Development 136, 3153-3160 (2009) doi:10.1242/dev.038034

Evolving enhancer-promoter interactions within the tinman complex of the flour beetle, *Tribolium castaneum*

Jessica Doran Cande, Vivek S. Chopra and Michael Levine*

Modifications of cis-regulatory DNAs, particularly enhancers, underlie changes in gene expression during animal evolution. Here, we present evidence for a distinct mechanism of regulatory evolution, whereby a novel pattern of gene expression arises from altered gene targeting of a conserved enhancer. The tinman gene complex (Tin-C) controls the patterning of dorsal mesodermal tissues, including the dorsal vessel or heart in *Drosophila*. Despite broad conservation of Tin-C gene expression patterns in the flour beetle (*Tribolium castaneum*), the honeybee (*Apis mellifera*) and the fruit fly (*Drosophila melanogaster*), the expression of a key pericardial determinant, *ladybird*, is absent from the dorsal mesoderm of *Tribolium* embryos. Evidence is presented that this loss in expression is replaced by expression of *C15*, the neighboring gene in the complex. This switch in expression from *ladybird* to *C15* appears to arise from an inversion within the tinman complex, which redirects a conserved *ladybird* 3′ enhancer to regulate *C15*. In *Drosophila*, this enhancer fails to activate *C15* expression owing to the activity of an insulator at the intervening *ladybird early* promoter. By contrast, a chromosomal inversion allows the cardiac enhancer to bypass the *ladybird* insulator in *Tribolium*. Given the high frequency of genome rearrangements in insects, it is possible that such enhancer switching might be widely used in the diversification of the arthropods.

KEY WORDS: Evolution, Heart, Cis-regulation

INTRODUCTION

Insect genomes are highly dynamic and subject to rapid chromosomal rearrangements. Unlike vertebrate genomes, which contain extensive synteny, insect genomes rarely retain similar linkage arrangements (Bolshakov et al., 2002; Zdobnov and Bork, 2007). A small handful of gene complexes represent the main exceptions to this rule of rapid chromosomal reorganization. However, even these are subject to a variety of chromosomal rearrangements among different insects. For example, the ancestral Hox complex is split into two smaller complexes (ANT-C and BX-C) in *Drosophila* (Negre and Ruiz, 2007). Here, we investigate the possibility that rearrangements within such complexes create an opportunity for evolving novel patterns of gene expression. We present evidence that an internal inversion within the tinman gene complex (Tin-C) redirects a conserved enhancer to different target genes in the flour beetle, *Tribolium castaneum* and in the fruit fly, *Drosophila melanogaster*.

The Tin-C contains a series of NK homeobox genes that are evolutionarily ancient, pre-dating even the Hox complex. Phylogenetic analysis of homeobox genes in the sponge *Amphimedon* indicates the existence of a proto-NK gene cluster prior to the eumetazoan radiation (Larroux et al., 2007). Gene linkage in the Tin-C has been conserved in protostomes such as flies, but is essentially lost in deuterostomes (Luke et al., 2003).

In *Drosophila*, all members of the Tin-C are involved in muscle cell differentiation, and many of the mesodermal patterning functions of Tin-C genes are conserved between flies, annelids and vertebrates. For example, the founding member, *tinman* (also known as *NK4*), is expressed in the cardiac mesoderm of all three major animal groups (Azpiazu and Frasch, 1993; Bodmer et al., 1990; Saudemont et al., 2008; Tanaka et al., 1998). Moreover, *bagpipe* (*NK3*) is involved in patterning both fly and vertebrate visceral mesoderm (Azpiazu and Frasch, 1993; Tribioli and Lufkin, 1999), whereas *ladybird*/*Lbx*, *slouch*/*Nk1*, *C15*/*Txl* and *Msh* (*Dr*)/*Msx* regulate the patterning of somatic muscle precursors in both flies and annelids (Jagla et al., 2001; Saudemont et al., 2008).

Here, we examine the role of the Tin-C in patterning the cardiac mesoderm, with a particular emphasis on the regulation of the *ladybird* gene in the heart field of the *Tribolium* embryo. In *Drosophila*, *ladybird* encodes a transcriptional repressor that is thought to be essential for the subdivision of the cardiac mesoderm into distinct pericardial and cardial lineages (Jagla et al., 1997). *ladybird* is expressed in the pericardial and cardial cells of *Drosophila* and *Apis mellifera* (honeybee) embryos, but is absent from the heart field of *Tribolium*. Evidence is presented that this loss in expression is replaced by the neighboring *C15* gene within the *Tribolium* Tin-C.

The replacement of *ladybird* expression by *C15* appears to result from altered enhancer-promoter interactions arising from an internal inversion within the *Tribolium* Tin-C. Evidence is presented that this inversion bypasses an insulator located in the *ladybird-C15* region. In *Drosophila*, the cardiac enhancer is located 3′ of the *ladybird early* and *ladybird late* genes, and is unable to activate *C15* expression owing to insulator activity at the *ladybird early* promoter. The chromosomal inversion in *Tribolium* relocates this enhancer so that the *ladybird* promoter is no longer positioned between the enhancer and *C15* gene. As a result, the cardiac enhancer is able to activate *C15* expression in *Tribolium*, but not *Drosophila*. Thus, a novel pattern of gene expression, *C15* expression in *Tribolium* pericardial and cardial cells, is not due to the modification of gene regulatory networks or the de novo evolution of enhancer sequences, but rather results from the interaction of a conserved enhancer with different target genes: *ladybird* in *Drosophila* and the neighboring *C15* gene in *Tribolium*. We propose that the redirection of conserved enhancers might be an important mechanism of regulatory evolution.

Department of Molecular and Cell Biology, Division of Genetics, Genomics and Development, Center for Integrative Genomics, University of California at Berkeley, Berkeley, CA 94720, USA.

*Author for correspondence (mlevine@berkeley.edu)

Accepted 22 July 2009

MATERIALS AND METHODS

Stocks

The *D. melanogaster* strain used for *P*-element-mediated transgenesis was *yw*^{67}, as described previously (e.g. Stathopoulos and Levine, 2002). pUAST overexpression constructs were balanced using the 612 double balancer (*yw; Sp/CyO; Pr Dr/TM3 Sb Ser*), and balanced or homozygous males were crossed to *twi-GAL4; 24B-GAL4* (Lockwood and Bodmer, 2002) virgin females. For promoter blocking, transgenic males were crossed to *yw* virgin females to prevent two copies of the insulator transgene from pairing in homozygous embryos, and RNA in situs were carried out to detect embryos with the *white* and *lacZ* transgenes. All experiments involving *P*-elements were performed with a minimum of three independent lines. ΦC31-mediated transgenesis was carried out using either *86Fb* (Bischof et al., 2007) or an *attP2* (Groth et al., 2004) pNos-integrase line provided by Michael Eisen (University of California at Berkeley, Berkeley, CA, USA), or usually both. *Tribolium castaneum* stocks were a gift from Nipam Patel (University of California at Berkeley, Berkeley, CA, USA), and were raised according to standard methods. *Apis mellifera* hives were obtained from Honeybee Genetics (Vacaville, CA, USA).

