, thus avoiding clumps of nuclei. These single nuclei were secondarily selected for mCherry fluorescence (594 nM excitation, 619-640 nM emission) stronger than 99.9% of nuclei from control (*Oregon R*) embryos. Recovered nuclei constituted between 1 and 6% of total single nuclei (Figure 1, figure supplement 1C), depending of the parasegment under selection. Recovered nuclei were kept at 0°C, and used for chromatin preparation within 5 hours.

Chromatin preparation

The yield of sorted nuclei in sort buffer (PBS) was determined with a hemacytometer, and 50-200K nuclei were used for each ChIP. Chromatin fragmentation was accomplished either with micrococcal nuclease (MNase) or bath sonication. For MNase fragmentation, nuclei in sort buffer were supplemented with PBS containing 0.1% Triton-X100 (PBS-Tx) to a volume of

400µl, and CaCl₂ to 1 mM. Nuclei were pre-warmed to 37°C and digested with 12U MNase (Worthington Biochemical) for 3 min. Digestion was stopped by moving the tubes to ice and adding 10 µl of 250 mM EDTA, 250 mM EGTA, and nuclei were briefly sonicated in a diagenode bioruptor (3 min, high intensity, 30 sec on, 30 sec off). For bath sonication, the nuclear suspension in sort buffer was adjusted to ChIP buffer conditions (10 mM Tris pH8, 100 mM NaCl, 0.1% sodium deoxycholate, 0.5% sarkosyl, 1% Triton-X100) in a volume of 500 µl. Nuclei were sonicated for 30 min at high intensity (30 sec on, 30 sec off) in a diagenode Bioruptor. Fragmented chromatin was snap frozen and stored at -80°C prior to ChIP.

ChIP-seq

MNase fragmented chromatin was adjusted to ChIP buffer conditions in a final volume of 500 µl, and incubated on ice for 5 min. Prior to ChIP, protease inhibitors (COMPLETE, Roche) were added to either MNase fragmented chromatin or sonicated chromatin. Chromatin was subjected to a high speed spin at 4°C for 10 min. Supernatant was moved to a new low retention tube, 1% input was removed and stored at 4°C, and the remaining chromatin was incubated overnight with the appropriate antibody at 4°C. After a high speed spin at 4°C for 10 min, supernatant was moved to a low retention tube with 10 µl prewashed protein A dynabeads (Life Technologies), and incubated at 4°C for 1-2 h with gentle rotation. After 6 rinses with ice cold ChIP buffer, immunoprecipitates were eluted from the beads with two successive additions of 125 µl freshly made elution buffer (0.2% SDS, 0.1M NaHCO₃, 5 mM DTT) incubated at 65°C for 10 min. Eluates were combined, and adjusted with Tris and EDTA to a final concentration of 10 mM Tris pH8 and 2 mM EDTA. 250 µl elution buffer was added to the reserved input and likewise adjusted. DNA cleanup began by incubating with RNase (DNase-free, Roche) at 37°C for 30 min, proteinase K (PCR-grade, Roche) at 55°C for 1 h, and crosslinks were reversed by incubating at 65°C for 1 h. After two phenol-chloroform extractions and one chloroform extraction, purified DNA was ethanol precipitated in the presence of sodium acetate and Glycoblue (Ambion). Sequencing library construction was performed according to (Bowman et

al., 2013). Libraries were sequenced on an Illumina HiSeq2000, HiSeq2500, or MiSeq according to manufacturers instructions.

ChIP antisera

anti-K27me3 (Active Motif 39155), anti-K27ac (Active Motif 39136), anti-K4me3 (Millipore 07-473), anti-CTCF (Smith et al., 2009), anti Ph (residues 772-984 (Oktaba et al., 2008)), anti-CP190 (Pai et al, 2004), anti-Pol-II (Abcam ab5131), anti-Pc (Schuettengruber et al 2009), anti-SU(Z)12 (Müller et al 2002).

