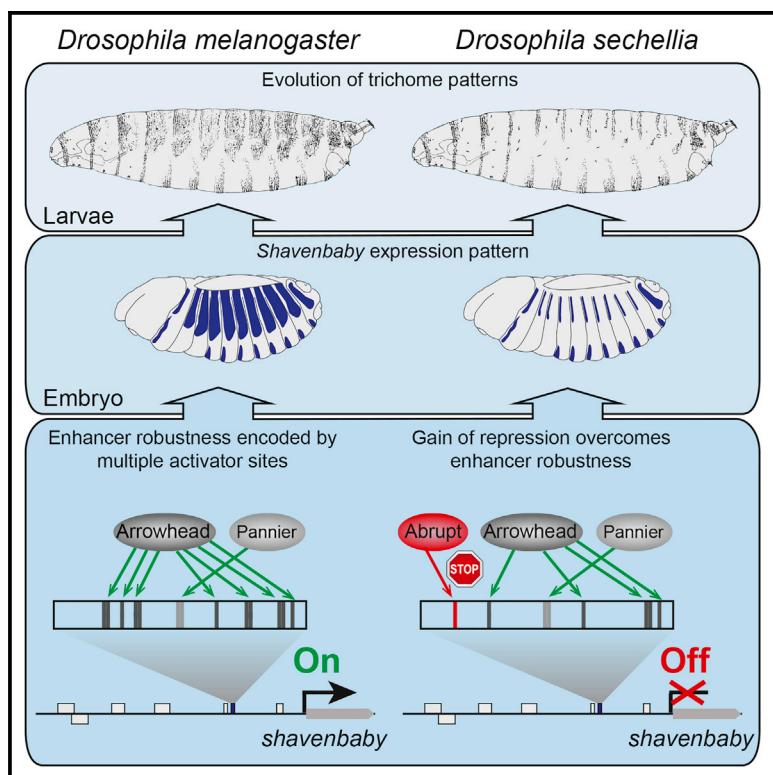


Developmental Cell

Evolved Repression Overcomes Enhancer Robustness

Graphical Abstract



Authors

Ella Preger-Ben Noon, Fred P. Davis,
David L. Stern

Correspondence

pregere@janelia.hhmi.org (E.P.-B.N.),
sternd@janelia.hhmi.org (D.L.S.)

In Brief

Preger-Ben Noon et al. deciphered the evolved regulatory changes in a robust enhancer of the *shavenbaby* gene and discovered that gain of a repressor binding site can overcome robustness encoded by multiple activator binding sites and contribute to morphological evolution.

Highlights

- Clusters of transcription factor binding sites encode enhancer robustness
- Evolutionary loss of multiple activator binding sites reduced enhancer activity
- Gain of a non-canonical binding site for the repressor Abrupt eliminated expression
- Gain of repression overcomes robustness encoded by multiple activator binding sites

Accession Numbers

GSE80790

Evolved Repression Overcomes Enhancer Robustness

Ella Preger-Ben Noon,^{1,*} Fred P. Davis,^{1,2} and David L. Stern^{1,3,*}

¹Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA

²Present address: Molecular Immunology and Inflammation Branch, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA

³Lead Contact

*Correspondence: pregere@janelia.hhmi.org (E.P.-B.N.), sternd@janelia.hhmi.org (D.L.S.)

<http://dx.doi.org/10.1016/j.devcel.2016.10.010>

SUMMARY

Biological systems display extraordinary robustness. Robustness of transcriptional enhancers results mainly from clusters of binding sites for the same transcription factor, and it is not clear how robust enhancers can evolve loss of expression through point mutations. Here, we report the high-resolution functional dissection of a robust enhancer of the *shavenbaby* gene that has contributed to morphological evolution. We found that robustness is encoded by many binding sites for the transcriptional activator Arrowhead and that, during evolution, some of these activator sites were lost, weakening enhancer activity. Complete silencing of enhancer function, however, required evolution of a binding site for the spatially restricted potent repressor Abrupt. These findings illustrate that recruitment of repressor binding sites can overcome enhancer robustness and may minimize pleiotropic consequences of enhancer evolution. Recruitment of repression may be a general mode of evolution to break robust regulatory linkages.

INTRODUCTION

Development is robust at many levels, including the development of organs (Abouchar et al., 2014; Debat et al., 2009), the expression of single genes (Cannavò et al., 2016; Degenhardt et al., 2010; Frankel et al., 2010; Perry et al., 2010), and the function of single transcriptional enhancers, where robustness can be encoded by clusters of binding sites for the same transcription factor (Crocker et al., 2015; Uhl et al., 2016). Robustness clearly evolved to facilitate life on a variable earth. But robustness also presents an evolutionary paradox. How can robust transcriptional systems evolve altered functions through mutations of relatively small effect—probably the major contributors to evolution (Rockman, 2012)—to drive morphological evolution?

The diversity of life on earth is prima facie evidence that evolution can overcome robustness to generate diversity, but it is not clear how this is accomplished. Biological robustness is encoded in part directly within transcriptional enhancers. Understanding how enhancer robustness can be overcome requires a reasonably comprehensive understanding of the regulatory

inputs into enhancers that have evolved new functions. Several enhancers have been functionally dissected to identify most or all of the regulatory inputs (Flores et al., 2000; Small et al., 1991; Stanojevic et al., 1991; Yuh and Davidson, 1996), but these enhancers appear not to have contributed to phenotypic evolution. In contrast, most previous examples of changes in transcription factor binding sites contributing to phenotypic evolution do not provide significant insight into the regulatory architecture of the entire enhancer (Gompel et al., 2005; Jeong et al., 2006; Van Moerkerke et al., 2011; Shim et al., 2012; Shirangi et al., 2009; Visser et al., 2012; Williams et al., 2008). Here, we provide a high-resolution dissection of the evolved sites at a robust enhancer that evolved to generate a morphological difference between two closely related species of *Drosophila*.

For many years, our laboratory has focused our evolutionary studies on an apparently simple anatomical feature of *Drosophila*, the pattern of microtrichia, non-sensory cuticular projections that are often called trichomes, which decorate the dorsal and lateral surface of first-instar larvae. Trichome patterns have provided a useful system for several reasons (Stern and Frankel, 2013), including the fact that trichomes represent an easily quantified output of gene expression and trichome patterns have evolved in multiple *Drosophila* species. The relative simplicity of this system has allowed extensive, detailed studies of the molecular mechanisms underlying evolutionary change (Frankel et al., 2010, 2011, 2012; McGregor et al., 2007; Sucena and Stern, 2000; Sucena et al., 2003). For example, the dorsal and lateral surface of first-instar larvae of *Drosophila melanogaster* is decorated with broad swathes of so-called quaternary trichomes (Figures 1A and 1B). In contrast, first-instar larvae of the closely related species *Drosophila sechellia* do not produce these quaternary trichomes (Figures 1D and 1E). This evolutionary transition resulted entirely from changes in the regulation of the *shavenbaby* (*svb*) gene (Sucena and Stern, 2000). *Svb* encodes a transcription factor that directs trichome morphogenesis in many epidermal cells (Arif et al., 2015; Chanut-Delalande et al., 2006; Payre et al., 1999). Seven enhancers in the *cis*-regulatory region upstream of the *svb* promoter (Figure 1H) control the complex embryonic expression of *svb* in *D. melanogaster* (Figure 1F; Frankel et al., 2010, 2011; McGregor et al., 2007). The enhancers drive expression in overlapping patterns that provide robustness of *svb* expression against both environmental and genetic variation (Frankel et al., 2010). In *D. sechellia*, five of these enhancers evolved reduced embryonic activity specifically in quaternary cells, causing loss of *svb* expression in these cells (Figure 1G) and the evolution of naked cuticle (Frankel et al., 2010; McGregor et al., 2007). Trichomes have evolved in a similar way in the distantly related *Drosophila*