Cloning and injection of transgenic constructs

DNA fragments were amplified from genomic DNA; primer sequences are available upon request. *Drosophila* enhancers and the *Tribolium ladybird* downstream enhancer were cloned into the *P*-element vector nE2G (Markstein et al., 2004). All remaining fragments were cloned into the *attB* equivalent of nE2G (*eve* minimal promoter, *lacZ* reporter) (Hare et al., 2008). DNA fragments tested for enhancer activity were generally ~2 kb in length. Full-length *Tribolium ladybird* and *C15* cDNAs were amplified from embryonic *Tribolium* cDNA generated using a BD Marathon RACE Kit (BD Biosciences). Full-length *Drosophila C15* cDNA was generated from an equivalent library using fusion PCR, and all constructs were cloned into pUAST (Brand and Perrimon, 1993). Constructs were introduced into the *Drosophila* germline as previously described (e.g. Rubin and Spradling, 1982; Groth et al., 2004).

Embryo fixation and staining

Drosophila, *Tribolium* and *Apis* embryo fixation and in situ hybridizations were done as described previously (Kosman et al., 2004; Zinzen et al., 2006). For RNA detection, embryos were hybridized with digoxigenin (Roche) or dinitrophenyl (PerkinElmer) labeled probes and visualized colorimetrically (Jiang et al., 1991) or fluorescently (Kosman et al., 2004), together with the DNA stain DRAQ5 (Biostatus). Probes were generated by in vitro transcription; primer sequences are available upon request. In general, probes were 2-3 kb in length, and encompassed the 5′ regions of the genes. Tc-*ladybird* and Tc-*C15* probes were made using full-length cDNAs, and *Apis pannier* was made with a 3′ RACE product (Ohara et al., 1989). Eve protein was visualized with the cross-reactive 2B8 antibody from Nipam Patel (Patel et al., 1994), which was diluted to 1:50 and detected using the mouse ABC Elite Kit (Vector Labs), together with fluorescein-conjugated tyramide signal amplification (PerkinElmer).

Microarray design and analysis

The *Tribolium* Tin-C sequence was obtained from the Tcas_2.0 Baylor HSGC assembly. The tiled region consists of ~240 kb on LG9 (12,440,000-12,680,000). NimbleGen designed 50 bp features covering these regions at a density of approximately one feature per 90 bp. Chromatin for immunoprecipitations (IPs) was prepared from ~15- to 40-hour *Tribolium* embryos (raised at 30°C) according to Zeitlinger et al. (Zeitlinger et al., 2007b), with the exception that *Tribolium* embryos were dechorionated for 2 minutes in 100% bleach and were crosslinked for 10 minutes in formaldehyde-buffered hexanes. *Tribolium* embryo stages were chosen to capture gastrulation and the onset of segmentation, a time when many developmental genes are active. Both *ladybird* and *C15* are expressed at this time, albeit not in large numbers of cells: *C15* is expressed in the dorsal ectoderm, *ladybird* in ectodermal segmental stripes. Chromatin IPs were performed according to NimbleGen's standard protocol (http://www.nimblegen.com), using a combination of anti-Pol II antibodies (H14, H5 and 8WG16, from Covance). Final IP samples were amplified using the WGA2 Kit (Sigma) before being submitted to NimbleGen for labeling and hybridization. Visualization and scaling of tiling data was done with the Integrated Genome Browser (Affymetrix). *Drosophila* Pol II data was obtained from Zeitlinger et al. (Zeitlinger et al., 2007a). *Tribolium* microarray data is available from ArrayExpress (Accession number E-TABM-758).

Chromosome conformation capture (3C) analysis

Late *Drosophila* embryos (8-10 hours) were chosen for 3C experiments because of the expression of Tin-C genes at this stage. Embryos were dechorionated and fixed in formaldehyde-saturated hexanes, as described for ChIP experiments (Zeitlinger et al., 2007b). The embryos were then treated according to a standard 3C protocol (Lanzuolo et al., 2007). In brief, the crosslinked embryos were resuspended in ice-cold lysis buffer and incubated on ice for ~15 minutes. The embryos were lysed on ice using a Dounce homogenizer, followed by washes in restriction buffer and the chromatin was then dissolved in the appropriate restriction buffer with SDS. The SDS helps to dissolve the insoluble chromatin for restriction enzyme digestion. This dissolved crosslinked chromatin was subjected to *Eco*RI digestion to completion. The reaction was stopped by adding an excess of SDS and by heat inactivation of the enzyme. The digested chromatin was diluted 16-fold in ligation buffer and subjected to ligation at 18°C for 2 hours, followed by reversal of the crosslinks and purification of the DNA. This 3C library was then subjected to PCR amplification using primer pairs flanking *Eco*RI sites near the enhancer and promoter sequences. Similar 3C libraries were constructed as experimental controls from non-crosslinked embryos, genomic DNA (to check primer pair efficiency), and crosslinked embryos but without ligation as essential controls.

Computational analysis

Apis and *Tribolium* orthologs of *Drosophila* heart genes were identified by a reciprocal BLAST-BLAST strategy, and the Clustal alignment of Tin-C genes was generated in BioEdit. ClusterDraw2 has been described previously (Zinzen et al., 2006). The program and the motifs used in this analysis are available online at http://flydev.berkeley.edu/cgi-bin/cld/submit.cgi, except for dTCF, which is described by Chang et al. (Chang et al., 2008).

RESULTS

Organization of the Tin-C

The Tin-C is conserved in all insect genomes sequenced to date (Diptera, Coleoptera and Hymenoptera), spanning three orders and an estimated 300 million years of evolutionary divergence (see Fig. S1 in the supplementary material). Tin-C orthologs are readily identified by homology in their homeodomains and TN domains (aligned in Fig. 1A and Fig. S1A in the supplementary material). The TN domain, named after Tinman, is an N-terminal protein interaction domain similar to the eh1 domain present in engrailed family proteins, and is required for interaction with the Groucho corepressor protein (Choi et al., 1999; Jagla et al., 2001). With the exception of *Drosophila*, there is only a single copy of *ladybird*, indicating that the *ladybird* duplication seen in *Drosophila melanogaster* is a relatively recent occurrence (Fig. 1, compare B and C).

The Tin-C appears to contain a chromosomal breakpoint between *C15* and *slouch* in *Tribolium*, and between *ladybird* and *C15* in *Apis*. However, this latter break might be an assembly error, as the two *Apis* genes are located on the ends of contigs in the current *Apis* genome assembly. Furthermore, the *ladybird* and *C15* linkage is conserved in the genome of the parasitic wasp and fellow hymenopteran *Nasonia vitripennis*. The orientation of *ladybird* is inverted in *Tribolium* as compared with *Drosophila* (Fig. 1, compare B and C). The tandem arrangement of the *ladybird* and *C15* genes seen in *Tribolium* is also present in *Nasonia* (not shown).