Data analysis

ChIP-seq reads were aligned to the dm3 genome using BWA and allowing up to one mismatch. Potential PCR duplicates, tags with multiple alignments, and paired-end reads with insert size greater than 1 kb were removed from further analysis. We calculated normalized positional coverage as previously described (Pinter et al., 2012; Yildirim et al., 2012). Briefly, experimental coverage was normalized by corresponding input coverage for each position: $n_{\text{norm}} = [(n+1)/(ni+1)] \cdot [N_i/N]$, where n , ni , N , and N_i are positional coverages in experiment and input, and total genome coverages in experiment and input, respectively.

To determine regions of enrichment, we summed the coverage within sliding windows of size 1 kb and step 50 bp across the entire genome. The same was calculated for tags assigned a random location within the chromosome. The resulting distribution of random window coverages was then used to determine the significance of observed window coverages. Individual p-value cutoffs were selected based on manual inspection of each experiment. For samples with replicates, only regions significant in both replicates were used.

For genome-wide analysis of coverage difference between two parasegments for H3K27me3, we first merged the significant regions of enrichment from the different parasegments to create an aggregate set of regions of interest. Regions greater than 50 kb were split into equally sized sub-regions. For each region we calculated the density, or total

normalized coverage divided by region length. Replicate density values were averaged. To find differences, region densities for two parasegments were scatter plotted against one another, and a line of best fit calculated. Points located furthest away from this fit line were identified as the candidate differing regions. This approach allows us to account for the variation in sequencing coverage and noise for each sample and parasegment.

SPP (Kharchenko et al., 2008) was used to generate tag density profiles for visualization. For each chip ChIP experiment, we followed standard methods to estimate the binding peak separation distance, remove low quality or anomalous tags, and generate an input-subtracted, Gaussian-smoothed tag density profile.

Figure Legends:

Figure 1. Marking single parasegments. **A.** Drivers for the Gal4 activator and the Gal80 repressor, each with a different anterior limit, are combined genetically. Gal4 activity is thus limited to a single parasegment, and is used to activate transcription of a fluorescent nuclear envelope protein. **B.** Expression patterns used for isolation of parasegment nuclei are shown, visualized with antibody to the FLAG epitope on the INTACT fusion protein. Each panel shows embryos at about 6, 8, 10, 12, 14, and 16 h after fertilization, with the stained cells marking the indicated parasegments. Embryos between 5 and 13 h old (or between 4 and 10 h for PS6) were harvested for analysis.

Figure 1, figure supplement 1: Tools for parasegment-specific ChIP-seq. **A.** A P element was constructed that initially expresses Gal4 from the P promoter, and transgenes were moved into the homeotic gene complexes by P element swapping. Recombination between the FRT sites flanking Gal4 convert the transgene into one expressing Gal80. **B.** A second P element expressed Gal4 or Gal80, driven by enhancers from the BX-C. The enhancers were positioned

upstream of a promoter and a PRE from the *engrailed* locus. The PRE helps to maintain the restricted expression by the enhancer in older embryos. A cluster of binding sites for the *suppressor of Hairy-wing* protein flanks the enhancer, to block potential position effects at random chromosomal locations. **C.** Nuclei were sorted using Hoechst 33342 (to select only single nuclei) and mCherry fluorescence. Nuclei from wild type (Oregon R) embryos were used to set the fluorescence threshold. **D.** The flow chart from embryos to sequencing libraries is diagramed; the procedure could be paused at either freezing step.

Figure 1, figure supplement 2: Close-up of parasegment-specific expression patterns. Embryos expressing the INTACT marker in parasegments 4-7 were obtained as in Fig. 1B. All are at ~10 h of development, prior to dorsal closure; they were dissected along the dorsal midline, and flattened to display the epidermis in one focal plane. The last panel shows an embryo stained for UBX (in blue) and ABD-A (in brown)(Karch et al., 1990), to illustrate the expression patterns of these two proteins in parasegments 4-7.

