

# Redundant and Cryptic Enhancer Activities of the *Drosophila yellow* Gene

Gizem Kalay,<sup>\*1</sup> Jennifer Lachowiec,<sup>\*2</sup> Ulises Rosas,<sup>†</sup> Mackenzie R. Dome,<sup>\*</sup> and Patricia Wittkopp<sup>\*,†,3</sup>

<sup>\*</sup>Molecular, Cellular, and Developmental Biology and <sup>†</sup>Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109 and <sup>†</sup>Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México, 04510 CDMX, Mexico

**ABSTRACT** *Cis*-regulatory sequences known as enhancers play a key role in regulating gene expression. Evolutionary changes in these DNA sequences contribute to phenotypic evolution. The *Drosophila yellow* gene, which is required for pigmentation, has emerged as a model system for understanding how *cis*-regulatory sequences evolve, providing some of the most detailed insights available into how activities of orthologous enhancers have diverged between species. Here, we examine the evolution of *yellow cis*-regulatory sequences on a broader scale, by comparing the distribution and function of *yellow* enhancer activities throughout the 5' intergenic and intronic sequences of *Drosophila melanogaster*, *D. pseudoobscura*, and *D. willistoni*. We find that *cis*-regulatory sequences driving expression in a particular tissue are not as modular as previously described, but rather have many redundant and cryptic enhancer activities distributed throughout the regions surveyed. Interestingly, cryptic enhancer activities of sequences from one species often drove patterns of expression observed in other species, suggesting that the frequent evolutionary changes in *yellow* expression observed among *Drosophila* species may be facilitated by gaining and losing repression of preexisting *cis*-regulatory sequences.

**KEYWORDS** *cis*-regulatory element; evolution; pigmentation; novelty; gene expression

**C**IS-REGULATORY elements known as enhancers affect development and physiology by controlling the time, place, and amount of mRNA that is transcribed from a gene. These DNA sequences range from hundreds to thousands of base pairs, are typically located in noncoding regions of the genome, and contain binding sites for transcription factors (Spitz and Furlong 2012; Long *et al.* 2016). In multicellular organisms such as *Drosophila*, most genes are regulated by multiple enhancers. Early studies suggested that each enhancer was responsible for a unique subset of a gene's expression (*e.g.*, Davidson 2001), but it is now known that multiple enhancers with varying degrees of functional redundancy can also contribute to the expression of a gene in a given cell (Frankel *et al.* 2010; Perry *et al.* 2010; Cannavò *et al.* 2016; Letelier *et al.* 2018; Osterwalder *et al.* 2018). Despite decades

of research into enhancer structure and function, predicting either the location or function of enhancers from DNA sequence alone remains challenging (Lim *et al.* 2018). Consequently, empirically testing DNA sequences for their ability to activate gene expression *in vivo* using transgenic reporter genes remains a critical step for identifying enhancers and determining their function (Wittkopp and Kalay 2012; Barrière and Ruvinsky 2014).

Evolutionary changes in enhancer sequences can alter an expression pattern within a tissue, eliminate enhancer function, or create novel expression patterns, all of which can contribute to phenotypic differences within or between species (Prud'homme *et al.* 2007; Wittkopp and Kalay 2012; Rubinstein and de Souza 2013; Rebeiz and Tsiantis 2017). Modifying expression within a tissue often results from genetic changes affecting one or more transcription factor-binding sites within an enhancer (Swanson *et al.* 2010; Rogers *et al.* 2013). Loss of enhancer activity can result from point mutations in transcription factor-binding sites (Frankel *et al.* 2011), as well as larger deletions or insertions that disrupt enhancer function (Chan *et al.* 2010). It can also result from the gain of binding sites for repressors within an enhancer (Galant and Carroll 2002; Preger-Ben Noon *et al.* 2016). Genetic changes that create new enhancers are less-well understood.

Copyright © 2019 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.119.301985>

Manuscript received January 29, 2019; accepted for publication February 25, 2019; published Early Online March 6, 2019.

Supplemental material available at <https://doi.org/10.25386/genetics.7140425>.

<sup>1</sup>Present address: Department of Evolution and Ecology, University of California, Davis, Davis, CA 95616.

<sup>2</sup>Present address: Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717.

<sup>3</sup>Corresponding author: University of Michigan, 4010 Biological Sciences Bldg., 1105 North University Ave., Ann Arbor, MI 48109. E-mail: [wittkopp@umich.edu](mailto:wittkopp@umich.edu)

They can evolve *de novo* from sequences that did not drive any enhancer activity previously (Eichenlaub and Ettwiller 2011; Emera *et al.* 2016) or can evolve from sequence modifications within an existing enhancer. Examples of this type of co-option have shown that sequence modifications can add activating elements (Rebeiz *et al.* 2011), eliminate repressing elements (Prabhakar *et al.* 2008; Sumiyama and Saitou 2011), or add both activating and repressing elements (Gompel *et al.* 2005; Arnoult *et al.* 2013). Co-option of existing enhancers might be more common than their *de novo* evolution because, with co-option new activities can arise from only a small number of mutations (Gompel *et al.* 2005; Rebeiz *et al.* 2011; Arnoult *et al.* 2013; Koshikawa *et al.* 2015).

The *Drosophila yellow* gene, which encodes a protein required for the production of black pigment, has emerged as a model for studying the evolution of enhancer sequences. *yellow* expression is divergent among species, with changes in *yellow* expression evolving in concert with changes in the distribution of black melanin affecting adult body color (Wittkopp *et al.* 2002), and well-characterized enhancers control its expression. In *Drosophila melanogaster*, enhancers controlling *yellow* expression during the pupal stages when adult pigmentation develops have been identified in the 5' intergenic sequence upstream of *yellow* that drive expression in the developing wings and body (head, thorax, and abdomen). An enhancer in the lone intron of *yellow* has been shown to drive *yellow* expression in bristles (Geyer and Corces 1987; Martin *et al.* 1989; Wittkopp *et al.* 2002; Jeong *et al.* 2006). Finally, a sequence required for sexually dimorphic expression in the abdomen has been identified within the "body enhancer" that contains binding sites for the Abdominal-B (Jeong *et al.* 2006) and Bric-a-brac (Roeske *et al.* 2018) transcription factors. Evolutionary changes in the body and wing enhancers have been identified that alter *yellow* expression in a manner that correlates with divergent pigmentation (Gompel *et al.* 2005; Jeong *et al.* 2006; Prud'homme *et al.* 2006; Kalay and Wittkopp 2010; Arnoult *et al.* 2013).

In addition to changes in individual enhancers, larger-scale reorganization of enhancers within *yellow cis*-regulatory sequences have also been described, with wing and body enhancer activities found in the 5' intergenic region of *D. melanogaster* also found in the introns of other *Drosophila* species (Kalay and Wittkopp 2010). This rapid evolution of the genomic organization of *yellow* enhancer activities was surprising given the collinearity of enhancers controlling expression of other genes among *Drosophila* species (e.g., Cande *et al.* 2009; Hare *et al.* 2008). DNA sequence analysis of *yellow cis*-regulatory regions suggested that enhancer synteny changes were due to gradual gain and loss of transcription factor-binding sites, rather than duplication or translocation events. Using parsimony, Kalay and Wittkopp (2010) inferred that the common ancestor of the contemporary *Drosophila* species examined most likely had wing and body enhancer activities in both the 5' intergenic and intronic regions of *yellow*.

Here, we investigate evolutionary changes in the organization of enhancer activities within *yellow cis*-regulatory sequences by using reporter genes to test the function of subfragments from the 5' intergenic and intronic sequences of *yellow* from *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni*. For each species, we compare enhancer activity of these subfragments to enhancer activity of the full region to determine which sequences are responsible for driving expression in specific parts of the fly during pupal development. We find that in all three species, redundant enhancers are common, with multiple fragments driving overlapping expression in the body and wings of developing pupae. We also find that some enhancer fragments drive cryptic expression patterns not seen in reporter genes containing larger *cis*-regulatory sequences from that species. Interestingly, these cryptic patterns are often similar to expression patterns seen in other species, suggesting that repressed enhancer activities might have contributed to the evolution of new enhancer activities. These data show that the architecture of *yellow cis*-regulatory sequences is less modular and more variable among species than previously described, with the observed organization of enhancer activities potentially making *yellow* expression more robust, as well as more amenable to evolutionary change.

## Materials and Methods

### Constructing reporter genes and transgenic lines

Putatively orthologous 5' intergenic and intronic regions of *yellow* from *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* described in Kalay and Wittkopp (2010) were subdivided using PCR into ~1000-bp fragments (except for mel\_A5, pse\_B6, and will\_C7, which were 423-, 641-, and 345-bp long, respectively). Each fragment overlapped the flanking fragments by ~100 bp. PCR was conducted using a mix of Taq DNA polymerase and Phusion High-Fidelity DNA Polymerase [New England Biolabs (NEB), Beverly, MA] to prevent PCR-introduced mutations. Recognition sites for the *AscI* (NEB) restriction enzyme were introduced to the ends of each PCR product using primers with 5' *AscI* tails. Supplemental Material, Table S1 contains a list of primers used to amplify each *yellow* 5' intergenic or intronic subfragment from the three species in this study. Also included in Table S1 are the exact amplicon length, length of overlap with preceding and following fragments, and coordinates of each fragment tested.

PCR products for *yellow* enhancer fragments were then subcloned into the sequencing vector pGEM-T (Promega, Madison, WI), and sequenced using M13 Forward and M13 Reverse primers. File S1 contains sequences of all fragments, with each sequence beginning and ending with the recognition sequence for the *AscI* restriction enzyme (GGCGCGCC) used for cloning. Sequence-confirmed *yellow* enhancer fragments were then subcloned into a piggyBac-attB vector [as described in Kalay and Wittkopp (2010)] using the *AscI* unique site. Next, we cloned the coding sequence for a

nuclear Enhanced Green Fluorescent Protein (nEGFP) downstream (3') of each enhancer fragment using the *FseI* (NEB) unique restriction site. Multiple attempts to construct this reporter gene for *pse\_B4* failed with these methods, thus we employed the GeneArt method (Thermo Fisher Scientific) to produce this construct. In all cases, the resulting construct was confirmed with diagnostic restriction digests, prepared at high concentration using a Zyppy Plasmid Maxi kit, reconfirmed with diagnostic digest, and sent to either Genetic Services (Cambridge, MA) or Genetivision (Houston, TX) for injection into the attP-40 line of *D. melanogaster*. This is the same attP site used to insert the reporter genes containing the full 5' intergenic and intronic regions from each species, as described in Kalay and Wittkopp (2010).

### Documenting reporter gene expression

*D. melanogaster* lines homozygous for the reporter gene were constructed as described in Kalay and Wittkopp (2010). Briefly, for each transgenic line, transformant flies were crossed to a second chromosome balancer line marked by the “curly wings” phenotype (Bloomington Stock Center #7197). Among the F<sub>1</sub> progeny, flies with green fluorescing eyes from the Pax6-GFP transformation marker and curly wings were crossed to each other. Among the F<sub>2</sub> progeny, flies that had green fluorescing eyes and flat wings were assumed to be homozygous for the transgene, and were crossed to each other to create a stable homozygous line for the transgene. For six of the *yellow* enhancer fragments—*mel\_A5*, *pse\_B4*, *pse\_B6*, *pse\_B7*, *pse\_B9*, and *will\_C8*—multiple crosses failed to produce flies with straight wings (*i.e.*, individuals homozygous for the transgene). Therefore, for those lines, we analyzed flies that were hemizygous for the transgene (*i.e.*, green fluorescing eyes with curly wings). Pupal bodies and wings from each line were prepared for microscopy ~80-hr after pupa formation as described in Kalay and Wittkopp (2010), and imaged immediately using a Leica SP5 confocal microscope. Images of the same tissue (body or wing) were processed identically for all lines in Adobe Photoshop CS6. All fragments except *mel\_A5*, *mel\_A7*, *pse\_B6*, and *will\_C8* showed expression in at least one of the structures scored under these conditions. However, only one of these fragments, *mel\_A7*, was assayed in flies homozygous for the reporter gene; heterozygosity of reporter genes driven by the other three fragments might have caused us to underestimate their activity.