ChIP-chip revealed a large spike of RNA polymerase II (Pol II) binding in the promoter region of the *Drosophila ladybird early* (*lbe*) gene in the dorsal ectoderm (Fig. 1C). As *lbe* and *ladybird late* (*lbl*) are not transcribed in this tissue, it appears that Pol II is stalled at the

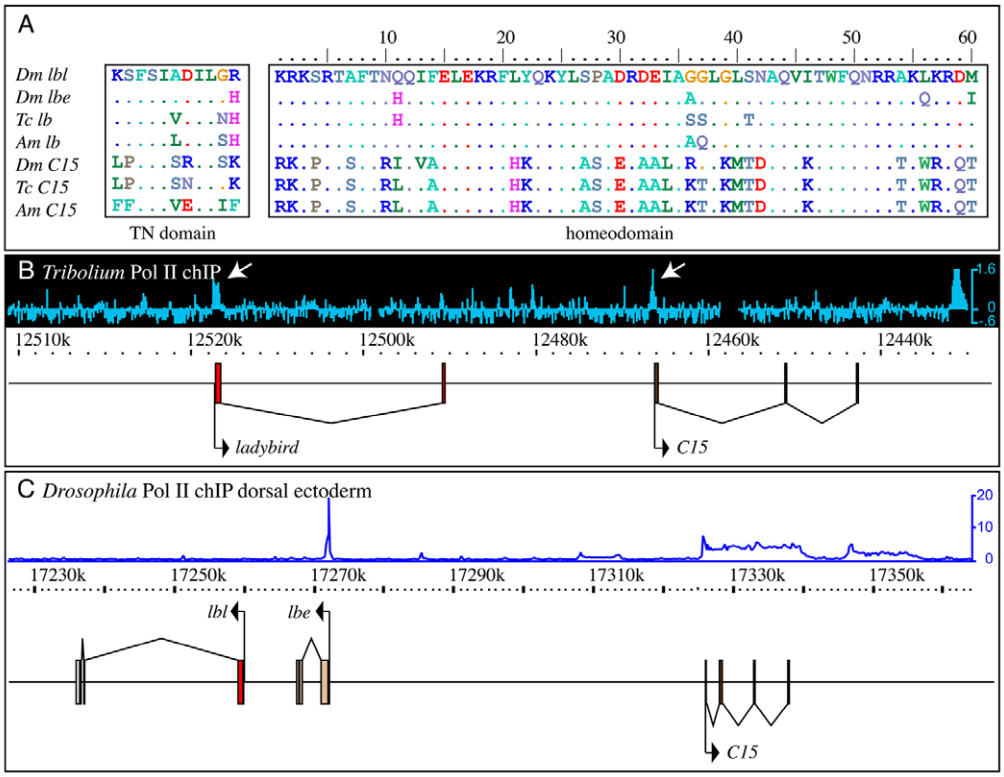

Fig. 1. The *ladybird* and *C15* locus in *Drosophila* and *Tribolium*. (**A**) An alignment of the TN domain and homeodomain of the *Apis mellifera*, *Tribolium castaneum* and *Drosophila melanogaster ladybird* and *C15* orthologs. (**B**) *Tribolium ladybird* and *C15* locus. Across the top, Pol II ChIP-chip from ~15- to 40-hour *Tribolium* embryos in which both *ladybird* and *C15* are transcribed. The *y*-axis indicates log2 of the ratio between the ChIP sample and the chromatin input control, indicating fold enrichment. Arrows indicate the promoters. (**C**) *Drosophila ladybird* and *C15* locus. Pol II ChIP-chip data from *Drosophila* 2- to 4-hour dorsal ectoderm tissue is taken from Zeitlinger et al. (Zeitlinger et al., 2007a). The *y*-axis indicates fold enrichment over input DNA. *lbe* and *lbl* are not expressed in the dorsal ectoderm, yet Pol II binding indicates stalled polymerase at the *lbe* promoter. *C15* is highly expressed and Pol II is bound across the gene. *Dm*, *Drosophila melanogaster*; *Tc*, *Tribolium castaneum*; *Am*, *Apis mellifera*; *lb*, *ladybird*; *lbe*, *ladybird early*; *lbl*, *ladybird late*.

lbe, but not the *lbl*, promoter (Zeitlinger et al., 2007a). Additional Pol II ChIP-chip assays were consistent with the occurrence of stalled Pol II at the *ladybird* and *C15* promoters in *Tribolium* (white arrows, Fig. 1B).

Tin-C gene expression patterns in *Tribolium* and *Apis*

The expression patterns of Tin-C genes are broadly conserved in *Apis*, *Tribolium* and *Drosophila* (see Figs S2-S4 in the supplementary material) (Janssen and Damen, 2008). For example, *tinman* is expressed in the developing heart in both *Apis* and *Tribolium*, whereas the duplicated *Msh* genes are expressed in overlapping neurogenic stripes and in specific dorsal somatic muscles, similar to the patterns seen in *Drosophila* (Azpiazu and Frasch, 1993; Bodmer et al., 1990; Lord et al., 1995). A striking exception to this general trend of conserved expression is seen for the neighboring *ladybird* and *C15* genes in *Tribolium*.

The *ladybird* expression patterns seen in *Drosophila* were found to be conserved in *Tribolium*, except for the cardiac component of the pattern, which was absent (arrow, Fig. 2B). In addition, *ladybird* was expressed in the developing larval limbs in *Tribolium*, an expression pattern that is frequently seen with genes expressed in *Drosophila* leg imaginal discs (J.D.C. and M.L., unpublished). The loss of *ladybird* expression in the *Tribolium* heart is surprising, as Ladybird is thought to play a crucial role in patterning pericardial and cardial cells in *Drosophila* (Jagla et al., 1997).

The *C15* gene was more broadly expressed in *Tribolium* than in *Drosophila* or *Apis* (Fig. 2, *Tribolium*; Fig. S3 in the supplementary material, *Apis*). In *Drosophila*, *C15* is expressed in the dorsal ectoderm (Lin et al., 2006) and in the leg discs, where it works with *aristaless* to pattern the distal tip of the leg (Campbell, 2005). *Drosophila C15* is additionally expressed in a subset of dorsal mesodermal cells outside the cardiac field (M. Frasch, personal communication). In *Tribolium*, *C15* was expressed in the dorsal ectoderm, starting at the onset of gastrulation (not shown), and in the tips of the larval legs, consistent with its role in limb patterning (Fig. 2C). *C15* was also broadly expressed along the edge of the germ band encompassing the presumptive heart (arrow, Fig. 2D).

The *Tribolium C15* expression pattern partially overlapped *tinman* staining (Fig. 2E) and was expressed in cells alongside the Eve-positive pericardial cells (Fig. 2F; see Fig. S5 in the supplementary material), as seen for the *Drosophila ladybird* expression pattern (Cripps and Olson, 2002). However, unlike either the *Drosophila ladybird* or *C15* pattern, *Tribolium C15* staining extended beyond the *Tribolium tinman* domain, and included the cardiac mesoderm and dorsal ectodermal cell types. By contrast, the *Apis C15* expression pattern was restricted to the imaginal discs and a small number of dorsal cells (arrows, Fig. S3C,D in the supplementary material). Thus, it appears that *ladybird* expression is replaced by *C15* in the cardiac mesoderm of *Tribolium*, although the remaining components of the expression patterns (such as expression in the central nervous system and imaginal discs) are conserved.

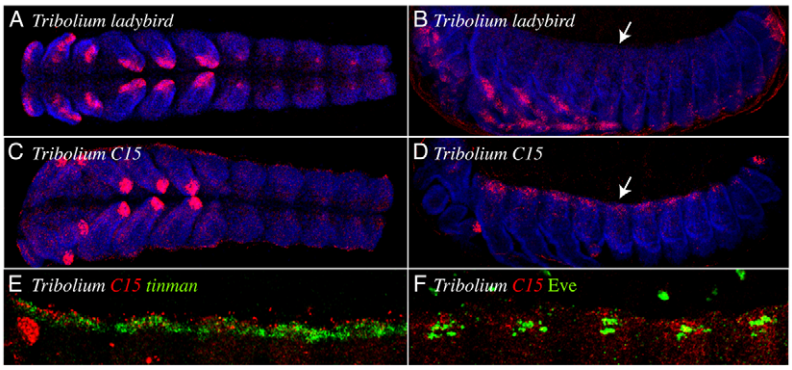

Fig. 2. *ladybird* and *C15* expression patterns in *Tribolium*. Anterior is to the left. Ventral views of germ band extended stage embryos (**A**,**C**) and lateral view of embryos after germ band retraction (**B**,**D-F**). Embryos are stained by in situ hybridization for Tc*lb* (A,B) or Tc*C15* (C,D) in red, with DNA in blue. Tc*lb* is expressed in the developing larval limbs, in clusters of cells in the nervous system and in the anal plate (A,B). Arrow in B indicates the presumptive heart field. Tc*C15* is clearly visible in the tips of the larval limbs (C), and is broadly expressed along the edge of the germ band and in the presumptive heart field (arrow in D). Staining for *C15* (red) and the heart marker *tinman* (green, E) or the pericardial marker Eve (green, F) reveals *C15* expression in the presumptive heart field. Dorsal fusion of the two heart fields has not yet occurred.