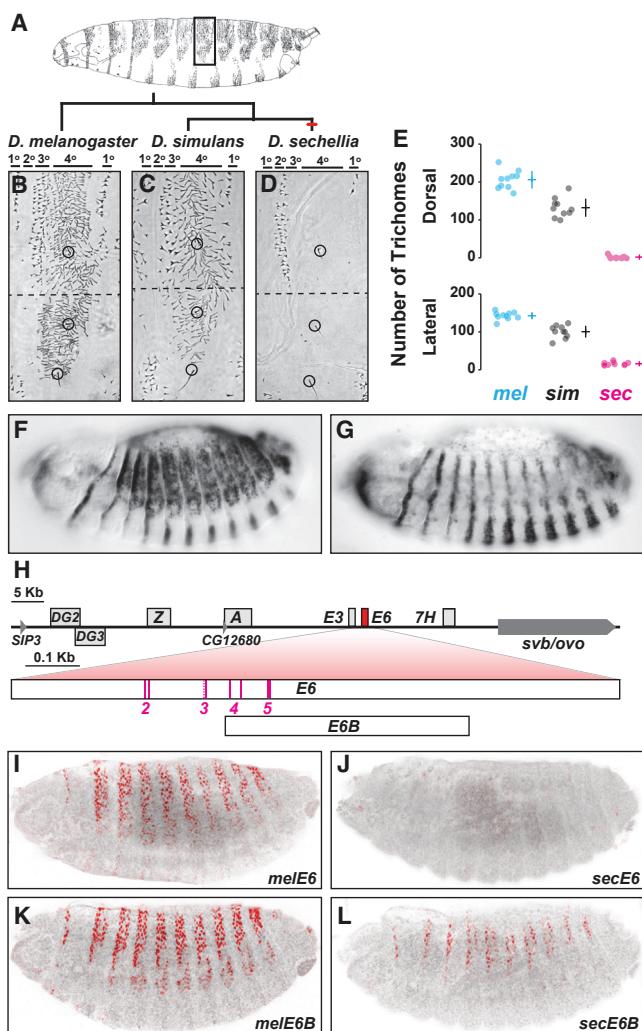


Figure 1. Trichome Patterns Have Evolved between *Drosophila* Species by Changes in the Regulatory Region of the *svb* Gene

(A) Lateral-view drawing of a first-instar larva of *D. melanogaster*. The dark rectangle demarcates the region shown in (B)–(D).

(B–D) Dorsolateral cuticle of the fifth abdominal segment of *D. melanogaster* (B), *D. simulans* (C), and *D. sechellia* (D). Dashed line indicates the border between the dorsal and lateral domains where trichomes were counted. The circles indicate three sensory bristles that are present in all three species. Adapted from McGregor et al. (2007).

(E) Trichome numbers in dorsal (top) and lateral (bottom) regions of the indicated species. Circles indicate counts for each individual ($n = 10$ for each species). Horizontal and vertical lines indicate the mean ± 1 SD.

(F and G) Expression patterns of *svb* mRNA in stage 14 embryos of *D. melanogaster* (F) and *D. sechellia* (G). Adapted from McGregor et al. (2007).

(H) Schematic representation of the *svb* locus, indicating embryonic enhancers (light gray boxes). The evolved *E6* enhancer is highlighted in red. Genes are indicated in dark gray. Vertical magenta lines in the *E6* box mark the position of *D. sechellia*-specific mutations that contributed to trichome loss. The numbers refer to the cluster in which these mutations reside. The position of the *E6B* fragment relative to *E6* is shown at the bottom of this panel.

(I–L) Expression of *D. melanogaster* *E6::LacZ* (I, melE6), *D. sechellia* *E6::LacZ* (J, secE6), *D. melanogaster* *E6B::LacZ* (K, melE6B), and *D. sechellia* *E6B::LacZ* (L, secE6B) reporter constructs in *D. melanogaster* stage 15 embryos.

See also Figure S1.

virilis clade as a result of changes in *svb* enhancers that are orthologous to the enhancers found in *D. melanogaster* and *D. sechellia* (Frankel et al., 2012). The *svb* enhancers are therefore “hot spots” for evolutionary changes in larval trichome patterns.

Detailed studies of one of these enhancers, called *E6*, revealed that the reduced embryonic expression of the *D. sechellia* *E6* enhancer resulted from many single-nucleotide substitutions and one single base-pair deletion located in four clusters (Figure 1H; Frankel et al., 2011). These observations suggested that reduced *E6* expression in *D. sechellia* resulted from changes in the binding of specific, but unknown, transcription factors.

It has traditionally been challenging to identify the transcription factors that bind to enhancers when, as is the case with these *E6* sequences, the relevant DNA sequences do not exhibit strong similarity to the known DNA binding sites for a single transcription factor. This problem is particularly challenging for developmentally regulated enhancers where it is not possible to extract large quantities of proteins from specific cells for potential discovery of protein-DNA binding through biochemical approaches. Here we overcame these roadblocks by combining genetic, genomic, and biochemical approaches to identify the transcription factors that bind differentially to evolved sites in the *D. melanogaster* and *D. sechellia* *E6* enhancers. We show that the transcriptional activators Arrowhead (Awh) and Pannier (Pnr) are positive regulators of the *svb* *E6* enhancer and that four Awh binding sites were lost in the *D. sechellia* *E6* enhancer. Despite the loss of these binding sites, the *D. sechellia* enhancer maintained functional activator sites that can drive expression. However, the *D. sechellia* enhancer also acquired a binding site for the potent repressive transcription factor Abrupt (Ab), leading to complete loss of embryonic expression. This repressor overcomes the residual expression driven by the remaining activator binding sites. The repressor Ab displays spatially restricted expression, which may limit the pleiotropic consequences of the gain of this repressor binding site. These results illustrate how loss and gain of transcription factor binding sites can overcome the robust function of a conserved enhancer to contribute to morphological evolution.

