Heredity (2006) 97, 139–147

www.nature.com/hdy

SHORT REVIEW

Evolution of *cis*-regulatory sequence and function in Diptera

PJ Wittkopp[1,2]
[1]*Department of Ecology and Evolutionary Biology, University of Michigan, 1061 Natural Science Building, 830 North University Ave., Ann Arbor, MI 48109-1048, USA;* [2]*Department of Molecular, Cellular, and Developmental Biology, University of Michigan, 1061 Natural Science Building, 830 North University Ave., Ann Arbor, MI 48109-1048, USA*

Cis-regulatory sequences direct patterns of gene expression essential for development and physiology. Evolutionary changes in these sequences contribute to phenotypic divergence. Despite their importance, *cis*-regulatory regions remain one of the most enigmatic features of the genome. Patterns of sequence evolution can be used to identify *cis*-regulatory elements, but the power of this approach depends upon the relationship between sequence and function. Comparative studies of gene regulation among Diptera reveal that divergent sequences can underlie conserved expression, and that expression differences can evolve despite largely similar sequences. This complex structure-function relationship is the primary impediment for computational identification and interpretation of *cis*-regulatory sequences. Biochemical characterization and *in vivo* assays of *cis*-regulatory sequences on a genomic-scale will relieve this barrier.
Heredity (2006) **97,** 139–147. doi:10.1038/sj.hdy.6800869; published online 19 July 2006

Keywords: *Drosophila*; gene regulation; enhancer; expression; divergence; conservation

Introduction

> 'The art of progress is to preserve order amid change and to preserve change amid order.' Alfred North Whitehead

Mutations are inevitable. Biological systems maintain their function in the face of genetic changes, while preserving flexibility that allows the system to adapt to new environments. Genomic regulatory networks that control gene expression are no exception. These networks are composed of highly conserved *trans*-regulatory proteins and *cis*-regulatory DNA sequences that specify gene expression patterns (Davidson, 2001). Comparisons of *cis*-regulatory elements among Diptera (ie 'true' flies) indicate that their sequences are robust to mutational changes, yet receptive to functional divergence.

Molecular mechanisms that control protein expression facilitate both developmental stability and evolutionary change. Proteins required for the development of characters shared among Diptera typically have conserved expression patterns (eg Averof and Patel, 1997; Panganiban *et al*, 1997). Identifying *cis*-regulatory sequences mediating conserved regulatory inputs helps unravel genomic regulatory networks. Traits that differ among Dipteran species, such as body coloration, bristle patterns, and larval hairs, often correlate with divergent expression of developmental proteins (Stern, 1998; Sucena and Stern, 2000; Wulbeck and Simpson, 2000; Pistillo *et al*, 2002; Wittkopp *et al*, 2002; Gompel and Carroll, 2003) (Figure 1). *Cis*-regulatory sequences that control transcription are a common source of divergent protein expression patterns and thus of phenotypic change (Carroll *et al*, 2001).

Here, I examine comparative studies of gene regulation among Diptera. For reviews of regulatory evolution that encompass more taxa, see Stern (2000), Tautz (2000), Ludwig (2002), Simpson (2002), Wray *et al* (2003). After providing an overview of *cis*-regulatory architecture and molecular evolution, I review case studies that compare the sequence and function of *cis*-regulatory elements among species. Properties of regulatory systems that allow *cis*-regulatory sequences and function to evolve at different rates are discussed, illustrated by case studies where available. Understanding the structure–function relationship of *cis*-regulatory regions is essential for comparative genomic studies of gene regulation. I conclude by examining computational approaches for identifying *cis*-regulatory regions and arguing that additional biochemical, genetic, and transgenic studies are sorely needed to improve computational tools.

Enhancers control patterns of gene expression

Expression of protein coding sequences is controlled by *cis*-regulatory regions, which include a 'basal promoter' and one or more 'enhancers' (Figure 2a). The basic structure of *cis*-regulatory regions is shared not only among Diptera but among all eukaryotes. For more comprehensive reviews of *cis*-regulatory architecture, see Arnone and Davidson (1997), Carroll *et al* (2001),

Correspondence: PJ Wittkopp, Department of Ecology and Evolutionary Biology, University of Michigan, 1061 Natural Science Building, 830 North University Ave., Ann Arbor, MI 48109-1048, USA.
E-mail: wittkopp@umich.edu
Received 10 November 2005; accepted 5 June 2006; published online 19 July 2006

Davidson (2001), Smale (2001), Arnosti (2003) and Wray *et al* (2003).

Basal (or 'core') promoters are necessary for transcription, but do not provide spatiotemporal information for gene expression. They contain binding sites for the general transcription machinery, including the TATA-binding protein and the RNA polymerase II protein complex. As these proteins must bind promoters throughout the genome, and because the assembly of the protein complex is strictly required for the production of mRNA, basal promoter sequences and the proteins that bind to them are under strong functional constraint. This constraint is visible as a reduced level of polymorphism and divergence in these regions among *Drosophila* species (Kohn *et al*, 2004). The few polymorphisms that do exist in basal promoters appear to contribute little to variable gene expression among strains of *D. melanogaster* (Brown and Feder, 2005). Genes can also contain multiple, alternative promoters that are active under different cellular conditions (Ayoubi and Van De Ven, 1996). The contribution of alternative promoters to regulatory divergence remains unclear.

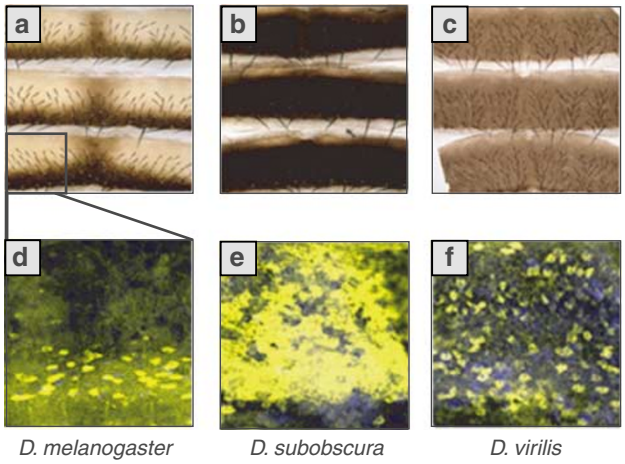

Figure 1 Divergent protein expression correlates with divergent phenotypes. Dorsal abdominal cuticle from adult flies (**a**–**c**) is shown above developing abdominal tissue (**d**–**f**) from *Drosophila melanogaster*, *D. subobscura*, and *D. virilis*, respectively. The distribution of the Yellow protein, which controls the production of black pigment, is shown in the color yellow in lower panels. Note that the expression of Yellow correlates with both the pattern and intensity of black pigment in adult flies (Wittkopp *et al*, 2002).