Drosophila and *Tribolium* are sister groups, whereas *Apis* is the outgroup of this comparison (Savard et al., 2006), and the gene duplication event that gave rise to *ladybird* and *C15* occurred before the protostome/deuterostome split (Larroux et al., 2007). The *Apis ladybird* and *C15* genes are orthologs of the vertebrate *Lbx* and *Tlx* genes, respectively (Saudemont et al., 2008). This observation, along with the conservation of expression between *Drosophila* and *Apis*, suggests that the loss of *ladybird* and gain of *C15* expression in the cardiac mesoderm is specific to *Tribolium*.

Ladybird and C15 function as repressors

ladybird and *C15* encode similar (but not identical) transcription factors, and contain well-conserved homeodomains, as well as related TN domains (Fig. 1A). It is therefore possible that one protein might be able to substitute for the other. In *Drosophila*, overexpression of Ladybird throughout the mesoderm represses its direct transcriptional target, *even-skipped* (*eve*) (Jagla et al., 2002), whereas *eve* is expanded in *ladybird* mutants (Jagla et al., 1997). *Tribolium* Ladybird, *Tribolium* C15 and *Drosophila* C15 were overexpressed throughout the mesoderm using the *GAL4*/*UAS* system (Lockwood and Bodmer, 2002). All three proteins repressed the late mesodermal *eve* expression pattern (arrows, Fig. 3). As an internal control, *eve* expression was largely unaffected in the neurogenic ectoderm, which resides outside the limits of the specific *GAL4* mesodermal driver used in these assays. However, there might be some divergence in protein function: *Drosophila* C15 attenuated Eve expression only in the dorsal somatic muscles and not in the pericardial cells when misexpressed with a weaker *GAL4* driver. By contrast, the *Tribolium* C15 protein repressed Eve equally well in both cell types (see Fig. S6 in the supplementary material).

Cis-regulation of *ladybird* and *C15* expression in *Drosophila*

To determine the basis for the switch in the *ladybird* and *C15* expression patterns in *Tribolium*, it was necessary to identify the cis-regulatory DNAs within the *ladybird-C15* region of the Tin-C in both *Drosophila* and *Tribolium*. In *Drosophila*, mesodermal expression of *lbe* and *lbl* is regulated by two enhancers, which were identified by ChIP-chip assays for the pan-mesodermal transcription factor Twist (Sandmann et al., 2007). One of the enhancers is located within the *lbl* intron (Fig. 4A, fragment 2), and the other is located downstream of the two *ladybird* genes (Fig. 4A, fragment 1). These enhancers drove nearly identical patterns of reporter gene expression in transgenic *Drosophila* embryos in the cardiac mesoderm (Fig. 4C,E), starting around stage 9 (see Fig. S7 in the supplementary material).

The expression patterns directed by the *ladybird* enhancers were complementary to the *eve* pattern (e.g. Fig. 4E; see Fig. S7 in the supplementary material), consistent with their repression by Eve (Jagla et al., 2002). The expression patterns of the reporter genes were slightly broader than the endogenous patterns (see Fig. S8 in the supplementary material), in that they persisted in most cardial cells (Fig. 4C,E), rather than being restricted to just two cardial cells per segment (Cripps and Olson, 2002). This discrepancy can be explained by the loss of repressor sites in the minimal enhancer sequences used in this analysis. Expression of both *ladybird* enhancers was also drastically reduced or absent in $tin^{-/-}$ mutant embryos (see Fig. S9 in the supplementary material), which lack the heart and associated dorsal mesodermal lineages (Azpiazu and Frasch, 1993; Bodmer, 1993; Yin and Frasch, 1998).

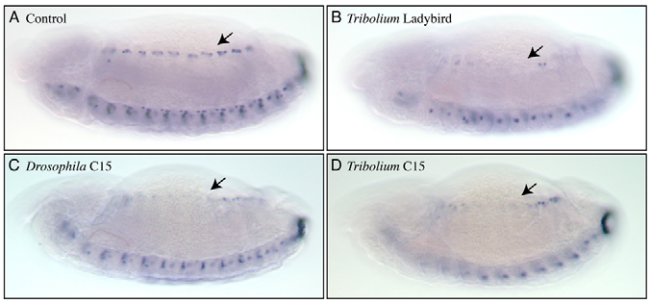

Fig. 3. Overexpression of Ladybird and C15 represses Eve. Anterior is to the left, dorsal is up. The *twi-GAL4; 24B-GAL4* mesodermal driver was used to ectopically express Tc*lb*, Tc*C15* and Dm*C15* full-length cDNAs in the *Drosophila* embryo mesoderm using the *GAL4*/*UAS* system. All embryos were stained for Eve expression; arrows indicate Eve expression in the dorsal and cardiac mesoderm. (**A**) An embryo with ectopic expression of a GFP control gene shows the wild-type Eve pattern. Embryos overexpressing Tc*lb* (**B**), Dm*C15* (**C**) and Tc*C15* (**D**) all have reduced mesodermal Eve expression starting around stage 13, whereas Eve expression in the anal plate and central nervous system is relatively unaffected.