RESULTS

The *D. melanogaster* *svb* *E6* (melE6) enhancer drives stripes of expression in the dorsal and lateral epidermis of stage 15 embryos. In contrast, the orthologous DNA sequence from *D. sechellia* (secE6) does not drive detectable levels of expression (Figures 1I and 1J), and this loss of expression contributed to the loss of quaternary trichomes in *D. sechellia* (Frankel et al., 2011; McGregor et al., 2007). We performed a systematic functional dissection of the 1 kb *melE6* enhancer and identified a 390 bp region of *melE6*, named *melE6B* (Figure 1H), that accurately recapitulated the expression pattern of the full *melE6* region (Figure 1K; Frankel et al., 2011). Surprisingly, the orthologous *D. sechellia* *E6B* region (secE6B) drove low levels of expression in the quaternary cells (Figure 1L). This observation indicates that secE6B encodes binding sites for transcriptional activators and that the full secE6 enhancer acquired one or more binding sites for a repressor. Two of the four evolved sites that contributed to trichome loss in *D. sechellia* are located

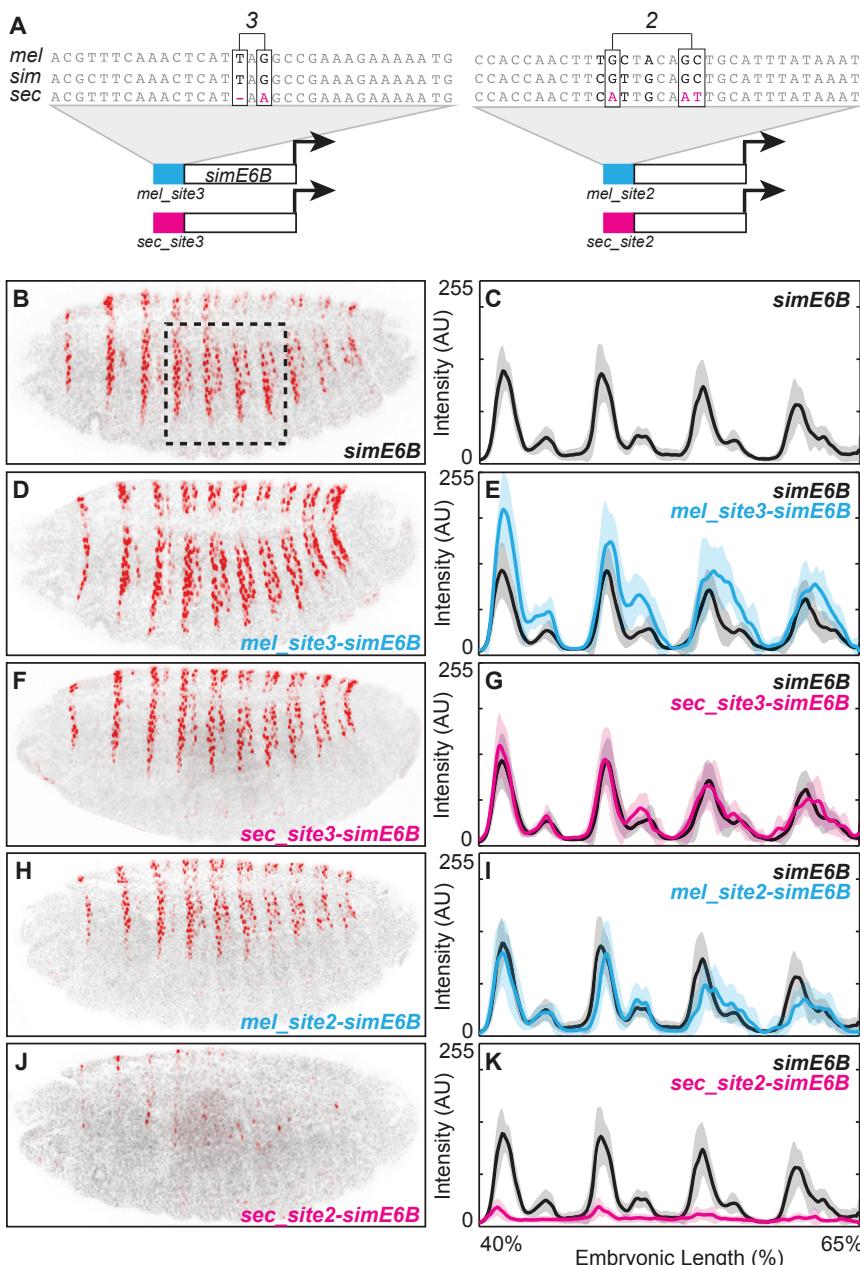


Figure 2. The *D. sechellia* E6 Enhancer Evolved through a Loss of Activation and a Gain of Repression

(A) Sequences of the evolved site 3 (left) and site 2 (right) with the aligned sequences from *D. melanogaster* (*mel*), *D. simulans* (*sim*), and *D. sechellia* (*sec*). The *D. sechellia*-specific changes are highlighted in magenta. The experimental design is illustrated at the bottom of the panel.

(B–K) Expression of *simE6B*:*LacZ* reporter constructs with the indicated sequence attached (dorsolateral view of representative embryos is shown), juxtaposed with plots of average expression in the region outlined in (B) ($n = 10$ for each genotype). In all plots, the black line denotes expression driven by the wild-type *simE6B*. Cyan and magenta lines denote average expression driven from constructs with the attached *D. melanogaster* and *D. sechellia* sequences, respectively. Shaded areas indicate ± 1 SD. AU, arbitrary units of fluorescence intensity. See also Figure S2.

~100 and 17 bp upstream of *secE6B* (Figure 1H). Of all the evolved sites these two, named sites 2 and 3, respectively, exhibited the largest effects on trichome number in functional assays (Frankel et al., 2011). We therefore hypothesized that at least one of these two sites created a binding site for a repressor in *D. sechellia*.

Loss of an Activator Binding Site and Gain of a Repressor Binding Site Caused Evolutionary Reduction in Enhancer Activity

To determine the effects of evolved sites 2 and 3 in isolation, we tested the ability of each site to modulate the activity of *E6B* in transgenic embryos. We first examined expression of the *E6B* region from the species most closely related to

D. sechellia, *Drosophila simulans*, and found that the *D. simulans* *E6B* region (*simE6B*) drives weaker expression than *melE6B* (Figure S1). Therefore, the *simE6B* enhancer serves as a sensitized system for testing the effects of attached small regions, and we used this fragment to maximize the sensitivity of our assays. We attached 32 bp surrounding site 2 or site 3 from either *D. melanogaster* or *D. sechellia* directly upstream of *simE6B* in reporter constructs (Figure 2A). If an evolved *D. sechellia* site caused loss of an activator binding site, then the *D. melanogaster* site should increase *simE6B* expression and the *D. sechellia* site should not alter *simE6B* expression. In contrast, if an evolved *D. sechellia* site caused gain of a repressor binding site, then the *D. melanogaster* site should not alter *simE6B* activity and the *D. sechellia* site should reduce *simE6B* expression.

Since we thought that we might observe subtle effects of these manipulations, we assayed the activity of the chimeric transgenes quantitatively, by measuring the fluorescence intensity in abdominal segments A2–A5 (Figures 2B–2K).

Placing the *D. melanogaster* site 3 (*mel_site3*) upstream of *simE6B*—to generate *mel_site3*-*simE6B*—caused a small elevation in the expression levels driven by *simE6B* (Figures 2D and 2E). In contrast, appending site 3 from *D. sechellia* (*sec_site3*) onto *simE6B*—to make *sec_site3*-*simE6B*—did not alter the activity of *simE6B* (Figures 2F and 2G). These results indicate that *mel_site3* contains an activator binding site, and that this site was lost in *D. sechellia*.

