The expression of Drosophila melanogaster Hox gene Ultrabithorax is not overtly regulated by the intronic long noncoding RNA IncRNA:PS4 in a wild-type genetic background

Anita Hermann,† Dave Kosman, William McGinnis, and Ella Tour *

*Corresponding author: Section of Cell and Developmental Biology, Division of Biological Sciences, University of California, San Diego, 3080E York Hall, 9500 Gilman Drive, La Jolla, CA 92093-0355, USA. Email: etour@ucsd.edu
†Present address: Section of Molecular Biology, Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA.

Abstract

Long noncoding RNAs (IncRNAs) have been implicated in a variety of processes in development, differentiation, and disease. In Drosophila melanogaster, the bithorax Hox cluster contains three Hox genes [Ultrabithorax (Ubx), abdominal-A, and Abdominal-B], along with a number of IncRNAs, most with unknown functions. Here, we investigated the function of a IncRNA, IncRNA:PS4 that originates in the second intron of Ubx and is transcribed in the antisense orientation to Ubx. The expression pattern of IncRNA:PS4 is complementary to Ubx in the thoracic primordia, and the IncRNA:PS4 coding region overlaps the location of the large insertion that causes the dominant homeotic mutation Contrabithorax-1 (UbxCbx-1), which partially transforms Drosophila wings into halteres via ectopic activation of Ubx. This led us to investigate the potential role of this IncRNA in regulation of Ubx expression. The UbxCbx-1 mutation dramatically changes the pattern of IncRNA:PS4, eliminating the expression of most IncRNA:PS4 sequences from parasegment 4 (where Ubx protein is normally absent) and ectopically activating IncRNA:PS4 at high levels in the abdomen (where Ubx is normally expressed). These changes, however, did not lead to changes in the Ubx embryonic transcription pattern. Targeted deletion of the two promoters of IncRNA:PS4 did not result in the change of Ubx expression in the embryos. In the genetic background of a UbxCbx-1 mutation, the IncRNA:PS4 mutation does slightly enhance the ectopic activation of Ubx protein expression in wing discs and also slightly enhances the wing phenotype seen in UbxCbx-1 heterozygotes.

Keywords: long noncoding RNA; IncRNA; Hox gene; Ultrabithorax; Ubx; Drosophila melanogaster; gene regulation; Contrabithorax (UbxCbx-1) homeotic allele

Introduction

Long noncoding RNAs (IncRNAs) are a ubiquitous group of RNA transcripts that are over 200 nucleotides long, do not code for potential proteins or peptides longer than 100 amino acids, and are transcribed from a variety of genomic locations, including enhancers, promoters, and intergenic regions (Rinn and Chang 2012; Yamamoto and Saijoh 2019). Such IncRNAs have diverse roles in development, differentiation, and disease (Yao et al. 2019). IncRNAs are primarily localized in the nucleus where they can regulate gene expression by a variety of mechanisms, including interactions with enzymatic complexes that modify and remodel chromatin, direct transcriptional regulation via interaction with Preinitiation Complex, and regulation of alternative splicing (Squillaro et al. 2020). One of the classic examples of regulatory IncRNAs is Xist which controls female X chromosome inactivation in mammals (Yao et al. 2019). Another example is the partial derepression of human and mouse HoxD genes by mutations in the IncRNA HOXATIR (Rinn and Chang 2012; Li et al. 2013), although other studies that ablated Hotair function in mice reported little or no changes in Hox gene embryonic patterning function (Schorderet and Duboule 2011; Lai et al. 2015; Amândio et al. 2016; Yao et al. 2019).

Numerous IncRNAs have been detected within the clusters of Hox genes—evolutionarily conserved complexes of genes coding for homeobox-containing transcriptional factors that control axial patterning in bilateral animals (McGinnis and Krumlauf 1992; Kumar and Krumlauf 2016). These IncRNAs are transcribed from both coding and noncoding strands of the Hox gene clusters and can be found within Hox introns, in intergenic regions, or flanking the Hox gene clusters (Kumar and Krumlauf 2016). They can function both in cis, by regulating the adjacent Hox genes and in trans, by regulating the expression of Hox genes in other Hox clusters or by regulating or modulating the activity of non-Hox genes (Kumar and Krumlauf 2016).
The Drosophila bithorax Hox cluster (BX-C) contains three Hox genes: Ultrabithorax (Ubx), abdominal-A, and Abdominal-B, as well as many noncoding RNAs (Lipshtiz et al. 1987; Sanchez-Herrero and Akam 1989; Cumberledge et al., 1990; Bae et al., 2002; Pease et al. 2013; Schor et al. 2018). Some of these noncoding RNAs are processed into microRNAs (miRNAs) that can regulate Hox expression (reviewed in Garaulet and Lai 2015), but most have no known function. An example of such a non-miRNA-containing lncRNA is bxd, which is encoded in upstream regulatory regions of the Ubx gene. Functionally, the bxd RNA has been reported to repress early transcription of Ubx in cis, perhaps by readthrough through the Ubx promoter (Petruk et al. 2006), however, the ablation of bxd RNA revealed that the early transcription repression of Ubx transcription in the bxd expressing region was very transient and no embryonic or adult Ubx mutant phenotypes were detected from bxd RNA elimination (Pease et al. 2013).