### Principal component analysis

We created an appearance model to evaluate the main trends of spatial variation in gene expression along the abdomen in five male and five female flies from each of the transgenic lines, except *pse\_B4* and the full *D. willistoni* intron (Full\_will\_int). For *pse\_B4*, four images of females and three images of males were used, and for *will\_int*, only a single image from each sex was used because strains carrying these two reporter genes went extinct during the course of this project. In all cases, we analyzed downsized 72-ppi images in jpg format. We created

a 30-point model template using the AAMToolbox in the MATLAB environment (Whibley *et al.* 2006). The point model was designed to capture the abdominal segments A2–A6. We then performed Procrustes superimposition by modifying rotation, translation, and size, so that all of the images were fitted into the mean abdomen shape, encapsulating 6000 pixels. This procedure allowed us to compare the patterns of RGB pixel values after removing size variation and/or subtle variation in image orientation. Finally, the warped abdominal images were used to perform principal component (PC) analysis, extracting the main trends of variation in RGB values for each of the 6000 pixels, giving an appearance space, and calculating PC values for each individual fly, using the “Shape Model Toolbox Software” as described in Whibley *et al.* (2006). Table S2 contains the values of PC1, PC2, and PC3 for each of the reporter genes, as well as negative control flies carrying a reporter gene without any enhancer sequences.

### Sequence analysis

The sequences of the *yellow* 5' intergenic and intronic regions from *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* were compared using the *promoterwise* program from the wise2.4 package (Ettwiller *et al.* 2005). This program allows alignments of sequences that are not necessarily colinear, allowing for inversions and translocations, in addition to insertions and deletions. A cut-off of at least 20 bits was used because bit scores of > 20 bits are extremely rare when comparing random sequences (Ettwiller *et al.* 2005). A custom R script was used to identify regions with significant sequence similarity on the schematic of *yellow* sequences shown in Figure 7. File S2 contains *promoterwise* output showing all significant local alignments between each pair of sequences as well as a custom R script used to process the *promoterwise* results from all pairwise comparisons.

To test for enrichment of potential transcription factor-binding site motifs among fragments with similar enhancer activity, we used the Analysis of Motif Enrichment tool (McLay and Bailey 2010) with the “Combined *Drosophila* Databases” from the MEME suite (meme-suite.org). In all, we tested 1419 motifs, ranging from 4 to 26 nucleotides long, which describe binding sites for ~60% of all *D. melanogaster* transcription factors (<http://meme-suite.org/db/motifs>, July 2018), for enrichment among sequences driving similar expression patterns. For sexually dimorphic expression, the 13 fragments shown as driving sexually dimorphic expression in Figure 7 were compared to the 16 subfragments from the 5' intergenic and intronic regions of *yellow* that did not. For bristle expression, sequences from the putatively orthologous *mel\_A6*, *pse\_B7*, and *will\_C9* fragments were compared to all other intronic fragments except *pse\_B10*. *pse\_B10* was excluded from the control set because it drove expression in wing margin bristles. For the flight muscle attachment sites, the putatively orthologous *mel\_A3*, *pse\_B4*, and *will\_C3* fragments were compared to all other 5' intergenic fragments except *mel\_B3*, *mel\_B5*, and *will\_C1*. This three fragments

were excluded from the control set because they also drove expression in this region of the thorax.

### Data availability

*Drosophila* strains are available upon request. File S1 contains sequences of the regions tested for enhancer activity. File S2 contains detailed sequence alignments used to produce the summary in Figure 7. Table S1 contains primer sequences used in this work. Table S2 contains detailed results from the PC analysis described in the manuscript. Supplemental material available at <https://doi.org/10.25386/genetics.7140425>.

## Results

### Dissecting the architecture of *Drosophila yellow* cis-regulatory sequences

*D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* are members of the *Sophophora* subgenus of *Drosophila* that have evolved distinct pigmentation since they diverged between 25 and 36 MYA (Figure 1A, Russo *et al.* 1995). Specifically, *D. pseudoobscura* has evolved an overall dark pigmentation on both its abdomen and thorax, whereas *D. melanogaster* and *D. willistoni* show much more limited thoracic pigmentation, and dark stripes near the posterior edge of each abdominal segment. These abdominal stripes are more lightly pigmented in *D. willistoni* than *D. melanogaster*. *D. melanogaster* males also show sexually dimorphic dark pigment throughout the A5 and A6 abdominal segments that is absent in females, whereas neither *D. willistoni* nor *D. pseudoobscura* are considered to have sexually dimorphic pigmentation (Camino *et al.* 2015). The A6 segment of *D. melanogaster* females shows broad pigmentation in some strains (Kopp *et al.* 2003). None of these three species display the dark melanotic spots of pigment found on the wings of some other *Drosophila* species (e.g., Gompel *et al.* 2005; Prud'homme *et al.* 2006; Werner *et al.* 2010; Koshikawa *et al.* 2015), but they do have dark pigment distributed evenly throughout the wing blades and wing veins. Finally, bristles throughout the body and along the wing margins are darkly pigmented in all three species. The 5' intergenic and intronic sequences of *yellow* from all three of these species were previously shown to drive expression of a reporter gene integrated into the *D. melanogaster* genome in pupae that correlated well with these adult pigment patterns (Kalay and Wittkopp 2010).

To determine how enhancer activities are distributed within the previously tested 5' intergenic and intronic regions of *yellow*, we subdivided each full region into ~1-kb fragments with adjacent fragments overlapping by ~100 bp (Figure 1B and File S1). Each of these fragments was cloned upstream of a minimal hsp70 promoter and the coding sequence for an nEGFP, and transformed into the same site of the *D. melanogaster* genome using the attB/attP system for targeted insertion (Bischof *et al.* 2007; Markstein *et al.* 2008). Enhancer activity was assessed using confocal microscopy in flies ~80-hr after puparium formation. We analyzed this

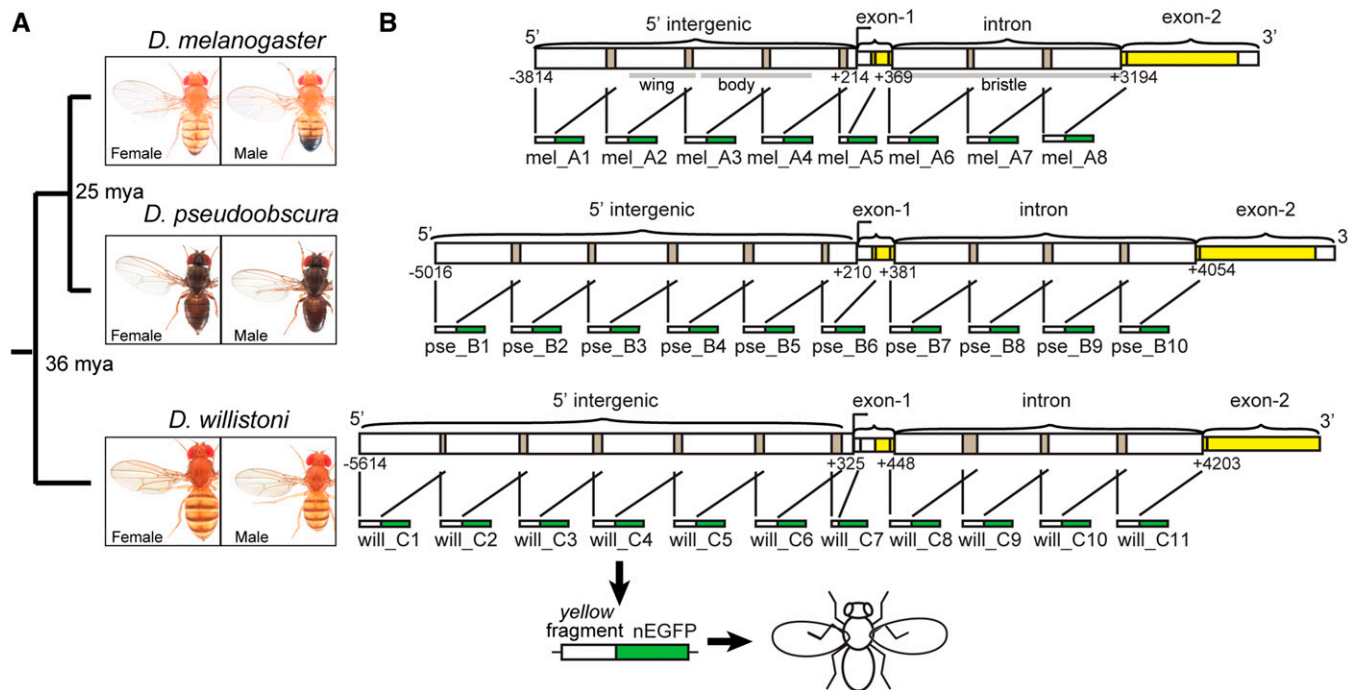
stage because it is when *yellow* expression is required for the development of adult pigmentation (Massey and Wittkopp 2016); additional expression might be driven by these enhancers at other developmental stages.

Using these confocal images of pupal bodies, we characterized enhancer activity in the epidermal and bristle cells of the head, thorax, and abdomen, recording whether or not the expression pattern among replicate flies consistently showed: (1) stripes of expression in abdominal segments, (2) broad expression within abdominal segments, (3) sexually dimorphic expression in abdominal segments A5 and A6, (4) expression marking the circumference of the head, (5) expression on the top of the head, (6) expression in the thoracic trident pattern, (7) expression in the thoracic flight muscle attachment sites (Fernandes *et al.* 1996), (8) broad thoracic expression, and/or (9) expression in bristles on the body (Figure 2A). In the wings, we recorded whether there was expression in: the (1) epidermal cells of the wing blade, (2) longitudinal and cross veins, and/or (3) bristles along the anterior margin (Figure 2A). Many of these elements are visible in the expression pattern driven by the full 5' intergenic sequence from *D. melanogaster* (Figure 2B). We inferred enhancer activity when we consistently saw higher expression driven by a reporter gene in a tissue than that driven by a control reporter gene lacking any *yellow* fragment (Figure S1A). Fluorescence in the control line was highest and most variable in the wing blade (Figure 1B), making it difficult to confidently infer enhancer activities that drove low expression in this part of the fly. In all cases where expression elements were difficult to interpret, we report the opinion of the majority of coauthors and note this ambiguity in the associated figure legend. All transformant flies showed high expression levels of green fluorescent protein (GFP) in the eyes and ocelli driven by the *Pax6-GFP* transformation marker (Figure S1A; Horn and Wimmer 2000).

For 29 of the 35 transformant lines, we imaged individuals homozygous for the transgene. For the six remaining lines (mel\_A5, pse\_B4, pse\_B6, pse\_B7, pse\_B9, and will\_C8, indicated with asterisks in Figure 3, Figure 4, and Figure 5), we were unable to recover homozygous flies and imaged heterozygous individuals instead. These heterozygous flies showed weaker expression of GFP in the eyes and ocelli, suggesting that our images for these lines underestimated expression relative to homozygous individuals in other tissues. Some variability in fluorescence levels among images was also expected to result from fluctuations in confocal laser intensity, despite the fact that we used the same laser and imaging settings each day. Expression of the green fluorescent reporter gene is shown using an inverted color scheme in Figure 3, Figure 4, and Figure 5 to make it easier to see low levels of expression. The fluorescent images, which show differences in regions of high expression more clearly, are also provided as Figures S2–S4.

