3352–3361 *Nucleic Acids Research, 2013, Vol. 41, No. 5* *Published online 17 January 2013*
doi:10.1093/nar/gks1445

Structure, evolution and function of the bi-directionally transcribed iab-4/iab-8 microRNA locus in arthropods

Jerome H. L. Hui[1,2], Antonio Marco[2], Suzanne Hunt[2], Janet Melling[2], Sam Griffiths-Jones[2] and Matthew Ronshaugen[2,*]

[1]School of Life Sciences, Chinese University of Hong Kong, Shatin, Hong Kong and [2]Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK

Received October 25, 2011; Revised December 10, 2012; Accepted December 12, 2012

ABSTRACT

In *Drosophila melanogaster*, the iab-4/iab-8 locus encodes bi-directionally transcribed microRNAs that regulate the function of flanking Hox transcription factors. We show that bi-directional transcription, temporal and spatial expression patterns and Hox regulatory function of the iab-4/iab-8 locus are conserved between fly and the beetle *Tribolium castaneum*. Computational predictions suggest iab-4 and iab-8 microRNAs can target common sites, and cell-culture assays confirm that iab-4 and iab-8 function overlaps on Hox target sites in both fly and beetle. However, we observe key differences in the way Hox genes are targeted. For instance, abd-A transcripts are targeted only by iab-8 in *Drosophila*, whereas both iab-4 and iab-8 bind to *Tribolium* abd-A. Our evolutionary and functional characterization of a bi-directionally transcribed microRNA establishes the iab-4/iab-8 system as a model for understanding how multiple products from sense and antisense microRNAs target common sites.

INTRODUCTION

Specification of segmental identity on the head–tail axis of all animals is accomplished, in large part, by the highly coordinated expression of an ancient conserved complex of related transcription factors collectively known as Hox genes (1). Since the first molecular characterization of the *Drosophila* Hox complex, it has been recognized that numerous non-coding RNAs are interspersed between the protein coding genes (2,3). Subsequent work in other animals has found that non-coding RNAs are a prevalent feature of many animal Hox complexes (4). Insight into the function of a subset of these RNAs was facilitated by the discovery of microRNAs (miRNAs): ~22-bp RNAs that direct translational repression. The Hox gene complex contains several miRNAs, which are themselves ancient genes, and their conserved positions within the complex suggest that this association arose near the time of the origin of the Hox complex itself (1). Only a small number of metazoan miRNAs conserved between distantly related metazoan animals date from this time (~600 million years ago). Indeed, the oldest shared conserved animal miRNA is miR-100, which was duplicated in an ancestor of bilaterians to give rise to the Hox miR-10 miRNA (5). Studies of Hox complex miRNA function find that they seem to act in part by modulating the translation of Hox genes themselves. As translational repression occurs after processing in the cytoplasm, there is no obvious reason why a miRNA should target genes located nearby in the genome.

Recent advances in sequencing technologies have enabled large-scale identification of miRNAs in animals (6,7). The available data suggest that only a tiny proportion of miRNA loci are transcribed in both sense and antisense directions. Indeed this bi-directional transcription has been shown for only a handful of loci in a variety of animals [e.g. *Drosophila melanogaster* (8) and human (9)]. In these cases, the two distinct primary transcripts are independently processed by Drosha and Dicer to produce distinct mature miRNA sequences. The origin and conservation of bi-directional miRNA transcription and the functional relationships of products from sense and antisense transcription have not been extensively studied.

The miRNAs produced from the *D. melanogaster* iab-4 locus in the Hox complex are currently the only functionally characterized sense/antisense miRNAs (10–13). The bi-directionally transcribed iab-4 locus produces two primary transcripts that fold into two distinct hairpin precursors, named dme-iab-4 and dme-iab-8 (the latter is also

*To whom correspondence should be addressed. Tel: +44 161 275 3685; Fax: +44 161 275 1505; Email: matthew.ronshaugen@manchester.ac.uk

sometimes referred to as dme-iab-4-as). The spatial and temporal expression of the sense and antisense transcripts differ during embryogenesis (10,12,13). The mature miRNA products regulate translation of the Hox genes *Ubx* and *abd-A* in the fly (10–13). Interestingly, the vertebrate mir-196 and the insect iab-4 miRNAs are located in analogous regions in the Hox cluster and regulate homologous Hox genes, yet there is no evidence that the two miRNAs are themselves homologues (14). This has been suggested as an example of 'genomic parallelism', where selective pressure acts on miRNAs that have independently arisen in this locus, resulting in functional convergence to target the flanking Hox gene (15). Here, we investigate the evolution of bi-directional transcription at the iab-4/iab-8 miRNA locus. We use computational analysis and *in vivo* assays to identify and confirm iab-4 and iab-8 target sites in insect Hox UTRs (untranslated regions). Furthermore, we explore the overlapping targeting properties of the four mature miRNA products to gain insight into the mechanisms that underlie their acquisition of homeotic function. This work establishes the iab-4/iab-8 system as a model for overlapping function for multiple miRNA products of sense and antisense transcription.

MATERIALS AND METHODS

Animal culture

Tribolium castaneum was cultured at 28°C with either whole-wheat flour for maintenance or white flour supplemented with yeast extract for embryo collection. *D. melanogaster* w1118 was reared on standard cornmeal molasses, and embryos were collected on juice agar supplemented with yeast extract at 25°C.

Analyses of available insect short RNA libraries

The following small RNAseq data sets were downloaded from the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) and the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra): GEO:GSE7448 and GEO:GSE11624 (*D. melanogaster*), GEO:GSE17965 (*Bombyx mori*), GEO:GSE26036 (*T. castaneum*), GEO:GSE12640 (*Locusta migratoria*) and SRA:SRA012371 (*Apis mellifera*). Reads were mapped using our sequential trimming strategy (16) to the corresponding reference genomes, except in the case of *L. migratoria*, which was mapped against the sequence of the *T. castaneum* iab-4 precursor. Read counts mapping to the iab-4/8 locus shown in Figure 1A were obtained from GEO:GSE11624, GEO:GSE26036 (17) and SRA:SRA012371.

Nascent transcript whole-mount *in situ* hybridization

Embryo fixation, probe labelling and multiplex nascent transcript RNA fluorescent *in situ* hybridization (FISH) were performed according to (18), with the following modifications: the RNA probes were labelled with digoxigenin-UTP and were detected with sheep anti-digoxigenin (Roche) and donkey anti-sheep Alexa Fluor 555 (Molecular Probes) as primary and secondary antibodies, respectively. Probes labelled with Biotin RNA labelling mix were detected with mouse anti-biotin (Roche) and donkey anti-mouse Alexa Fluor 488 (Molecular Probes) as primary and secondary antibodies, respectively.