In 2013, Pease et al. (2013) discovered novel lncRNAs in a 143-kb region of the Ubx locus that included the Ubx transcription unit and its upstream bxd cis-regulatory region. One of these new lncRNAs originates in the second intron of Ubx and is transcribed in the antisense direction relative to Ubx (Pease et al. 2013), placing it among the Natural Antisense Transcript (NAT) group of noncoding RNAs (Faghihi and Wahlestedt 2009). The early embryonic domain of expression of this lncRNA was in parasegment 4; therefore, Pease and colleagues named it PS4. In accord with standard Drosophila gene nomenclature (Flybase 2020), we call this transcript lncRNA:PS4. Previous studies have shown that in numerous instances the NAT members of the sense–antisense gene pairs regulate the genes expressed from the sense strand (Faghihi and Wahlestedt 2009), raising the question whether this antisense transcript has a role in Ubx regulation. Intriguingly, the lncRNA:PS4 promoter mapped in the vicinity of the Contrabithorax (UbxCbx-1) insitional mutation (Lewis 1955; Bender et al. 1983).

Contrabithorax is a dominant mutation that results in a transformation of some adult structures of the second thoracic segment (T2) into those of the third thoracic segment (T3), e.g., transforming wings into halteres (Lewis 1963). The UbxCbx-1 allele results from an insertion of 17 kb from an Ubx upstream regulatory element into the second intron of Ubx and is the only known Contrabithorax allele that maps to that region (Bender et al. 1983; Lewis 1955). This inverted insertion contains parts of two Ubx upstream regulatory regions: 16 kb of the phx region and 1 kb of bxd. In Contrabithorax mutants, Ubx transcript and protein expression are abnormally activated in the wing discs, primarily in the posterior compartment of the disc, which results in a partial transformation of wing toward haltere morphology (White and Akam 1985); however, the mechanisms by which Cbx mutations ectopically activate Ubx expression are not understood.

Taken together, the lncRNA:PS4 domain of early expression that is adjacent to the Ubx expression domain and the proximity of the Ubx-deregulating mutation UbxCbx-1 to lncRNA:PS4 gene lead us to investigate the following questions:

- Does lncRNA:PS4 have a role in the regulation of Ubx transcription?
- To what extent does the UbxCbx-1 mutation change lncRNA:PS4 transcription?
- If such changes occur, do they play a role in modifying the morphological functions provided by Ubx?

## Materials and methods

### Determination of the extent of lncRNA: PS4 transcription unit

We used 5′ and 3′ RACE [5′/3′ RACE kit, second generation (Roche)] to determine the start and the end of lncRNA:PS4 transcription. The extent of the lncRNA:PS4 transcription unit was determined as follows: the region between the lncRNA:PS4 promoter and Ubx transcription start site was virtually subdivided into regions of about 1 kb long. These regions were cloned by PCR amplification from the genomic DNA of w1118 flies. Antisense RNA probes were generated using in vitro transcription and the presence of transcripts from each region was assessed using whole-mount in situ hybridization. To determine if any of the transcribed regions were spliced, RT-PCR of 900–1100 bp amplicons spanning the entire predicted transcript was produced, followed by sequencing. For RT-PCR template generation, total RNA was extracted with RNAeasy mini kit (Qiagen) and treated with Turbo DNA-free kit (Ambion) to remove DNA.

### Determination of lncRNA: PS4 expression pattern

Whole-mount in situ hybridization was performed as previously described (Kosman et al. 2004). The Ubx cDNA probe was generated from the BSUbx1a plasmid (Harding et al. 1985). Another probe to detect Ubx nascent transcripts was prepared with a pBluescript plasmid clone that included the 5′-most 1.5 kb of the first intron of Ubx. The wingless probe was generated from a 3 kb clone in pBluescript plasmid (a gift from B. Cohen). lncRNA:PS4 transcription was detected via the M probe (upstream of UbxCbx-1 insertion), or the X probe (downstream of UbxCbx-1 insertion, Figure 1B). Images were obtained using Leica SP4, Leica SP5, and Leica SP8 fluorescent confocal microscopes and processed in Adobe Photoshop software.