Changes in enhancer sequence are a common cause of *cis*-regulatory divergence (eg Fang and Brennan (1992), Ross *et al* (1994), Wittkopp *et al* (2002). Enhancer sequences specify when, where, and how much mRNA will be transcribed from the associated coding sequence. They are modular (ie function independently) and many genes contain more than one enhancer element, with each directing a subset of the total gene expression pattern. Enhancers are also composed of binding sites for transcription factor proteins, but, unlike basal promoters, each enhancer contains binding sites for a unique combination of transcription factors. Once bound, transcription factor proteins interact with each other and the polymerase protein complex assembled on the basal promoter to activate and sustain transcription. The specific combination of transcription factors assembled on an enhancer determines its activity.

As enhancers are modular, their activity can be determined using transgenic 'reporter genes' (Barolo *et al*, 2000). These constructs contain putative *cis*-regulatory

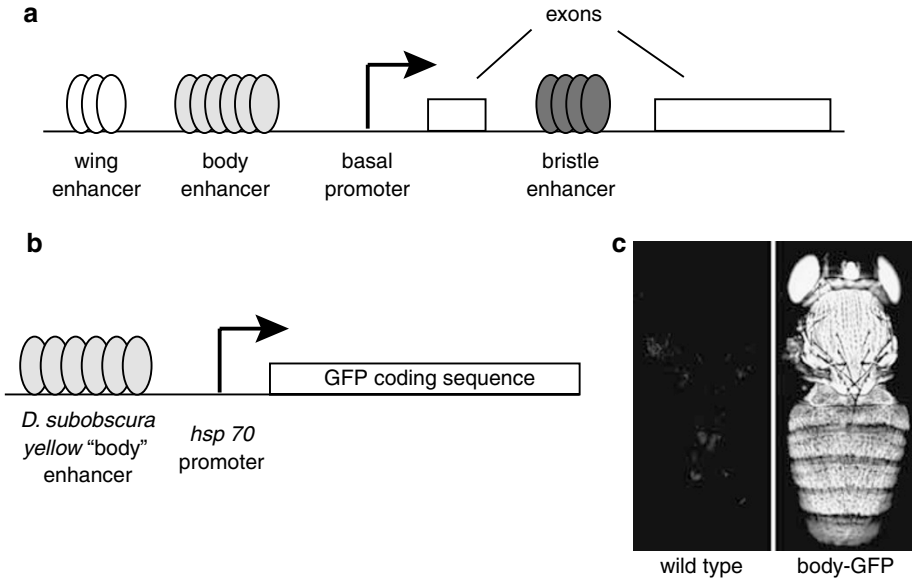

Figure 2 Reporter genes are used to ascertain enhancer activity. (**a**) A schematic of the *D. melanogaster yellow* gene is shown, containing two exons, a basal promoter, and multiple tissue-specific enhancers. (**b**) Reporter genes contain enhancer sequences, a generic basal promoter, and coding sequences for an easily visualized protein, such as the Green Fluorescent Protein (GFP) or β-galactosidase (β-gal). The composition of a GFP reporter gene for the *D. subobscura* 'body' enhancer is illustrated. (**c**) When transformed into a fly, the reporter gene provides a read-out of enhancer activity. GFP expression driven by the *D. subobscura* body enhancer at a late pupal stage is shown in a *D. melanogaster* transformant fly on the right, with a wild type fly on the left. Note that reporter gene expression within each abdominal segment (bracket) is comparable to expression of the endogenous *D. subobscura* Yellow protein shown in Figure 1e. (Expression in eyes and ocelli is due to Pax6-GFP transformation marker, not the *D. subobscura* body enhancer.)

sequences and a basal promoter that drives expression of an easily visualized reporter protein (Figure 2b). After transforming a reporter gene into a host species, its expression is determined (Figure 2c). *D. melanogaster* P-elements (Spradling and Rubin, 1982) are the most commonly used transformation system for assaying Dipteran enhancers.

When transformed into *D. melanogaster*, heterologous *cis*-regulatory sequences are regulated in *trans* by *D. melanogaster* transcription factors. This has both advantages and disadvantages. Assaying orthologous *cis*-regulatory elements in a common *trans*-regulatory background allows their functions to be directly compared. However, if properties of the *trans*-regulators (eg transcription factors) have diverged between the donor and host species, the activity of the *cis*-regulatory element in *D. melanogaster* will differ from its activity in the species from which it was derived. Comparing the activity of *cis*-regulatory sequences transformed into multiple species provides the most complete view of regulatory evolution (Cavener, 1992; Christophides *et al*, 2000; Wittkopp *et al*, 2002; Lombardo *et al*, 2005).

Molecular evolution of *cis*-regulatory sequences

Enhancers and basal promoters are subject to the same process of molecular evolution as all other regions of the genome. Nucleotide substitutions, insertions, deletions, and rearrangements arise, and the balance of selection and drift determines their survival over time (Li, 1997). Studies of polymorphism and divergence in *cis*-regulatory regions of Diptera provide evidence both for and against models of neutral sequence evolution (Ludwig and Kreitman, 1995; Hancock *et al*, 1999; Kohn *et al*, 2004; Phinchongsakuldit *et al*, 2004; Andolfatto, 2005). Each of these studies employs different population genetic models and tests for selection; the most appropriate model for the neutral evolution of *cis*-regulatory regions is not yet established.

cis-Regulatory mutations can influence phenotypes by altering gene expression. Therefore, selection coefficients for *cis*-regulatory changes should be related to their effects on expression. Mutations that do not alter expression are assumed to be neutral (ie 'silent'), whereas mutations that disrupt sequences essential for *cis*-regulatory function are assumed to be deleterious. Sequences comprising transcription factor binding sites may thus be more constrained than sequences not used as binding sites. Surprisingly, the pattern of nucleotide substitutions is similar within characterized binding sites and in surrounding regions of DNA, suggesting this may not be the case (Emberly *et al*, 2003; Costas *et al*, 2004; Phinchongsakuldit *et al*, 2004; Balhoff and Wray, 2005). It remains to be seen how often 'surrounding' sequences contain unidentified binding sites, and how sequence divergence affects gene expression.