In the following three sections, we describe expression in the head, thorax, abdomen, and wing, driven by sequences from the *yellow* 5' intergenic and intronic regions, in detail for





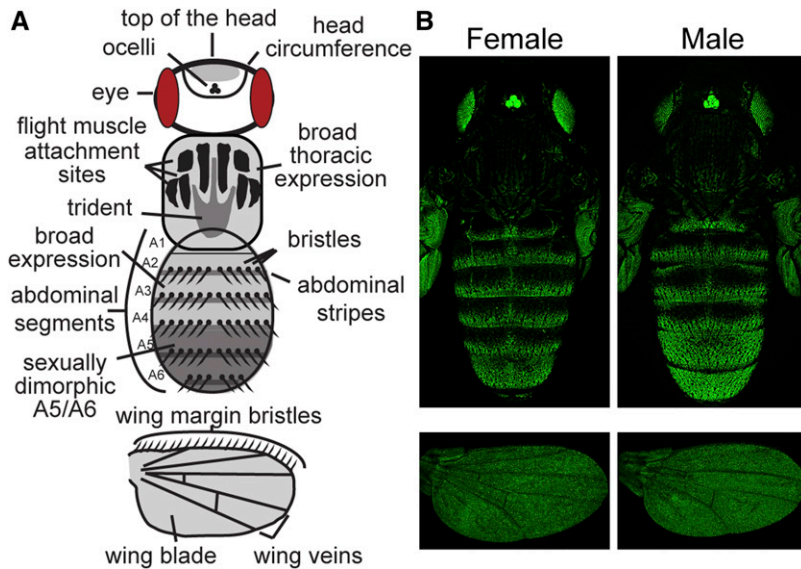
**Figure 1** Fragments of *yellow* 5' intergenic and intronic sequences tested for enhancer activity. (A) Divergence times among *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* are shown, along with images of adult males and females from each species. (B) Schematics show the overlapping fragments from regions of 5' intergenic and intronic sequence of *yellow* from each species tested for enhancer activity. Each of these ~1 kb fragments was cloned upstream of a nuclear Enhanced Green Fluorescent Protein (nEGFP) to form a reporter gene, and each reporter gene was inserted into the attP40 landing site on chromosome arm 2L. Gray bars under the *D. melanogaster* schematic show the borders of previously identified tissue-specific enhancers of *yellow* driving expression in the wing, body, and bristles (Geyer and Corces 1987; Wittkopp *et al.* 2002; Kalay and Wittkopp 2010). This schematic was modified from Kalay *et al.* (2016). Pictures of *Drosophila* species generously provided by Nicolas Gompel (Ludwig-Maximilians-University of Munich).

each species. We then present results from PC analysis used to objectively identify elements of the abdominal expression patterns likely to be controlled independently during development. Finally, we describe regions of sequence similarity among *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni*, and compare these regions of sequence similarity to enhancer activity, revealing how these enhancer activities might have evolved. Lessons learned about the modularity, redundancy, cryptic nature, and evolution of these enhancers from the syntheses of these data are then presented in the *Discussion*.

### Pupal enhancer activities of *D. melanogaster yellow*

The 4-kb intergenic sequence located 5' of the *D. melanogaster yellow* gene drove expression in the thoracic trident and abdominal stripes (Figure 3A, Kalay and Wittkopp 2010). Sexually dimorphic expression was also observed in the abdomen, with expression expanded and elevated in the A5 and A6 segments of males relative to females. Females also showed broader expression in the A6 segment than segments A1–A5. Among the subfragments from the 5' intergenic sequence tested (Figure 3A), the mel\_A2 fragment drove stripes of expression in the abdomen most similar to the full 5' intergenic sequence, but broad expression was limited only to the A6 segment of males. Sexually dimorphic expression in abdominal segments A5 and A6 was most similar to

that observed with the full enhancer for fragments mel\_A1 and mel\_A4, but these fragments also drove broad expression throughout segments A2–A4. The mel\_A2 and mel\_A6 fragments also drove higher levels of expression throughout segments A5 and/or A6 in males than females. In the thorax, a trident pattern of expression similar to that driven by the full 5' intergenic sequence was observed for mel\_A1, mel\_A2, and mel\_A4. For mel\_A3, strong expression was driven throughout the thorax. In the head, expression in the circumference was most prominent for the mel\_A2 and mel\_A3 fragments, with lower levels of expression also observed on the top of the head for mel\_A2, mel\_A3, and mel\_A4 (Figure 3A). In the wings (Figure 3B), the full 5' intergenic sequence drove expression in the epidermal cells of the wing blade. Expression in the wing blade was also driven by mel\_A1, mel\_A2, and mel\_A3, with the highest expression driven by mel\_A2. Expression in the wing blade was consistently elevated in the region surrounding and posterior to the L5 wing vein for the mel\_A1 fragment (arrows, Figure 3B). Finally, one fragment, mel\_A4, also drove expression at low levels in the wing veins. Fragment mel\_A4 might also drive low levels of expression in the wing (Figure 3B), but this enhancer activity was not called because some control flies showed similar levels of fluorescence (Figure S1B). In all of the tissues examined, the mel\_A5 fragment, which was the shortest, was



**Figure 2** Regions of the pupal body scored for *yellow* enhancer activity. (A) This schematic represents a *Drosophila* pupa and has the regions scored for reporter gene expression on the head, thorax, abdomen, and wings indicated. Abdominal segments A1–A6 are also labeled. Eyes and ocelli are also shown because the transformation marker (Pax6-GFP) inserted with the *yellow* fragment reporter genes drives expression in these tissues. (B) Activity of the reporter gene driven by the full 5' intergenic sequence from *D. melanogaster yellow*. The fluorescence observed in the eyes and ocelli is driven by the Pax6-GFP transformation marker, which was integrated with the empty reporter gene (Figure S1B). These are the same images shown for the full *mel\_5'* fragment in Figure 3, A and B without the inverted color scheme.

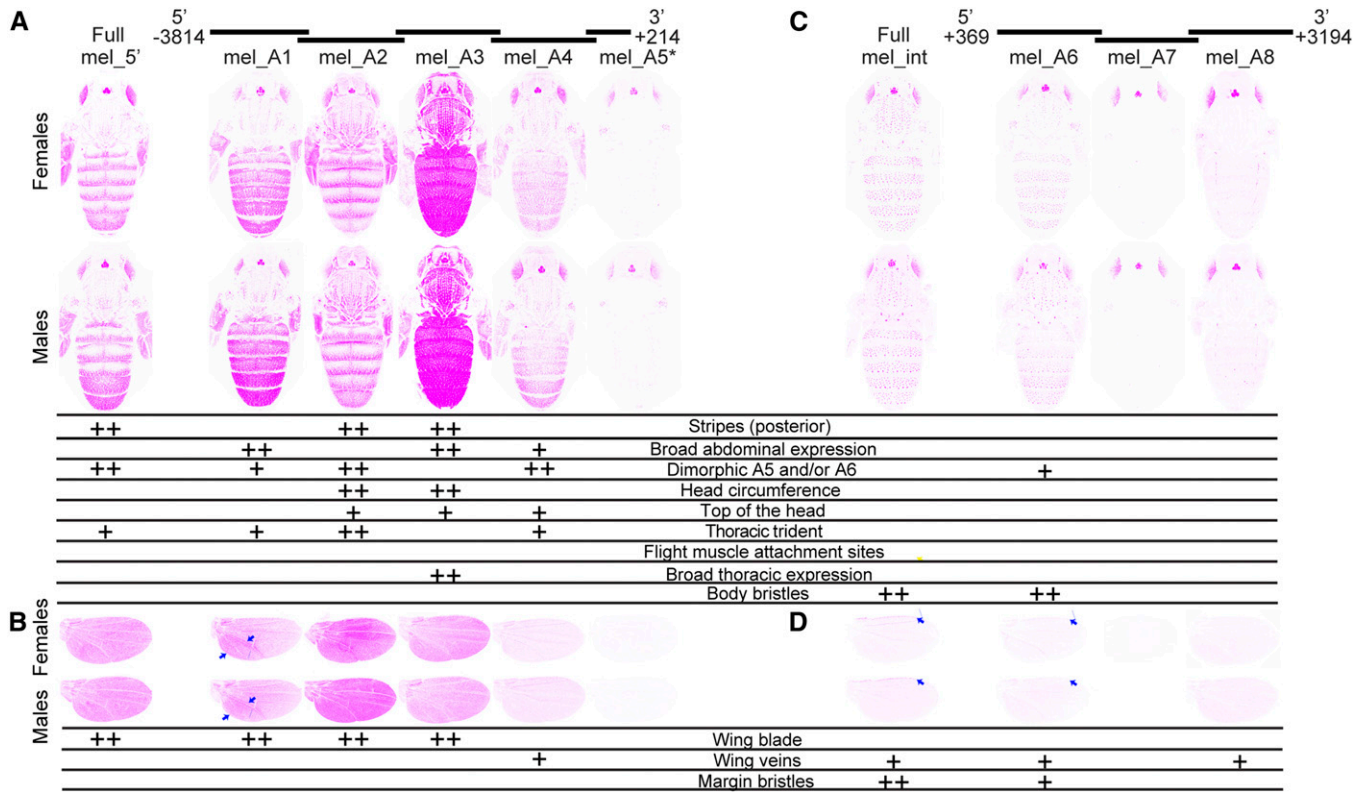
assessed in flies heterozygous for the transgene, and closest to the transcription start site of *yellow*, failed to drive expression of the reporter gene above background levels.

The 2.8-kb intronic sequence of the *D. melanogaster yellow* gene drove expression in the bristles of the body, but not epidermal cells (Figure 3C, Kalay and Wittkopp 2010). Among the intronic fragments (Figure 3C), sequences included in the *mel\_A6* fragment were sufficient to drive expression in this pattern, although expression seemed to be lower (especially in A2 and A3). The *mel\_A6* fragment also appeared to drive low levels of sexually dimorphic expression in A5 and A6 in males. Neither the *mel\_A7*, nor *mel\_A8* fragments drove expression in any of the body regions scored. In the wings (Figure 3D), the full intronic sequence drove low levels of expression in the veins, with higher levels of expression in the margin bristles (arrows, Figure 3D). Fragments *mel\_A6* and *mel\_A8* also drove low levels of expression in the wing veins, with *mel\_A6* driving expression in the margin bristles as well (arrows, Figure 3D). The other intronic fragment, *mel\_A7*, did not drive any detectable expression in the wing.

#### **Pupal enhancer activities of *D. pseudoobscura yellow***

The 5.2-kb intergenic sequence located 5' of the *D. pseudoobscura yellow* gene drove expression throughout the developing head, thorax, and abdomen (Figure 4A, Kalay and Wittkopp 2010). Although pigmentation of this species is generally considered to be sexually monomorphic (Kopp *et al.* 2000; Jeong *et al.* 2006; Salomone *et al.* 2013; Camino *et al.* 2015), expression was elevated in the A5 and A6 abdominal segments relative to A2–A4 in males, and decreased in A6 relative to A2–A5 in females. Among the subfragments from the 5' intergenic sequence tested (Figure 4A), *pse\_B1*, *pse\_B3*, and *pse\_B5* all drove broad expression in the body similar to the full 5' intergenic region; however, of these three fragments, only *pse\_B1* drove sexually dimorphic

expression similar to the full 5' intergenic sequence. Sexually dimorphic expression in abdominal segments A5 and A6 was also driven by *pse\_B2*, but without the accompanying broader abdominal expression. This fragment also drove expression at the lateral edges of segment A4 in males, and at the posterior half of segments A5 and A6 in females. The *pse\_B3* fragment also drove elevated expression along the posterior edge of each abdominal segment near the midline in females (arrows, Figure 4A), but we did not score this expression as “stripes” because it was not throughout the segment nor present in males. In addition, the *pse\_B3* fragment drove a lower level of expression in the A6 abdominal segments of both males and females; expression was also reduced to a lesser extent in the A5 abdominal segment of males. In the thorax, the full 5' intergenic sequence drove low levels of expression in the trident as well as higher levels of expression along the lateral edges of the thorax. Among the intergenic subfragments, the *pse\_B3* fragment drove expression in the most similar pattern, albeit with lower expression in the middle of the thorax and with expression also in flight muscle attachment sites. The *pse\_B1* fragment also consistently drove expression along the lateral edges of the thorax, but expression in the trident pattern was much more variable among replicate flies. Stronger expression was driven by *pse\_B4* and *pse\_B5* broadly throughout the thorax. Although the only detectable expression driven by *pse\_B4* was in the thorax, it might also drive low levels of expression in the abdomen that were missed by assessing activity of this fragment only in a heterozygous state. In the head, *pse\_B5* drove expression in the head circumference, whereas *pse\_B1* and *pse\_B3* both drove expression on the top of the head more similar to the full 5' intergenic sequence. In the wings (Figure 4B), the full 5' intergenic fragment drove low levels of epidermal cell expression throughout the wing blade without any expression in the wing veins or margin bristles. Similar epidermal expression throughout the wing blade was