### CRISPR/Cas9-generated mutations in lncRNA:PS4 promoters

To generate a targeted deletion of the sequences encompassing both putative promoters of lncRNA:PS4, two gRNA sequences were designed using CRISPR Optimal Target Finder (Gratz et al. 2014) and cloned into pCFD4-U6.1_U6:3tandem-gRNAs (Addgene plasmid No. 49411), using the protocol described in Port et al. (2014). The first guide RNA sequence targeting the site 182 bp upstream of the “upstream” promoter was placed under the control of the U6.1 promoter in pCFD4 vector and its sequence was GGAGTAAATTTATCTGGCTCT (in the genome, this sequence would be followed by a 3′ PAM: CGG). The guide RNA sequence targeting the “downstream” promoter was placed under the control of the U6.3 promoter and its sequence was GGTTGATTTCTATTGGCTCT (in the genome, this sequence would be followed by a 3′ PAM: CGG). gRNA cloning was verified by sequencing.

Transgenic lines carrying the gRNA construct were generated by BestGene Inc (Chino Hills, CA, USA). Males of this gRNA line [genotype yISCµv1; P(CaryP+I]+attP40/CyO] were crossed to females carrying Vasa-Cas9 transgene either on the X chromosome (Bloomington stock No. 52669) or on the third chromosome (Bloomington stock No. 51324) and the progeny of these crosses was back-crossed. To screen for the presence of the desired deletions, we PCR-amplified the region surrounding the lncRNA:PS4 promoters from whole-fly genomic DNA, followed by sequencing of the PCR fragments. One line of transgenic flies was isolated from the cross to a female with a Vasa-Cas9 transgene on the
third chromosome. This line has a 324 bp deletion that eliminates both the upstream promoter and the downstream promoter of lncRNA:PS4. Homozygous for this deletion chromosome survive to adulthood at 5% the expected frequency, although this effect on viability does not map to the 324 bp deletion or to the Ubx region, as the chromosome with the small deletion is fully viable when crossed to a chromosome bearing the Ubx\(^{109}\) mutation [D\(\text{f}(3\text{R})\text{Ub} \text{x}^{109}\)], which deletes the entire Ubx locus.

**Drosophila stocks and crosses**

The adult wing phenotype of lncRNA:PS4 heterozygote mutants was assessed in lncRNA:PS4\(^{P}\)/TM3Sb flies, maintained as a stock. Ubx protein expression in the wing discs was assessed in lncRNA:PS4\(^{P}\)/TM3Sb, Kr-GFP flies, maintained as a stock. Ubx\(^{Cbx-1}\) mutant flies [Ubx\(^{Cbx-1}\)/\(T(2; 3)apXa, apXa\)] were obtained from Bloomington Drosophila Stock Center (stock No. 3433). This fly strain has had the deletion for the \(pxb/\text{bd}\) region crossed away from the inselational Ubx\(^{Cbx-1}\) mutation, so is wild type for sequences upstream of the Ubx transcription unit. Ubx\(^{Cbx-1}\) adult wing phenotypes were assessed in Ubx\(^{Cbx-1}\)/TM3Sb flies that were obtained by crossing Ubx\(^{Cbx-1}\)/\(T(2; 3)apXa, apXa\) flies to lncRNA:PS4\(^{P}\)/TM3Sb flies; the resulting Ubx\(^{Cbx-1}\)/TM3Sb flies were maintained as a stock. Ubx protein expression in the wing discs of Ubx\(^{Cbx-1}\) heterozygotes was assessed in Ubx\(^{Cbx-1}\)/TM3Sb, Kr-GFP, obtained via the cross of Ubx\(^{Cbx-1}\)/TM3Sb, Kr-GFP flies. Ubx\(^{Cbx-1}\)/TM3Sb and Ubx\(^{Cbx-1}\)/TM3Sb, Kr-GFP flies had identical wing phenotypes. Pc\(^2\) mutant flies [\(\text{In}(3\text{R})\text{P}(\text{Pc}^2), \text{Pc}^2/\text{TM1}\)] were obtained from Bloomington Drosophila Stock Center (stock No. 106475). Ubx\(^{Cbx-1}\)/Pc\(^2\) flies were obtained via a cross between In(3R)P(Pc\(^2\), Pc\(^2\)/TM1 and Ubx\(^{Cbx-1}\)/apXa, apXa; the resulting Ubx\(^{Cbx-1}\)/Pc\(^2\) flies were maintained as a stock.

**Results**

**Characterization of lncRNA:PS4 transcription unit**

lncRNA:PS4 was identified by Pease and colleagues in a systematic survey of noncoding RNAs originating from the Ubx gene and its upstream regulatory elements (Pease et al. 2013). Pease and colleagues characterized the lncRNA:PS4 transcription unit as spanning ~15 kb and including nonprotein coding sequences from the first and the second Ubx introns (Pease et al. 2013). Here, we characterized the exact boundaries of the lncRNA:PS4 transcription unit, using 5' and 3' RACE, followed by verification using fluorescent in situ hybridization with ~1 kb probes that mapped throughout the lncRNA:PS4 region (Figure 1).