With few exceptions (Erives and Levine, 2004; Markstein *et al*, 2004; Senger *et al*, 2004), the architecture of binding sites in an enhancer, and the nature of interactions among transcription factors that regulate its activity, are not understood well enough to predict the consequences of specific *cis*-regulatory changes from sequence alone. Empirical tests are required to determine *cis*-regulatory function and to assess the impact of sequence divergence on gene expression.

Conserved sequence and function

Expression patterns conserved across species are specified by *cis*-regulatory elements that have preserved their function over time. *Cis*-regulatory sequences from other Dipteran species often retain their activity when introduced into *D. melanogaster* using transgenes. This is true for sequences taken from other *Drosophila* species, distantly related flies (including the house fly *Musca domestica* and the black fly *Simulium vittatum*) and even animals outside Diptera (Mitsialis and Kafatos, 1985; Martin *et al*, 1988; Langeland and Carroll, 1993; Magoulas *et al*, 1993; Lukowitz *et al*, 1994; Pan *et al*, 1994; Xiong and Jacobs-Lorena, 1995; Tortiglione and Bownes, 1997; Ludwig *et al*, 1998; Wolff *et al*, 1999; Wittkopp *et al*, 2002).

The simplest mechanism for maintaining activity of a *cis*-regulatory element is to conserve the sequences that determine its function. Sequence comparisons of orthologous *cis*-regulatory elements generally show blocks of conserved sequence surrounded by more divergent sequences (Kassis *et al*, 1985; Wilde and Akam, 1987; Kassis *et al*, 1989; Langeland and Carroll, 1993; Lukowitz *et al*, 1994; Pan *et al*, 1994; Sackerson, 1995; Ludwig *et al*, 1998; Wolff *et al*, 1999; Kim, 2001; Dellino *et al*, 2002; Emberly *et al*, 2003; Berman *et al*, 2004; Costas *et al*, 2004). Sequence similarity to *D. melanogaster cis*-regulatory regions has been used to identify enhancers in other Dipterans, including *Anopheles gambiae* (Papatsenko and Levine, 2005), *Scaptodrosophila lebanonensis* (Papaceit *et al*, 2004), and *Calliphora vicina* (Gibert and Simpson, 2003).

Evolutionary comparisons of well-characterized *D. melanogaster* enhancers were used to help motivate the sequencing of genomes from other Dipterans. The expectation *a priori* was that most *cis*-regulatory regions would be easily identified based on sequence similarity to noncoding sequences of the *D. melanogaster* genome (Hardison, 2000; Bergman *et al*, 2002). Unfortunately, this does not appear to be the case. With the completion of the *D. pseudoobscura* genome sequence, researchers found that computational searches for conserved sequences only identified a small fraction of the enhancers in the genome (Richards *et al*, 2005). Ascertainment bias in early empirical studies of *cis*-regulatory regions may have overestimated the requirement for sequence conservation; Dipteran enhancers were often identified based on sequence similarity to *D. melanogaster cis*-regulatory regions, and many *D. melanogaster* enhancers were recognized precisely because they evolved slower than surrounding sequences.

Conserved function despite divergent sequence

Comparisons of *cis*-regulatory elements among Diptera demonstrate that enhancer activity can be maintained despite extensive sequence divergence (Martin *et al*, 1988; Magoulas *et al*, 1993; Tortiglione and Bownes, 1997; Wolff *et al*, 1999; Ludwig *et al*, 2000). This phenomenon has been most clearly illustrated in the stripe 2 enhancer of the *even-skipped* gene, described below.

Even-skipped (*eve*) encodes a transcription factor that plays a similar role in embryonic patterning of

Drosophila, *Anopheles*, and presumably all Diptera (Goltsev *et al*, 2004). Eve protein is expressed in seven transverse stripes along the embryo, which are controlled by five independent enhancers. The *D. melanogaster eve* stripe 2 enhancer includes binding sites for five transcription factors (including two activators and three repressors) that are required for expression of the Eve protein in stripe 2 (Stanojevic *et al*, 1991). Although these binding sites were required for activity in the *D. melanogaster* enhancer that was dissected experimentally, sequence variation within and between species is comparable to other noncoding regions and fits a model of neutral sequence evolution (Ludwig and Kreitman, 1995).

To determine the functional consequences of sequence divergence in the *eve* stripe 2 enhancer (Figure 3), Ludwig *et al* isolated DNA orthologous to the *D. melanogaster* enhancer from *D. yakuba*, *D. erecta*, and *D. pseudoobscura*, and assayed its activity in transgenic *D. melanogaster* using reporter genes. Despite little sequence similarity (including divergence of binding sites essential for expression of the *D. melanogaster* enhancer) orthologous enhancers were able to drive gene expression in a pattern comparable to the *D. melanogaster eve* stripe 2 enhancer (Ludwig *et al*, 1998). Chimeric enhancers were constructed between the *D. pseudoobscura* and *D. melanogaster* alleles, each containing the 5′ and 3′ regions from different species, and introduced into *D. melanogaster* (Ludwig *et al*, 2000). The chimeric enhancers did not function properly, indicating that compensatory changes have evolved since the split of *D. melanogaster* and *D. pseudoobscura*. Orthologous enhancers that produce the same expression pattern despite differences in the arrangement of binding sites have presumably evolved under stabilizing selection.

Recently, the *D. melanogaster*, *D. yakuba*, *D. erecta*, and *D. pseudoobscura eve* stripe 2 enhancers were tested for their ability to rescue an *eve* mutant phenotype (Ludwig *et al*, 2005). The *D. yakuba* and *D. pseudoobscura eve* alleles restored a wild-type phenotype, but the stripe 2 enhancer from *D. erecta*, a species which is more closely related to *D. melanogaster* than is *D. pseudoobscura*, failed to complement the mutation. Sequence divergence of the *D. erecta* stripe 2 enhancer may be such that the *D. erecta* allele requires sequences outside of the region orthologous to the *D. melanogaster* enhancer. Alternatively, the activity and/or expression level of transcription factors regulating the element may have diverged between species. Further experimentation will distinguish among these possibilities.