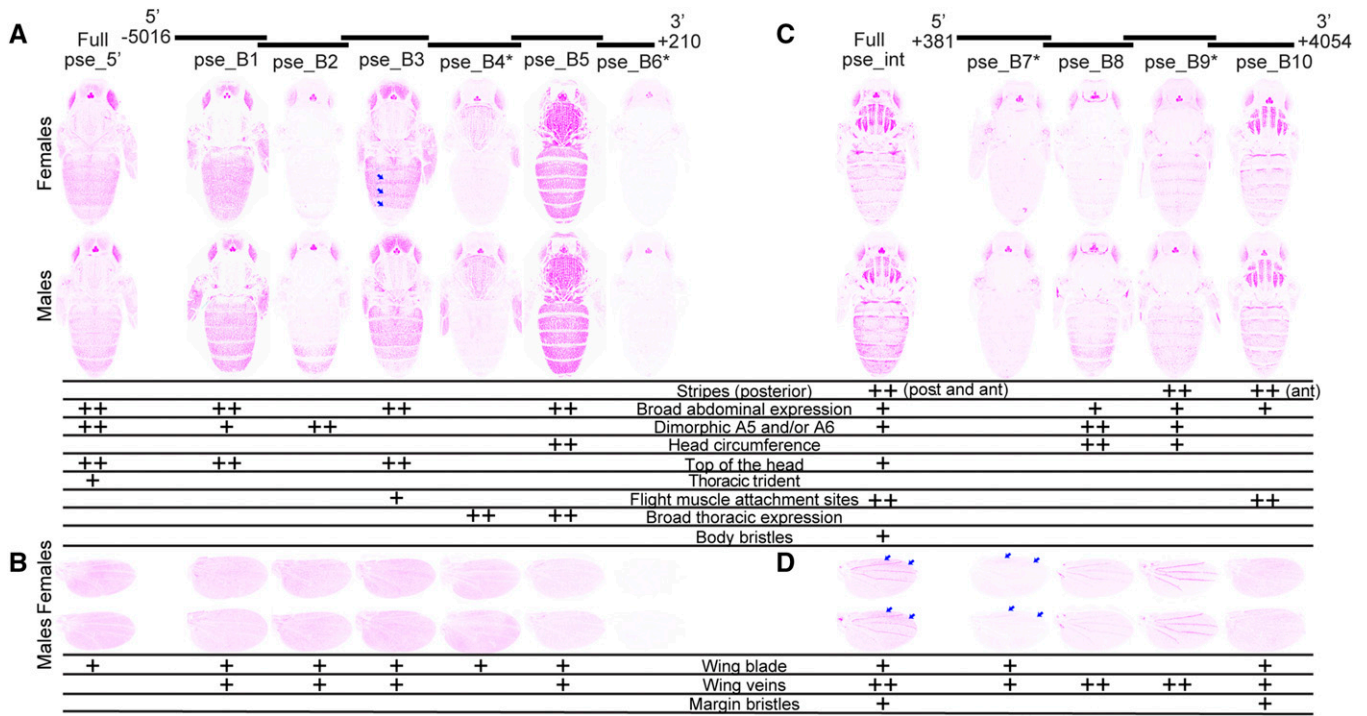
**Figure 3** Enhancer activity of *D. melanogaster yellow* fragments. Below the schematic of overlapping sequence regions from the 5' intergenic and intronic regions of the *D. melanogaster yellow* gene tested for enhancer activity are images of transgenic pupae, which show expression from the GFP reporter genes in magenta. Expression driven by fragments from the 5' intergenic region (A and B) and the intron (C and D) is shown in the body (A and C) and wings (B and D), for both female (top row) and male (bottom row) flies. Summary tables below each pair of pupal bodies (A and C) and wings (B and D) show our interpretation of these images: “++” indicates strong fluorescence observed in the body region, whereas “+” indicates weaker fluorescence. Blue arrows highlight elevated expression in a region of the wing driven by the mel\_A1 fragment, and expression in the margin bristles driven by the full *D. melanogaster* intron (“Full mel\_int”) and the mel\_A6 fragment. The asterisk next to the mel\_A5 fragment indicates that activity of this element is shown for flies heterozygous for the reporter gene. All other images show GFP expression in flies homozygous for the reporter gene. The magenta color used in this figure makes it easier to see low expression levels; a copy of this figure with GFP expression shown in the more traditional green is provided as Figure S2. The dimorphic expression driven by mel\_A6 and the wing blade expression driven by mel\_A4 were equivocal, and went through a secondary inspection before reaching a consensus among authors.

driven by pse\_B1, pse\_B2, pse\_B3, pse\_B4, and pse\_B5. All of these fragments except pse\_B4 also drove ectopic expression in the veins. The pse\_B6 fragment, which was the shortest, was assessed in flies heterozygous for the transgene, and was closest to the transcription start site of *yellow*, did not drive expression in any of the pupal structures scored (Figure 4, A and B).

The 3.7-kb intronic sequence of the *D. pseudoobscura yellow* gene drove expression in the flight muscle attachment sites of the thorax, on top of the head, and throughout each abdominal segment, with expression elevated at the anterior and posterior edges of each abdominal segment (Figure 4C, Kalay and Wittkopp 2010). Broad, male-specific expression in abdominal segments A5 and A6, as well as along the lateral edges of segments A2–A4, was also observed. Expression in bristles on the body was driven by the full intron too, but this expression was much more subtle than the bristle expression driven by the *D. melanogaster* intron (Figure 3C). Among the intronic subfragments (Figure 4C), pse\_B8, pse\_B9, and pse\_B10 all drove subsets of the abdominal expression pattern

driven by the full intron. Specifically, pse\_B8 drove broad expression elevated along the lateral edges of A2–A4, and throughout segments A5 and A6 of males. The pse\_B9 fragment drove this sexually dimorphic expression pattern as well, although this driving was weaker, and also drove elevated expression at the posterior edge of each abdominal segment. The pse\_B10 fragment drove expression throughout the abdomen, with expression elevated in stripes along the anterior edge of each abdominal segment. None of the intronic fragments drove expression in the body bristles similar to that driven by the full intron. In contrast, the thoracic expression driven by the full intron was mimicked by pse\_B10, with strong expression in the flight muscle attachment sites. Expression on the top of the head driven by the full intron was not consistently reproduced by any of the subfragments; however, fragments pse\_B8 and pse\_B9 drove strong expression in the head circumference. Low levels of expression driven by pse\_B7 may have been missed because this reporter gene was heterozygous in the flies assayed. In the wings, the full intronic fragment drove expression in all of





**Figure 4** Enhancer activity of *D. pseudoobscura yellow* fragments. Below the schematic of overlapping sequence regions from the 5' intergenic and intronic regions of the *D. pseudoobscura yellow* gene tested for enhancer activity are images of transgenic pupae that show expression from the GFP reporter genes in magenta. Expression driven by fragments from the 5' intergenic region (A and B) and the intron (C and D) is shown in the body (A and C) and wings (B and D), for both female (top row) and male (bottom row) flies. Summary tables below each pair of pupal bodies (A and C) and wings (B and D) show our interpretation of these images: “++” indicates strong fluorescence observed in the body region, whereas “+” indicates weaker fluorescence. Note that in addition to the stripes seen at the posterior edge (post) of abdominal segments in many cases, the full *D. pseudoobscura* intron (“Full pse\_int”) as well as fragment pse\_B10 drove elevated expression in stripes at the anterior edge of each abdominal segment (indicated by the “ant” notation in the table). Blue arrows highlight elevated expression along the posterior edge of abdominal segments in females driven by the pse\_B3 fragment, and between the L1 and L2, as well as L2 and L3, veins in the wing blade driven by full pse\_int and the pse\_B7 fragment. The asterisks next to the pse\_B4, pse\_B6, pse\_B7, and pse\_B9 fragments indicate that activities of these elements are shown for flies heterozygous for the reporter gene. All other images show GFP expression in flies homozygous for the reporter gene. The magenta color used in this figure makes it easier to see low expression levels; a copy of this figure with GFP expression shown in the more traditional green is provided as Figure S3. The wing blade expression driven by pse\_B5 was equivocal and went through a secondary inspection before reaching a consensus among authors.

the structures scored (blade, veins, and bristles), with elevated expression in the wing blade between the L1 and L2, as well as L2 and L3, veins (arrows, Figure 4D). As in the body, expression in bristles along the wing margin was more subtle than that driven by the *D. melanogaster* intron. Among the subfragments (Figure 4D), pse\_B7 drove expression between the L1 and L2, as well as L2 and L3, veins similar to the regions of highest expression driven by the full intron (arrows, Figure 4D), and pse\_B10 drove low levels of expression in the wing blade (Figure 4D). Expression in the veins was driven by all four intronic subfragments, with the highest expression driven by the pse\_B8 and pse\_B9 fragments. Finally, expression in the wing margin bristles was driven by the pse\_B10 fragment.

**Pupal enhancer activities of *D. willistoni yellow***

The 5.9-kb *D. willistoni yellow* 5' intergenic region drove expression in the abdomen highest in stripes along the posterior edge of segments A2–A5 in both males and females, with expression reduced in A5 and undetectable in A6 (Figure 5A, Kalay and Wittkopp 2010). Among the subfragments

from the 5' intergenic sequence (Figure 5A), will\_C1 and will\_C3 drove stripes of expression elevated at the posterior edge of each segment. The will\_C4 fragment also drove elevated expression in stripes at the posterior edge of most segments, with some elevated expression also seen near the anterior edges of some segments. Broad expression throughout abdomen segments was driven by will\_C1, will\_C2, and will\_C4, as well as (to a lesser extent) will\_C6 and will\_C7. Surprisingly, sexually dimorphic patterns of elevated expression in the A5 and/or A6 of males relative to females were driven by will\_C1, will\_C2, will\_C3, and will\_C4, despite the absence of sexually dimorphic pigmentation or expression driven by the full 5' intergenic sequence. In the thorax, the full 5' intergenic sequence drove expression broadly throughout the thorax. Similar patterns of expression were driven at higher levels by the will\_C1 and will\_C3 fragments, with will\_C4 driving lower expression detected only in the trident. No expression in the head appeared to be driven by the full 5' intergenic fragment, yet will\_C2 drove strong expression on the top of the head, will\_C4 drove some low expression in the head circumference, and will\_C1 drove expression in both of



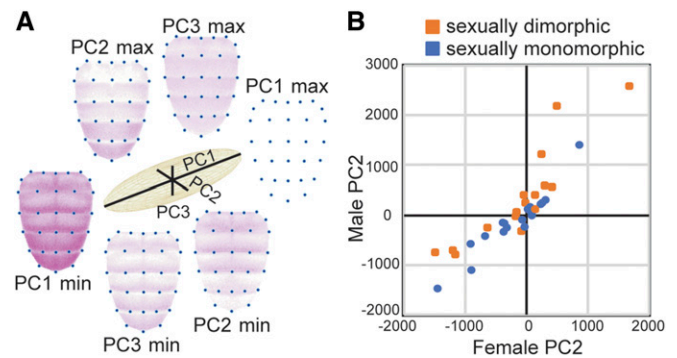


30 predefined landmarks in the abdomen (blue dots in Figure 6A), used these landmarks to align the images, and extracted comparable fluorescence information from each of the images. PC analysis was then run on these fluorescence data and we analyzed the first three PCs (Table S2), which together explained 55% of the variation.

The first PC (PC1) explained 34% of the variation in expression among lines and primarily captured the overall fluorescence intensity (Figure 6A). Variability in this trait likely reflects differences in the average expression level driven by different fragments, as well as some day-to-day variation in confocal conditions and the difference-in-ploidy for the six heterozygous lines. The second PC (PC2), an orthogonal axis to PC1, explained 12% of the variance, and primarily captured the intensity of fluorescence in abdominal segments A5 and A6 relative to A2–A4 (Figure 6A). Consistent with this interpretation, fragments classified as driving sexually dimorphic expression tended to have greater values of PC2 in males than females relative to fragments driving expression patterns not described as sexually dimorphic (Figure 6B, one-sided Student's *t*-test,  $P = 0.0009$ ). We note two exceptions to this relationship. First, the *mel\_A3* fragment drove high expression in abdominal segments A5 and A6 in both sexes, and was classified as sexually monomorphic (Figure 3A and Figure S2), but PC2 was greater in males than females suggesting that this fragment drives subtle differences between the sexes. Second, expression driven by the *mel\_A6* fragment was scored as sexually dimorphic (Figure 3A and Figure S2), but PC2 was more similar in males and females than in other cases scored as dimorphic (Figure 3A and Figure S2). In addition to sexual dimorphism in A5 and A6, PC2 also captured the prominence of stripes along the posterior edge of abdominal segments A2–A4, presumably because most fragments driving elevated expression in A5 and A6 also drove broad expression in A2–A4 (Figure 3, Figure 4, and Figure 5). The third PC (PC3) explained 9% of the variance and seemed to capture whether expression was elevated at the posterior edge of each segment (Figure 6A). The independent variation of these three traits (overall expression level, expression in A5 and A6, and abdominal stripes) suggests that they are controlled by distinct developmental processes.