In contrast to site 3, placing the *D. melanogaster* site 2 (*mel_site2*) upstream of *simE6B*—to generate *mel_site2*-*simE6B*—did

not alter expression of *simE6B* (Figures 2H and 2I). Strikingly, attaching the orthologous sequence from *D. sechellia* (*sec_site2*) to *simE6B*—to generate *sec_site2-simE6B*—almost completely abolished *simE6B* expression (Figures 2J and 2K). Thus, the *secE6* enhancer acquired a binding site, *sec_site2*, for a potent transcriptional repressor. Upon binding, this repressor can overcome all the positive inputs that drive *E6B* expression.

To test whether the evolved repressor binding site, *sec_site2*, can function in other contexts, we placed it upstream of the *D. melanogaster* ventral *svb* enhancer *melE3N* to generate *sec_site2-melE3N* (Figure S2A). The wild-type *melE3N* drives expression in ventral rows of segments A1–A8 (Crocker et al., 2015) and in the dorsal amnioserosa (Figures S2B and S2C). The *sec_site2-melE3N* construct, on the other hand, did not drive any ventral expression, but showed amnioserosa expression similar to that driven by *melE3N* (Figures S2D and S2E). Thus, the repressor binding site can function in multiple contexts and the repressor that binds to *sec_site2* functions both in the dorsal and ventral epidermis but not in the amnioserosa.

Sec_site2 contains three substitutions specific to *D. sechellia* (Figure 2A). Mutating each of these positions separately from the *D. sechellia* state back to their ancestral state revealed that the two rightmost nucleotide substitutions (GC to AT) are required for the repression activity of *sec_site2* (Figures S2G–S2P). In fact, the 8 bp encompassing these substitutions (GCAATTGC), which we call *sec_site2_8*, are sufficient to recapitulate the repression activity caused by the 32 bp *sec_site2* (Figures S2Q and S2R). Taken together, these results indicate that the loss of an activator binding site at site 3 and the gain of a repressor binding site at site 2 caused most of the reduction in *secE6* function.

Cell Type-Specific Transcriptome Profiling of *svb*-Expressing Cells

To clarify how these evolved sites confer new functions, we sought to identify the transcription factors that bind to *mel_site3* and *sec_site2*. Bioinformatics analysis using JASPAR (Mathelier et al., 2015) predicted binding of 40 different homeobox-containing proteins to *mel_site3* (Table S1). *Sec_site2_8*, on the other hand, contains an E-box motif (CANNTG), but is not predicted to bind any of the HLH proteins—which bind E-box motifs—encoded in the fly genome. To narrow down the candidate transcription factor(s) that might bind to *mel_site3* and to identify the transcription factor(s) that bind to *sec_site2_8*, we combined cell type-specific transcriptomics with a reverse genetic screen.

Previous genetic work demonstrated that all of the evolved genetic changes influencing *svb* expression are encoded in the *svb* upstream *cis*-regulatory region and that the *trans*-regulatory environment that influences *svb* expression is conserved between *D. sechellia* and *D. melanogaster* (McGregor et al., 2007; Sucena and Stern, 2000). Thus, identification of the transcription factors expressed in *D. melanogaster* *svb* cells should identify all factors that bind to both the *D. melanogaster* and *D. sechellia* enhancers.

The fly genome encodes approximately 700 transcription factors, most of which are expressed during embryogenesis (Hammonds et al., 2013). To narrow down the list of potential regulators we sought to identify all of the transcription factors expressed specifically in *svb*-expressing cells in early stage 14

D. melanogaster embryos. We used a *D. melanogaster* strain carrying a 3 kb enhancer that includes both *melE3* and *melE6* (*melE10*, Frankel et al., 2011) to drive EGFP in a subset of *svb*-expressing cells (Figures 3A and S3A). The cells in early stage 14 embryos carrying this reporter construct were dissociated and EGFP-positive cells were isolated on a fluorescence-activated cell sorter (FACS). RNA from EGFP-positive cells (representing approximately 1% of the cells in the embryo, Figure S3B) was used for transcript profiling by RNA sequencing (RNA-seq). As predicted, the *melE10*-positive cells were enriched for *svb* transcripts and also for transcripts derived from most of the known *svb* target genes (Figures 3B and S3C; Menoret et al., 2013). We estimated the number of transcripts expressed per cell by spiking quantitative controls into our libraries to normalize the transcript counts and by dividing these normalized transcript counts by the number of cells used to make each library. We determined that 169 transcription factors were expressed in *melE10* cells at levels greater than half a transcript per cell, which we used as a conservative cutoff to identify genes likely to be expressed at functional levels in *svb* cells (Figure 3C and Table S2). These results allowed a focused search for transcription factors that bind differentially to evolved sites between *D. melanogaster* and *D. sechellia*.

The Evolved Site 3 Encodes a Binding Site for the Transcriptional Activator Arrowhead

We first sought to identify the transcriptional activator that binds *mel_site3* and whose binding was lost in *sec_site3*. Only 3 of the 40 homeobox-containing transcription factors that bioinformatics analysis predicted to bind to *mel_site3*, but not to *sec_site3*, are expressed in *melE10* cells (*tailup*, *Arrowhead*, and *C15*; Table S1). Among these, *Arrowhead* (*Awh*) was a good candidate for two reasons. First, *mel_site3* contains two predicted *Awh* binding sites in opposite orientations that were both lost in *sec_site3* (Figure 4A and Table S1). Second, *Awh* is expressed in the embryo in a pattern similar to the *svb* expression pattern (compare Figures 4B and 1F).

We first tested whether the *Awh* protein can bind directly to *mel_site3* using electrophoretic mobility shift assays (EMSA) with the purified *Awh* homeodomain (HD). *Awh* showed concentration-dependent binding to *mel_site3*, with a double-band shift that is consistent with the prediction of two *Awh* sites in this oligonucleotide (Figure 4C). In contrast, the nucleotide changes in *sec_site3* abolished *Awh* binding (Figure 4C). Therefore, two *Awh* binding sites were lost in *sec_site3* through a single base-pair substitution and a single bp deletion.

We next tested whether the increase in expression observed for *mel_site3-simE6B* compared with *simE6B* requires *Awh*. Downregulation of *Awh* using two independent double-strand RNA lines significantly reduced expression driven by *mel_site3-simE6B* (Figure S4). In addition, we found that complete loss of *Awh* function, in *Awh* null embryos, almost completely abrogated expression of *mel_site3-simE6B* (Figures 4D and 4E). These results indicate that *Awh* acts as a transcriptional activator and that *E6* enhancer activity requires *Awh*.