Using 5' RACE, we determined that lncRNA:PS4 has two transcription start sites (P1 and P2), separated by 116 bp and located 14,299 and 14,182 bp downstream of the transcription start site of Ubx, respectively. Promoter site 2 is preceded by a good match, TCATT, to the Drosophila Inr promoter consensus sequence of TCAGTY (Ngoc et al. 2019, Supplementary Figure S1). Promoter 2 (P2, Figure 1) also includes a putative DPE promoter sequence (Supplementary Figure S1). Promoter 1 (P1) contains a weaker match, TCAGTG, to the Inr promoter consensus and no matches to a DPE consensus sequence (Supplementary Figure S1). In our 5' RACE reactions, the genomic coordinate for the first 5' nucleotide of the P1 transcripts is 16,720,127, and for P2 the initiation site maps to nucleotide 16,720,245 (Drosophila melanogaster genome release r6.37, Larkin et al. 2021). The 3' RACE results indicated that lncRNA:PS4 transcripts are polyadenylated and located the 3' end of lncRNA:PS4 transcripts to nucleotide 16,729,960 on Chromosome 3R (D. melanogaster genome release r6.37, Larkin et al. 2021). The predicted length of lncRNA:PS4 transcription unit was 9833 bp for transcripts starting from promoter P1 and 9715 bp for transcripts starting from promoter P2.

To determine if any of the sections of the predicted transcript were spliced, we performed RT-PCR of 900–1100 bp amplicons spanning the entire predicted transcript, followed by sequencing. All RNA-generated sequences were the same size and sequence as genomic DNA, indicating that the lncRNA:PS4 transcript contains no introns. The only potential protein coding region identified by PhyloCSF in the lncRNA:PS4 transcript could in theory encode a 39 amino acid peptide, MLQMPYKTVPFCCFMNCFCSVGKSTTTLATHKVLPS (Lin et al. 2011). We were unable to detect the conservation of this protein coding region in other Drosophila species. No palindromic miRNA precursor sequences were identified in the lncRNA:PS4 transcription unit.

**lncRNA:PS4 expression**

Pease et al. (2013) described the blastoderm expression of lncRNA:PS4 as a broad stripe, with PS4 and PS10 as its anterior
and posterior borders, respectively. Here, we show that lncRNA:PS4 RNA is first detected by in situ hybridization during stage 5 of Drosophila embryonic development (Figure 2A). The expression pattern of lncRNA:PS4 between stages 6 and 12 was previously characterized (Pease et al. 2013). During stages 6 and 7, the lncRNA:PS4 pattern of expression resolves into a broad and strong anterior domain and weaker posterior stripes (Figure 2B and Supplementary Figure S2). Containing with a wingless probe at stage 6 allowed us to map the broad anterior domain of expression to the posterior part of parasegment 4 and the entirety of parasegment 5 (segment T2 and the anterior compartment of T3; Supplementary Figure S3), thus abutting the anterior boundary of Ubx expression. lncRNA:PS4 is excluded from the ventral mesoderm primordia at this stage (Pease et al. 2013).

We were also interested in precisely mapping the lncRNA:PS4 transcription domain relative to that of Ubx, in whose intron lncRNA:PS4 resides. Throughout stages 6–10, Ubx and lncRNA:PS4 transcription domains were largely nonoverlapping; however, in a few nuclei, both lncRNA:PS4 and Ubx transcription were detected in parasegments 5 through 12 (Supplementary Figures S2 and S3). High-resolution imaging at stage 7 (Figure 2, B–E) shows that in the domain of strong Ubx transcription, lncRNA:PS4 transcripts are absent from almost all Ubx transcribing nuclei (Figure 2, C–E).

To better understand the cellular localization of lncRNA:PS4 transcripts, we performed high-resolution imaging of stage 6 embryos after in situ hybridization with lncRNA:PS4 and Ubx probes and costaining with lamin antibodies that outline the nuclear envelope (Figure 2, F and G). lncRNA:PS4 signals are localized almost exclusively in the nuclear compartment (Figure 2F). There is a prominent signal corresponding to lncRNA:PS4 nascent transcripts (Figure 2F, arrowheads), as well as other small lncRNA:PS4 nuclear RNA signal “specks” that are not associated with the lncRNA:PS4 locus, presumably originating from polyadenylated transcripts that accumulate in that compartment (Figure 2F, arrows). Ubx, which was detected with a probe against its first intron, was mostly observed as the nascent transcripts being made from Ubx loci on the two homologous chromosomes (Figure 2G, arrowheads).