Figure 2. H3K27me3, CTCF, and CP190 profiles. **A.** ChIP-seq profiles across 27 Mb of chromosome 3R are virtually identical, except at the BX-C, in the middle of the chromosome arm. **B.-F.** H3K27me3 profiles across 380 kb encompassing the BX-C. **G.-H.** CTCF binding sites. **I.-J.** CP190 binding sites. Panels B, G, and I were prepared from unsorted nuclei; parasegment-specific nuclei were used for the other panels, as indicated. Transcription units of coding genes are shown below the profiles. The *Drosophila* reference sequence includes a 6.1 kb Diver retroposon insertion at the indicated position; it was not present in the strains used for this analysis. At the bottom are shown the three regulatory domains defined by this analysis.

Figure 2, figure supplement 1: Genome-wide comparisons of H3K27me3 patterns. The most striking differences in H3K27me3 enrichment across different parasegments occur at the BX-C. Each plot compares the density of H3K27me3 between two parasegments over all H3K27me3

peaks (see supplementary methods for additional details). A linear line of fit was drawn. H3K27me3 density across the entire genome was largely unchanging between parasegments, with the exception of a few points corresponding to the BX-C that are markedly far off from the line of fit. Within each plot, the comparison of five BX-C segments (labeled a-e) are indicated with pink or red circles, the color of which corresponds to the Z-score (number of observed standard deviations from the line of fit).

Figure 2, figure supplement 2: H3K27me3, CTCF, and CP190 profiles in PS7 for the Antennapedia complex, illustrated as in Figure 2. The H3K27me3 profiles in parasegments 4 through 7 are very similar. ANTP is expressed at high levels in PS4 and PS5 in wild type embryos, and in the absence of the BX-C, parasegments 4-12 all show similar expression patterns (Carroll et al., 1986). By this criterion, *Antp* appears to be free of Pc-G repression in PS4-12, which is consistent with the lack of H3K27 methylation over its transcription unit in PS7, as shown. CTCF and CP190 profiles look identical between mixed and PS7 nuclei, as they are in the BX-C. CP190 marks several potential domain boundaries in the Antennapedia complex that lack apparent CTCF binding.

Figure 3. Additional features of the BX-C in PS7. **A.** The H3K27me3 distribution, repeated from Figure 2F, marks the border between active and repressed domains. **B., C.** H3K27ac profiles are shown for mixed and PS7 nuclei; the acetylation mark in PS7 is largely restricted to the active domains. **D., E.** Pol-II profiles show peaks at the promoters of *Ubx*, *abd-A*, and *Abd-B* both in whole embryo and in PS7 nuclei. In PS7, polymerase is distributed across the transcription units of *Ubx* and *abd-A*, but not *Abd-B*. The arrow in D marks the PS5/PS6 border; the arrow in E marks the major *Abd-B* promoter. **F., G.** H3K4me3 profiles show prominent peaks over the *Ubx* and *abd-A* promoters, both of which drive strong transcription in PS7. **H., I.** POLYCOMB profiles show sharp peaks over all known Polycomb Response Elements. In PS7, there is a reduction of the PC peak at the "bx PRE" in the active PS5 domain. There is also a PC

peak at the *abd-A* promoter, which is somewhat reduced in PS7, where *Abd-A* is transcribed. **J.,**
K. POLYHOMEOTIC protein shows peaks at known PREs and the *abd-A* promoter in both the
active and repressed domains. As with PC, the PH peak at the bx PRE is reduced in PS7. **L., M.**
SU(Z)12 profiles also mark PREs, but in PS7, the PRE peaks are reduced in all three active
domains. The ranges for the vertical axes are indicated above the left edge of each trace.

References:

- Bae, E., Calhoun, V. C., Levine, M., Lewis, E. B., and Drewell, R. A. (2002) Characterization of the intergenic RNA profile at *abdominal-A* and *Abdominal-B* in the *Drosophila* bithorax complex. *Proc. Natl. Acad. Sci. USA* *99*, 16847-16852.
- Bender, W., and Hudson, A. (2000). P element homing to the *Drosophila* bithorax complex. *Development* *127*, 3981-3992.
- Bender, W., and Lucas, M. (2013). The border between the *Ultrabithorax* and *abdominal-A* regulatory domains in the *Drosophila* bithorax complex. *Genetics* *193*, 1135-1147.
- Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., McGouran, J. F., et al. (2014). Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. *Cell* *157*, 1-15.
- Boulet, A. M., Lloyd, A., and Sakonju, S. (1991) Molecular definition of the morphogenetic and regulatory functions and the *cis*-regulatory elements of the *Drosophila Abd-B* homeotic gene. *Development* *111*, 393-405.
- Bowman, S.K., Simon, M.D., Deaton, A.M., Tolstorukov, M., Borowsky, M.L., and Kingston, R.E. (2013). Multiplexed Illumina sequencing libraries from picogram quantities of DNA. *BMC genomics* *14*, 466.
- Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* *118*, 401-415.
- Carroll, S. B., Laymon, R. A., McCutcheon, M. A., Riley, P. D., and Scott, M. P. (1986). The Localization and Regulation of *Antennapedia* Protein Expression in *Drosophila* Embryos. *Cell* *47*, 113-122.
- Casares, F., Bender, W., Merriam, J., and Sanchez-Herrero, E. (1997). Interactions of *Drosophila Ultrabithorax* regulatory regions with native and foreign promoters. *Genetics* *145*, 123-137.
- Celniker, S. E., S. Sharma, Keelan, D. J., and Lewis, E. B. (1990) The molecular genetics of the bithorax complex of *Drosophila*: *cis*-regulation in the *Abdominal-B* domain. *EMBO J.* *9*,

4277-4286.

Chopra, V. S., Hong, J-W, and Levine, M. (2009) Regulation of Hox Gene Activity by Transcriptional Elongation in *Drosophila*. *Current Biology* 19, 688-693.

Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., et al. (2014) Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Reports* 7, 1-15.

Deal, R.B., and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nature protocols* 6, 56-68.

DeVido, S.K., Kwon, D., Brown, J.L., and Kassis, J.A. (2008). The role of Polycomb-group response elements in regulation of *engrailed* transcription in *Drosophila*. *Development* 135, 669-676.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & development* 17, 1870-1881.

Gonzalez, I., Mateos-Langerak, J., Thomas, A., Cheutin, T., and Cavalli, G. (2014). Identification of Regulators of the Three-Dimensional Polycomb Organization by a Microscopy-Based Genome-wide RNAi Screen. *Mol. Cell* 54, 1-15.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108-112.

Kalb R., Latwiel, S., Baymaz, H. I., Jansen, P. W. T. C., Müller, C. W., Vermeulen, M., and Müller, J. (2014) Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nature Struc & Mol Biol* 21, 569-571.

Karch, F., Bender, W., and Weiffenbach, B. (1990). *abdA* expression in *Drosophila* embryos. *Genes & development* 4, 1573-1587.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485.

Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nature biotechnology* 26, 1351-1359.

Kwong, C., Adryan, B., Bell, I., Meadows, L., Russell, S., Manak, J. R., and White, R. (2008). Stability and Dynamics of Polycomb Target Sites in *Drosophila* Development. *PLoS Genetics* 4, e1000178.

- Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570.
- Lewis, E.B. (1981). Developmental genetics of the bithorax complex in *Drosophila*. *Developmental Biology Using Purified Genes ICN-UCLA Symposia on Molecular and Cellular Biology XXIII*, D. D. Brown and C. F. Fox, ed., Academic Press, New York.
- Lewis, P.H. (1947). *Melanogaster*-New mutants: Report of Pamela H. Lewis. *Dros Inf Serv* 21, 69.
- Maeda, R.K., and Karch, F. (2009). The bithorax complex of *Drosophila*: an exceptional *Hox* cluster. *Current topics in developmental biology* 88, 1-33.
- Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. *Nature genetics* 37, 1090-1097.
- Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heinl, N., *et al.* (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.* 26, 4203-4214.
- Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., *et al.* (2002). Histone Methyltransferase Activity of a *Drosophila* Polycomb Group Repressor Complex. *Cell* 111, 1-20.
- Nègre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., *et al.* (2010). A Comprehensive Map of Insulator Elements for the *Drosophila* Genome. *PLoS Genetics* 6, e1000814.
- Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., *et al.* (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* 471, 527-531.
- Ohlsson, R., Bartkuhn, M., and Renkawitz, R. (2010). CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. *Chromosoma* 119, 351-360.
- Oktaba, K., Gutierrez, L., Gagneur, J., Girardot, C., Sengupta, A.K., Furlong, E.E., and Müller, J. (2008). Dynamic Regulation by Polycomb Group Protein Complexes Controls Pattern Formation and the Cell Cycle in *Drosophila*. *Developmental Cell* 15, 877-889.
- Pai, C. Y., Lei, E. P., Ghosh, D., and Corces, V. G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* 16, 737-748.
- Papp, B., and Müller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes & development* 20, 2041-2054.