Finally, we examined whether *Awh* is required for trichome production in *D. melanogaster*. The number of trichomes on larvae heterozygous for null mutation in *Awh* was similar to the number of trichomes on wild-type larvae (Figures 4F, 4G, and

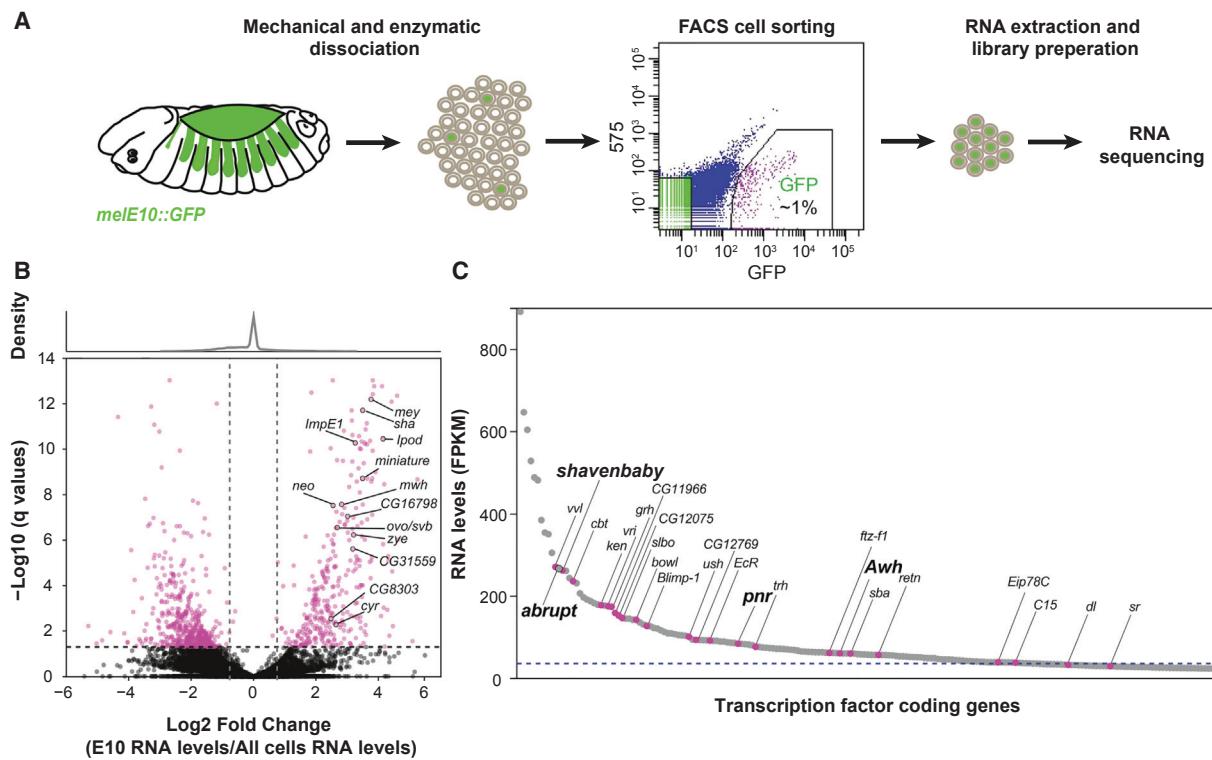


Figure 3. Cell Type-Specific Transcript Profiling of svb Cells

(A) Schematic representation of the cell type-specific transcript profiling procedure. Early stage 14 embryos carrying the *melE10* enhancer tagged with EGFP (*melE10::EGFP*) were disassociated and EGFP-positive cells were isolated on FACS. The *melE10*-positive cells represented ~1% of the cells in the embryo. RNA from the *melE10* cell population was used for transcript profiling by RNA-seq.

(B) Volcano plot representing differential gene expression between *melE10* cells and the FSC cell population (all cells in the embryo). The data were derived from four independent biological replicates and each point represents a gene. Differentially expressed genes (fold change ≥ 1.5 , $q \leq 0.05$) are shown in magenta. Enriched known svb target genes (Menoret et al., 2013) are indicated.

(C) Abundance of all of the transcription factors expressed in *melE10* cells plotted from most to least abundant. The blue dashed line represents the threshold (half a transcript per cell) that we used as a cutoff to generate a list of 169 transcription factors that are expressed in svb-expressing cells. The magenta points indicate genes that are enriched in svb-expressing cells compared with the rest of the embryo. These points are labeled with the corresponding gene. In addition, the point corresponding to *abrupt* is labeled, although this gene is not differentially expressed in svb-expressing cells.

See also Figure S3 and Table S2.

4I). In contrast, embryos that lacked *Awh* produced significantly fewer trichomes than wild-type embryos, exhibiting 30% and 54% reduction in dorsal and lateral trichome number, respectively (Figures 4F, 4H, and 4I). The quantitative reduction in trichome number in the *Awh* null, as opposed to a complete loss of trichomes, is consistent with our previous findings that multiple enhancers regulate *svb* expression in quaternary cells (Frankel et al., 2010, 2011; McGregor et al., 2007). It is possible that *Awh* regulates the expression of *svb* exclusively through binding to *E6* and that other *svb* enhancers respond to different inputs.

Arrowhead Regulates the *E6* Enhancer through Multiple Binding Sites

We noticed that complete loss of *Awh* function, in *Awh* null embryos, caused a stronger reduction in *mel_site3-simE6B* activity (Figure 4E) than we observed with loss of two *Awh* binding sites when *mel_site3* was changed to *sec_site3* (compare Figures 2E and 2G). This result suggests that *simE6B* (and probably *melE6B*) contains additional *Awh* binding site(s). Furthermore,

mel_site3-simE6B drove lower expression in *Awh* null embryos (Figure 4E) than *secE6B* drove in a wild-type background (Figure 1L), suggesting that *secE6B* contains intact *Awh* site(s). To test whether *Awh* regulates the activity of *E6B* from all three species, we assayed expression of all three *E6B* orthologs in *Awh* null embryos (Figures 5A–5O). Loss of *Awh* function dramatically reduced expression driven by *melE6B* and *simE6B* (Figures 5D, 5E, 5I, and 5J) and completely eliminated expression driven by *secE6B* (Figures 5N and 5O). These results suggest that *melE6B* and *simE6B* contain *Awh* binding sites and that some of these sites are conserved in *D. sechellia*.

The residual expression driven by *melE6B* and *simE6B* in *Awh* null embryos (Figures 5D and 5I) suggests that these enhancers respond to additional inputs. Indeed, we identified two conserved Pannier (Pnr) binding sites (Figure S5A) that are required for the proper function of *simE6B*. Site-directed mutagenesis of these Pnr binding sites reduced *simE6B* expression (Figures S5B–S5E), suggesting that Pnr, which is enriched in *svb*-expressing cells (Figure 3C), is another activator of *simE6B*. The perfect conservation of these two Pnr sites across all