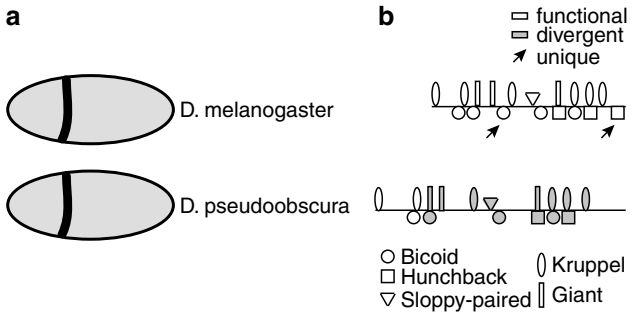

Figure 3 Enhancer activity is conserved despite divergent binding sites. (**a**) The function of the *even-skipped* (*eve*) stripe 2 enhancer is conserved between *D. melanogaster* and *D. pseudoobscura* (Ludwig *et al*, 1998, 2000, 2005). Expression patterns of reporter genes containing *eve* stripe 2 enhancers from both species are depicted in schematic *D. melanogaster* embryos. (**b**) Despite functional conservation of *eve* enhancers, >80% of characterized binding sites have diverged between species. Locations of binding sites for the Bicoid (circle), Hunchback (square), Kruppel (oval), Giant (rectangle), and Sloppy-paired (inverted triangle) transcription factor proteins are shown for the *D. melanogaster* and *D. pseudoobscura eve* stripe 2 enhancers. Binding sites with divergent sequences are shaded; arrows indicate binding sites unique to the *D. melanogaster* enhancer. Adapted from Ludwig *et al* (2005).

If transcription factor binding sites can diverge between species while maintaining enhancer function, then polymorphisms in binding sites may also be segregating within species. Indeed, analysis of another enhancer of the *eve* gene revealed an experimentally confirmed and phylogenetically conserved binding site segregating in natural populations (A Palsson, M Ludwig, and M Kreitman, personal communication), indicating that empirically validated biding sites are not necessarily fixed within species. A cluster of binding sites was also recently found to be polymorphic in a sea urchin *cis*-regulatory element (Balhoff and Wray, 2005). Such intraspecific variation provides raw material for changing enhancer sequences while maintaining enhancer function.

Uncoupling enhancer sequence and function

How can enhancer activity be maintained despite overall sequence divergence? Molecular mechanisms that translate *cis*-regulatory sequences into gene expression patterns allow them to evolve at different rates. Features of regulatory mechanisms that can separate the evolution of *cis*-regulatory sequence and function include: biochemical properties of transcription factors, redundant binding sites and enhancers, changes in transcription factor inputs, and coevolution of transcription factors and their binding sites.

Properties of transcription factors

Flexibility in transcription factor binding (ie 'degeneracy') as well as flexibility in the arrangement and spacing of transcription factors permit many sequence changes to evolve without altering enhancer function (Arnone and Davidson, 1997). Degenerate binding sites allow a transcription factor to continue regulating an enhancer despite sequence divergence. Flexible *cis*-regulatory architecture allows binding sites to be reshuffled while maintaining *cis*-regulatory function.

Redundant binding sites and enhancers

Redundant transcription factor binding sites within an enhancer also facilitate sequence divergence. If individual binding sites can be mutated with minimal disruption to enhancer function, compensatory binding sites can produce a fluid restructuring of *cis*-regulatory regions. For example, expression of the *Drosophila spalt* and *knot* genes is repressed in the developing haltere by the Ultrabithorax (Ubx) homeodomain protein. Multiple Ubx binding sites are present in enhancers for both genes and the loss of individual binding sites has minimal effect on *cis*-regulatory activity (Galant *et al*, 2002; Hersh and Carroll, 2005). Distinct Ubx binding sites have evolved to repress *knot* expression in the halteres of

D. melanogaster and *D. pseudoobscura* (Hersh and Carroll, 2005). Redundant binding sites may also promote sequence divergence and reorganization of the *eve* stripe 2 enhancer (Ludwig *et al*, 2000) and *yolk protein* genes (Piano *et al*, 1999). Redundancy among enhancer modules (Buttgereit, 1993; Piano *et al*, 1999; Pappu *et al*, 2005) also permits sequence changes to accumulate with out affecting *cis*-regulation; if one element is altered, the redundant element can compensate for its function.

Changing transcription factor inputs

Developmental system drift (DSD, True and Haag, 2001) can create enhancers with conserved functions but greatly diverged sequences. DSD occurs when the output of a developmental system remains the same despite the evolution of underlying developmental mechanisms. An apparent case of DSD is embryonic patterning between the mosquito, *Anopheles gambiae*, and the fruit fly, *D. melanogaster*. Expression of the Even-skipped protein appears to be conserved between species, but expression patterns of the genes that regulate *eve* in *D. melanogaster* are different in *Anopheles* (Goltsev *et al*, 2004). These data suggest that *Anopheles cis*-regulatory elements of *eve* are controlled by different transcription factors than the *D. melanogaster eve* enhancers. DSD may also contribute to the restructuring of the *D. erecta eve* stripe 2 enhancer (Ludwig *et al*, 2005).

Coevolution of transcription factors and binding sites

Evolutionary changes in the DNA binding domains of transcription factors promote the divergence of *cis*-regulatory sequences. An example from Diptera is the coevolution of the binding domain of the Bicoid transcription factor and the *cis*-regulatory sequences of the *hunchback* (*hb*) enhancer. The developmental function of the Hunchback protein in early embryonic patterning is conserved among *Drosophila*, the housefly *Musca domestica*, and blowflies *Lucilia sericata* and *Calliphora vicina* (Sommer and Tautz, 1991; Bonneton *et al*, 1997; McGregor *et al*, 2001a). Despite this conservation, *hb cis*-regulatory elements have undergone changes in primary sequence that affect the number and organization of binding sites, especially for Bicoid (Hancock *et al*, 1999; McGregor *et al*, 2001b). Biochemical and transgenic experiments demonstrate that the DNA binding domain of the Bicoid protein coevolved with the binding sites in the *hb* promoter to maintain their regulatory interaction (Shaw *et al*, 2002). Similar coevolution of transcription factor specificity and *cis*-regulatory binding sites has been invoked to explain divergent bristle locations between *D. melanogaster* and *D. simulans* (Skaer *et al*, 2002) and as a possible cause for the divergence of the *D. erecta eve* stripe 2 enhancer (Ludwig *et al*, 2005). However, DNA binding domains of transcription factors are among the most conserved sequences in animal genomes, and it is unclear whether coevolution of transcription factors and their binding sites is a common feature of regulatory evolution.