Prediction of putative iab-4 binding sites in *Tribolium* Hox 3′-UTR sequences

RNAs from 0 to 5 days, and adult *T. castaneum* were extracted with TRIzol (Invitrogen), and reverse transcriptions were carried out with SuperScript reverse transcriptase (Invitrogen). The ends of the Hox gene 3′-UTRs were confirmed by reverse transcriptase–polymerase chain reaction (RT–PCR). The cloned portions of 3′-UTR from the tested Hox genes were scanned for putative iab-4 canonical binding sites and 6mer marginal sites as defined in (19), and a more relaxed predicted target set was also produced allowing one wobble pairing (U:G) between the miRNA seed and the target. In addition, the *T. castaneum* Hox 3′-UTRs were scanned with the miRanda target prediction tool (20). The significance of target overlap between miRNA products was assessed using the randomization-based approach described in (21).

Construction of miRNA, 3′-UTR and perfect target site reporter constructs

Drosophila Hox 3′-UTRs were retrieved from FlyBase (http://flybase.org/). *Tribolium* Hox 3′-UTRs were annotated by manual inspection of next generation transcriptome data available from the iBeetle genome browser (http://bioinf.uni-greifswald.de/tcas). All previous UTR annotations were supported except for *Tribolium* Utx (the beetle orthologue of Ubx). The novel Utx UTR prediction was validated with RT–PCR using primers in the homeodomain of Utx and at the end of the predicted UTR. All UTRs were cloned into the multiple cloning sites of the psi-check-2 vector using standard cloning procedures. Predicted iab-4 and iab-8 target sites were mutagenized with the QuikChange site-directed mutagenesis kit (Stratagene) following the manufacturer's instructions. Mutagenesis sites are shown in Figure 3B, and primer information are shown in supplementary information. Constructs containing perfect complementary sequences to the mature miRNAs from both arms of dme-iab-4, dme-iab-8, tca-iab-4 and tca-iab-8 were also constructed. The perfectly complementary sites were cloned into the XhoI site of the dual luciferase vector psi-check-2 (Promega). Approximately 200 bp of *Drosophila* and *Tribolium* iab-4 and iab-8 was amplified by PCR from genomic DNA and cloned into pAc5.1 vectors (Invitrogen) using standard cloning methods. Primer information is listed in Supplementary Table S1. All constructs were subjected to ABI sequencing to confirm the sequences of the inserts.

Quantification of mature miRNA production in S2 cells

In all, 1.5 μg of pAc5.1 plasmids containing the *Drosophila* and *Tribolium* iab-4 and iab-8 constructs were transfected into S2 cells using Effectene (Qiagen) according to the

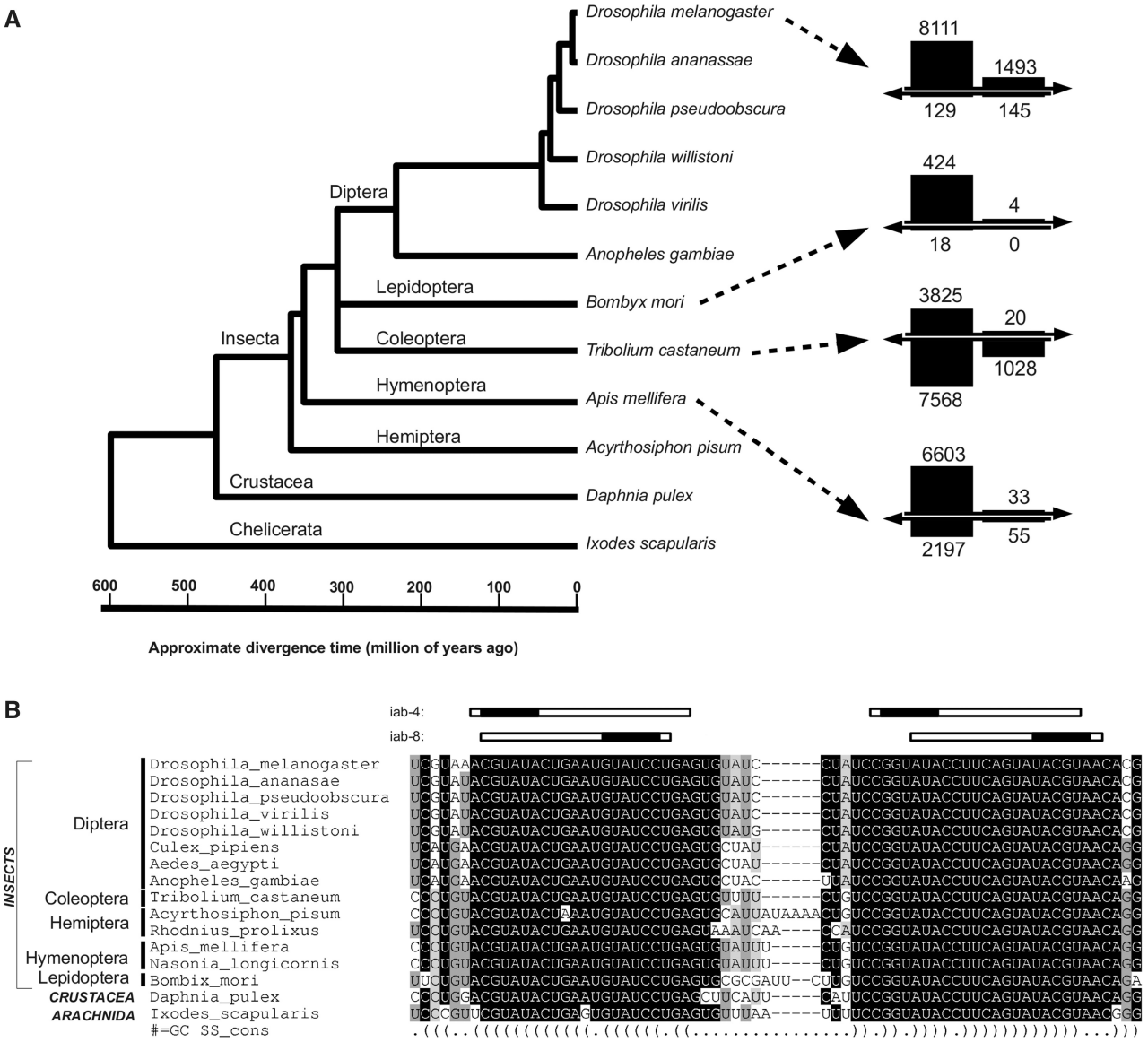

Figure 1. Expression and conservation of iab-4 in *T. castaneum* and other arthropods. (**A**) Phylogenetic relationships and divergence times (27) of the arthropod species with conserved iab-4/8 sequences. For species where deep sequencing data are present, the relative abundance of the four mature products from the iab-4 locus is plotted on the right for *D. melanogaster*, *B. mori*, *T. castaneum* and *A. mellifera*. Arrows indicate the direction of transcription (i.e. bars above the lines from left to right are the expression data for iab-4-5p and iab-4-3p, whereas bars beneath the lines from right to left indicate mapped deep sequencing reads for iab-8-5p and iab-8-3p). (**B**) Alignment of arthropod iab-4 sequences. Shadowing is based on sequence conservation. Above the alignment, white boxes represent the mature sequences experimentally determined for *Tribolium*. Black boxes indicate the putative seed sequences (18). Asterisks below the alignment mark columns with at least one nucleotide substitution in any species.

manufacturer's instructions. Small RNAs from both transfected and control S2 cells were isolated 48 h post-transfection, using the miRVana miRNA isolation kit (Applied Biosystems). Quantification of mature miRNAs derived from both arms of the hairpin precursors was carried out with locked nucleic acid (LNA) miRNA real-time PCR assays (Exiqon). Samples from three independent transfections were assayed in triplicate using the U6 spliceosomal RNA for normalization.