- Pease, B., Borges, A.C., and Bender, W. (2013). Noncoding RNAs of the *Ultrabithorax* domain of the *Drosophila* bithorax complex. *Genetics* 195, 1253-1264.
- Peifer, M., Karch, F., and Bender, W. (1987). The bithorax complex: control of segmental identity. *Genes & development* 1, 891-898.
- Pengelly, A.R., Copur, O., Jackle, H., Herzig, A., and Müller, J. (2013). A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* 339, 698-699.
- Pinter, S.F., Sadreyev, R.I., Yildirim, E., Jeon, Y., Ohsumi, T.K., Borowsky, M., and Lee, J.T. (2012). Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. *Genome research* 22, 1864-1876.
- Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Molecular cell* 25, 15-30.
- Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A., and Cavalli, G. (2009). Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos. *PLoS biology* 7, e13.
- Sepp, K.J., and Auld, V.J. (1999). Conversion of lacZ enhancer trap lines to GAL4 lines using targeted transposition in *Drosophila melanogaster*. *Genetics* 151, 1093-1101.
- Simon, J.A., and Kingston, R.E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular cell* 49, 808-824.
- Smith, S.T., Wickramasinghe, P., Olson, A., Loukinov, D., Lin, L., Deng, J., Xiong, Y., Rux, J., Sachidanandam, R., Sun, H., *et al.* (2009). Genome wide ChIP-chip analyses reveal important roles for CTCF in *Drosophila* genome organization. *Developmental biology* 328, 518-528.
- Soshnikova, N., and Duboule, D. (2009). Epigenetic Temporal Control of Mouse *HOX* Genes in Vivo. *Science* 324, 1320-1323.
- Steiner, F.A., Talbert, P.B., Kasinathan, S., Deal, R.B., and Henikoff, S. (2012). Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome research* 22, 766-777.
- Straub, T., Zabel, A., Gilfillan, G. D., Feller, C., and Becker, P. B. (2013). Different chromatin interfaces of the *Drosophila* dosage compensation complex revealed by high-shear ChIP-seq. *Genome Res.* 23, 473-485.
- Thibault, S.T., Singer, M.A., Miyazaki, W.Y., Milash, B., Dompe, N.A., Singh, C.M., Buchholz, R., Demsky, M., Fawcett, R., Francis-Lang, H.L., *et al.* (2004). A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. *Nature genetics* 36, 283-287.

695
696
697
698
699
700
701
702
703
704

Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* *136*, 3131-3141.

Yildirim, E., Sadreyev, R.I., Pinter, S.F., and Lee, J.T. (2012). X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. *Nature structural & molecular biology* *19*, 56-61.