Sources of divergent *cis*-regulatory activity

Although many *cis*-regulatory elements maintain their function over time, changes in gene expression among Dipteran species are also common. During the 2 million years since the divergence of the *D. melanogaster* and *D. simulans* lineages, up to half of the genes in the genome have evolved differences in their expression level (Ranz *et al*, 2003; Rifkin *et al*, 2003). The majority of these changes appear to be caused by functional divergence of *cis*-regulatory sequences associated with the affected gene (Wittkopp *et al*, 2004). *Cis*-regulatory elements that specify new or altered expression patterns can evolve (1) *de novo*, (2) by divergence of parologous enhancers following duplication, or (3) through the modification of existing enhancers. Studies providing evidence for these modes of *cis*-regulatory divergence are reviewed below.

Enhancer evolution *de novo*

Theoretically, *cis*-regulatory elements controlling evolutionarily novel patterns of gene expression may arise *de novo*. Stone and Wray (2001) simulated neutral sequence evolution by point mutations, while MacArthur and Brookfield (2004) simulated enhancer evolution using a model that incorporates positive selection. Both studies concluded that transcription factor binding sites appear frequently and can be fixed in a population over relatively short periods of time. Currently, there is no empirical evidence of an enhancer evolving *de novo*, but lineage-specific enhancers derived from neutral sequences may be very difficult to identify.

Duplication and divergence

The function of an enhancer can be altered following gene duplication. When a gene, including its *cis*-regulatory sequences, is duplicated, the two copies are redundant and one is free to change its expression pattern (Li and Noll, 1994; Lynch and Force, 2000). Indeed, Gu *et al* (2004) found that duplicated genes are more likely to have evolved expression differences between *Drosophila* species than single copy genes. Parologous genes have been identified in *D. melanogaster* for which the expression change seems to be the primary difference among duplicates; the protein functions remain interchangeable (Rodriguez *et al*, 1990; Li and Noll, 1994). A pair of duplicated genes with divergent expression has also been identified in the medfly, *Cereatitis capitata* (Christophides *et al*, 2000). In a rare experiment using transgenic flies other than *D. melanogaster*, the authors showed that despite extensive sequence similarity, the parologous regulatory elements had evolved differences in tissue-specific and sex-specific expression by altering *cis*-regulatory activity. The two, ~280 basepair (bp) *cis*-regulatory sequences differ by only 12 divergent bases and seven deleted nucleotides, indicating that patterns of gene expression can be dramatically altered with minimal differences in enhancer sequence.

Modified enhancer activities

Gene expression can diverge by altering the function of existing *cis*-regulatory elements. For example, the Alcohol dehydrogenase (Adh) gene shows differences in its spatiotemporal expression pattern among *Drosophila* species that are caused by modifications of *cis*-regulatory sequences (Dickinson *et al*, 1984; Fang *et al*, 1991; Papaceit *et al*, 2004). Similarly, changes in expression of the Yellow protein that correlate with differences in abdominal pigmentation among *D. melanogaster, D. subobscura*, and

D. virilis, are caused by functional divergence of orthologous enhancers (Wittkopp *et al*, 2002). *Cis*-regulatory changes are also responsible for differences in the expression of the *glucose dehydrogenase* (*gld*) gene among flies (Schiff *et al*, 1992). The presence and absence of a five base pair sequence (TTAGA) in the *gld* enhancer correlates with the expression of Gld protein in the ejaculatory duct among *Drosophila* species (Ross *et al*, 1994), implying that enhancer functions may be modified by only a few changes in *cis*-regulatory sequence.

Enhancers controlling lineage-specific patterns of gene expression can also evolve from existing *cis*-regulatory elements, taking advantage of transcription factor binding sites already present. The *D. biarmipes* enhancer regulating expression of the pigmentation gene *yellow* in a male-specific wing spot was recently identified (Gompel *et al*, 2005). This enhancer sequence is orthologous to the *cis*-regulatory element controlling ubiquitous wing expression in *D. melanogaster*. Only 6 bp of the 675 bp enhancer are essential for activation in the wing spot (B Prud'homme and S Carroll, personal communication). Transcription factor(s) that bind to these sequences have not yet been identified. Repression of the spot expression in the posterior compartment is controlled by two, 8 bp binding sites for the Engrailed transcription factor that have evolved in the *D. biarmipes* lineage. Again, changing only a handful of nucleotides produces major changes in *cis*-regulatory activity.

Computational enhancer prediction

Understanding the relationship between sequence and function is essential for developing computational methods to study *cis*-regulatory elements. Comparative genomic approaches hold great promise for accelerating studies of gene regulation in Diptera. Current methods can be broken into two general classes (Bulyk, 2003): (1) phylogenetic footprinting, which uses sequence conservation among two or more species to identify putative *cis*-regulatory sequences (eg Moses *et al*, 2004; Siepel *et al*, 2005) and (2) motif detection, in which statistical models are used to recognize binding sites for specific transcription factors or sequences shared among co-regulated genes (eg Markstein *et al*, 2002). Phylogenetic footprinting may miss many *cis*-regulatory elements because, as discussed in this review, sequence conservation is not strictly required for *cis*-regulatory function. Indeed, the first comparison of two *Drosophila* genome sequences failed to uncover many known *cis*-regulatory elements (Richards *et al*, 2005). Motif detection algorithms, especially those that look for clusters of experimentally defined binding sites, can identify new enhancers without searching for linear sequence conservation (eg Bergman *et al*, 2002; Markstein *et al*, 2002; Berman *et al*, 2004; Erives and Levine, 2004; Papatsenko and Levine, 2005). However, this approach is limited to finding enhancers regulated by transcription factors with known binding sites. The two strategies can also be combined to improve the accuracy of enhancer prediction (Grad *et al*, 2004).

Back to the bench

Despite the promise of computational approaches for studying *cis*-regulatory evolution, analyses of additional enhancers using biochemical, genetic, and transgenic tools is essential for refining these methods. Empirical studies that elucidate enhancer structure and function will allow powerful motif finding algorithms to be used more broadly. The experimental characterization of *cis*-regulatory regions has historically been time-consuming and labor-intensive. Fortunately, high-throughput techniques for studying transcription factor binding on a genomic scale are now available (Sun *et al*, 2003; van Steensel *et al*, 2003; Bergman *et al*, 2005), putting a complete list of transcription factor binding sites in the *D. melanogaster* genome within reach. Functional, *in vivo*, tests of enhancer function in transgenic flies remain the rate-limiting step. Methods that automate the production, isolation and characterization of transgenic flies carrying reporter genes will expedite this work. In addition, tests of enhancer function in *D. melanogaster* must be supplemented with tests of enhancer function in other Dipteran hosts to uncover changes in *trans*-regulation (Cavener, 1992; Christophides *et al*, 2000; Wittkopp *et al*, 2002; Lombardo *et al*, 2005), *minos* (Loukeris *et al*, 1995), *piggyBac* (Handler *et al*, 1998), *mariner* (Coates *et al*, 1998), and *Hermes* (Jasinskiene *et al*, 1998) transposable elements, engineered with dominant fluorescent transformation markers (Berghammer *et al*, 1999), are now available for transforming diverse insect species (Handler, 2002; Wimmer, 2003).