Functional analysis of iab-4 and iab-8 by luciferase assays

A dual luciferase miRNA target reporter assay was used as previously reported (22,23). *Drosophila* S2 cells were transfected with psi-check-2 sensor (1 μg) and pAc5.1 plasmids containing the iab-4 or iab-8 hairpin of either *Drosophila* or *Tribolium* (1.5 μg) using Effectene (Qiagen). The sensors consisted of firefly and *Renilla reniformis* luciferase enzymes expressed from a single plasmid, with the target sequence of the iab-4/iab-8, the Hox UTRs, or the mutated Hox UTRs fused with the UTR of the *Renilla* luciferase. Cells were lysed and assayed for luciferase activity using the dual luciferase assay kit (Promega) 48-h post-transfection. Data were collected using a MicroLumat Plus LB96V Microplate Luminometer. Experiments were performed in three biological replicates with *Renilla*:firefly luciferase ratios averaged for three technical replicates each.

RESULTS

The miRNA iab-4 locus encodes an arthropod-specific miRNA

Our recent genome-wide analysis of small RNAs in the flour beetle *T. castaneum* identified five miRNAs with evidence of antisense transcription and processing by Drosha and Dicer (24). Three of these miRNAs seem to be unique to *Tribolium* and have not been identified in any other animal. The other two are iab-4 and miR-307. miR-307 is conserved throughout the protostomes, and there is experimental evidence for conserved antisense transcription in two other drosophilid species (13,25). This observation, in conjunction with data from other deeply sequenced arthropods, suggests that sense/antisense transcription of miRNAs occurs infrequently and is normally species specific, that is, miRNAs that exhibit sense and antisense transcription and processing are normally evolutionarily young, or sense/antisense transcription is transient and rapidly lost. Conservation of sense and antisense transcription likely indicates that mature miRNA products from both transcripts have acquired biological functions. Thus, the iab-4 locus provides an excellent opportunity to gain insight into the evolution of a conserved antisense miRNA locus and the acquisition of target mRNAs. We searched available genomes for iab-4 homologues (see 'Materials and Methods' section) using all iab-4 miRNA sequences deposited in the miRBase database (v.16) (26). iab-4 homologues, containing a highly conserved hairpin sequence and structure, were identified in all available insect genomes, as well as in the crustacean *Daphnia pulex* and in the arachnid *Ixodes scapularis* (Figure 1B). No homologues were identified in lophotrochozoan or deuterostome genomes (7). In all cases, the location of the iab-4 miRNA was syntenically conserved within the putative Hox gene complex between the homologues of *abd-A* and *Abd-B*.

To determine whether sense iab-4 and antisense iab-8 loci were both transcribed and processed into mature miRNAs, we re-mapped and re-analysed all available deep sequencing data sets from insects (see 'Materials and Methods' section). These data show that mature miRNAs derived from both iab-4 and iab-8 transcripts are detected in *D. melanogaster* (8,25), *B. mori* (26,28), *T. castaneum* (24) and *A. mellifera* (29) (Figure 1A). Indeed, all four possible mature miRNAs derived from the sense and antisense transcripts are detected in three of these data sets at levels comparable with other miRNA loci (Figure 1A). In the absence of a complete genome, we mapped sequence reads from *L. migratoria* to the *T. castaneum* iab-4 locus. We find reads mapping to both sense and antisense sequences (23). We also identify the four miRNA products from the iab-4 and iab-8 loci in the hemimetabolous pea aphid, *Acyrthosiphon pisum* [(30) S. Jaubert-Possamai and F. Legeai, personal communication]. Bi-directional transcription of the iab-4 locus is, therefore, a common feature of all studied insects. Furthermore, we note almost perfect conservation of all four putative mature miRNA sequences between insects and crustaceans (~500 million years ago), suggesting substantial constraint on sequences from both arms of the hairpin precursor. In general, miRNAs exhibit an increased constraint on the hairpin arm giving rise to the dominant mature miRNA with the alternate arm more variable. The data suggest that the antisense transcription of iab-4 dates from the origin of the locus (Figure 1A), and that functional interactions with target mRNAs have lead to constraint on all four potential mature miRNAs, from both arms of the hairpin precursors expressed from both strands.

Conserved spatial and temporal expression of iab-4 and iab-8

Both iab-4 and iab-8 transcripts in *Drosophila* are expressed during embryogenesis in the posterior of the presumptive abdomen (10,12,13,34). iab-4 transcription has been demonstrated in a similar domain in *Tribolium* (35). We have investigated the spatial expression patterns of iab-8 transcripts using nascent transcript *in situ* hybridization in *T. castaneum* embryos. Early development and segmentation in flies and beetles is significantly different (33). *Drosophila* is a long-germ insect where segmentation occurs relatively synchronously, whereas *Tribolium* is a short-germ insect that forms posterior segments sequentially. Despite these significant differences in embryonic development, we find that the spatial domains of expression, relative to the underlying pattern of Hox gene expression, are largely conserved in the two animals. The relative domains of iab-4 and iab-8 transcription are also similar (Figure 2A–C) (35). Transcription of tca-iab-4 is detected earlier than tca-iab-8 (Figure 2A and C), similar to previous observations in *D. melanogaster* (10,12,13). We also find that the anterior boundary of tca-iab-4 expression is posterior to that of *Utx*/*Ubx* (Figures 2A and 3B).

Once iab-8 transcription is initiated, there is a one to two cell wide domain of overlap between the anterior boundary of iab-8 and the posterior boundary of iab-4 (Figure 2D). This region of overlap is similar to that seen in *D. melanogaster* (Figure 2E and F) (10). Later, the expression of tca-iab-4 is reduced, and it is completely absent near the end of segmentation when the expression domain of tca-iab-8 expands to include the last segment (Figure 2D). This pattern is then maintained until later stages when the thoracic segments and abdominal segments are developed. Thus, the embryonic expression patterns of both sense and antisense transcripts of iab-4 in *D. melanogaster* and *T. castaneum* seem to be largely conserved in terms of broad spatial domains and temporal occurrence.