Using in situ hybridization, we made multiple attempts to detect spatially localized transcription from lncRNA:PS4 in wing, haltere and leg discs, using our most sensitive method of tyramide amplification. We never observed a localized signal within discs, or between different discs. The positive controls were in situ hybridizations with engrailed and Ubx cDNA probes, which had easily detectable localized signals. When using quantitative RT-PCR with gene-specific primers covering region X (Figure 1B), we did detect low-level transcripts from this region in the discs we tested (wing, leg, and haltere; Supplementary Table S1). The abundance of these low-level transcripts was slightly higher in haltere and leg discs than in wing discs, the opposite of what we expected from the embryonic expression pattern of lncRNA:PS4. The positive control was Ubx expression levels using quantitative RT-PCR on RNA from the same discs.

Figure 2 Embryonic pattern and nuclear localization of lncRNA:PS4 transcripts. (A) lncRNA:PS4 transcription is first detected at stage 5 by in situ hybridization with probe. (B) A stage 7 embryo was hybridized with lncRNA:PS4 (probe X, green) and Ubx (first intron antisense probe, red) probes and imaged at 20× magnification. In (C–E), the embryo in (B) was imaged at 63× magnification, focusing on the anterior stripes of lncRNA:PS4 and Ubx, maximal projections spanning most of the depth of the epidermis are shown. (C) lncRNA:PS4 transcripts. (D) Ubx transcripts. (E) an overlay of lncRNA:PS4 and Ubx signals. (F, G) Stage 6 wild-type embryo, probed with lncRNA:PS4 (probe X, green) and Ubx (first intron antisense probe, red) probes, with the addition of a lamin antibody stain (cyan) that outlines nuclei. (F) Nuclei in the anterior stripe of lncRNA:PS4 expression, in a region where Ubx expression is absent. (G) Same embryo as in F, nuclei where lncRNA:PS4 and Ubx expressions overlap.

Arrowheads point to nascent transcripts, whereas arrows point to nuclear “specks,” which we attribute to nuclear transcripts (in the case of lncRNA:PS4) that are not associated with the transcription site on the lncRNA:PS4 locus or to transient products of the spliced first intron (in the case of Ubx).

lncRNA:PS4 location relative to Ubx^{Chx-1} insertion

In order to determine the exact location of Ubx^{Chx-1} insertion in the lncRNA:PS4 region, we performed genomic PCR amplification followed by sequencing analysis of the approximate region of Ubx^{Chx-1} insertion in Ubx^{Chx-1}/T(2; 3)ap^A3, ap^A3 flies (Bloomington). This allowed us to map the Ubx^{Chx-1} insertion location to position 16,721,339 of Chromosome 3R (D. melanogaster genome release r6.37, Larkin et al. 2021). Thus, the Ubx^{Chx-1} insertion is located within the lncRNA:PS4 transcribed region, 1.06 kb downstream of the lncRNA:PS4 P2 transcription start site.
The effect of the UbxCbx-1 insertion on transcription of the lncRNA:S4 region

The UbxCbx-1 insertion profoundly changes the extent and the pattern of lncRNA:PS4 embryonic transcription (Figure 3, A–C). The insertion splits the lncRNA:PS4 transcription unit and results in two separate transcripts, as shown schematically in Figure 1B: a transcript containing ~1 kb of 5’ lncRNA:PS4 sequences fused to an unknown amount of Cbx insertion sequence, and another transcript initiated within UbxCbx-1 insertion that is fused to the 3’ most 8.6 kb of lncRNA:PS4. The homozygous UbxCbx-1 mutant embryo shown in Figure 3, A–C demonstrates the different expression patterns of these two transcripts. In this in situ hybridization, lncRNA:PS4 transcription was imaged with two probes (Figure 1B): probe M contains 1.4 kb of sequence upstream and 0.4 kb downstream of the UbxCbx-1 insertion (Figure 3B) and probe X, located downstream of the insertion, thus detecting the transcription of the transcript that initiates within the UbxCbx-1 insertion and has 3’ sequences of lncRNA:PS4 (Figure 3A). To simplify, we will refer to this transcript as the Cbx hybrid transcript.

The strongest domain of the Cbx hybrid expression was posterior to parasegment 6 and its expression extended uniformly to the posterior boundary of parasegment 12 (Figure 3A). This transcript was also observed in a few nuclei located in parasegments 4–6 (Figure 3, A and C). In addition to the abundant posterior expression relative to wild-type lncRNA:PS4, the Cbx hybrid transcript was not excluded from embryonic mesoderm, unlike wild-type lncRNA:PS4 (Figure 3, A and C, arrow). The Cbx hybrid transcript pattern mimics the normal expression pattern of the bxd transcript (Lipshitz et al. 1987; Pease et al. 2013; Petruk et al. 2006). In UbxCbx-1 mutants, the bxd promoter and regulatory sequences lie within the 17 kb insertional mutation, in the opposite orientation to their normal direction in the BX-C complex. The expression of the Cbx hybrid transcript, therefore, appears to be driven by bxd enhancers located in the insertion. In the hybrid transcript, transcription of the 3’ region of lncRNA:PS4 extends beyond its normal termination site to at least the vicinity of the Ubx first exon, since Cbx hybrid transcripts can be detected with a probe corresponding to the 5’ most 1.5 kb of the first intron of Ubx (Supplementary Figure S4, A and B).