The complex relationship between *cis*-regulatory sequence and function is only beginning to be uncovered. With the sequence of *Anopheles* and 12 *Drosophila* genomes completed or in progress – and advanced transgenic and functional genomic tools already available – studies of Dipteran *cis*-regulatory sequences are poised to continue providing valuable insights into the process of regulatory evolution.

Acknowledgements

I thank A Palsson, M Ludwig, M Kreitman, B Prud'homme, and S Carroll for sharing unpublished data, B Prud'homme, A Palsson and B Haerum for insightful comments that significantly improved this manuscript, and the University of Michigan for financial support.

References

Andolfatto P (2005). Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**: 1149–1152.

Arnone MI, Davidson EH (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**: 1851–1864.

Arnosti DN (2003). Analysis and function of transcriptional regulatory elements: insights from *Drosophila*. *Annu Rev Entomol* **48**: 579–602.

Averof M, Patel NH (1997). Crustacean appendage evolution associated with changes in *Hox* gene expression. *Nature* **388**: 682–686.

Ayoubi TA, Van De Ven WJ (1996). Regulation of gene expression by alternative promoters. *FASEB J* **10**: 453–460.

Balhoff JP, Wray GA (2005). Evolutionary analysis of the well characterized *endo16* promoter reveals substantial variation within functional sites. *Proc Natl Acad Sci USA* **102**: 8591–8596.

Barolo S, Carver LA, Posakony JW (2000). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. *Biotechniques* **29**: 726, 728, 730, 732.

Berghammer AJ, Klingler M, Wimmer EA (1999). A universal marker for transgenic insects. *Nature* **402**: 370–371.

Bergman CM, Carlson JW, Celniker SE (2005). *Drosophila* DNase I footprint database: a systematic genome annotation of transcription factor binding sites in the fruitfly, *Drosophila melanogaster*. *Bioinformatics* **21**: 1747–1749.

Bergman CM, Pfeiffer BD, Rincon-Limas DE, Hoskins RA, Gnirke A, Mungall CJ *et al* (2002). Assessing the impact of comparative genomic sequence data on the functional annotation of the *Drosophila* genome. *Genome Biol* **3**: RESEARCH0086.

Berman BP, Pfeiffer BD, Laverty TR, Salzberg SL, Rubin GM, Eisen MB *et al* (2004). Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol* **5**: R61.

Bonneton F, Shaw PJ, Fazakerley C, Shi M, Dover GA (1997). Comparison of *bicoid*-dependent regulation of *hunchback* between *Musca domestica* and *Drosophila melanogaster*. *Mech Dev* **66**: 143–156.

Brown RP, Feder ME (2005). Reverse transcriptional profiling: non-correspondence of transcript level variation and proximal promoter polymorphism. *BMC Genom* **6**: 110.

Bulyk ML (2003). Computational prediction of transcription-factor binding site locations. *Genome Biol* **5**: 201.

Buttgereit D (1993). Redundant enhancer elements guide *beta 1 tubulin* gene expression in apodemes during *Drosophila* embryogenesis. *J Cell Sci* **105**(Pt 3): 721–727.

Carroll SB, Grenier JK, Weatherbee SD (2001). *From DNA to diversity: molecular genetics and the evolution of animal design*. Blackwell Science: Oxford, Malden, Mass.

Cavener DR (1992). Transgenic animal studies on the evolution of genetic regulatory circuitries. *Bioessays* **14**: 237–244.

Christophides GK, Livadaras I, Savakis C, Komitopoulou K (2000). Two medfly promoters that have originated by recent gene duplication drive distinct sex, tissue and temporal expression patterns. *Genetics* **156**: 173–182.

Coates CJ, Jasinskiene N, Miyashiro L, James AA (1998). Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci USA* **95**: 3748–3751.

Costas J, Pereira PS, Vieira CP, Pinho S, Vieira J, Casares F (2004). Dynamics and function of intron sequences of the *wingless* gene during the evolution of the *Drosophila* genus. *Evol Dev* **6**: 325–335.

Davidson EH (2001). *Genomic regulatory systems: development and evolution*. Academic Press: San Diego.

Dellino GI, Tatout C, Pirrotta V (2002). Extensive conservation of sequences and chromatin structures in the bxd polycomb response element among Drosophilid species. *Int J Dev Biol* **46**: 133–141.

Dickinson WJ, Rowan RG, Brennan MD (1984). Regulatory gene evolution: adaptive differences in expression of *alcohol dehydrogenase* in *Drosophila melanogaster* and *Drosophila simulans*. *Heredity* **52**(Pt 2): 215–225.

Emberly E, Rajewsky N, Siggia ED (2003). Conservation of regulatory elements between two species of *Drosophila*. *BMC Bioinformatics* **4**: 57.

Erives A, Levine M (2004). Coordinate enhancers share common organizational features in the *Drosophila* genome. *Proc Natl Acad Sci USA* **101**: 3851–3856.

Fang XM, Brennan MD (1992). Multiple *cis*-acting sequences contribute to evolved regulatory variation for *Drosophila Adh* genes. *Genetics* **131**: 333–343.

Fang XM, Wu CY, Brennan MD (1991). Complexity in evolved regulatory variation for *alcohol dehydrogenase* genes in Hawaiian *Drosophila*. *J Mol Evol* **32**: 220–226.

Galant R, Walsh CM, Carroll SB (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**: 3115–3126.

Gibert JM, Simpson P (2003). Evolution of *cis*-regulation of the proneural genes. *Int J Dev Biol* **47**: 643–651.

Goltsev Y, Hsiong W, Lanzaro G, Levine M (2004). Different combinations of gap repressors for common stripes in *Anopheles* and *Drosophila* embryos. *Dev Biol* **275**: 435–446.

Gompel N, Carroll SB (2003). Genetic mechanisms and constraints governing the evolution of correlated traits in Drosophilid flies. *Nature* **424**: 931–935.

Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005). Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**: 481–487.

Grad YH, Roth FP, Halfon MS, Church GM (2004). Prediction of similarly acting *cis*-regulatory modules by subsequence profiling and comparative genomics in *Drosophila melanogaster* and *D. pseudoobscura*. *Bioinformatics* **20**: 2738–2750.