The predominant mature iab-4/8 products change during evolution

The deep sequencing data sets reveal an interesting property of the iab-4 locus miRNAs. Using read counts as estimates of expression levels, we can clearly see that different mature products dominate in different species. For example, across 42 different deep-sequencing experiments catalogued in the miRBase database, the dominant mature miRNA in *D. melanogaster* is iab-4-5p, which is

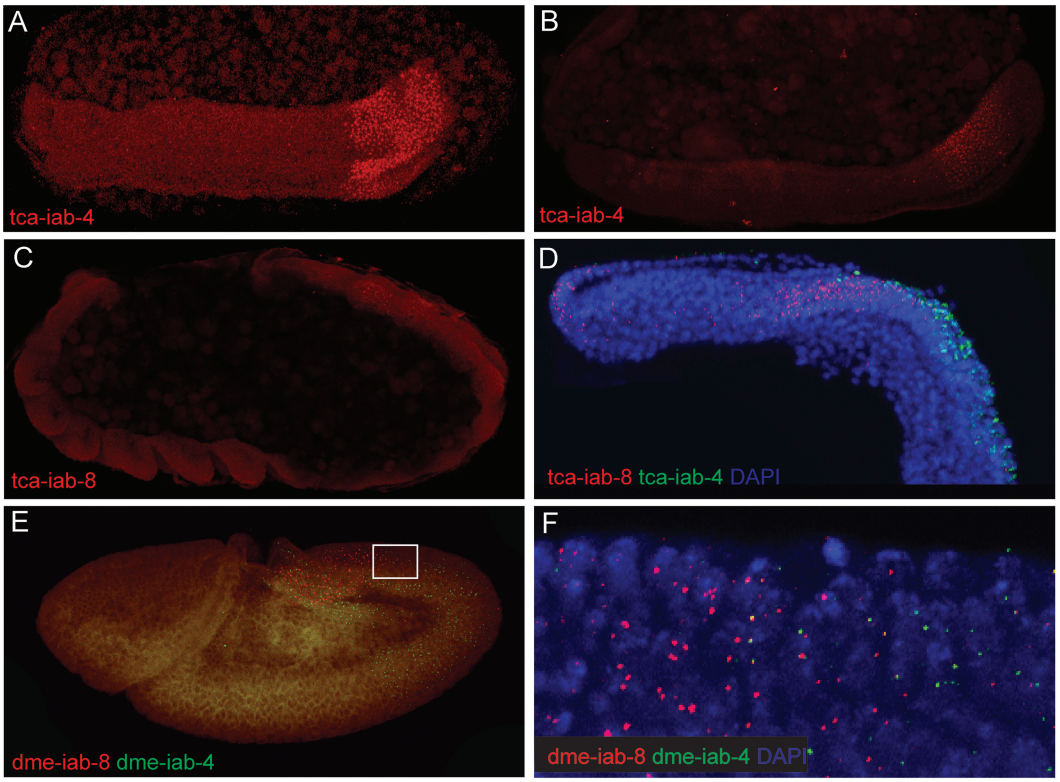

Figure 2. Spatial and temporal expression of iab4/iab-8 transcription in the beetle *T. castaneum* and fly *D. melanogaster*. Nascent transcript FISH detects a posterior stripe of tca-iab-4 expression in early (**A**) and late (**B**) *Tribolium* embryos. (**C**) The antisense iab-8 miRNA is expressed more posteriorly and later during development. (**D**) Two-colour FISH identifies a 1–2 cell overlapping domain of co-expression of tca-iab-4 (green) and tca-iab-8 (red). (**E**) Two-colour FISH of dme-iab-4 (green) and dme-iab-8 (red) on *Drosophila* embryos. (**F**) Higher magnification image (white box in Figure 2E) also identifies a 1–2 cell overlapping domain of co-expression of dme-iab-4 (green) and dme-iab-8 (red) as indicated by the yellow signals.

expressed at 1–2 orders of magnitude greater than iab-4-3p or either of the iab-8 sequences (see Figure 1A) (25). However, in both *T. castaneum* and *A. mellifera*, both iab-4-5p and iab-8-3p are highly abundant, again 1–2 orders of magnitude higher than the alternative products (Figure 1A). As the temporal and spatial expression of primary transcripts is similar in *T. castaneum* and *D. melanogaster*, it is likely that this difference is accomplished through differential processing of the pri-miRNA. Thus, it is clear that the choice of the dominant mature miRNA product generated from the bi-directional iab-4 locus is subject to significant change during evolution. For most miRNAs, mature miRNAs derived from opposite arms of the hairpin precursor will have different targeting properties, thus, impacting on miRNA function (21). However, the near perfect conservation of the mature iab-4/iab-8 sequences and their expression domains implies conservation of function between all four possible mature miRNA products. Therefore, we investigated the target specificity of the iab-4 and iab-8 miRNAs and the evolution of their target sites.

Mature iab-4 and iab-8 miRNA sequences have partially overlapping targeting properties

The observed flexibility in iab-4/8 miRNA arm usage in animals could be explained by common targets of the mature products. Indeed, previous studies have observed common Hox gene targets of iab-4 and iab-8 miRNAs (12,13), and it has been observed that the sequences of the 5′-arms of the sense and antisense miRNAs are similar because of the constraints on hairpin structure (specifically, hairpin formation requires that the 5′- and 3′-arms derived from the same precursor are imperfect reverse complements, and the arrangement of the iab-4 and iab-8 transcripts in the genome specifies that the opposite arms of antisense sequences are exact reverse complements) (13). Beyond the expected 5′-arm similarity, the iab-4/8 locus has an unusual property: the mature miRNA sequences derived from 5′- and 3′-arms of the same hairpin are also similar in sequence because of two occurrences of a 6-nt palindrome within the mature miRNA (Figure 4A). However, the 5′- and 3′-miRNA ends resulting from processing by Drosha and Dicer cleavage are not symmetrical. Thus, the so-called ‘seed’ regions of all four mature miRNAs, as defined by nucleotide position from the 5′-end, are different (Figure 4A). Indeed, the four mature products are canonical seed-shifted variants of each other (7). Nevertheless, an identical 6mer sequence in the seed region of three of the mature miRNA products (iab-4-5p, iab-4-3p and iab-8-5p—see Figure 4A) suggests some potential to target common sites. Although the use of only 6mer seeds in target prediction leads to an increased false-positive rate, use of 6mer seeds also excludes significantly fewer verifiable targets than use of 8mer seeds.