The M probe primarily detected the expression of the lncRNA:PS4 that is encoded upstream of the Cbx1 insertion (Figure 3B). This 5’ 1 kb region of lncRNA:PS4 continues to be transcribed in parasegments 4 and 5 and continues to be largely excluded from the mesoderm, similar to the lncRNA:PS4 in wild-

![Figure 3](image-url)
type embryos (Figure 3B, arrow). Because M probe has a small amount of homology to sequences downstream of the Cbx insertion (Figure 1B), it also weakly detects the Cbx hybrid transcript (see the weak mesodermal signal, probably resulting from codetection of the Cbx hybrid transcript in Figure 3, B and C). Thus, in UbxCbx-1 mutants, the IncRNA:PS4 transcription unit is split into two transcripts: the 1 kb upstream of the insertion that is expressed in the largely wild-type IncRNA:PS4 pattern and another, containing over 8 kb of IncRNA:PS4 downstream of the insertion that is expressed in bxd pattern (Figure 3C).

As the UbxCbx-1 insertion eliminates the parasegment 4 and 5 expression of the 3′-most 8.6 kb of IncRNA:PS4 transcribed region, we next investigated whether this change was associated with the expansion of Ubx transcription into that region. We detected no expansion of embryonic Ubx transcription into parasegments 4 and 5 of UbxCbx-1 mutants (Supplementary Figure S4, C and D), suggesting that PS4–PS5 expression of IncRNA:PS4 was not necessary for the repression of Ubx transcription anterior to its normal domain in embryos. Furthermore, the abundant and broadened transcription of the 3′ portion of IncRNA:PS4 sequences in parasegments 6–12 of UbxCbx-1 mutants were not associated with suppression of Ubx transcription in those parasegments (Supplementary Figure S4, C and D).

Using high-resolution microscopy, we investigated whether the presence of the Cbx hybrid transcript interferes with transcription of the Ubx gene from the same locus. To that end, we detected Ubx transcription using a probe for the 5′-most 7.5 kb-long Ubx intron, which should detect the first ~10 min of Ubx transcription (Figure 3, D–G). In some nuclei we detected Ubx and Cbx hybrid transcripts that appeared to derive from the same DNA molecule (Figure 3, F and G), thus arguing against transcriptional interference by Cbx hybrid transcript. One caveat to this statement is that it is possible that some of the imaged chromosomes contained replicated sister chromatids and so the overlap of Cbx hybrid and Ubx transcript signals could be due to adjacent DNA molecules in cell cycle G2.

CRISPR-mediated knockout of IncRNA:PS4 has no effect on Ubx expression, but it interacts genetically with the UbxCbx-1 mutation in developing wings

To further test the function of the IncRNA:PS4 transcript, we generated a CRISPR/Cas9-mediated deletion of 324 bp that removed both of its promoters. Whole-mount in situ hybridization demonstrated that IncRNA:PS4 transcription was largely eliminated in the homozygous mutants (Supplementary Figure S5). Next, we investigated the effects of the IncRNA:PS4 mutation on Ubx expression. In embryos homozygous for the promoter deletion, the abundance and pattern of Ubx transcripts were the same as in wild-type embryos. Only 5–6% of the IncRNA:PS4 homozygous mutant flies emerged as adults (with most homozygotic dying in late larva and pupa stage), apparently because of an additional mutation elsewhere on the third chromosome, since trans-heterozygotes of the IncRNA:PS4 promoter deletion over a deletion of the Ubx locus [D(3R)Ubx109], resulted in viable adults at approximately wild-type frequencies. We then examined if the deletion of IncRNA:PS4 promoters resulted in ectopic activation of Ubx protein in larval wing discs, or in wing abnormalities. The wing discs and the wings of the surviving homozygous adults were similar to those of flies heterozygous for the deletion and contained no wing deformations (Figure 4A), or ectopic activation of Ubx protein in wing discs (Figure 4B).