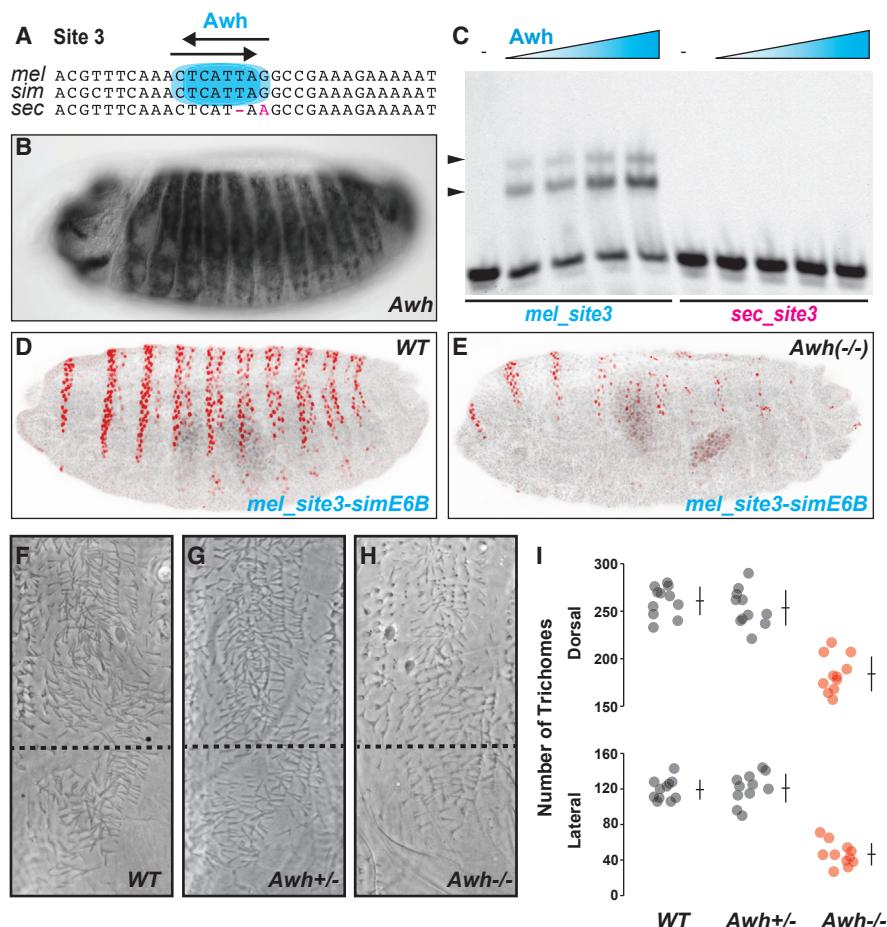


Figure 4. *D. melanogaster* Site 3 Encodes an Awh Binding Site that Was Lost in *D. sechellia*

(A) Site 3 contains two predicted binding sites for Arrowhead (Awh), highlighted with cyan ovals, that are lost in the *D. sechellia* sequence. Arrows mark the directions of the two predicted sites.

(B) Awh mRNA expression in a stage 14 embryo (adopted from Berkeley Drosophila Genome Project).

(C) EMSAs with increasing amounts of Awh-HD protein show that Awh bound specifically to the *D. melanogaster* site 3 (*mel_site3*), but not to the *D. sechellia* site 3 (*sec_site3*).

(D and E) Expression of the *mel_site3-simE6B::LacZ* reporter construct in wild-type (D) and Awh mutant (E) embryos. Awh is required for the normal expression of *mel_site3-simE6B*.

(F–H) Cuticle images showing the quaternary trichomes in the dorsal and lateral domains of wild-type (F), Awh heterozygous (G), and Awh null (H) first-instar larvae of *D. melanogaster*. Dashed line indicates the border between the dorsal and lateral domains.

(I) Number of trichomes in dorsal (top) and lateral (bottom) regions of the fifth abdominal segments of first-instar larvae of the indicated genotypes. Circles indicate counts for each individual. Horizontal and vertical lines indicate the mean \pm 1 SD.

See also Figure S4 and Table S1.

Drosophila species we examined (Figure S6A) suggests that this activator input is conserved in all of these species, including *D. sechellia*.

Motif search analysis predicted that the *melE6B* sequence encodes 15 Awh binding sites (Figures 5P and S6A). To test whether Awh actually binds to these predicted sites, and to search for additional Awh binding sites, we systematically screened the entire *E6B* sequence with EMSAs (Figures 5P and S6). In addition to site 3 (included in *E6B1*), we identified four regions—*E6B2*, *E6B6*, *E6B8*, and *E6B10*—that bound Awh with varying affinities (Figures 5P, 5Q, and S6B). We systematically dissected and mutated the predicted Awh sites in these *melE6B* subfragments (Figures S6C–S6F) and found

that Awh binds to 9 of the 15 predicted sites (Figures 5P and 5Q). Most of the bound Awh sites had a similar architecture, with two overlapping sites, one on each DNA strand in opposite orientations (Figure S6A).

Two of the Awh binding sites in *E6B* were lost specifically in *D. sechellia* as a result of a single base-pair substitution (Figure S6A). This *D. sechellia*-specific nucleotide substitution corresponds to evolved site 4 (Figure 1H), and evolution of this site in *D. sechellia* caused a weak decrease in *E6* function (Frankel et al., 2011). Comparison with sequences from outgroup species (Figure S6A) indicates that two Awh binding sites were gained in *D. melanogaster* (in fragment *E6B8*). The five remaining Awh binding sites—one in *E6B1* (see fragment *E6B1.2*, Figures S6C

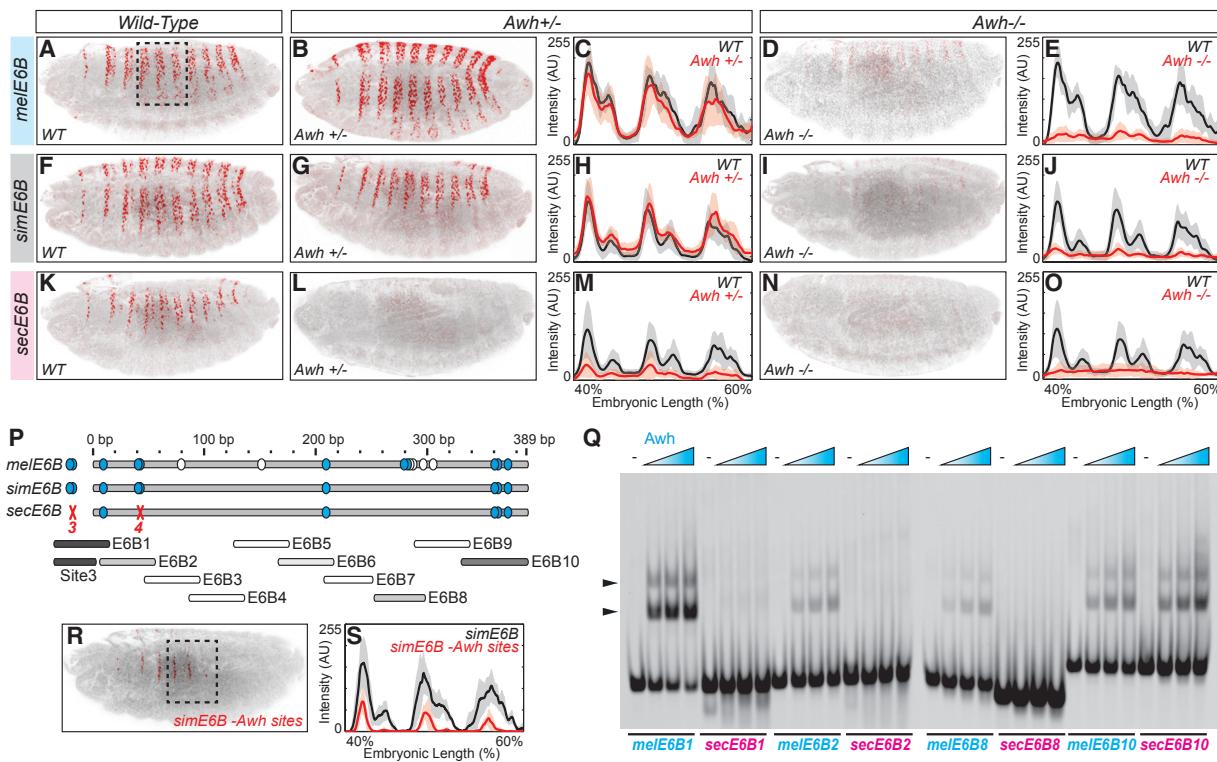


Figure 5. Awh Regulates the *E6B* Enhancer through Multiple Awh Binding Sites that Confer Robustness