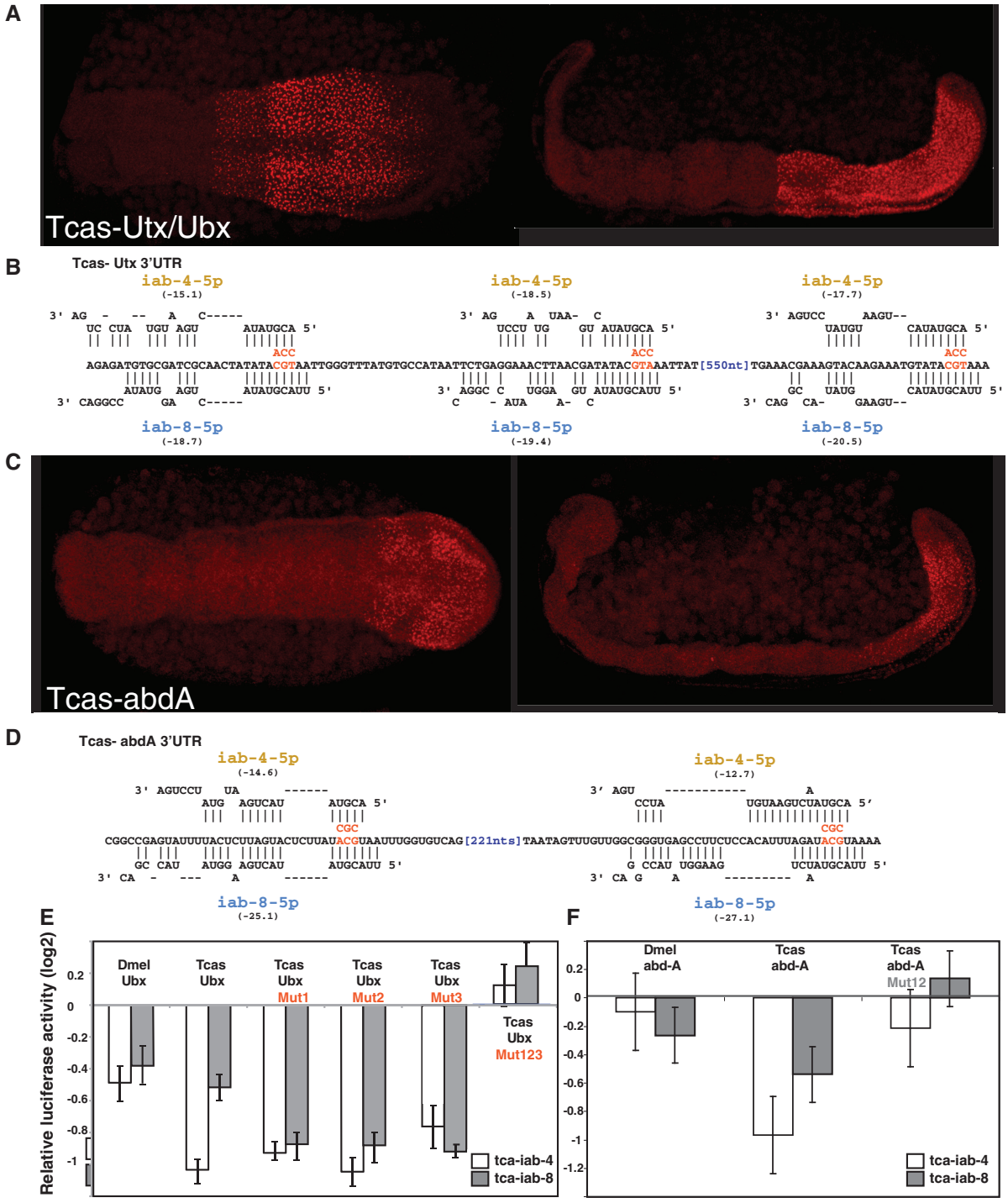

Figure 3. Functional analysis of iab-4 and iab-8 in regulating *T. castaneum* and *D. melanogaster* Hox genes Ubx and Abd-A. Expression patterns and iab-4/8 targeted sites of (**A** and **B**) tca-Utx/Ubx, (**C** and **D**) tca-Abd-A. Note that the expression domains of these genes follow the rules of spatial and temporal collinearity according their positions within the intact *T. castaneum* Hox cluster. Mutated residues are highlighted in red. Numbers in parentheses are hybridization energies. (**E** and **F**) Relative firefly:Renilla luciferase activity of the Ubx and Abd-A 3′-UTRs and mutated 3′-UTRs (Mut1-3) in the presence of dme-iab-4, dme-iab-8, tca-iab-4 or tca-iab-8. Both sense and anti-sense transcripts of iab-4 in *T. castaneum* are able to regulate the posterior Hox genes to different degrees but not the correct mutated 3′-UTRs. Error bars correspond to standard error of the mean ($n = 3$).