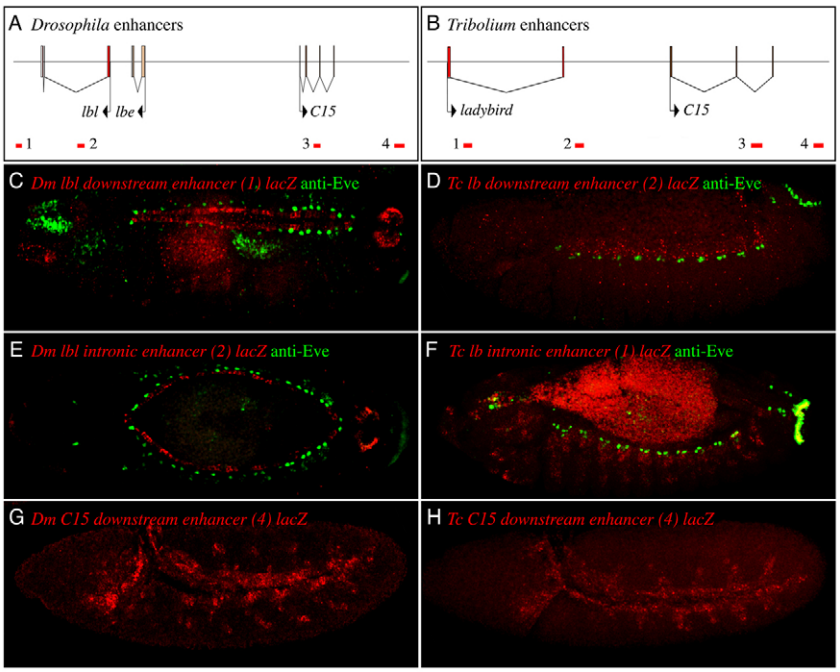

Fig. 4. *ladybird* and *C15* enhancers in *Drosophila* and *Tribolium*. (**A**,**B**) Schematics of the locus in *Drosophila* (A) and *Tribolium* (B), indicating the location of the enhancers as red bars. Numbers next to the red bars correspond to the transgenic embryos in the other panels. (**C-H**) Transgenic *Drosophila* embryos carrying *lb* or *C15* enhancers from *Drosophila* or *Tribolium*, driving a *lacZ* reporter (red), co-stained with anti-Eve (green, C-F). In *Drosophila*, enhancers downstream of the *lbl* gene (fragment 1) and in the *lbl* intron (fragment 2) drive a *lacZ* reporter in the cardiac mesoderm like the *lb* genes (C,E). A published enhancer in the *C15* intron (fragment 3) (Lin et al., 2006) and an additional downstream enhancer (fragment 4) drive expression in the dorsal ectoderm (G and Fig. S7 in the supplementary material). *Tribolium* has a cardiac enhancer downstream of *ladybird* (fragment 2; D), a dorsal mesodermal enhancer in the *lb* intron (fragment 1; F), and *C15* dorsal ectodermal enhancers in the same positions as the *Drosophila* enhancers (fragments 3 and 4 in B; embryos in H and Fig. S7 in the supplementary material).

Previous Twist ChIP-chip assays (Sandmann et al., 2007) identified a putative enhancer downstream of the *C15* transcription unit (Fig. 4A, fragment 4; Fig. 4G). When tested in transgenic embryos, this downstream enhancer appeared to recapitulate the endogenous expression pattern in the dorsal ectoderm (see Fig. S7 in the supplementary material), although it was weaker than the previously published intronic *C15* enhancer (Fig. 4A, fragment 3) (Lin et al., 2006).

Cis-regulation of *ladybird* and *C15* in *Tribolium*

Computational methods were used to identify potential *ladybird* and *C15* enhancers within the *Tribolium* Tin-C. Particular efforts focused on the identification of binding site clusters for transcription factors known to be involved in the differentiation of the cardiac mesoderm, including Tinman, Eve and GATA factors such as Pannier, Smads, TCF (also known as Pan), Twist and Mef2 (Fig. S10 in the supplementary material shows the cluster analysis) (Cripps and Olson, 2002). Putative enhancers were tested for activity in transgenic *Drosophila* embryos (Fig. 4D,F,H; see Fig. S8 in the supplementary material). Genomic DNA fragments 3 and 4 (indicated in Fig. 4B) appeared to function as *C15* enhancers (Fig. 4H; see Fig. S7C in the supplementary material). The downstream *Tribolium C15* enhancer (fragment 4) directed reporter expression in the *Drosophila* dorsal ectoderm and a small number of dorsal mesodermal cells, just like the endogenous *Drosophila C15* gene (Fig. 4H; compare with 4G). The intronic enhancer (fragment 3) directed expression primarily in the dorsal ectoderm (see Fig. S7C in the supplementary material). These enhancers are located in roughly the same relative positions as the two *Drosophila C15* enhancers (Fig. 4B; compare with 4A). Neither enhancer appeared to direct reporter gene expression in the *Drosophila* cardiac field, so it would appear that the expression of *C15* in the *Tribolium* heart might depend on enhancers located elsewhere in the Tin-C.

Two potential enhancers were also identified in the vicinity of the *Tribolium ladybird* gene. One is located downstream of *ladybird* and upstream of *C15* (Fig. 4B, fragment 2), and the other is located within the *ladybird* intron (Fig. 4B, fragment 1). This is similar to the arrangement seen in *Drosophila* (Fig. 4B; compare with 4A). The *Tribolium ladybird* intronic enhancer mediated broad reporter expression throughout the developing *Drosophila* dorsal somatic musculature (Fig. 4F). This intronic enhancer might be a bona fide non-cardiac mesodermal *Tribolium ladybird* enhancer that is misexpressed in *Drosophila* embryos due to evolutionary changes in crucial upstream regulators. The downstream 3′ enhancer (fragment 2) directed expression throughout the cardial cells (Fig. 4D), as well as in additional cells located dorsal to the developing heart. These expression profiles closely resemble the endogenous *Tribolium C15* expression pattern, and those produced by the corresponding *lbe* and *lbl* enhancers in *Drosophila* (Fig. 4D; compare with 4C,E). An additional six tested genomic DNA fragments failed to produce specific patterns of expression in transgenic *Drosophila* embryos (gray bars in Fig. S10 in the supplementary material; data not shown).

In principle, the *Tribolium* cardiac enhancer located between *ladybird* and *C15* (fragment 2, Fig. 4B), along with the dorsal ectodermal enhancers within the *C15* locus, can account for the complete *Tribolium C15* expression pattern. These enhancers are organized in a similar manner in *Drosophila* and *Tribolium*. However, the *ladybird* locus in *Tribolium* is inverted with respect to *Drosophila* (Fig. 4B; compare with 4A). This flips the orientation of the *Tribolium ladybird* gene with respect to *C15*, although the order of the genes in the Tin-C is unchanged. As a result, the cardiac enhancer that resides downstream of *ladybird* is now located 5′ of the *C15* gene in *Tribolium*. We performed a number of assays to determine whether the novel cardiac *C15* expression pattern seen in *Tribolium* could arise from the interaction of the '*ladybird*' downstream enhancer with the *Tribolium C15* promoter.

Regulation of enhancer-promoter interactions

In *Drosophila*, the 3′ *lbe* enhancer might not be able to activate *C15* expression in the cardiac mesoderm owing to an insulator DNA sequence located somewhere in the *lbe-C15* genomic interval (see summary in Fig. 4A). A recent study suggests that promoters containing stalled Pol II can function as insulators (Chopra et al., 2009). Consequently, the *lbe* promoter, which contains stalled Pol II

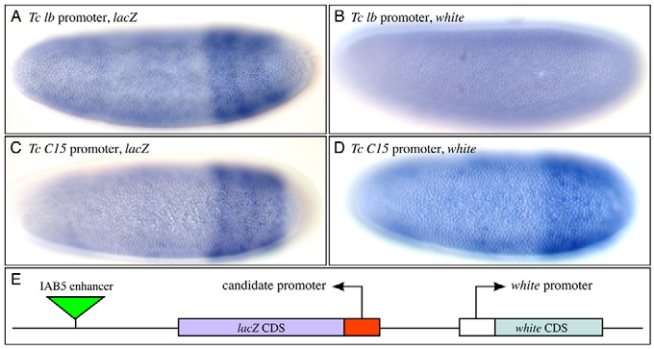

Fig. 5. The *Tribolium ladybird* promoter is an insulator. (**A-D**) *Tribolium ladybird* and *C15* promoter sequences were cloned upstream of *lacZ* in the vector shown in **E** (Ohtsuki and Levine, 1998). In the absence of an insulator, the IAB5 enhancer is capable of driving expression of the *white* transgene in the posterior embryo in addition to *lacZ*, as is the case with the Tc*C15* promoter (C,D). When a promoter that is also an insulator, such as Tc*lb*, is placed upstream of *lacZ*, it prevents IAB5 from activating the *white* transcription unit (B), whereas *lacZ* expression is unaffected (A). CDS, coding sequence.

(Fig. 1C), might be responsible for blocking *C15* expression in the cardiac mesoderm. Thus, the inversion of *ladybird* in *Tribolium* might be sufficient to account for the activation of *C15* in the cardiac mesoderm, as the downstream *ladybird* enhancer is no longer positioned between the *ladybird* promoter and the *C15* transcription start site (see summary in Fig. 4B). This model for the switch in the *ladybird* and *C15* expression patterns in *Drosophila* and *Tribolium* was tested using a combination of enhancer-blocking and chromosome conformation capture (3C) assays.