(A–O) Wild-type (A, F, and K), Awh null heterozygote (B, C, G, H, L, and M), and Awh null homozygote (D, E, I, J, N, and O) embryos carrying *E6B::LacZ* reporter constructs from *D. melanogaster* (*melE6B*), *D. simulans* (*simE6B*), and *D. sechellia* (*secE6B*), juxtaposed with plots of average expression in the region outlined in (A) ($n = 10$ for each genotype). In all plots, the black and red lines denote expression driven by the reporter constructs in wild-type and mutant embryos, respectively. Shaded areas indicate ± 1 SD. AU, arbitrary units of fluorescence intensity. The *D. melanogaster* and *D. simulans* sequences are robust to the reduced dose of Awh in Awh heterozygotes (B, C, G, and H), while the *D. sechellia* sequence is sensitive to such reduction (L and M). All three *E6B* enhancers lose expression in Awh null embryos (D, E, I, J, N, and O).

(P) A schematic representation of the binding sites for Awh, indicated in cyan, in *melE6B*, *simE6B*, and *secE6B* identified by systematic EMSAs using oligos from the regions represented on the bottom. The open ellipses on the *melE6B* schematics represent predicted sites that did not bind Awh in EMSAs (see Figure S6). Site 3 is shown upstream of *E6B1*. The *melE6B* region contains nine additional binding sites for Awh. Two of these sites were gained specifically in *D. melanogaster* (see Figure S6). In addition to the loss of two Awh sites in site 3, *D. sechellia* lost two binding sites for Awh in site 4 (marked with red X).

(Q) Awh binds to multiple sites in *melE6B*, clustered in five regions (*melE6B1*, *melE6B2*, *melE6B6*, *melE6B8*, and *melE6B10*), as demonstrated with EMSAs. Awh binding was greatly reduced in the *D. sechellia* *E6B1* fragment, due to the mutations in site 3. The residual binding results from a conserved Awh site in *E6B1* (see Figure S6). Awh did not bind to the *D. sechellia* orthologous fragments of *E6B2* and *E6B8* (*secE6B2* and *secE6B8*), but bound to conserved sites on *secE6B10*.

(R and S) Disruption of six Awh sites within *simE6B* (*simE6B*-Awh sites) caused almost complete elimination of reporter expression (R) compared with wild-type (F). Region indicated by dashed box in (R) is quantified in (S).

See also Figures S5 and S6; Table S2.

and S6D), one in *E6B6*, and three in *E6B10*—are conserved in *D. sechellia* (Figure 5P).

We next tested the effect of this sequence divergence and conservation on the binding of Awh to the *D. sechellia* fragments (Figure 5Q). As predicted, Awh did not bind to the *secE6B2* and *secE6B8* fragments, and binding to *secE6B1.2* (Figure S6D) and *secE6B10* was conserved (Figure 5Q). Taken together, these results indicate that the *E6* enhancer encodes multiple binding sites for the transcriptional activator Awh and that some of these sites are conserved and four were lost in *D. sechellia*.

To test whether *E6B* activity requires the identified Awh sites in vivo, we mutated six of the identified Awh sites in *simE6B*. Disruption of these sites almost completely abolished the expression driven from this reporter construct (Figures 5R and 5S), consistent with the reduced expression of *simE6B* observed in Awh null embryos (Figures 5I and 5J). Taken together, these

results demonstrate that Awh directly regulates the expression of *svb* by binding to multiple sites within the *E6* enhancer and that evolutionary loss of Awh binding at sites 3 and 4 in the *D. sechellia* lineage contributed to reduced embryonic function of the *secE6* enhancer.

Loss of Awh Sites in *secE6B* Reduced Its Robustness to Genetic Variability

Homotypic clusters of transcription factor binding sites have been documented in many developmental enhancers (Crocker et al., 2015; Driever et al., 1989; Giorgetti et al., 2010; Nachman et al., 2015; Rowan et al., 2010; Struhl et al., 1989; Uhl et al., 2016; Williams et al., 2008) and, at least in some contexts, can provide regulatory robustness (Crocker et al., 2015; Uhl et al., 2016). Since *secE6B* lost two binding sites for Awh compared with *simE6B* (Figure 5P), we asked whether *secE6B* is less robust

to genetic variability than *melE6B* and *simE6B*. To test this hypothesis we assayed expression driven by these enhancers in embryos heterozygous for an *Awh* null allele. While *melE6B* and *simE6B* drove normal levels of expression in *Awh* heterozygotes (Figures 5B, 5C, 5G, and 5H), the *secE6B* enhancer drove dramatically lower levels of expression in *Awh* null heterozygotes compared with wild-type embryos (Figures 5L and 5M). These results are in agreement with our findings that the number of trichomes produced by *Awh* heterozygous larvae are similar to the number produced by wild-type *D. melanogaster* larvae (Figure 4I) and indicate that *secE6B* is less robust to variable genetic backgrounds than are *melE6B* and *simE6B*. The loss of four *Awh* binding sites may have reduced the robustness of the *D. sechellia* *E6B* enhancer to genetic variability, but these changes were not sufficient to completely abolish *E6B* expression.

Abrupt Is a Repressor that Binds and Regulates

D. sechellia Site 2

We next sought to identify the transcriptional repressor that binds to *D. sechellia* site 2. Since this site did not match the known binding sites for any *Drosophila* transcription factor, we searched for the repressor with an RNAi-based screen. We used fly lines carrying *UAS::RNAi* constructs generated by colleagues (Dietzl et al., 2007; Ni et al., 2011) and drove RNAi in post-gastrulation epidermal cells with a *Kni::Gal4* construct (Figure S7A; Jenett et al., 2012). Using these lines, we systematically targeted all the transcription factors and co-repressors expressed at greater than an estimated half transcript per *melE10* cell (Figures 3C and Table S2). We then assayed embryos for derepression of the *sec_site2-simE6B* reporter expression. Of the 347 lines screened (Figure 6A and Table S3), we observed derepression in only three lines (Figures 6B, 6C, and S7), all of which target the mRNA for the BTB-zinc finger protein Ab. Ab protein is expressed uniformly in the embryonic epidermis starting from stage 11 through stage 15 (Figures 6D and 6E; Hu et al., 1995). Consistent with our observation that *sec_site2* does not mediate repression in the amnioserosa (Figure S2), Ab was not detected in the amnioserosa.

To test whether Ab represses *secE6* by binding to *sec_site2*, we assayed *sec_site2-simE6B* expression in *ab* null embryos. The expression driven by *sec_site2-simE6B* was significantly increased in embryos homozygous for an *ab* null allele (Figures 6F–6I), suggesting that Ab mediates the repression of *secE6*, probably through binding to *sec_site2*.