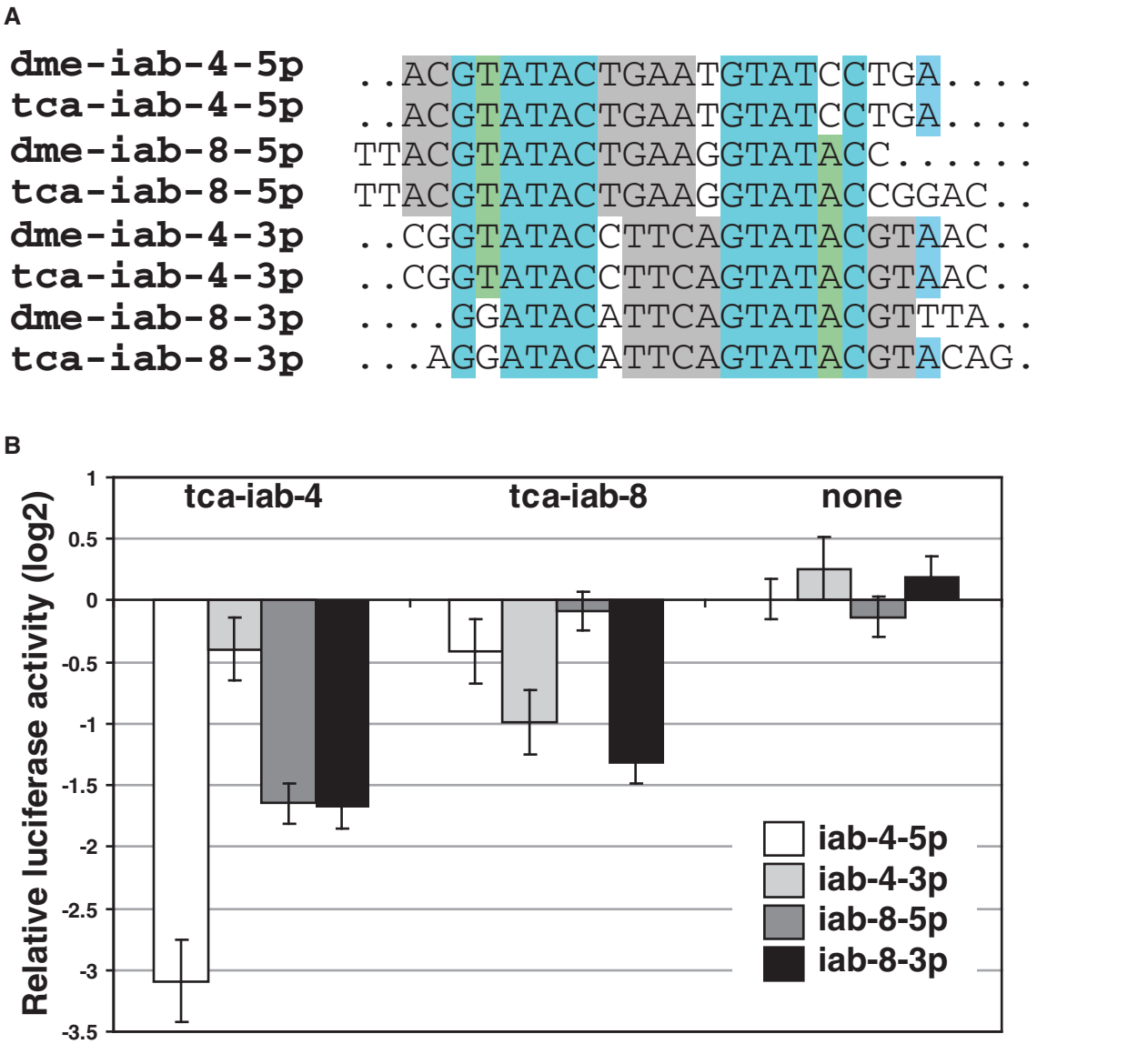

Figure 4. Comparison of iab-4/iab-8 miRNA sequences and cross-regulation of perfect targets. (**A**) Sequence alignment of the mature miRNA product of dme-iab-4, dme-iab-8, tca-iab-4 and tca-iab-8 showing the palindromic nature of these sequences. (**B**) Dual-luciferase assay to access perfect complementary target sites repression by iab-4 and iab-8. A decrease in relative luciferase activity indicates an interaction between the tested miRNA and the perfect complementary target sensor. Relative activities of firefly and Renilla luciferases show dme-iab-4, dme-iab-8, tca-iab-4 and tca-iab-8 miRNAs can all cross-regulate/suppress the perfect targets in the S2 cell culture. Error bars correspond to standard error of the mean ($n = 3$).

We investigated targeting properties of iab-4/iab-8 miRNAs *in silico* and *in vitro*. *In silico* miRNA target prediction methods are predicated on constructing an informative model of miRNA-target interactions derived from measurable features of verified interactions (19). Although these methods differ and none is error-proof, they do establish a framework for investigating common targeting features of miRNAs. Prediction of putative 3′-UTR targets of all four products was performed using two methods: canonical seeds (19) and the miRanda software (20). We used the lists of predicted targets to estimate the probability that each pair of miRNAs target common genes (see 'Material and Methods' section) (5). The canonical seed approach finds that iab-4-5p and iab-8-5p target more common UTRs than is expected by chance, although miRanda does not identify any significant overlap (Supplementary Table S3).

To test the common targeting properties of iab-4 and iab-8 miRNAs *in vitro*, a set of luciferase reporters containing perfect target sites (see 'Material and Methods' section) for each of the mature miRNAs were constructed. These reporters were co-transfected into S2 cells with expression constructs driving the production of either iab-4 or iab-8 pri-miRNAs. We then analysed the ability of iab-4 and iab-8 miRNAs to repress luciferase translation (Figure 4B). As expected, sites that are perfectly complementary to the mature miRNA products from a given hairpin were the most highly repressed. Repression of target sites for the 5′- and 3′-arms of each hairpin is correlated with their relative production as determined

by next-generation sequencing (Figure 1A). For example, sites for iab-4-3p are weakly repressed on transfection with iab-4, whereas iab-4-5p sites are strongly repressed (~6-fold greater repression than iab-4-3p). However, iab-4 also strongly represses iab-8-5p and iab-8-3p target sites (~2-fold greater than iab-4-3 p). Translation of mRNAs containing iab-4 target sites is also significantly repressed by iab-8. Thus, there is a substantial capacity for iab-4 and iab-8 miRNAs to target common sites.

Predicted Hox gene targets of *Drosophila* and *Tribolium* iab-4 miRNAs

Previous computational analysis has predicted that the Hox genes *Antp*, *Ubx* and *abd-A* are putative targets of miRNAs from the iab-4/8 locus in drosophilds and many other arthropods [see, for example, (12,31,32)]. Subsequent validation established that *Ubx* was a common target of both iab-4 and iab-8, and that *abd-A* was a target of iab-8 in *D. melanogaster* (10–12).

To determine whether the insect iab-4/8 locus has conserved function and targets, we have investigated Hox gene targets of iab-4/8 miRNAs in *T. castaneum* using both computational and experimental methods. We used available whole-genome transcriptome analysis to refine annotations of *Tribolium* Hox 3′-UTRs (see 'Materials and Methods' section, and Supplementary Figure S3), and we validated a novel Ubx 3′-UTR prediction by RT–PCR. We note that previous analysis of iab-4 Hox targets in *Drosophila* showed differential regulation mediated by alternate 3′-UTRs (34). Our use of the longest known UTR for each *Tribolium* Hox gene ensures that we include all possible target sites. Computational analyses of validated UTRs for putative canonical and 6mer marginal sites for all four possible iab-4/8 miRNAs identified multiple candidates (see 'Materials and Methods' section). Our predictions suggest that *T. castaneum abd-A* has canonical sites for all four iab-4 products, in agreement with previous findings (35). The other four transcripts are predicted to be weaker targets. For example, the *Abd-B* UTR contains up to five potential iab-4/8 binding sites, but only one is a canonical seed site (Supplementary Figure S2H). We also predict multiple target sites in the putative *T. castaneum Antp/ptl* and *Scr/Cx* UTRs, similar to previous computational predictions for *Antp* in *D. melanogaster* (12).

Hox target site validation

To validate the predicted repression of Hox targets, we used a dual-luciferase reporter approach to assay the UTRs of *Antp*, *Ubx* and *abd-A* from *D. melanogaster*, and *Scr/Cx*, *Antp/ptl*, *Ubx/Utx*, *abd-A* and *Abd-B* from *T. castaneum* for functional iab-4/8 target sites. We used S2 cells for the assay; deep sequencing data and LNA-qRT–PCR show that iab-4 and iab-8 are not endogenously expressed in S2 cells (Supplementary Figure S1) (36). Vectors constitutively expressing *Drosophila* or *Tribolium* iab-4 or iab-8 precursors were co-transfected into S2 cells along with the luciferase vectors containing Hox or mutated 3′-UTR fusions.

The results of the luciferase assays are summarized in Figure 3E and F. In agreement with previous data (12), we find iab-4 and iab-8 miRNAs can both downregulate translation via the Ubx UTR in *D. melanogaster*, whereas iab-8 alone regulates *abd-A*. Contrary to previous predictions, we find that the *D. melanogaster Antp* UTR is not a target of iab-4/8 in our luciferase assays, even though canonical target sites are present (Supplementary Figure S2F and I) (10,12,32). We also find that *T. castaneum* Scr/Cx, Antp/ptl and Abd-B 3′-UTRs are not targeted by either iab-4 or iab-8. This supports the argument that the sole use of a seed match is not a reliable predictor of functional miRNA target sites (Supplementary Figure S2A–I).