### Conservation and divergence of enhancer activities are not well reflected in the conservation and divergence of cis-regulatory sequences

To search for DNA sequences driving specific elements of the expression patterns shown in Figure 3, Figure 4, and Figure 5, we examined sequence similarity among the 5' intergenic and intronic regions from *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* (Figure 7). The program *promoterwise*, which was designed to search for sequences regulating transcription in orthologous noncoding sequences (Ettwiller *et al.* 2005), was used for this analysis. This program uses a local alignment method to identify regions of sequence with greater similarity than expected by chance, and is robust to

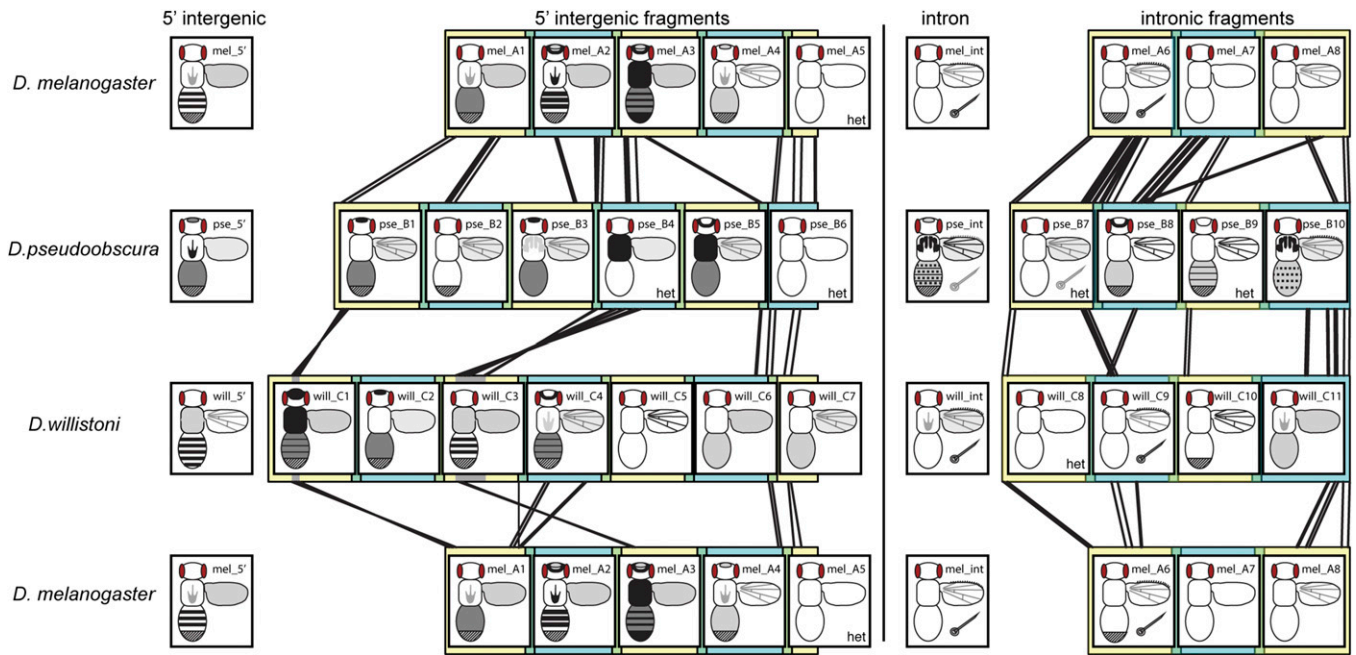


**Figure 6** Principal component analysis of abdominal expression patterns. (A) Relative proportions of variation in abdominal expression patterns explained by principal components 1, 2, and 3 (PC1, PC2, and PC3) are shown with black lines in the center of the figure. Abdominal images at the ends of each of these black lines show inferred expression patterns that correspond to the minimum (min) and maximum (max) values of each of the principal components. Comparing the images with min and max values for each PC shows the elements of abdominal expression captured by that principal component. (B) The scatterplot compares values of PC2, which primarily captures elevated expression in abdominal segments A5 and A6, between males and females for each reporter gene. PC2 values for reporter genes driving expression in abdominal segments A5 and A6 interpreted as sexually dimorphic are shown in orange, and PC2 values for reporter genes driving sexually monomorphic expression in abdominal segments A5 and A6 are shown in blue.

inversions and translocations, as well as insertions and deletions. Sequences found to be more similar between species than expected by chance (File S2) are connected with black lines in Figure 7.

As expected from the phylogenetic relationships among these three species (Figure 1A), the greatest sequence similarity was observed between *D. melanogaster* and *D. pseudoobscura*. We found no evidence of large translocations or rearrangements between any pair of species, consistent with (Kalay and Wittkopp 2010); the orthologous 5' intergenic and intronic sequences appeared to be syntenic (Figure 7). Despite this overall synteny, we observed what appear to be two small inversions (~20 and ~300 bp) that are specific to the *D. willistoni* lineage (gray regions in *will\_C1* and *will\_C3* fragments, Figure 7). Sequence similarity was detected between all pairs of species near the 5' and 3' ends of both the 5' intergenic and intronic fragments, consistent with the regions assayed being orthologous.

Comparing regions of sequence similarity between species to enhancer activities, we observed two cases where similar enhancer activities were driven by fragments with shared sequences. First, *mel\_A3*, *pse\_B4*, and *will\_C3* all drove expression broadly throughout the thorax, and contained regions of sequence similarity suggesting a conserved enhancer inherited from the common ancestor shared by all three species. The *pse\_B5* fragment also drove broad thoracic expression and showed sequence similarity with *mel\_A3*, but not *will\_C3*. Thoracic expression was driven by *will\_C1* as well, but this fragment did not show sequence similarity with any other fragment driving broad expression in the thorax,



**Figure 7** Synthesizing enhancer activity and sequence similarity for *yellow* fragments. Overlapping fragments from the 5' intergenic and intronic regions of *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* *yellow* are shown in alternating yellow and blue square blocks, with regions of overlap shown in green. Black lines between fragments from different species show regions of significant sequence similarity detected with the *promoterwise* program, as described in the *Materials and Methods* section. Gray regions in the will\_C1 and will\_C3 fragments indicate sequences that appear to be inverted in *D. willistoni* relative to the other two species. Enhancer activity observed from each of these fragments, as well as the full 5' intergenic and intronic sequences from each species, are represented in schematics similar to that shown in Figure 2A. Regions of the body shaded dark gray or black showed higher expression levels than regions of the body shaded light gray. The checkered region at the posterior end of the abdomen indicates sexually dimorphic expression with elevated expression in abdominal segments A5 and A6 of males. The posterior region is fully darkened in the schematic for fragment mel\_A3 because it drove elevated expression in abdominal segments A5 and A6 of both males and females. The six fragments assayed only in flies heterozygous for the reporter gene are marked with "het."

suggesting that this similar enhancer activity may have evolved independently.

Second, in the intron, the two fragments that drove high levels of expression in bristles on both the body and wing (mel\_A6 and will\_C9) showed sequence similarity, suggesting the presence of a conserved bristle enhancer inherited from a common ancestor. These two fragments also showed sequence similarity with the pse\_B7 fragment, but we did not score pse\_B7 as driving bristle expression in the body or wing (Figure 4C). We hypothesized that bristle expression driven by pse\_B7 might have been missed because of heterozygosity of the pse\_B7 reporter gene. Using digital image enhancements, we did indeed find evidence of bristle expression in the body driven by pse\_B7 (Figure S5), and thus conclude that mel\_A6, pse\_B7, and will\_C9 contain an orthologous bristle enhancer. Even with digital enhancements, bristle expression along the wing margin was detected only with the pse\_B10 fragment (Figure S5), which did not show any sequence similarity to mel\_A6 or will\_C9, suggesting that this activity evolved independently.

Enhancers from *D. melanogaster* and *D. willistoni* driving expression of abdominal stripes might also have a common origin, but this is less obvious from the sequence comparisons. The mel\_A2 and will\_C3 fragments drove the clearest abdominal stripes but shared no identified regions of sequence similarity.

However, sequences similar to regions of will\_C3 as well as the neighboring will\_C4, which also drove expression in stripes, mapped to the mel\_A1 and mel\_A3 fragments flanking mel\_A2, suggesting that enhancer activities in this larger region driving abdominal stripes might be homologous. Sequence similarity of mel\_A2 and mel\_A3 to pse\_B3 and pse\_B4, as well as similarity of pse\_B4 to will\_C3, further supports the orthology of these stripe enhancers. In addition, although the full *D. pseudoobscura* 5' intergenic sequence did not drive elevated expression in abdominal stripes, there was some evidence of elevated expression driven by the pse\_B3 fragment in females (Figure 4). Enhancer activities driving elevated expression along the posterior edge of abdominal segments were also observed for fragments will\_C1 and pse\_B9, but these sequences shared no similarity with mel\_A2, pse\_B4, or will\_C3, and thus appear to be independently evolved.

Expression in all of the other structures scored (*i.e.*, top of head and head circumference, broad abdominal expression, sexually dimorphic A5/A6 expression, thoracic trident expression, and expression in the wing blade and veins) except the flight muscle attachment sites was driven by multiple fragments within each species, without any clear patterns of sequence similarity linking these fragments (Figure 7). Of these traits, enhancers driving sexually dimorphic expression

have been studied most extensively in prior work with two transcription factors, Abdominal-B (Abd-B) and Bric-a-brac (Bab), shown to directly regulate a sexually dimorphic enhancer of *yellow* in *D. melanogaster* (Jeong *et al.* 2006; Camino *et al.* 2015; Roeske *et al.* 2018). We tested for enrichment of motifs representing binding sites for Abd-B, Bab, or any of ~400 other transcription factors within the 13 fragments that drove sexually dimorphic expression in A5 and A6 relative to the 16 fragments that did not. None of the 1419 motifs tested were enriched among the fragments that drove sexually dimorphic expression. We also failed to find significant enrichment of any of these motifs among fragments containing the putatively orthologous bristle or flight muscle attachment site enhancers described above.

Finally, we compared results from this study to data from a prior study using a yeast one-hybrid assay to test 670 transcription factors for evidence of binding to these same *yellow* 5' intergenic and intronic fragments (Kalay *et al.* 2016). Only one transcription factor, Hr78, showed evidence of binding to fragments from all three species in this prior work (mel\_A2, mel\_A3, mel\_A4, mel\_A5, pse\_B1, and will\_C1). The pse\_B1 and will\_C1 fragments showed some sequence similarity, but none was detected between these fragments and any of the four fragments from *D. melanogaster* (Figure 7). Furthermore, we did not observe any expression patterns that were unique to this set of fragments (Figure 7).

## Discussion

By systematically testing enhancer activity of overlapping fragments in the 5' intergenic and intronic regions of *yellow* from three *Drosophila* species, we have revealed more complexity in the *cis*-regulatory sequences of *yellow* than suggested by prior work. Furthermore, by comparing these enhancer activities to sequence similarity among the fragments tested, we have identified enhancer activities that are likely to have been derived from a common ancestral sequence as well as enhancer activities that are more likely to have evolved convergently. As described below, synthesizing these results (1) identifies modular enhancers within larger *cis*-regulatory regions, (2) suggests many redundant enhancers with overlapping function, and (3) uncovers cryptic enhancer activities in all 5' intergenic and intronic regions tested except the *D. pseudoobscura* intron. Taken together, our data suggest that the *cis*-regulatory sequences of *yellow* have properties that might have facilitated the frequent changes in expression observed among *Drosophila* species for this gene.

### ***yellow* enhancer modules refined and redefined**

Most previous studies of *Drosophila yellow cis*-regulatory sequences have focused on tissue-specific enhancer modules [reviewed in Massey and Wittkopp (2016) and Rebeiz and Williams (2017)], which are continuous regions of *cis*-regulatory DNA that drive a particular subset of a gene's expression pattern. For *D. melanogaster yellow*, the locations of such

modules were first suggested by deletion studies showing that specific regions of 5' intergenic and intronic sequence were necessary for the pigmentation of specific body parts (Geyer and Corces 1987; Martin *et al.* 1989). Wittkopp *et al.* (2002) later demonstrated that these regions identified as necessary for expression in the body and wing were also sufficient to drive expression in these tissues, consistent with the idea that *yellow* was regulated by a collection of modular, tissue-specific enhancers; however, studies of another *D. melanogaster* pigmentation gene, *ebony* (Rebeiz *et al.* 2009), suggest that *cis*-regulatory architecture might not always be this simple.