We also examined whether the IncRNA:PS4 promoter deletion mutation interacted genetically with the UbxCbx-1 mutation. In wing discs heterozygous for UbxCbx-1, we observed ectopic Ubx protein in the posterior region of the disc (Figure 4D), as previously described (White and Akam 1985). In our hands, adults that were heterozygous for the UbxCbx-1 insertion had only slight defects in wing morphology (Figure 4C), likely because of accumulation of modifiers in this strain. As reported previously (Castelli-Gair et al. 2019), Polycorm (Pc) mutants enhance the amount and extent of Ubx expression in the wing disc of UbxCbx-1 mutants (Figure 4H), and have a stronger transformation of wings toward halteres (Figure 4G). The IncRNA:PS4 promoter deletion also enhances the adult wing phenotype of UbxCbx-1 (Figure 4E), albeit much more mildly than Pc. The ectopic expression of Ubx protein in the posterior compartment of wing discs from UbxCbx-1/ IncRNA:PS4 larvae was also enhanced (Figure 4F), compared with that seen in discs of larvae heterozygous for UbxCbx-1.

Discussion

Although in embryos the antisense IncRNA:PS4 transcripts are expressed in cells that are largely nonoverlapping with Ubx, the function of IncRNA:PS4 seems to have little if any effect on the pattern of Ubx expression in otherwise wild-type animals. Two lines of evidence support this conclusion. First, a small deletion that eliminates almost all IncRNA:PS4 transcription has no detectable effect on Ubx expression in embryos and in imaginal discs. Second, the abundant transcription of 3′ IncRNA:PS4 sequences driven by the UbxCbx-1 insertion in posterior regions of the embryo has no detectable repressive effect on Ubx transcription in such posterior regions.

In animals that are heterozygous for UbxCbx-1 and IncRNA:PS4 promoter mutations there is an enhancement of the UbxCbx-1 phenotype, and one can interpret this as noncomplementation between these two alleles, suggesting that IncRNA:PS4 has a very subtle repressive effect on Ubx transcription in wing discs which can only be observed in a UbxCbx-1 mutant background. One possible mechanism by which the IncRNA:PS4 region may provide this subtle repression is by helping to set a repressive chromatin state in the cells in which it is expressed, as was suggested for the bxd IncRNA (Pease et al. 2013).

One puzzle is how the enhancer/promoter for IncRNA:PS4 that resides in the Ubx locus is able to activate transcription within parasegment 4, whereas the rest of the Ubx locus is repressed in parasegment 4. This seems potentially inconsistent with the “open for business” model, in which the Ubx locus as a whole has been observed to be associated with H3K27 trimethylation repressive modifications in the bulk of embryonic parasegment 4 cells (Bowman et al. 2014; Maeda and Karch 2015). Those modifications recruit Polycomb group proteins that maintain the off state of Ubx in those cells. However the expression of IncRNA:PS4 is transient in most cells of parasegment 4, and may help recruit Polycomb group complexes at very early embryonic stages that then quantitatively assist in the maintenance of Ubx repression in the cells of that parasegment through later stages of embryogenesis and beyond.

The IncRNA:PS4 promoters are both contained within a predicted Polycomb Repression Element (Supplementary Figure S1; Négré et al. 2011). It may be that either IncRNA:PS4 transcripts themselves contribute to the function of this PRE or that they are merely a signal associated with the PRE regulatory element, as noncoding RNAs are often produced in the vicinity of cis-regulatory elements (Natoli and Andrau 2012). If this is the case, it is
possible that the Ubx\textsuperscript{Cbx\textendash}1-dependent effect of lncRNA:PS4 promoter deletions on Ubx expression in the wing is due to the sequences removed from the predicted PRE, and not due to elimination of lncRNA:PS4 transcripts per se. It is also possible that lncRNA:PS4 transcripts have a repressive effect on Ubx in some cells that we did not test, but an important effect of lncRNA:PS4 on early embryonic patterning via regulation of Ubx is not supported by our results.

In some respects, the Hox-encoded lncRNA:PS4 transcripts resemble that of the Hox cluster-encoded HOTAIR transcripts of mice. Different studies have proposed that HOTAIR either has an important (Rinn et al. 2007; Li et al. 2013), or largely unimportant (Amândio et al. 2016; Schorderet and Duboule 2011) role in regulating Hox gene expression and embryonic patterning. As proposed by Amândio et al. (2016), these inconsistent findings might be explained by the different genetic backgrounds in which HOTAIR mutant alleles were tested. As described here, the function of lncRNA:PS4 transcripts also seem sensitive to genetic background, as lncRNA:PS4 mutant animals only exhibit ectopic expression of Ubx and abnormal wing shapes in the background of a Ubx\textsuperscript{Cbx\textendash}1 mutant allele.

**Data availability**

Strains and plasmids are available upon request. Reagents and primer sequences not included in the Materials and Methods section are listed in the Reagents Table. The GenBank accession number of lncRNA:PS4 is OK501974.

Supplementary material is available at G3 online.

**Funding**

This research was supported by NIH grant R01HD28315 to W.M.
Conflicts of interest

The authors declare that there is no conflict of interest.