Both iab-4 and iab-8 miRNAs can mediate translational repression through the *Ubx/Utx* UTR in *Tribolium*. The *Ubx/Utx* UTR is, therefore, a conserved target of miRNAs expressed from both strands in *Drosophila* and *Tribolium*. We also find that abd-A is a conserved target of iab-8 in *Drosophila* and *Tribolium*. However, unlike *Drosophila*, the abd-A UTR is also targeted by iab-4 miRNAs in *Tribolium* showing that target site miRNA interactions have evolved either through functional gain in *Tribolium* or loss in *Drosophila*. One particularly interesting observation is that the predicted target sites for iab-4 miRNAs overlapped predictions for iab-8 sites (Figure 3B and D). This is consistent with our seed analysis of the miRNAs, which shows significant overlap in sense/antisense miRNA activities on common target sites. To further investigate and characterize the sites through which iab-4/8 mediate repression, we performed mutagenesis of the predicted sites, starting with the highest scoring sites (see 'Materials and Methods' section; Supplementary Table S2). The assay demonstrates that repression of abd-A by both *Tribolium* iab-4 and iab-8 can be eliminated through mutation of the two overlapping iab-4/8 sites and a perfect 8-mer iab-8-5p site (Figure 3F). Our mutagenesis of individual *Tribolium* Ubx target sites failed to significantly reduce the translational repression mediated by either iab-4 or iab-8 (see Figure 3E). Only simultaneous mutagenesis of all three targets relieves the regulatory effect of both iab-4 and iab-8 (Figure 3E). Therefore, there is redundancy in the miRNA target sites, such that a single site is sufficient to repress Ubx.

In summary, repression assays clearly identify *abd-A* as a target of miRNAs from the iab-8 locus in both *D. melanogaster* and *T. castaneum*. However, the key regulatory sequences seem to differ significantly between the two genomes. In *Drosophila*, our data and previous studies show that abd-A is regulated primarily by the iab-8 transcript (10). In contrast, computational target prediction, luciferase assays and target site mutagenesis all confirm that *Tribolium abd-A* is regulated by miRNAs processed from both the sense iab-4 sequence and the antisense iab-8 sequence (Figure 3F).

DISCUSSION

The iab-4 miRNA locus has some unusual properties: the locus is transcribed in both directions, producing two

primary miRNA transcripts and two hairpin precursors. Each precursor is processed to produce two mature miRNAs, one from each arm of each precursor hairpin. Only a handful of other miRNAs have been shown by deep sequencing data to be transcribed in both directions; currently, the miRBase database has only 27 animal examples [v17 (26)]. We show that bi-directional transcription of the iab-4/8 locus and production of miRNAs from both transcripts is conserved in insects. However, the relative abundance of the four mature miRNAs varies significantly between fly and beetle.

The four mature miRNAs produced from the iab-4 locus are extremely similar. Indeed, they are all seed-shifted variants of each other. This state is possible only because the mature sequences are partially palindromic. Thus, sense and antisense sequences are highly similar, as are partially complementary mature sequences from opposite arms of the same hairpin. As the mature sequences are closely related, the predicted targets of the four mature products overlap significantly. Our previous work suggests that this is an unusual situation: the targets of alternate miRNAs derived from the 5′- and 3′-arms of almost all miRNAs are largely different (21). We have shown that iab-4-5p and iab-8-5p have more common targets that expected by chance. This functional overlap of antisense products may have facilitated the maintenance of the bi-directionality in the iab-4/iab-8 locus. Indeed, we observed the same pattern in mir-307, the other miRNA locus that produces mature miRNAs from both genomic strands (Supplementary Table S4).

Our analysis of the repression of engineered perfect target sites clearly shows significant cross-targeting for three of the four mature miRNAs. Furthermore, we have shown the Hox gene *Ubx*/*Utx* is a conserved target of both iab-4 and iab-8 miRNAs in both *Drosophila* and *Tribolium*. However, between fly and beetle, we find differences in Hox gene targets of iab-4/8 miRNAs and differences in the sites that mediate those targets. For example, *abd-A* is regulated only by iab-8 miRNAs in *Drosophila*, whereas both iab-4 and iab-8 miRNAs target *abd-A* transcripts in *Tribolium*. There are, therefore, both conserved and variable aspects of the targeting properties of the four mature miRNAs produced from the iab-4/8 locus in insects. The conservation of the Hox genes *Ubx* and *abd-A* as targets of the iab-4/8 miRNAs further establishes the ancient connection between the miRNAs of the Hox complex and their role in modulating the function of the Hox genes themselves. No other intergenic miRNA has been linked by genomic position to its function, yet all Hox complex miRNAs (iab-4, mir-196 and mir-10) have been found to modulate Hox gene function (22).

The iab-4/8 locus provides us with fundamental insight into the mechanisms of evolution and the function of sense/antisense miRNA pairs. The production of functional products from both strands of the same locus may impose an evolutionary trade-off, driven on one hand by sequence conservation because of structural constraints, and on the other hand by constraints imposed by target specificity. We propose that the deep conservation can be explained in part by the common targeting properties of the multiple mature products generated from these two transcripts. Given the functional similarity of the miRNA products of iab-4 and iab-8, the antisense transcription of the locus can be considered analogous to the acquisition of an enhancer by the sense transcript to drive expression and miRNA production in the additional domain. Furthermore, the palindromic nature of the iab-4/iab-8 mature sequences determines that the novel antisense miRNA will share targets with the pre-existing sense miRNA. We suggest that this explains the apparent contradiction between extreme conservation of mature miRNA sequences on both arms, yet significant plasticity between organisms as to which arm is the dominant product. However, the evolution of target sites in *abd-A* demonstrates that functional target sites can be differentially regulated between even closely related species.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–4 and Supplementary Figures 1–3.

ACKNOWLEDGEMENTS

The authors thank Tom Millard for supplying pAc5.1 vectors, Caroline Milner for S2 cells, Keith Brennan and Andrew Sharrocks for use of the microplate luminometer and the DFG research unit FOR1234 iBeetle for providing access to sequence data before publication. They also thank Yimei Cai, Stéphanie Jaubert-Possamai, Fabrice Legeai, Susan Brown and Gregor Bucher for sharing unpublished data, Casey Bergman for constructive discussion and Sue Crosthwaite for comments on the manuscript.

FUNDING

Biotechnology and Biological Sciences Research Council [BB/G011346/1 and BB/H017801/1]; Wellcome Trust [097820/Z/11/Z]; FLS Career Development Award from the University of Manchester (to J.H.L.H.). Funding for open access charge: Wellcome Trust [097820/Z/11/Z].

Conflict of interest statement. None declared.

REFERENCES

1. Pearson,J.C., Lemons,D. and McGinnis,W. (2005) Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.*, **6**, 893–904.
2. Lempradl,A. and Ringrose,L. (2008) How does noncoding transcription regulate Hox genes? *Bioessays*, **30**, 110–121.
3. Lemons,D. and McGinnis,W. (2006) Genomic evolution of Hox gene clusters. *Science*, **313**, 1918–1922.
4. Rinn,J.L., Kertesz,M., Wang,J.K., Squazzo,S.L., Xu,X., Brugmann,S.A., Goodnough,L.H., Helms,J.A., Farnham,P.J., Segal,E. *et al.* (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, **129**, 1311–1323.
5. Griffiths-Jones,S., Hui,J.H., Marco,A. and Ronshaugen,M. (2011) MicroRNA evolution by arm switching. *EMBO Rep.*, **12**, 172–177.