The *Tribolium ladybird* promoter possessed a robust enhancer-blocking activity when positioned between the IAB5 enhancer and distal *white* gene (Fig. 5A,B; see also Fig. S11 in the supplementary material). The IAB5 enhancer (Busturia and Bienz, 1993) activates *lacZ* expression in the presumptive abdomen (Fig. 5A), but fails to activate *white* (Fig. 5B). By contrast, the *Tribolium C15* promoter did not function as an insulator, and consequently, the shared IAB5 enhancer activated both the *lacZ* and *white* reporter genes in the presumptive abdomen (Fig. 5C,D). We also tested the *Drosophila lbl*, *lbe* and *C15* promoters for insulator activity using the same assay. In *Drosophila*, the *lbe* promoter acted as an insulator (Fig. 6C,D), whereas the *lbl* (Fig. 6A,B) and *C15* (Fig. 6E,F) promoters did not. This is consistent with the occurrence of stalled Pol II at the *lbe* promoter, but not at the *lbl* or *C15* promoters (Fig. 1C). However, there is not always a perfect correlation between stalled Pol II and enhancer blocking.

The *Tribolium ladybird* and *C15* promoters appear to contain stalled Pol II (see Fig. 1B). Nonetheless, only *ladybird* exhibits insulator activity in these assays. Similar results were observed in a recent study using the same *white*/*lacZ* expression vectors (Chopra et al., 2009). Most, but not all, stalled promoters were found to possess an insulator activity. For example, the *gooseberry* promoter contains stalled Pol II, whereas the linked *gooseberry-neuro* promoter does not. Neither promoter was found to contain insulator activity, suggesting that only a subset of stalled promoters mediate enhancer blocking (Chopra et al., 2009) (see Discussion).

Chromosome conformation capture (3C) permits the direct detection of long-range enhancer-promoter interactions (Dekker et al., 2002; Lanzuolo et al., 2007). This technique involves crosslinking embryos, restriction digestion of the crosslinked chromatin, ligation of noncontiguous sequences and PCR amplification. We used 3C assays to examine the regulation of the

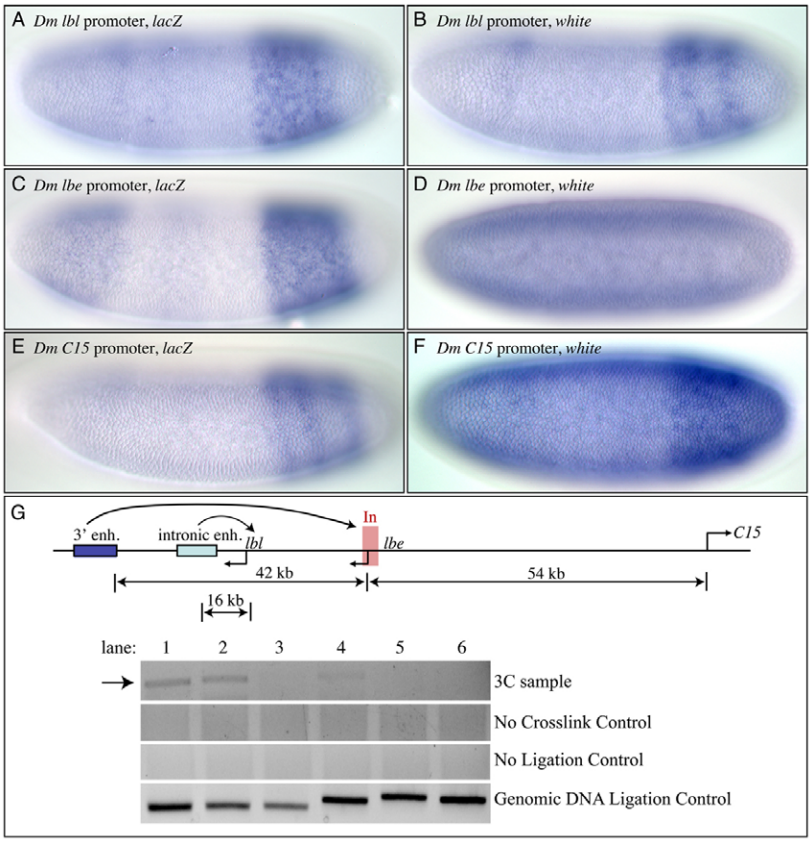

Fig. 6. The *Drosophila ladybird early* promoter is an insulator. *Drosophila lbe*, *lbl* and *C15* promoter sequences were cloned upstream of *lacZ* in the vector shown in Fig. 5E. (**A-F**) The *Drosophila lbe* promoter prevented *white* expression (compare C and D), whereas *lbl* (A,B) and *C15* (E,F) promoters did not act as insulators. (**G**) Chromosome conformation capture (3C) analysis of *Drosophila ladybird* and *C15* promoter and enhancer interactions in 8- to 10-hour embryos (all three genes are expressed). Distances between primer pairs are indicated in the diagram at top. The top gel is the *Eco*RI-digested and ligated chromatin sample. The second and third gels are negative controls in which the chromatin was either not crosslinked or not ligated, and so loops were not preserved. The bottom row is *Eco*RI-digested and ligated naked genomic DNA, and is a positive control for the primer pairs. The intronic enhancer interacts with the *lbl* promoter, the downstream enhancer primarily with the *lbe* promoter, and neither enhancer interacts with *C15*, consistent with the insulator activities of the *lbe* promoter. In, insulator. Primer pairs in each lane were as follows: 1, *lbl* promoter and *lbl* intronic enhancer; 2, *lbe* promoter and *lbl* 3′ enhancer; 3, *lbe* promoter and *lbl* intronic enhancer; 4, *lbl* promoter and *lbl* 3′ enhancer; 5, *C15* promoter and *lbl* 3′ enhancer; 6, *C15* promoter and *lbl* intronic enhancer. Arrow indicates PCR products from chromatin loops.

two *ladybird* genes in *Drosophila*. These assays suggest that the *ladybird* enhancers interact with the *lbe* and *lbl* promoters, but not the *C15* promoter (Fig. 6G). The *lbl* intronic enhancer appeared to activate *lbl* (Fig. 6G, lane 1), whereas the 3′ enhancer bypassed the *lbl* promoter (which lacks insulator activity) to interact with the distal *lbe* promoter (Fig. 6G, lane 2). Similar preferential enhancer-promoter interactions have been documented in the *Drosophila* Hox complexes (Ohtsuki et al., 1998).

The lack of a blocking activity by *Drosophila lbl* can account for the long-range activation of *lbe* by the 3′ enhancer (Fig. 6G). Moreover, this enhancer should be 'captured' by the *lbe* promoter/insulator, and therefore fail to activate the neighboring *C15* gene. The inversion of the *ladybird-C15* region in *Tribolium* causes the 3′ cardiac enhancer to be positioned on the other side of the *ladybird* promoter/insulator, where it is equidistant between the promoters for the *ladybird* and *C15* genes. As a result, the *Tribolium ladybird* promoter is not expected to block activation of *C15* expression in the cardiac mesoderm (Fig. 7).