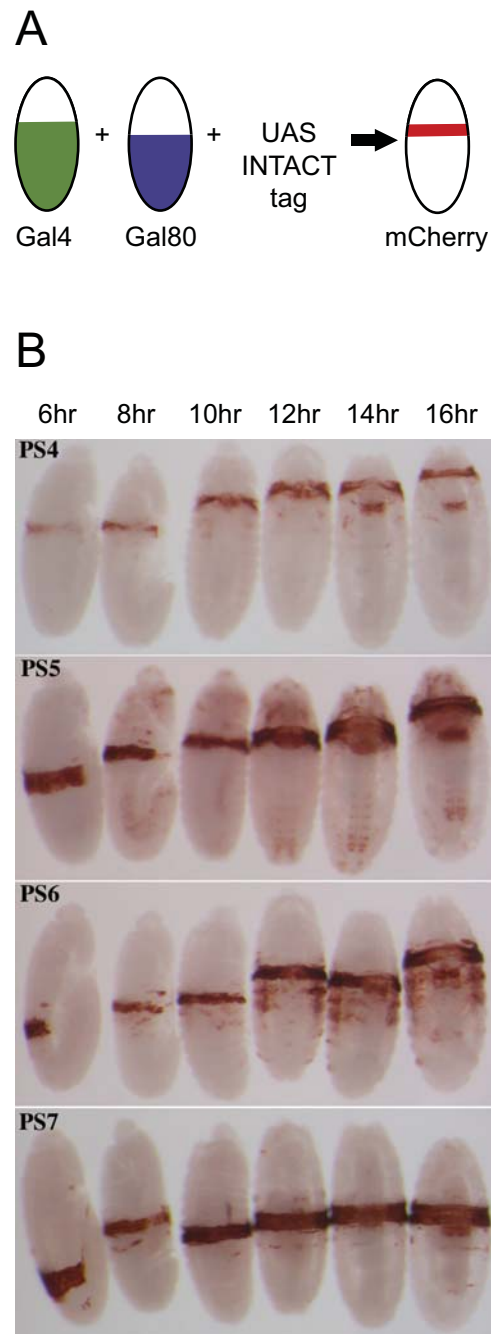


Figure 1. Marking single parasegments. **A.** Drivers for the Gal4 activator and the Gal80 repressor, each with a different anterior limit, are combined genetically. Gal4 activity is thus limited to a single parasegment, and is used to activate transcription of a fluorescent nuclear envelope protein. **B.** Expression patterns used for isolation of parasegment nuclei are shown, visualized with antibody to the FLAG epitope on the INTACT fusion protein. Each panel shows embryos at about 6, 8, 10, 12, 14, and 16 h after fertilization, with the stained cells marking the indicated parasegments. Embryos between 5 and 13 h old (or between 4 and 10 h for PS6) were harvested for analysis.