6. Grimson,A., Srivastava,M., Fahey,B., Woodcroft,B.J., Chiang,H.R., King,N., Degnan,B.M., Rokhsar,D.S. and Bartel,D.P. (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature*, **455**, 1193–1197.
7. Wheeler,B.M., Heimberg,A.M., Moy,V.N., Sperling,E.A., Holstein,T.W., Heber,S. and Peterson,K.J. (2009) The deep evolution of metazoan microRNAs. *Evol. Dev.*, **11**, 50–68.
8. Ruby,J.G., Stark,A., Johnston,W.K., Kellis,M., Bartel,D.P. and Lai,E.C. (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res.*, **17**, 1850–1864.
9. Landgraf,P., Rusu,M., Sheridan,R., Sewer,A., Iovino,N., Aravin,A., Pfeffer,S., Rice,A., Kamphorst,A.O., Landthaler,M. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401–1414.
10. Ronshaugen,M., Biemar,F., Piel,J., Levine,M. and Lai,E.C. (2005) The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev.*, **19**, 2947–2952.
11. Bender,W. (2008) MicroRNAs in the *Drosophila* bithorax complex. *Genes Dev.*, **22**, 14–19.
12. Stark,A., Bushati,N., Jan,C.H., Kheradpour,P., Hodges,E., Brennecke,J., Bartel,D.P., Cohen,S.M. and Kellis,M. (2008) A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. *Genes Dev.*, **22**, 8–13.
13. Tyler,D.M., Okamura,K., Chung,W.J., Hagen,J.W., Berezikov,E., Hannon,G.J. and Lai,E.C. (2008) Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Genes Dev.*, **22**, 26–36.
14. Yekta,S., Shih,I.H. and Bartel,D.P. (2004) MicroRNA-mediated cleavage of HOXB8 mRNA. *Science*, **304**, 594–596.
15. Campo-Paysaa,F., Sémon,M., Cameron,R.A., Peterson,K.J. and Schubert,M. (2011) microRNA complements in deuterostomes: origin and evolution of microRNAs. *Evol. Dev.*, **13**, 15–27.
16. Marco,A. and Griffiths-Jones,S. (2012) Detecting microRNAs in color-space. *Bioinformatics*, **28**, 318–323.
17. Jagadeeswaran,G., Zheng,Y., Sumathipala,N., Jiang,H., Arrese,E.L., Soulages,J.L., Zhang,W. and Sunkar,R. (2010) Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. *BMC Genomics*, **11**, 52.
18. Kosman,D., Mizutani,C.M., Lemons,D., Cox,W.G., McGinnis,W. and Bier,E. (2004) Multiplex detection of RNA expression in *Drosophila* embryos. *Science*, **305**, 46.
19. Bartel,D.P. (2009) MicroRNAs, target recognition and regulatory functions. *Cell*, **136**, 215–233.
20. Enright,A.J., John,B., Gaul,U., Tuschl,T., Sander,C. and Marks,D.S. (2003) MicroRNA targets in *Drosophila*. *Genome Biol.*, **5**, R1.
21. Marco,A., MacPherson,J.I., Ronshaugen,M. and Griffiths-Jones,S. (2012) MicroRNAs from the same precursor have different targeting properties. *Silence*, **3**, 8.
22. Yekta,S., Tabin,C.J. and Bartel,D.P. (2008) MicroRNAs in the Hox network, an apparent link to posterior prevalence. *Nat. Rev. Genet.*, **9**, 789–796.
23. Wei,Y., Chen,S., Yang,P., Ma,Z. and Kang,L. (2009) Characterization and comparative profiling of the small RNA transcriptomes in two phases of locust. *Genome Biol.*, **10**, R6.
24. Marco,A., Hui,J.H., Ronshaugen,M. and Griffiths-Jones,S. (2010) Functional shifts in insect microRNA evolution. *Genome Biol. Evol.*, **2**, 686–696.
25. Chung,W.J., Agius,P., Westholm,J.O., Chen,M., Okamura,K., Robine,N., Leslie,C.S. and Lai,E.C. (2011) Computational and experimental identification of mirtrons in *Drosophila melanogaster* and *Caenorhabditis elegans*. *Genome Res.*, **21**, 286–300.
26. Kozomara,A. and Griffiths-Jones,S. (2011) miRBase, integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, **39**, D152–D157.
27. Wiegmann,B.M., Kim,J.W. and Trautwein,M.D. (2009) Holometabolous insects (Holometabola). In: Hedges,S.B. and Kumar,S. (eds), *The TimeTree of Life*. Oxford University Press, New York, pp. 260–263.
28. Liu,S., Li,D., Li,Q., Zhao,P., Xiang,Z. and Xia,Q. (2010) MicroRNAs of *Bombyx mori* identified by Solexa sequencing. *BMC Genomics*, **11**, 148.
29. Chen,X., Yu,X., Cai,Y., Zheng,H., Yu,D., Liu,G., Zhou,Q., Hu,S. and Hu,F. (2010) Next-generation small RNA sequencing for microRNAs profiling in the honey bee *Apis mellifera*. *Insect Mol. Biol.*, **19**, 799–805.
30. Legeai,F., Rizk,G., Walsh,T., Edwards,O., Gordon,K., Lavenier,D., Leterme,N., Mereau,A., Nicolas,J., Tagu,D. *et al.* (2010) Bioinformatic prediction, deep sequencing of microRNAs and expression analysis during phenotypic plasticity in the pea aphid, *Acyrthosiphon pisum*. *BMC Genomics*, **11**, 281.
31. Patraquim,P., Warnefors,M. and Alonso,C.R. (2011) Evolution of Hox post-transcriptional regulation by alternative polyadenylation and microRNA modulation within twelve *Drosophila* genomes. *Mol. Biol. Evol.*, **28**, 2453–2460.
32. Miura,S., Nozawa,M. and Nei,M. (2011) Evolutionary changes of the target sites of two microRNAs encoded in the Hox gene cluster of *Drosophila* and other insect species. *Genome Biol. Evol.*, **3**, 129–139.
33. Davis,G.K. and Patel,N.H. (2002) Short, long, and beyond: molecular and embryological approaches to insect segmentation. *Annu. Rev. Entomol.*, **47**, 669–699.
34. Thomsen,S., Azzam,G., Kaschula,R., Williams,L.S. and Alonso,C.R. (2010) Developmental RNA processing of 3′UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. *Development*, **137**, 2951–2960.
35. Shippy,T.D., Ronshaugen,M., Cande,J., He,J., Beeman,R.W., Levine,M., Brown,S.J. and Denell,R.E. (2008) Analysis of the *Tribolium* homeotic complex, insights into mechanisms constraining insect Hox clusters. *Dev. Genes. Evol.*, **218**, 127–139.
36. Cherbas,L., Willingham,A., Zhang,D., Yang,L., Zou,Y., Eads,B.D., Carlson,J.W., Landolin,J.M., Kapranov,P., Dumais,J. *et al.* (2011) The transcriptional diversity of 25 *Drosophila* cell lines. *Genome Res.*, **21**, 301–314.