Early studies defined the *yellow* body enhancer as a region containing most of the sequence from the mel\_A3 fragment and the 5' half of sequence from the mel\_A4 fragment (Wittkopp *et al.* 2002, Figure S6). However, we found that sequences in the mel\_A2 fragment provided the best reproduction of body expression driven by the full 5' intergenic sequence (Figure 3A) and seen for the native Yellow protein (Hinaux *et al.* 2018). The mel\_A3 sequence drove broader and stronger expression than the full 5' intergenic sequence throughout the abdomen, thorax, and head, and the mel\_A4 sequence drove lower levels of broad expression without pronounced abdominal stripes (Figure 3A and Figure S2A). These data are consistent with Jeong *et al.* (2006), who found that extending the originally described body enhancer to include mel\_A2 sequences and an additional mel\_A4 sequence (Figure S6) produced a more faithful representation of *D. melanogaster yellow* expression in the abdomen.

The prominent abdominal stripes of *D. melanogaster yellow* expression were driven by both mel\_A2 and mel\_A3 (Figure 8A), but because these two fragments overlap by 100 bp, this expression pattern might be controlled by a single, continuous, modular abdominal stripe enhancer, as suggested by prior work (Figure S6). Prominent abdominal stripes of expression were also driven by orthologous sequences in the *D. willistoni* 5' intergenic region (Figure 8A), suggesting that this enhancer activity is conserved between these two species. Despite being more closely related to *D. melanogaster* than *D. willistoni*, orthologous sequences from *D. pseudoobscura* did not drive expression in abdominal stripes other than a faint pattern limited to females driven by the pse\_B3 females (Figure 8A). Rather, they drove expression broadly throughout the abdomen, which suggests that this sequence retained enhancer activity in the abdomen but evolved a new expression pattern within this tissue. Sequences in these orthologous regions (mel\_A2 and mel\_A3; pse\_B3 and pse\_B4; and will\_C3 and will\_C4) also appear to have enhancer activity driving expression in the thorax and throughout the wing blade that has been conserved in all three species (Figure 8A). This observation is consistent with prior studies of the *yellow* “wing enhancer” in *D. melanogaster*, which defined the wing enhancer as a region of sequence including most of the mel\_A2 fragment with some sequence from the 5' end of mel\_A3 (Wittkopp *et al.* 2002; Gompel *et al.* 2005) (Figure S6). It is also consistent with the observation



in Kalay and Wittkopp (2010) that enhancers controlling expression in the wing blade and body tend to be located in the same genomic region.

A modular enhancer driving sexually dimorphic expression that is higher in the A5 and A6 abdominal segments of males relative to females has also been reported for *D. melanogaster*. This activity was localized to a region of *yellow* 5' intergenic sequence overlapping both mel\_A3 and mel\_A4 in two prior studies (Jeong *et al.* 2006; Camino *et al.* 2015), with binding sites for Abd-B (Jeong *et al.* 2006) and Bab-1 (Roeske *et al.* 2018) identified in, and near, the region of overlap between mel\_A3 and mel\_A4 (Figure S6). Consistent with these data, we observed sexually dimorphic expression driven by the mel\_A4 fragment (Figure 3A), as well as elevated expression in abdominal segments A5 and A6 driven by the mel\_A3 fragment, which also showed signs of being sexually dimorphic in the PC analysis (Figure 3A, Figure 8B, and Table S2). Fragments from the orthologous region of *D. pseudoobscura* (pse\_B4) did not drive sexually dimorphic expression, but those from *D. willistoni* (wil\_C3) did (Figure 8B). This sexually dimorphic expression driven by fragments from *D. willistoni* changes our understanding of when the sexually dimorphic enhancer of *yellow* most likely evolved, as discussed further in the cryptic enhancer section below.

Finally, we identified a modular enhancer controlling expression in bristles. In all three species, bristle expression was driven by the *yellow* intron, with orthologous fragments from *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni* (mel\_A6, pse\_B7, and will\_C9) all driving expression in bristles on the body (Figure 8C). These *D. melanogaster* and *D. willistoni* fragments (mel\_A6 and will\_C9) also drove expression in bristles on the wing. Surprisingly, no wing margin bristle expression appeared to be driven by pse\_B7 (Figure 8C); rather, expression in wing bristles was driven by a nonorthologous fragment from the *D. pseudoobscura* intron, pse\_B10 (Figure 4C). The colocalization of enhancers driving bristle expression in the body and wings of both *D. melanogaster* and *D. willistoni* suggests that a common set of transcription factor-binding sites might be activating expression in the bristles of both tissues; however, the data from *D. pseudoobscura* suggest that it might also be possible to independently regulate expression in body and wing bristles.

### **Redundant enhancers are common for the *yellow* gene**

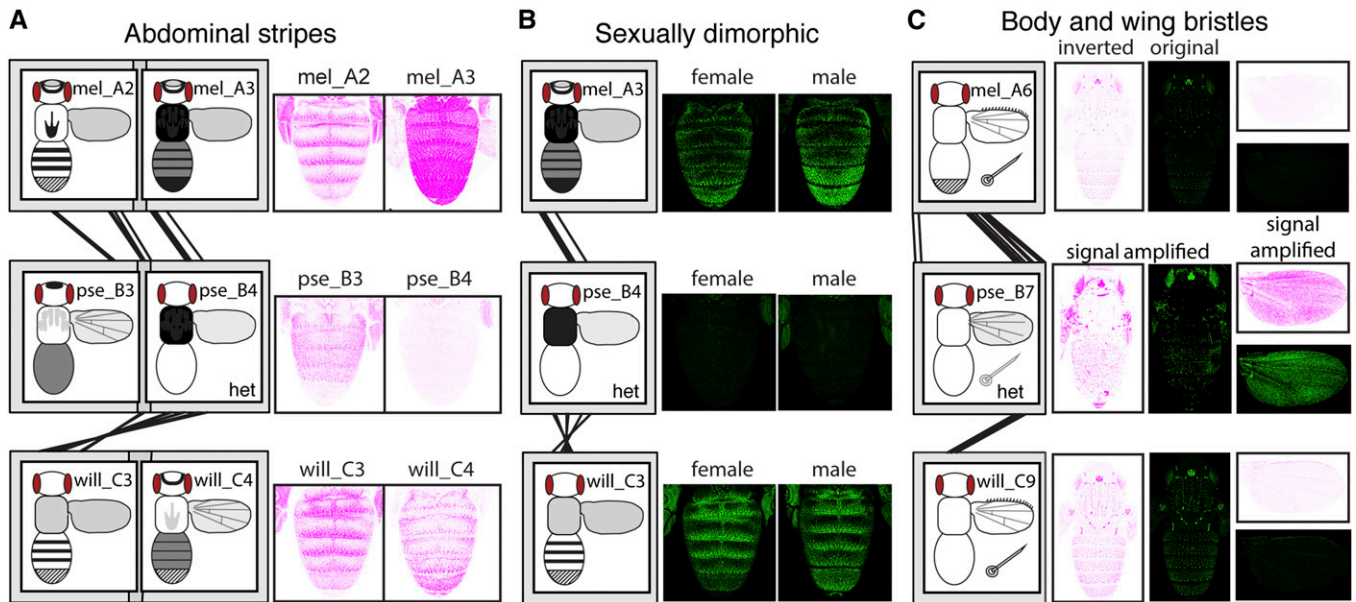
Although studies of enhancers have focused primarily on tissue-specific modules (Arnone and Davidson 1997; Carroll 2008; Rebeiz *et al.* 2009; Lorberbaum *et al.* 2016), it has become clear in the last decade that multiple enhancers driving overlapping expression patterns, known as redundant or “shadow” enhancers (Hong *et al.* 2008; Hobert 2010; Barolo 2012), are also common (Zeitlinger *et al.* 2007; Lagha *et al.* 2012; Miller *et al.* 2014; Cannavò *et al.* 2016; Osterwalder *et al.* 2018). Such redundant enhancers can be evolutionarily conserved because they confer robustness in the face of genetic variation or environmental variation, or have different functions in different tissues (Frankel

*et al.* 2010; Perry *et al.* 2010; Fujioka and Jaynes 2012; Lam *et al.* 2015; Osterwalder *et al.* 2018). Redundant enhancers might also play an important role during evolution because they provide opportunities for evolutionary novelty (Hong *et al.* 2008; Perry *et al.* 2009; Levine 2010; Long *et al.* 2016).

We found that redundant enhancers were common for the *yellow* gene: expression in all structures scored, except bristles and flight muscle attachment sites, was driven by at least two nonoverlapping fragments in at least one of the three species (Figure 7). For example, broad expression in the abdomen was driven by three fragments from *D. melanogaster*, six fragments from *D. pseudoobscura*, and six fragments from *D. willistoni* (Figure 9A). Redundant enhancers driving expression throughout the wing blade, in abdominal stripes, and in segments A5 and A6 of *D. melanogaster* males (Figure 3) were especially surprising, given that these patterns were previously described as being controlled by a single enhancer module (Wittkopp *et al.* 2002; Gompel *et al.* 2005; Jeong *et al.* 2006; Prud'homme *et al.* 2006; Camino *et al.* 2015; Roeske *et al.* 2018). In the wing blade, expression was driven by three, seven, and seven fragments from *D. melanogaster* (Figure 3, B and D), *D. pseudoobscura* (Figure 4, B and D), and *D. willistoni* (Figure 5, B and D), respectively (Figure 7). In *D. willistoni*, abdominal stripes of expression were driven not only by fragments orthologous to the *D. melanogaster* stripe enhancer, but also by the will\_C1 fragment that did not show any sequence similarity to other fragments driving expression in abdominal stripes (Figure 5A and Figure 7). Finally, in *D. melanogaster*, sexually dimorphic expression was driven not only by the previously characterized region overlapping the mel\_A3 and mel\_A4 fragments (Jeong *et al.* 2006; Camino *et al.* 2015; Roeske *et al.* 2018), but also by sequences in the mel\_A1 fragment (Figure 3A and Figure 7). Taken together, these data suggest that redundant enhancers controlling *yellow* expression are much more common than suggested by prior work. Such redundancy of enhancer activity might make mutations capable of altering activity of specific enhancers insufficient to cause detectable changes in pigmentation.

### **Cryptic enhancer activities are also common for *Drosophila yellow***

In some cases, we observed expression patterns driven by fragments from within the 5' intergenic or intronic region that were not driven by the full fragment from the corresponding region. This ectopic expression might be caused by genomic sequence surrounding the transgene insertion site, but a negative control reporter gene lacking any enhancer fragment failed to drive expression in these patterns (Figure S1A). Ectopic expression could also be caused by new transcription factor-binding sites created at junctions between *yellow* fragments and the reporter gene; however, the similarity of ectopic expression patterns among reporter genes, and to expression patterns driven by the full 5' intergenic and intronic regions of *yellow* from other species, suggests that they are unlikely to be caused by randomly generated