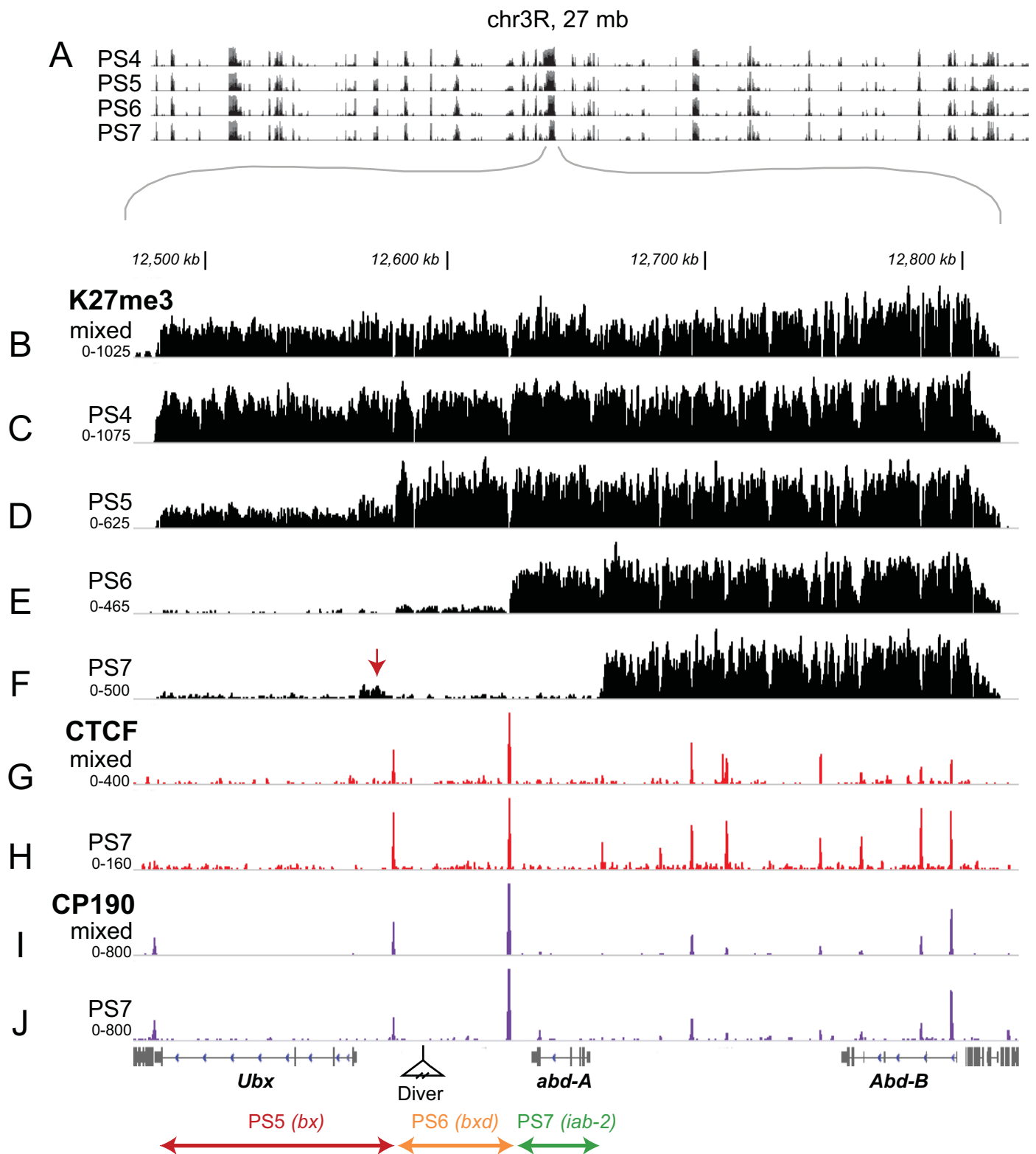


Figure 2. H3K27me3, CTCF, and CP190 profiles. **A.** ChIP-seq profiles across 27 Mb of chromosome 3R are virtually identical, except at the BX-C, in the middle of the chromosome arm. **B.-F.** H3K27me3 profiles across 380 kb encompassing the BX-C. **G.-H.** CTCF binding sites. **I.-J.** CP190 binding sites. Panels B, G, and I were prepared from unsorted nuclei; parasegment-specific nuclei were used for the other panels, as indicated. Transcription units of coding genes are shown below the profiles. The *Drosophila* reference sequence includes a 6.1 kb Diver retroposon insertion at the indicated position; it was not present in the strains used here. At the bottom are shown the three regulatory domains defined by this analysis.