Gu Z, Rifkin SA, White KP, Li WH (2004). Duplicate genes increase gene expression diversity within and between species. *Nat Genet* **36**: 577–579.

Hancock JM, Shaw PJ, Bonneton F, Dover GA (1999). High sequence turnover in the regulatory regions of the developmental gene *hunchback* in insects. *Mol Biol Evol* **16**: 253–265.

Handler AM (2002). Use of the piggyBac transposon for germline transformation of insects. *Insect Biochem Mol Biol* **32**: 1211–1220.

Handler AM, McCombs SD, Fraser MJ, Saul SH (1998). The lepidopteran transposon vector, piggyBac, mediates germline transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci USA* **95**: 7520–7525.

Hardison RC (2000). Conserved noncoding sequences are reliable guides to regulatory elements. *Trends Genet* **16**: 369–372.

Hersh BM, Carroll SB (2005). Direct regulation of *knot* gene expression by Ultrabithorax and the evolution of *cis*-regulatory elements in *Drosophila*. *Development* **132**: 1567–1577.

Jasinskiene N, Coates CJ, Benedict MQ, Cornel AJ, Rafferty CS, James AA *et al* (1998). Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the Hermes element from the housefly. *Proc Natl Acad Sci USA* **95**: 3743–3747.

Kassis JA, Desplan C, Wright DK, Ofarrell PH (1989). Evolutionary conservation of homeodomain-binding sites and other sequences upstream and within the major transcription unit of the *Drosophila* segmentation gene *engrailed*. *Molecular And Cellular Biology* **9**: 4304–4311.

Kassis JA, Wong ML, Ofarrell PH (1985). Electron-microscopic heteroduplex mapping identifies regions of the *engrailed* locus that are conserved between *Drosophila melanogaster* and *Drosophila virilis*. *Molecular And Cellular Biology* **5**: 3600–3609.

Kim J (2001). Macro-evolution of the *hairy* enhancer in *Drosophila* species. *J Exp Zool* **291**: 175–185.

Kohn MH, Fang S, Wu CI (2004). Inference of positive and negative selection on the 5′ regulatory regions of *Drosophila* genes. *Mol Biol Evol* **21**: 374–383.

Langeland JA, Carroll SB (1993). Conservation of regulatory elements controlling *hairy* pair-rule stripe formation. *Development* **117**: 585–596.

Li W-H (1997). *Molecular Evolution*. Sinauer Associates: Sunderland, MA.

Li X, Noll M (1994). Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different *cis*-regulatory regions. *Nature* **367**: 83–87.

Lombardo F, Nolan T, Lycett G, Lanfrancotti A, Stich N, Catteruccia F *et al* (2005). An *Anopheles gambiae* salivary gland promoter analysis in *Drosophila melanogaster* and *Anopheles stephensi*. *Insect Mol Biol* **14**: 207–216.

Loukeris TG, Livadaras I, Arca B, Zabalou S, Savakis C (1995). Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. *Science* **270**: 2002–2005.

Ludwig MZ (2002). Functional evolution of noncoding DNA. *Curr Opin Genet Dev* **12**: 634–639.

Ludwig MZ, Bergman C, Patel NH, Kreitman M (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**: 564–567.

Ludwig MZ, Kreitman M (1995). Evolutionary dynamics of the enhancer region of *even-skipped* in *Drosophila*. *Mol Biol Evol* **12**: 1002–1011.

Ludwig MZ, Palsson A, Alekseeva E, Bergman CM, Nathan J, Kreitman M (2005). Functional evolution of a *cis*-regulatory module. *PLoS Biol* **3**: e93.

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

Lukowitz W, Schroder C, Glaser G, Hulskamp M, Tautz D (1994). Regulatory and coding regions of the segmentation gene *hunchback* are functionally conserved between *Drosophila virilis* and *Drosophila melanogaster*. *Mech Dev* **45**: 105–115.

Lynch M, Force A (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**: 459–473.

MacArthur S, Brookfield JF (2004). Expected rates and modes of evolution of enhancer sequences. *Mol Biol Evol* **21**: 1064–1073.

Magoulas C, Loverre-Chyurlia A, Abukashawa S, Bally-Cuif L, Hickey DA (1993). Functional conservation of a glucose-repressible amylase gene promoter from *Drosophila virilis* in *Drosophila melanogaster*. *J Mol Evol* **36**: 234–242.

Markstein M, Markstein P, Markstein V, Levine MS (2002). Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the *Drosophila* embryo. *Proc Natl Acad Sci USA* **99**: 763–768.

Markstein M, Zinzen R, Markstein P, Yee KP, Erives A, Stathopoulos A *et al* (2004). A regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development* **131**: 2387–2394.

Martin CH, Mayeda CA, Meyerowitz EM (1988). Evolution and expression of the *Sgs*-3 glue gene of *Drosophila*. *J Mol Biol* **201**: 273–287.

McGregor AP, Shaw PJ, Dover GA (2001a). Sequence and expression of the *hunchback* gene in *Lucilia sericata*: a comparison with other Dipterans. *Dev Genes Evol* **211**: 315–318.

McGregor AP, Shaw PJ, Hancock JM, Bopp D, Hediger M, Wratten NS *et al* (2001b). Rapid restructuring of bicoid-dependent hunchback promoters within and between Dipteran species: implications for molecular coevolution. *Evol Dev* **3**: 397–407.

Mitsialis SA, Kafatos FC (1985). Regulatory elements controlling chorion gene expression are conserved between flies and moths. *Nature* **317**: 453–456.

Moses AM, Chiang DY, Pollard DA, Iyer VN, Eisen MB (2004). Monkey: identifying conserved transcription-factor binding sites in multiple alignments using a binding site-specific evolutionary model. *Genome Biol* **5**: R98.

Pan D, Valentine SA, Courey AJ (1994). The bipartite *D. melanogaster twist* promoter is reorganized in *D. virilis*. *Mech Dev* **46**: 41–53.

Panganiban G, Irvine SM, Lowe C, Roehl H, Corley LS, Sherbon B *et al* (1997). The origin and evolution of animal appendages. *Proc Natl Acad Sci USA* **94**: 5162–5166.

Papaceit M, Orengo D, Juan E (2004). Sequences upstream of the homologous *cis*-elements of the *Adh* adult enhancer of *Drosophila* are required for maximal levels of *Adh* gene transcription in adults of *Scaptodrosophila lebanonensis*. *Genetics* **167**: 289–299.