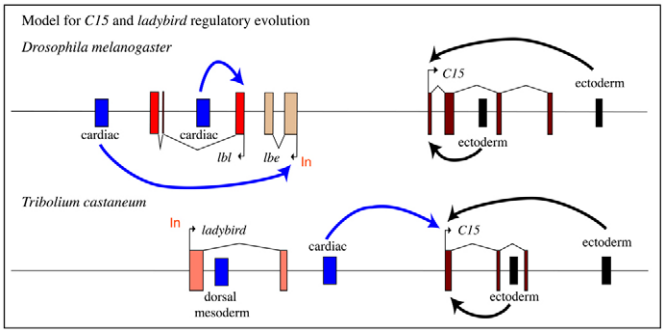

Fig. 7. Regulatory evolution in the insect *ladybird* and *C15* locus. Experimentally verified *ladybird* and *C15* enhancers and insulators in *Drosophila* and *Tribolium*. Arrows indicate which enhancers are likely to be interacting with which promoters (blue for *ladybird* proximal enhancers, black for *C15*). In *Tribolium*, rearrangement of Tc*lb* with respect to Tc*C15* brings the downstream heart enhancer into proximity with the Tc*C15* promoter, adding a heart component to the Tc*C15* expression pattern and removing the heart element of the Tc*lb* pattern. In, insulator.

DISCUSSION

There are a growing number of examples of enhancer modifications resulting in changes in gene expression and morphology during animal evolution. For example, regulatory changes in *pitx1* lead to a reduction in the pelvic fins of certain populations of stickleback fish (Shapiro et al., 2004). Moreover, regulatory changes in the *yellow* locus have been implicated in the diverse patterns of wing spots among divergent drosophilids (Prud'homme et al., 2006). Additionally, pre-existing enhancers are known to evolve through binding site turnover. Such changes can be neutral and cause little or no alterations in gene expression (Ludwig et al., 1998). Alternatively, such changes can lead to novel patterns of gene expression as binding sites for new transcription factors are gained or lost, as seen for the honeybee *sim* midline enhancer (Zinzen et al., 2006).

In the present study, we have presented evidence for a distinctive mechanism of regulatory evolution. Namely, a chromosomal inversion within the Tin-C results in a novel pattern of *C15* gene expression in the cardiac mesoderm of *Tribolium* (summarized in Fig. 7). Previous studies provide evidence that some of the most dramatic developmental mutants result from chromosomal rearrangements. For example, the dominant *Antennapedia* mutation in *Drosophila* is caused by a chromosomal inversion that results in the ectopic expression of Antennapedia in the head, thereby transforming antennae into legs (Frischer et al., 1986).

What are the consequences of replacing Ladybird with C15 in the Tribolium heart? Although *ladybird* and *C15* probably arose from the duplication of a single ancestral gene, the two have diverged significantly in function and protein sequence in flies and vertebrates (Jagla et al., 2001). Outside the DNA-binding and TN domains, the two proteins are extensively different, and might interact with different protein partners. It is therefore likely that exchanging one protein for the other is not a neutral change, even though *Tribolium* C15 can repress the *Drosophila* Ladybird target *eve* in the fly cardiac mesoderm (see Fig. 3). Expression arrays in *Drosophila* have demonstrated that Ladybird has a wide variety of transcriptional targets, including not only transcription factors involved in specifying cell identity during heart patterning, but also genes required for cell shape changes, adhesion and motility, and late-acting genes involved in myogenesis (Junion et al., 2007). Thus, the substitution of Ladybird with C15 might subtly change the functional or morphological properties of the *Tribolium* heart, such as alterations in contractile strength, beating frequency, cell migration or attachment.

We have presented evidence that the switch in expression in the *Tribolium* heart field is due to a chromosomal inversion in the *ladybird-C15* region of the Tin-C (Fig. 7). This inversion appears to foster the activation of *C15* by a cardiac enhancer located 3′ of the *ladybird* locus. In *Drosophila*, this interaction appears to be inhibited by an insulator located at the intervening *lbe* promoter. In *Tribolium*, the altered orientation of *ladybird* and *C15* places the conserved cardiac enhancer between the two genes. As a result, the *ladybird* insulator cannot block enhancer-*C15* interactions, as it no longer resides between the enhancer and the *C15* promoter, as seen in *Drosophila*. In principle, the cardiac enhancer should be able to activate both *ladybird* and *C15* expression because it is located at an equidistant position between the two genes. It is possible that preferential activation of *C15* expression results from promoter competition, whereby *C15* sequesters the cardiac enhancer (e.g. Choi and Engel, 1986). Alternatively, the *C15* promoter region might contain 'tethering' elements that preferentially recruit the cardiac enhancer (e.g. Calhoun and Levine, 2003).

We do not believe that the switch in enhancer-promoter interactions observed in this study is a unique attribute of the Tin-C. Inversions are quite common in insect genomes. By a recent estimate, only 10% of all genes share the same neighbors in the *Drosophila* and *Apis* genomes (Zdobnov and Bork, 2007). In principle, genome rearrangements have the potential to create novel patterns of gene expression by bringing new promoters under the influence of pre-existing enhancers. Unlike the more popularly documented examples of enhancer evolution through binding site turnover, changes in enhancer-promoter interactions not only produce novel patterns of gene expression, but also have the potential to introduce a new protein into a pre-existing developmental gene network in a single step.

Acknowledgements

We thank Nipam Patel for providing the Eve antibody, Rolf Bodmer for the *twi-GAL4; 24B-GAL4* driver and Manfred Frasch for the *tin* mutant flies. We also thank Chris Lotfabodhi and Grace Lim for technical help, and Matthew Ronshaugen, Manfred Frasch and Lionel Christiaen for insightful comments and discussions. J.D.C. was funded in part by a Chang Ling Tien fellowship for the environmental sciences. This work was funded by grants from the Moore Foundation and the NIH (GM46638). Deposited in PMC for release after 12 months.

Supplementary material

Supplementary material for this article is available at http://dev.biologists.org/cgi/content/full/136/18/3153/DC1

References

Azpiazu, N. and Frasch, M. (1993). tinman and bagpipe: two homeobox genes that determine cell fates in the dorsal mesoderm of *Drosophila*. *Genes Dev.* **7**, 1325-1340.

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* **104**, 3312-3317.

Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Development* **118**, 719-729.

Bodmer, R., Jan, L. Y. and Jan, Y. N. (1990). A new homeobox-containing gene, msh-2, is transiently expressed early during mesoderm formation of *Drosophila*. *Development* **110**, 661-669.

Bolshakov, V. N., Topalis, P., Blass, C., Kokoza, E., della Torre, A., Kafatos, F. C. and Louis, C. (2002). A comparative genomic analysis of two distant diptera, the fruit fly, *Drosophila melanogaster*, and the malaria mosquito, *Anopheles gambiae*. *Genome Res.* **12**, 57-66.

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.

Busturia, A. and Bienz, M. (1993). Silencers in abdominal-B, a homeotic *Drosophila* gene. *EMBO J.* **12**, 1415-1425.

Calhoun, V. C. and Levine, M. (2003). Long-range enhancer-promoter interactions in the Scr-Antp interval of the *Drosophila* Antennapedia complex. *Proc. Natl. Acad. Sci. USA* **100**, 9878-9883.

Campbell, G. (2005). Regulation of gene expression in the distal region of the *Drosophila* leg by the Hox11 homolog, C15. *Dev. Biol.* **278**, 607-618.

Chang, M. V., Chang, J. L., Gangopadhyay, A., Shearer, A. and Cadigan, K. M. (2008). Activation of wingless targets requires bipartite recognition of DNA by TCF. *Curr. Biol.* **18**, 1877-1881.

Choi, C. Y., Lee, Y. M., Kim, Y. H., Park, T., Jeon, B. H., Schulz, R. A. and Kim, Y. (1999). The homeodomain transcription factor NK-4 acts as either a transcriptional activator or repressor and interacts with the p300 coactivator and the Groucho corepressor. *J. Biol. Chem.* **274**, 31543-31552.