Literature cited

Bender W, Akam M, Karch F, Beachy PA, Peifer M, et al. 1983. Molecular genetics of the bithorax complex in D. melanogaster. Science. 221:23–29.

Bowman SK, Deaton AM, Domingues H, Wang PI, Sadreyev RI, et al. 2014. H3K27 modifications define segmental regulatory domains in the. Elife. 3:e02833.

Cumberledge S, Zaratzian A, Sakonju S. 1990. Characterization of two RNAs transcribed from the cis-regulatory region of the abd-A domain within the Drosophila bithorax complex. Proceedings of the National Academy of Sciences of the United States of America. 87:3259–3263. Doi: 10.1073/pnas.87.9.3259.

Castelli-Gair J, Muller J, Bienz M. 1992. Function of an Ultrabithorax minigene in imaginal cells. Development. 114:877–886. https://doi.org/10.1242/dev.114.4.877

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, et al. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 196:961–971.

Harding K, Wedeen C, McGinnis W, Levine M. 1985. Spatially regulated expression of homeotic genes in Drosophila. Science. 229:151–159.

Kosman D, Mizutani CM, Lemons D, Cox WG, McGinnis W, et al. 2004. Multiplex detection of RNA expression in Drosophila embryos. Science. 305:846.

Kumar B, Krumlauf R. 2016. Molecular biology: HOXs and lincRNAs: two sides of the same coin. Sci Adv. 2:1–8.

Lai K-MV, Gong G, Atanasio A, Rojas J, Quispe J, et al. 2015. Diverse domains with ablated lincRNAs. PLoS One. 10:e0125522.

Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, et al.; FlyBase Consortium. 2021. FlyBase: updates to the Drosophila melanogaster knowledge base. Nucleic Acids Res. 49: D899–D907.

Lewis EB. 1955. Some aspects of position pseudoallelism. Am Nat. 89: 73–89.

Lewis E. 1963. Genes and developmental pathways. Am Zool. 3: 33–56.

Li H, Liu B, Wapinski OL, Tsai MC, Qu K, et al. 2013. Targeted disruption of Hotair leads to homeotic transformation and gene derepression. Cell Rep. 5:3–12.

Lin MF, Jungeire I, Kellis M. 2011. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics. 27:1275–1282.

Lipshitz HD, Peatitie DA, Hogness DS. 1987. Novel transcripts from the Ultrabithorax domain of the bithorax complex. Genes Dev. 1: 307–322.

Maeda RK, Karch R. 2015. The open for business model of the bithorax complex. Chromosoma. 124:293–307.

McGinnis W, Krumlauf R. 1992. Homeobox genes and axial patterning. Cell. 68:283–302.

Natoli G, Andraud J-C. 2012. Noncoding transcription at enhancers: general principles and functional models. Annu Rev Genet. 46: 1–19.

Négé N, Brown CD, Ma L, Bristow CA, Miller SW, et al. 2011. A cis-regulatory map of the Drosophila genome. Nature. 471: 527–531.

Ngoc LV, Kassavetis GA, Kadonaga JT. 2019. The RNA polymerase II core promoter in Drosophila. Genetics. 212:13–24.

Pease B, Borges AC, Bender W. 2013. Noncoding RNAs of the Ultrabithorax domain of the Drosophila bithorax complex. Genetics. 195:1253–1264.

Petruk S, Sedkov Y, Riley KM, Hodgson J, Schweisguth F, et al. 2006. Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. Cell. 127: 1209–1221.

Port F, Chen H-M, Lee T, Bullock SL. 2014. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc Natl Acad Sci U S A. 111:E2967–E2976.

Rinn JL, Chang HY. 2012. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 81:145–166.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 129:1311–1323.

Sanchez-Herrero E, Akam M. 1989. Spatially ordered transcription of regulatory DNA in the bithorax complex of Drosophila. Development. 107:321–329.

Schor I, Bussotti G, Malei M, Forneris M, Viales R, et al. 2018. Non-coding RNA expression, function, and variation during Drosophila embryogenesis. Curr Biol. 28:3547–3561.e9.

Schorderet P, Duboule D. 2011. Structural and functional differences in the long non-coding RNA hotair in mouse and human. PLoS Genet. 7:e1002071.

Squillaro T, Peluso G, Galderisi U, Di Bernardo G. 2020. Long non-coding RNAs in regulation of adipogenesis and adipose tissue function. Elife. 9:1–15.

White RAH, Akam ME. 1985. Contrabithorax mutations cause inappropriate expression of Ultrabithorax products in Drosophila. Nature. 318:567–569.

Yamamoto T, Saitho N. 2019. Non-coding RNAs and chromatin domains. Curr Opin Cell Biol. 58:26–33.

Yao RW, Wang Y, Chen LL. 2019. Cellular functions of long noncoding RNAs. Nat Cell Biol. 21:542–551.

Communicating editor: B. Oliver