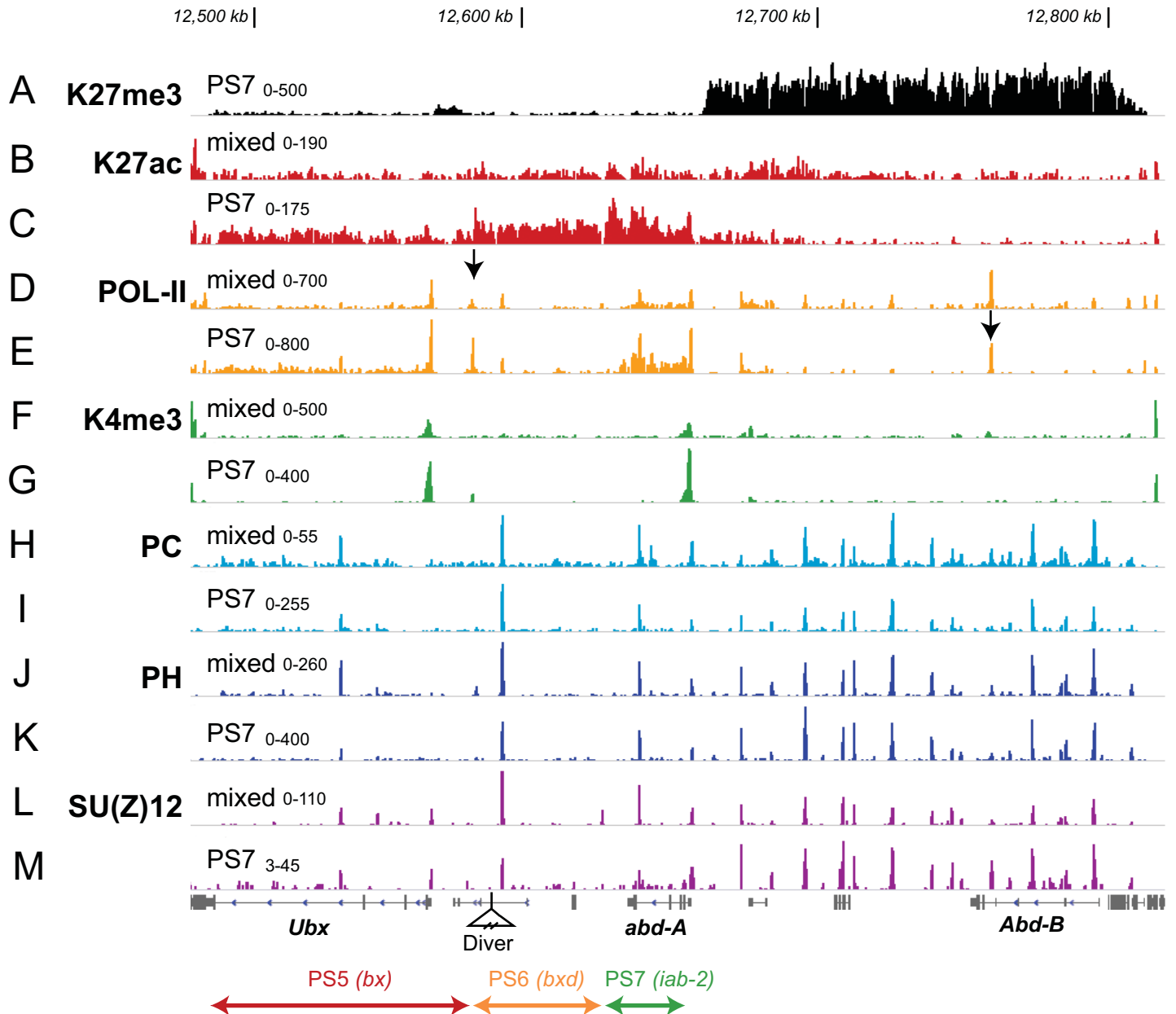


Figure 3. Additional features of the BX-C in PS7. **A.** The H3K27me3 distribution, repeated from Fig. 2F, marks the border between active and repressed domains. **B., C.** H3K27ac profiles are shown for mixed and PS7 nuclei; the acetylation mark in PS7 is largely restricted to the active domains. **D., E.** Pol-II profiles show peaks at the promoters of *Ubx*, *abd-A*, and *Abd-B* both in whole embryo and in PS7 nuclei. In PS7, polymerase is distributed across the gene bodies of *Ubx* and *abd-A*, but not *Abd-B*. Arrowhead indicates the PS5/PS6 border. **F., G.** H3K4me3 profiles show prominent peaks over the *Ubx* and *abd-A* promoters, both of which drive strong transcription in PS7. **H., I.** POLYCOMB profiles show sharp peaks over all known PREs. In PS7, there is a reduction of the PC peak at the “bx PRE” in the active PS5 domain. There is also a PC peak at the *abd-A* promoter, which is somewhat reduced in PS7, where *abd-A* is transcribed. **J., K.** POLYHOMEOTIC protein shows peaks at known PREs and the *abd-A* promoter in both the active and repressed domains. As with PC, the PH peak at the bx PRE is reduced in PS7. **L., M.** SU(Z)12 profiles also mark PREs, but in PS7, the peaks are reduced in all three active domains. The ranges for the vertical axes are indicated above the left edge of each trace.