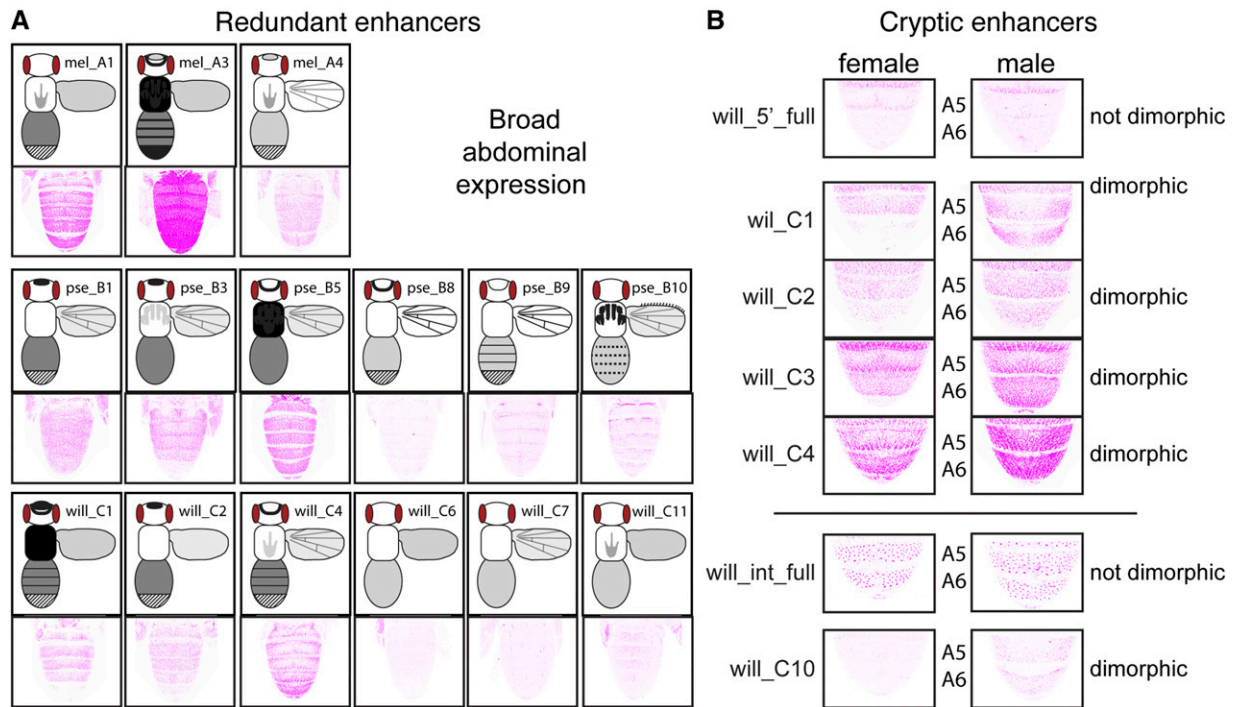
**Figure 8** Evolution of orthologous enhancers. (A) Abdominal expression driven by the mel\_A2, mel\_A3, pse\_B3, pse\_B4, and will\_C3 fragments, which share sequence similarity suggesting that they have evolved from a common ancestral sequence (i.e., are orthologous), is shown. Note that these fragments from *D. melanogaster* and *D. willistoni*, but not *D. pseudoobscura*, drive expression in abdominal stripes. Expression of the will\_C4 fragment is also shown because it overlaps with will\_C3 and also drives expression in abdominal stripes. Expression patterns of the green fluorescent protein are shown with an inverted color scheme. All images shown are from female flies. (B) Abdominal expression driven by the orthologous mel\_A3, pse\_B4, and will\_C3 fragments is shown, for both male and female flies. Note that expression in abdominal segments A5 and A6 differs between the two sexes in *D. melanogaster* and *D. willistoni*, whereas the orthologous *D. pseudoobscura* fragment does not drive abdominal expression. (C) Body and wing expression driven by the orthologous mel\_A6, pse\_B7, and will\_C9 fragments is shown, using both the original and inverted color schemes for the green fluorescent protein. Note that the fragments from *D. melanogaster* and *D. willistoni* drive strong expression in bristle cells of the body and wing margin, whereas the *D. pseudoobscura* fragment drives expression in bristles at much lower levels. Images shown for *D. pseudoobscura* have been digitally enhanced to show this weak bristle expression. All images shown are from male flies; bristle expression is not dimorphic. In (A–C), schematic representations of the full expression pattern for each fragment are the same as shown in Figure 7. Fragments assayed only in flies heterozygous for the reporter gene are marked with “het.”

transcription factor-binding sites. We therefore interpret this ectopic expression as being driven by cryptic enhancers, which are enhancer sequences that have the ability to drive expression in a particular pattern but are repressed by surrounding sequences in their native context. Such cryptic enhancers have been shown to be the source of novel expression patterns in fruit flies (Rebeiz *et al.* 2011) and humans (Prabhakar *et al.* 2008), and may facilitate the evolution of novel enhancer activities more generally.

We found that multiple cryptic enhancer activities were present in the *yellow* 5' intergenic and intronic regions that drove expression in the abdomen, thorax, wings, and/or head. For example, we did not observe expression in the wing veins driven by the full 5' intergenic sequence from either *D. melanogaster* or *D. pseudoobscura*, but did observe expression in the wing veins driven by multiple subfragments from these regions (Figure 3B, Figure 4B, and Figure 7). Similarly, the 5' intergenic regions of *D. melanogaster* and *D. willistoni* drove expression in abdominal segments A2–A4, most prominently in stripes at the posterior edge of each segment, but multiple 5' intergenic subfragments from both species drove expression much more broadly within these abdominal segments (Figure 3A, Figure 5A, and Figure 7). These broad abdominal expression patterns were similar to the broad expression

driven by the full 5' intergenic sequence of *D. pseudoobscura* (Figure 4A), suggesting that *D. pseudoobscura yellow* expression might have evolved by inactivating repressive elements inherited from the common ancestor shared with *D. melanogaster* and *D. willistoni*.

Perhaps the most interesting of these cryptic enhancers were those from *D. willistoni* driving sexually dimorphic expression in abdominal segments A5 and A6 of males similar to the sexually dimorphic pigmentation seen in *D. melanogaster*. Pigmentation is not sexually dimorphic in *D. willistoni* (Figure 1A), and neither the full 5' intergenic nor intronic regions of *D. willistoni yellow* drove expression in this pattern (Figure 9B). Nonetheless, four fragments from the *D. willistoni* 5' intergenic sequence, and one fragment from the *D. willistoni* intronic sequence, drove higher expression in segments A5 and/or A6 of males than females (Figure 9B). This observation indicates that sequences capable of driving sexually dimorphic abdominal expression (at least in *D. melanogaster*) were already present in the common ancestor shared by all three species. Therefore, the sexually dimorphic enhancer activity seems to have evolved earlier than described in Jeong *et al.* (2006) and Camino *et al.* (2015), which concluded that it arose in the lineage shared by *D. melanogaster* and *D. pseudoobscura* after it diverged from



**Figure 9** Examples of redundant and cryptic enhancers. (A) All fragments driving broad abdominal expression are shown for *D. melanogaster*, *D. pseudoobscura*, and *D. willistoni*. Although similar expression patterns driven by adjacent fragments might be caused by their overlapping sequences, similar expression patterns driven by nonadjacent fragments (e.g., mel\_A1 and mel\_A3) must be driven by independent sequences. Schematic representations of the full expression pattern for each fragment are the same as shown in Figure 7. (B) Expression in the A5 and A6 abdominal segments, driven by the full 5' intergenic ("will\_5\_full") and full intronic ("will\_int\_full") sequences from *D. willistoni*, as well as the subfragments will\_C1, will\_C2, will\_C3, will\_C4, and will\_C10, are shown. Note that the full 5' intergenic and intronic regions drive expression in these segments that is similar between males and females, yet the subfragments shown all drive sexually dimorphic expression in abdominal segment A5 and/or A6.

the lineage leading to *D. willistoni*. This element was inferred to be present in *D. pseudoobscura* despite its monomorphic pigmentation because sequences from the 5' intergenic region of *D. pseudoobscura*, as well as those from its close relative *D. subobscura*, which also has monomorphic pigmentation (Jeong *et al.* 2006), drove elevated abdominal expression in males when transformed into *D. melanogaster* (Camino *et al.* 2015). Our data are consistent with these observations, as we found that elevated expression in male A5 and/or A6 segments was driven by two 5' intergenic, and two intronic, fragments from *D. pseudoobscura* (Figure 4, A and C).

Although pigmentation of *D. pseudoobscura* is not dimorphic, we do not describe these sexually dimorphic enhancer activities in *D. pseudoobscura* as cryptic enhancers because similar patterns were driven by the full 5' intergenic and intronic sequences from *D. pseudoobscura* (Figure 4, A and C). Similarly, we do not describe sequences in pse\_B9 and pse\_B10 driving stripes of expression at the anterior and posterior edges of each abdominal segment (Figure 4C) as cryptic enhancers, even though they are not reflected in *D. pseudoobscura* pigmentation. This is because similar expression patterns were also driven by the full *D. pseudoobscura* intron (Figure 4C). It remains to be seen whether these sexually dimorphic and abdominal stripe enhancers from *D. pseudoobscura* also drive expression in these patterns in their

native species. If not, this observation would suggest that these enhancer activities are seen in transgenic *D. melanogaster* because of *trans*-regulatory divergence between *D. melanogaster* and *D. pseudoobscura* (Wittkopp *et al.* 2003). Reciprocal transformation of these reporter genes into *D. pseudoobscura* is needed to resolve this issue.

#### **Cis-regulatory architecture of yellow might facilitate expression divergence**

Expression of the *Drosophila yellow* gene has become a model system for understanding how gene regulation evolves, in part because it often shows differences among species that correlate with divergent pigmentation. Selection for divergent pigmentation is assumed to be driving this expression divergence, but the *cis*-regulatory architecture of *yellow* itself might make it especially amenable to expression divergence. Like many genes, *yellow* is controlled by multiple, tissue-specific enhancers that allow genetic changes to affect expression in one tissue without altering expression in others. But we also found that there are many redundant enhancer activities within *yellow cis*-regulatory sequences that increase the number of opportunities for expression to be modified. In addition, we found that many *yellow cis*-regulatory sequences have cryptic enhancer activities that are repressed by flanking sequences. Perhaps most interestingly, we found that cryptic enhancer activities from one species often drive expression in

patterns similar to those seen in other species, suggesting that these cryptic enhancers might be latent enhancers with evolutionary potential. Additional work is needed to determine whether these putatively latent enhancer activities have indeed been the source of what appear to be novel expression patterns in other species.

## Acknowledgments

We thank Jonathan Massey and Henry Ertl for comments on the manuscript. Funding for this work was provided by National Institutes of Health (NIH) National Research Service Award fellowship 5-F32 GM-119203 to J.L., a Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica grant from the Universidad Nacional Autónoma de México (PAPIIT-UNAM IA200217) to U.R., and grants from the National Science Foundation (MCB-1021398) and the NIH (R01 GM-089736 and 1R35 GM-118073) to P.W.

## Literature Cited

- Arnone, M. I., and E. H. Davidson, 1997 The hardwiring of development: organization and function of genomic regulatory systems. *Development* 124: 1851–1864.
- Arnoult, L., K. F. Y. Su, D. Manoel, C. Minervino, J. Magrina *et al.*, 2013 Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* 339: 1423–1426. <https://doi.org/10.1126/science.1233749>
- Barolo, S., 2012 Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *Bioessays* 34: 135–141. <https://doi.org/10.1002/bies.201100121>
- Barrière, A., and I. Ruvinsky, 2014 Pervasive divergence of transcriptional gene regulation in *Caenorhabditis* nematodes. *PLoS Genet.* 10: e1004435. <https://doi.org/10.1371/journal.pgen.1004435>
- Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for *Drosophila* using germ-line-specific C31 integrases. *Proc. Natl. Acad. Sci. USA* 104: 3312–3317. <https://doi.org/10.1073/pnas.0611511104>
- Camino, E. M., J. C. Butts, A. Ordway, J. E. Vellky, M. Rebeiz *et al.*, 2015 The evolutionary origination and diversification of a dimorphic gene regulatory network through parallel innovations in cis and trans. *PLoS Genet.* 11: e1005136. <https://doi.org/10.1371/journal.pgen.1005136>
- Cande, J., Y. Goltsev, and M. S. Levine, 2009 Conservation of enhancer location in divergent insects. *Proc. Natl. Acad. Sci. USA* 106: 14414–14419. <https://doi.org/10.1073/pnas.0905754106>
- Cannavò, E., P. Khoueiry, D. A. Garfield, P. Geeleher, T. Zichner *et al.*, 2016 Shadow enhancers are pervasive features of developmental regulatory networks. *Curr. Biol.* 26: 38–51. <https://doi.org/10.1016/j.cub.2015.11.034>
- Carroll, S. B., 2008 Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134: 25–36. <https://doi.org/10.1016/j.cell.2008.06.030>
- Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal, M. D. Shapiro *et al.*, 2010 Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327: 302–305. <https://doi.org/10.1126/science.1182213>
- Davidson, E. H., 2001 *Genomic Regulatory Systems*. Elsevier, Amsterdam.
- Eichenlaub, M. P., and L. Ettwiller, 2011 De novo genesis of enhancers in vertebrates. *PLoS Biol.* 9: e1001188. <https://doi.org/10.1371/journal.pbio.1001188>
- Emera, D., J. Yin, S. K. Reilly, J. Gockley, and J. P. Noonan, 2016 Origin and evolution of developmental enhancers in the mammalian neocortex. *Proc. Natl. Acad. Sci. USA* 113: E2617–E2626. <https://doi.org/10.1073/pnas.1603718113>
- Ettwiller, L., B. Paten, M. Souren, F. Loosli, J. Wittbrodt *et al.*, 2005 The discovery, positioning and verification of a set of transcription-associated motifs in vertebrates. *Genome Biol.* 6: R104. <https://doi.org/10.1186/gb-2005-6-12-r104>
- Fernandes, J. J., S. E. Celniker, and K. VijayRaghavan, 1996 Development of the indirect flight muscle attachment sites in *Drosophila*: role of the PS integrins and the stripe gene. *Dev. Biol.* 176: 166–184. <https://doi.org/10.1006/dbio.1996.0125>
- Frankel, N., G. K. Davis, D. Vargas, S. Wang, F. Payre *et al.*, 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466: 490–493. <https://doi.org/10.1038/nature09158>
- Frankel, N., D. F. Erezylmaz, A. P. McGregor, S. Wang, F. Payre *et al.*, 2011 Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474: 598–603. <https://doi.org/10.1038/nature10200>
- Fujioka, M., and J. B. Jaynes, 2012 Regulation of a duplicated locus: *Drosophila* sloppy paired is replete with functionally overlapping enhancers. *Dev. Biol.* 362: 309–319. <https://doi.org/10.1016/j.ydbio.2011.12.001>
- Galant, R., and S. B. Carroll, 2002 Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* 415: 910–913. <https://doi.org/10.1038/nature717>
- Geyer, P. K., and V. G. Corces, 1987 Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in *Drosophila melanogaster*. *Genes Dev.* 1: 996–1004. <https://doi.org/10.1101/gad.1.9.996>
- Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll, 2005 Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433: 481–487. <https://doi.org/10.1038/nature03235>
- Hare, E. E., B. K. Peterson, V. N. Iyer, R. Meier, and M. B. Eisen, 2008 Sepsid even-skipped enhancers are functionally conserved in *Drosophila* despite lack of sequence conservation. *PLoS Genet.* 4: e1000106. <https://doi.org/10.1371/journal.pgen.1000106>
- Hinaux, H., K. Bachem, M. Battistara, M. Rossi, Y. Xin *et al.*, 2018 Revisiting the developmental and cellular role of the pigmentation gene yellow in *Drosophila* using a tagged allele. *Dev. Biol.* 438: 111–123. <https://doi.org/10.1016/j.ydbio.2018.04.003>
- Hobert, O., 2010 Gene regulation: enhancers stepping out of the shadow. *Curr. Biol.* 20: R697–R699. <https://doi.org/10.1016/j.cub.2010.07.035>
- Hong, J. W., D. A. Hendrix, and M. S. Levine, 2008 Shadow enhancers as a source of evolutionary novelty. *Science* 321: 1314. <https://doi.org/10.1126/science.1160631>
- Horn, C., and E. A. Wimmer, 2000 A versatile vector set for animal transgenesis. *Dev. Genes Evol.* 210: 630–637. <https://doi.org/10.1007/s004270000110>
- Jeong, S., A. Rokas, and S. B. Carroll, 2006 Regulation of body pigmentation by the abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* 125: 1387–1399. <https://doi.org/10.1016/j.cell.2006.04.043>
- Kalay, G., and P. J. Wittkopp, 2010 Nomadic enhancers: tissue-specific cis-regulatory elements of yellow have divergent genomic positions among *Drosophila* species. *PLoS Genet.* 6: e1001222. <https://doi.org/10.1371/journal.pgen.1001222>