Choi, O. R. and Engel, J. D. (1986). A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. *Nature* **323**, 731-734.

Chopra, V. S., Cande, J., Hong, J. W. and Levine, M. (2009). Stalled Hox promoters as chromosomal boundaries. *Genes Dev.* **23**, 1505-1509.

Cripps, R. M. and Olson, E. N. (2002). Control of cardiac development by an evolutionarily conserved transcriptional network. *Dev. Biol.* **246**, 14-28.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306-1311.

Frischer, L. E., Hagen, F. S. and Garber, R. L. (1986). An inversion that disrupts the Antennapedia gene causes abnormal structure and localization of RNAs. *Cell* **47**, 1017-1023.

Groth, A. C., Fish, M., Nusse, R. and Calos, M. P. (2004). Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775-1782.

Hare, E. E., Peterson, B. K., Iyer, V. N., Meier, R. and Eisen, M. B. (2008). Sepsid even-skipped enhancers are functionally conserved in *Drosophila* despite lack of sequence conservation. *PLoS Genet.* **4**, e1000106.

Jagla, K., Frasch, M., Jagla, T., Dretzen, G., Bellard, F. and Bellard, M. (1997). ladybird, a new component of the cardiogenic pathway in *Drosophila* required for diversification of heart precursors. *Development* **124**, 3471-3479.

Jagla, K., Bellard, M. and Frasch, M. (2001). A cluster of *Drosophila* homeobox genes involved in mesoderm differentiation programs. *BioEssays* **23**, 125-133.

Jagla, T., Bidet, Y., Da Ponte, J. P., Dastugue, B. and Jagla, K. (2002). Cross-repressive interactions of identity genes are essential for proper specification of cardiac and muscular fates in *Drosophila*. *Development* **129**, 1037-1047.

Janssen, R. and Damen, W. G. (2008). Diverged and conserved aspects of heart formation in a spider. *Evol. Dev.* **10**, 155-165.

Jiang, J., Kosman, D., Ip, Y. T. and Levine, M. (1991). The dorsal morphogen gradient regulates the mesoderm determinant twist in early *Drosophila* embryos. *Genes Dev.* **5**, 1881-1891.

Junion, G., Bataille, L., Jagla, T., Da Ponte, J. P., Tapin, R. and Jagla, K. (2007). Genome-wide view of cell fate specification: ladybird acts at multiple levels during diversification of muscle and heart precursors. *Genes Dev.* **21**, 3163-3180.

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**, 846.

Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F. and Orlando, V. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.* **9**, 1167-1174.

Larroux, C., Fahey, B., Degnan, S. M., Adamski, M., Rokhsar, D. S. and Degnan, B. M. (2007). The NK homeobox gene cluster predates the origin of Hox genes. *Curr. Biol.* **17**, 706-710.

Lin, M. C., Park, J., Kirov, N. and Rushlow, C. (2006). Threshold response of C15 to the Dpp gradient in *Drosophila* is established by the cumulative effect of Smad and Zen activators and negative cues. *Development* **133**, 4805-4813.

Lockwood, W. K. and Bodmer, R. (2002). The patterns of wingless, decapentaplegic, and tinman position the *Drosophila* heart. *Mech. Dev.* **114**, 13-26.

Lord, P. C., Lin, M. H., Hales, K. H. and Storti, R. V. (1995). Normal expression and the effects of ectopic expression of the *Drosophila* muscle segment homeobox (msh) gene suggest a role in differentiation and patterning of embryonic muscles. *Dev. Biol.* **171**, 627-640.

Ludwig, M. Z., Patel, N. H. and Kreitman, M. (1998). Functional analysis of eve stripe 2 enhancer evolution in *Drosophila*: rules governing conservation and change. *Development* **125**, 949-958.

Luke, G. N., Castro, L. F., McLay, K., Bird, C., Coulson, A. and Holland, P. W. (2003). Dispersal of NK homeobox gene clusters in amphioxus and humans. *Proc. Natl. Acad. Sci. USA* **100**, 5292-5295.

Markstein, M., Zinzen, R., Markstein, P., Yee, K. P., Erives, A., Stathopoulos, A. and Levine, M. (2004). A regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development* **131**, 2387-2394.

Negre, B. and Ruiz, A. (2007). HOM-C evolution in *Drosophila*: is there a need for Hox gene clustering? *Trends Genet.* **23**, 55-59.

Ohara, O., Dorit, R. L. and Gilbert, W. (1989). One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci. USA* **86**, 5673-5677.

Ohtsuki, S. and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. *Genes Dev.* **12**, 3325-3330.

Ohtsuki, S., Levine, M. and Cai, H. N. (1998). Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev.* **12**, 547-556.

Patel, N. H., Condron, B. G. and Zinn, K. (1994). Pair-rule expression patterns of even-skipped are found in both short- and long-germ beetles. *Nature* **367**, 429-434.

Prud'homme, B., Gompel, N., Rokas, A., Kassner, V. A., Williams, T. M., Yeh, S. D., True, J. R. and Carroll, S. B. (2006). Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050-1053.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348-353.

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V. and Furlong, E. E. (2007). A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* **21**, 436-449.

Saudemont, A., Dray, N., Hudry, B., Le Gouar, M., Vervoort, M. and Balavoine, G. (2008). Complementary striped expression patterns of NK homeobox genes during segment formation in the annelid *Platynereis*. *Dev. Biol.* **317**, 430-443.

Savard, J., Tautz, D., Richards, S., Weinstock, G. M., Gibbs, R. A., Werren, J. H., Tettelin, H. and Lercher, M. J. (2006). Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. *Genome Res.* **16**, 1334-1338.

Shapiro, M. D., Marks, M. E., Peichel, C. L., Blackman, B. K., Nereng, K. S., Jonsson, B., Schluter, D. and Kingsley, D. M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**, 717-723.

Stathopoulos, A. and Levine, M. (2002). Linear signaling in the Toll-Dorsal pathway of *Drosophila*: activated Pelle kinase specifies all threshold outputs of gene expression while the bHLH protein Twist specifies a subset. *Development* **129**, 3411-3419.

Tanaka, M., Kasahara, H., Bartunkova, S., Schinke, M., Komuro, I., Inagaki, H., Lee, Y., Lyons, G. E. and Izumo, S. (1998). Vertebrate homologs of tinman and bagpipe: roles of the homeobox genes in cardiovascular development. *Dev. Genet.* **22**, 239-249.

Tribioli, C. and Lufkin, T. (1999). The murine Bapx1 homeobox gene plays a critical role in embryonic development of the axial skeleton and spleen. *Development* **126**, 5699-5711.

Yin, Z. and Frasch, M. (1998). Regulation and function of tinman during dorsal mesoderm induction and heart specification in *Drosophila*. *Dev. Genet.* **22**, 187-200.

Zdobnov, E. M. and Bork, P. (2007). Quantification of insect genome divergence. *Trends Genet.* **23**, 16-20.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J. W., Nechaev, S., Adelman, K., Levine, M. and Young, R. A. (2007a). RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* **39**, 1512-1516.

Zeitlinger, J., Zinzen, R. P., Stark, A., Kellis, M., Zhang, H., Young, R. A. and Levine, M. (2007b). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* **21**, 385-390.

Zinzen, R. P., Cande, J., Ronshaugen, M., Papatsenko, D. and Levine, M. (2006). Evolution of the ventral midline in insect embryos. *Dev. Cell* **11**, 895-902.