Papatsenko D, Levine M (2005). Quantitative analysis of binding motifs mediating diverse spatial readouts of the Dorsal gradient in the *Drosophila* embryo. *Proc Natl Acad Sci USA* **102**: 4966–4971.

Pappu KS, Ostrin EJ, Middlebrooks BW, Sili BT, Chen R, Atkins MR *et al* (2005). Dual regulation and redundant function of two eye-specific enhancers of the *Drosophila* retinal determination gene *dachshund*. *Development* **132**: 2895–2905.

Phinchongsakuldit J, MacArthur S, Brookfield JF (2004). Evolution of developmental genes: molecular microevolution of enhancer sequences at the *Ubx* locus in *Drosophila* and its impact on developmental phenotypes. *Mol Biol Evol* **21**: 348–363.

Piano F, Parisi MJ, Karess R, Kambysellis MP (1999). Evidence for redundancy but not *trans* factor-*cis* element coevolution in the regulation of *Drosophila Yp* genes. *Genetics* **152**: 605–616.

Pistillo D, Skaer N, Simpson P (2002). *scute* expression in *Calliphora vicina* reveals an ancestral pattern of longitudinal stripes on the thorax of higher Diptera. *Development* **129**: 563–572.

Ranz JM, Ponce AR, Hartl DL, Nurminsky D (2003). Origin and evolution of a new gene expressed in the *Drosophila* sperm axoneme. *Genetica* **118**: 233–244.

Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S, Nielsen R *et al* (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and *cis*-element evolution. *Genome Res* **15**: 1–18.

Rifkin SA, Kim J, White KP (2003). Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nat Genet* **33**: 138–144.

Rodriguez I, Hernandez R, Modolell J, Ruiz-Gomez M (1990). Competence to develop sensory organs is temporally and spatially regulated in *Drosophila* epidermal primordia. *EMBO J* **9**: 3583–3592.

Ross JL, Fong PP, Cavener DR (1994). Correlated evolution of the *cis*-acting regulatory elements and developmental expression of the *Drosophila Gld* gene in seven species from the subgroup *melanogaster*. *Dev Genet* **15**: 38–50.

Sackerson C (1995). Patterns of conservation and divergence at the *even-skipped* locus of *Drosophila*. *Mech Dev* **51**: 199–215.

Schiff NM, Feng Y, Quine JA, Krasney PA, Cavener DR (1992). Evolution of the expression of the *Gld* gene in the reproductive tract of *Drosophila*. *Mol Biol Evol* **9**: 1029–1049.

Senger K, Armstrong GW, Rowell WJ, Kwan JM, Markstein M, Levine M (2004). Immunity regulatory DNAs share common organizational features in *Drosophila*. *Mol Cell* **13**: 19–32.

Shaw PJ, Wratten NS, McGregor AP, Dover GA (2002). Coevolution in *bicoid*-dependent promoters and the inception of regulatory incompatibilities among species of higher Diptera. *Evol Dev* **4**: 265–277.

Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K *et al* (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**: 1034–1050.

Simpson P (2002). Evolution of development in closely related species of flies and worms. *Nat Rev Genet* **3**: 907–917.

Skaer N, Pistillo D, Gibert JM, Lio P, Wulbeck C, Simpson P (2002). Gene duplication at the *achaete-scute* complex and morphological complexity of the peripheral nervous system in Diptera. *Trends Genet* **18**: 399–405.

Smale ST (2001). Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* **15**: 2503–2508.

Sommer R, Tautz D (1991). Segmentation gene expression in the housefly *Musca domestica*. *Development* **113**: 419–430.

Spradling AC, Rubin GM (1982). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**: 341–347.

Stanojevic D, Small S, Levine M (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science* **254**: 1385–1387.

Stern DL (1998). A role of Ultrabithorax in morphological differences between *Drosophila* species. *Nature* **396**: 463–466.

Stern DL (2000). Evolutionary developmental biology and the problem of variation. *Evol Int J Org Evol* **54**: 1079–1091.

Stone JR, Wray GA (2001). Rapid evolution of *cis*-regulatory sequences via local point mutations. *Mol Biol Evol* **18**: 1764–1770.

Sucena E, Stern DL (2000). Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by

cis-regulatory evolution of *ovo/shaven-baby*. *Proc Natl Acad Sci USA* **97**: 4530–4534.

Sun LV, Chen L, Greil F, Negre N, Li TR, Cavalli G *et al* (2003). Protein-DNA interaction mapping using genomic tiling path microarrays in *Drosophila*. *Proc Natl Acad Sci USA* **100**: 9428–9433.

Tautz D (2000). Evolution of transcriptional regulation. *Curr Opin Genet Dev* **10**: 575–579.

Tortiglione C, Bownes M (1997). Conservation and divergence in the control of *yolk protein* genes in dipteran insects. *Dev Genes Evol* **207**: 264–281.

True JR, Haag ES (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev* **3**: 109–119.

van Steensel B, Delrow J, Bussemaker HJ (2003). Genomewide analysis of *Drosophila* GAGA factor target genes reveals context-dependent DNA binding. *Proc Natl Acad Sci USA* **100**: 2580–2585.

Wilde CD, Akam M (1987). Conserved sequence elements in the 5′ region of the Ultrabithorax transcription unit. *EMBO J* **6**: 1393–1401.

Wimmer EA (2003). Innovations: applications of insect transgenesis. *Nat Rev Genet* **4**: 225–232.

Wittkopp PJ, Haerum BK, Clark AG (2004). Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430**: 85–88.

Wittkopp PJ, Vaccaro K, Carroll SB (2002). Evolution of *yellow* gene regulation and pigmentation in *Drosophila*. *Curr Biol* **12**: 1547–1556.

Wolff C, Pepling M, Gergen P, Klingler M (1999). Structure and evolution of a pair-rule interaction element: *runt* regulatory sequences in *D. melanogaster* and *D. virilis*. *Mech Dev* **80**: 87–99.

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV *et al* (2003). The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* **20**: 1377–1419.

Wulbeck C, Simpson P (2000). Expression of *achaete-scute* homologues in discrete proneural clusters on the developing notum of the medfly *Ceratitis capitata*, suggests a common origin for the stereotyped bristle patterns of higher Diptera. *Development* **127**: 1411–1420.

Xiong B, Jacobs-Lorena M (1995). Gut-specific transcriptional regulatory elements of the *carboxypeptidase* gene are conserved between black flies and *Drosophila*. *Proc Natl Acad Sci USA* **92**: 9313–9317.