- Kalay, G., R. Lusk, M. Dome, K. Hens, B. Deplancke *et al.*, 2016 Potential direct regulators of the *Drosophila* yellow gene identified by yeast one-hybrid and RNAi screens. *G3* (Bethesda) 6: 3419–3430. <https://doi.org/10.1534/g3.116.032607>
- Kopp, A., I. Duncan, D. Godt, and S. B. Carroll, 2000 Genetic control and evolution of sexually dimorphic characters in *Drosophila*. *Nature* 408: 553–559. <https://doi.org/10.1038/35046017>
- Kopp, A., R. M. Graze, S. Xu, S. B. Carroll, and S. V. Nuzhdin, 2003 Quantitative trait loci responsible for variation in sexually dimorphic traits in *Drosophila melanogaster*. *Genetics* 163: 771–787.
- Koshikawa, S., M. W. Giorgianni, K. Vaccaro, V. A. Kassner, J. H. Yoder *et al.*, 2015 Gain of cis-regulatory activities underlies novel domains of wingless gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 112: 7524–7529. <https://doi.org/10.1073/pnas.1509022112>
- Lagha, M., J. P. Bothma, and M. Levine, 2012 Mechanisms of transcriptional precision in animal development. *Trends Genet.* 28: 409–416. <https://doi.org/10.1016/j.tig.2012.03.006>
- Lam, D. D., F. S. J. de Souza, S. Nasif, M. Yamashita, R. López-Leal *et al.*, 2015 Partially redundant enhancers cooperatively maintain mammalian Pomc expression above a critical functional threshold. *PLoS Genet.* 11: e1004935 (erratum: *PLoS Genet.* 11: e1005133). <https://doi.org/10.1371/journal.pgen.1004935>
- Letelier, J., E. de la Calle-Mustienes, J. Pieretti, S. Naranjo, I. Maeso *et al.*, 2018 A conserved Shh cis-regulatory module highlights a common developmental origin of unpaired and paired fins. *Nat. Genet.* 50: 504–509. <https://doi.org/10.1038/s41588-018-0080-5>
- Levine, M., 2010 Transcriptional enhancers in animal development and evolution. *Curr. Biol.* 20: R754–R763. <https://doi.org/10.1016/j.cub.2010.06.070>
- Lim, L. W. K., H. H. Chung, Y. L. Chong, and N. K. Lee, 2018 A survey of recently emerged genome-wide computational enhancer predictor tools. *Comput. Biol. Chem.* 74: 132–141. <https://doi.org/10.1016/j.compbiolchem.2018.03.019>
- Long, H. K., S. L. Prescott, and J. Wysocka, 2016 Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* 167: 1170–1187. <https://doi.org/10.1016/j.cell.2016.09.018>
- Lorberbaum, D. S., A. I. Ramos, K. A. Peterson, B. S. Carpenter, D. S. Parker *et al.*, 2016 An ancient yet flexible cis-regulatory architecture allows localized Hedgehog tuning by patched/Ptch1. *Elife* 5: e13550. <https://doi.org/10.7554/eLife.13550>
- Markstein, M., C. Pitsouli, C. Villalta, S. E. Celniker, and N. Perrimon, 2008 Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* 40: 476–483. <https://doi.org/10.1038/ng.101>
- Martin, M., Y. B. Meng, and W. Chia, 1989 Regulatory elements involved in the tissue-specific expression of the yellow gene of *Drosophila*. *Mol. Gen. Genet.* 218: 118–126. <https://doi.org/10.1007/BF00330574>
- Massey, J. H., and P. J. Wittkopp, 2016 The genetic basis of pigmentation differences within and between *Drosophila* species. *Curr. Top. Dev. Biol.* 119: 27–61. <https://doi.org/10.1016/bs.ctdb.2016.03.004>
- McLeay, R. C., and T. L. Bailey, 2010 Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11: 165. <https://doi.org/10.1186/1471-2105-11-165>
- Miller, S. W., M. Rebeiz, J. E. Atanasov, and J. W. Posakony, 2014 Neural precursor-specific expression of multiple *Drosophila* genes is driven by dual enhancer modules with overlapping function. *Proc. Natl. Acad. Sci. USA* 111: 17194–17199. <https://doi.org/10.1073/pnas.1415308111>
- Osterwalder, M., I. Barozzi, V. Tissières, Y. Fukuda-Yuzawa, B. J. Mannon *et al.*, 2018 Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* 554: 239–243. <https://doi.org/10.1038/nature25461>
- Perry, M. W., J. D. Cande, A. N. Boettiger, and M. Levine, 2009 Evolution of insect dorsoventral patterning mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* 74: 275–279. <https://doi.org/10.1101/sqb.2009.74.021>
- Perry, M. W., A. N. Boettiger, J. P. Bothma, and M. Levine, 2010 Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* 20: 1562–1567. <https://doi.org/10.1016/j.cub.2010.07.043>
- Prabhakar, S., A. Visel, J. A. Akiyama, M. Shoukry, K. D. Lewis *et al.*, 2008 Human-specific gain of function in a developmental enhancer. *Science* 321: 1346–1350. <https://doi.org/10.1126/science.1159974>
- Preger-Ben Noon, E., F. P. Davis, and D. L. Stern, 2016 Evolved repression overcomes enhancer robustness. *Dev. Cell* 39: 572–584. <https://doi.org/10.1016/j.devcel.2016.10.010>
- Prud'homme, B., N. Gompel, A. Rokas, V. A. Kassner, T. M. Williams *et al.*, 2006 Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* 440: 1050–1053. <https://doi.org/10.1038/nature04597>
- Prud'homme, B., N. Gompel, and S. B. Carroll, 2007 Emerging principles of regulatory evolution. *Proc. Natl. Acad. Sci. USA* 104: 8605–8612. <https://doi.org/10.1073/pnas.0700488104>
- Rebeiz, M., and M. Tsiantis, 2017 Enhancer evolution and the origins of morphological novelty. *Curr. Opin. Genet. Dev.* 45: 115–123. <https://doi.org/10.1016/j.gde.2017.04.006>
- Rebeiz, M., and T. M. Williams, 2017 Using *Drosophila* pigmentation traits to study the mechanisms of cis-regulatory evolution. *Curr. Opin. Insect Sci.* 19: 1–7. <https://doi.org/10.1016/j.cois.2016.10.002>
- Rebeiz, M., J. E. Pool, V. A. Kassner, C. F. Aquadro, and S. B. Carroll, 2009 Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326: 1663–1667. <https://doi.org/10.1126/science.1178357>
- Rebeiz, M., N. Jikomes, V. A. Kassner, and S. B. Carroll, 2011 Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. *Proc. Natl. Acad. Sci. USA* 108: 10036–10043. <https://doi.org/10.1073/pnas.1105937108>
- Roeske, M. J., E. M. Camino, S. Grover, M. Rebeiz, and T. M. Williams, 2018 Cis-regulatory evolution integrated the Bric-à-brac transcription factors into a novel fruit fly gene regulatory network. *Elife* 7: e32273. <https://doi.org/10.7554/eLife.32273>
- Rogers, W. A., J. R. Salomone, D. J. Tacy, E. M. Camino, K. A. Davis *et al.*, 2013 Recurrent modification of a conserved cis-regulatory element underlies fruit fly pigmentation diversity. *PLoS Genet.* 9: e1003740. <https://doi.org/10.1371/journal.pgen.1003740>
- Rubinstein, M., and F. S. J. de Souza, 2013 Evolution of transcriptional enhancers and animal diversity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368: 20130017. <https://doi.org/10.1098/rstb.2013.0017>
- Russo, C. A., N. Takezaki, and M. Nei, 1995 Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* 12: 391–404.
- Salomone, J. R., W. A. Rogers, M. Rebeiz, and T. M. Williams, 2013 The evolution of Bab paralogs expression and abdominal pigmentation among *Sophophora* fruit fly species. *Evol. Dev.* 15: 442–457. <https://doi.org/10.1111/ede.12053>
- Spitz, F., and E. E. M. Furlong, 2012 Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13: 613–626. <https://doi.org/10.1038/nrg3207>
- Sumiyama, K., and N. Saitou, 2011 Loss-of-function mutation in a repressor module of human-specifically activated enhancer HACNS1. *Mol. Biol. Evol.* 28: 3005–3007. <https://doi.org/10.1093/molbev/msr231>

- Swanson, C. I., N. C. Evans, and S. Barolo, 2010 Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. *Dev. Cell* 18: 359–370. <https://doi.org/10.1016/j.devcel.2009.12.026>
- Werner, T., S. Koshikawa, T. M. Williams, and S. B. Carroll, 2010 Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* 464: 1143–1148. <https://doi.org/10.1038/nature08896>
- Whibley, A. C., N. B. Langlade, C. Andalo, A. I. Hanna, A. Bangham *et al.*, 2006 Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science* 313: 963–966. <https://doi.org/10.1126/science.1129161>
- Wittkopp, P. J., and G. Kalay, 2012 Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13: 59–69. <https://doi.org/10.1038/nrg3095>
- Wittkopp, P. J., K. Vaccaro, and S. B. Carroll, 2002 Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Curr. Biol.* 12: 1547–1556. [https://doi.org/10.1016/S0960-9822\(02\)01113-2](https://doi.org/10.1016/S0960-9822(02)01113-2)
- Wittkopp, P. J., S. B. Carroll, and A. Kopp, 2003 Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends Genet.* 19: 495–504. [https://doi.org/10.1016/S0168-9525\(03\)00194-X](https://doi.org/10.1016/S0168-9525(03)00194-X)
- Zeitlinger, J., R. P. Zinzen, A. Stark, M. Kellis, H. Zhang *et al.*, 2007 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. <https://doi.org/10.1101/gad.1509607>

*Communicating editor: A. Moses*