We next tested whether the DNA binding domain of the Ab protein can bind specifically to the *sec_site2* sequence. The repressor binding site, GCAATTGC, does not resemble the consensus Ab sequence that was defined previously in a bacterial one-hybrid study (Zhu et al., 2011). Nevertheless, we found that the *sec_site2* sequence, but not its orthologous sequence from *D. melanogaster*, bound the Ab zinc finger domain in vitro (Figure 6J). This binding was specific, and could be competed with an unlabeled *sec_site2* oligonucleotide, but not with an unlabeled *mel_site2* oligonucleotide (Figures 6J [lane 11] and S7G). In addition, a single base pair mutation in the DNA binding domain of Ab, which disrupts its activity in vivo (*ab*^{clu2}, R575C; Hu et al., 1995), abolished binding to *sec_site2* (Figure 6J, lanes 12–15).

To verify that Ab binds specifically to the repressor binding site within the *sec_site2* oligonucleotide, we reverted each of the *D. sechellia*-specific mutations to the *D. melanogaster* state in an unlabeled *sec_site2* oligonucleotide and used these mutated oligonucleotides in competition assays (Figures S7F and S7G). Mutating either the A or T in the center of the repressor binding site to G or C, respectively, reduced the ability of these probes to compete the Ab binding to *sec_site2* (Figure S7G). These results are consistent with the in vivo experiments (Figure S2) and indicate that Ab binds specifically to the eight base pairs encoding a repressor binding site in *sec_site2*, resulting in repression of the *secE6* enhancer region. Furthermore, our results demonstrate that Ab has a novel DNA binding property that does not match the published Ab motif data (Zhu et al., 2011). Other studies have also demonstrated that transcription factors can function in vivo through non-canonical binding sites (Crocker et al., 2015; Farley et al., 2015).

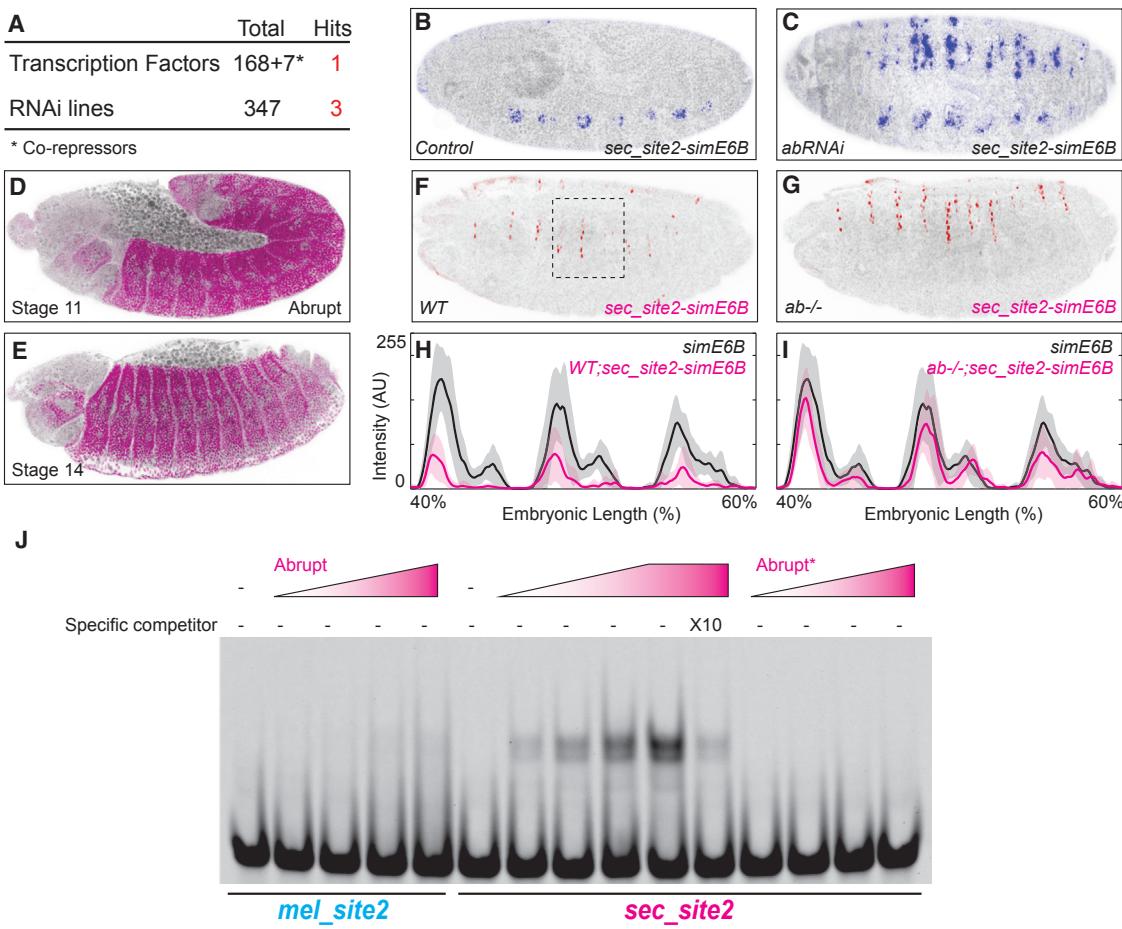
Taken together, our results indicate that the Ab transcription factor acts as a potent repressor that binds directly to the evolved *sec_site2* through a non-canonical binding site. The results for both sites 2 and 3 are consistent with the previously published findings that mutating these sites from the *D. sechellia* sequence to the *D. melanogaster* sequence in the context of the *D. sechellia* *E6* enhancer results in gain of trichomes and vice versa (Frankel et al., 2011).

DISCUSSION

We have deciphered most of the regulatory changes in the *svb E6* enhancer that contributed to morphological evolution between *D. sechellia* and closely related species (Figure 7). The transcription factor *Awh* binds at multiple sites in the *D. melanogaster* *E6* enhancer and *Pnr* binds at two sites to promote gene expression in quaternary cells. Four of the *Awh* binding sites were lost in the evolutionary lineage leading to *D. sechellia* through single base pair mutations and a single base pair deletion, resulting in a partial loss of *E6* function. The *D. sechellia* *E6* enhancer also acquired a non-canonical binding site for the Ab repressor protein through two single base pair substitutions, which causes complete suppression of *E6* activity. These data provide a detailed understanding of how one enhancer has evolved altered function and provide insight into fundamental unanswered questions about enhancer evolution.

Individual transcriptional enhancers are often robust (Payne and Wagner, 2015) and encode pleiotropic expression patterns (Cheng et al., 2014; Hiller et al., 2012; Jackman and Stock, 2006; Uhl et al., 2016). These two broad patterns raise two currently unresolved issues about enhancer evolution. First, it is not clear how individual nucleotide changes can overcome enhancer robustness. Second, it is not clear how nucleotide changes can generate specific changes in pleiotropic enhancers. We address these two issues in turn.

We have previously demonstrated that *svb* expression is buffered against genetic and environmental variation at two levels. First, the embryonic expression of *svb* is regulated by the activity of seven enhancers that drive overlapping patterns of expression (Frankel et al., 2010, 2011; McGregor et al., 2007). Expression from multiple partially redundant enhancers is required to confer phenotypic robustness under stressful conditions (Frankel et al.,



*R575C DNA binding domain mutant

Figure 6. The *D. sechellia* E6 Region Acquired a Binding Site for the Transcriptional Repressor Abrupt

(A) Table summarizes the results of an RNAi screen for the repressor that binds to *D. sechellia* E6 site 2. The screen resulted in three hits, all targeting the transcription factor *abrupt* (*ab*).

(B and C) Expression of *sec_site2-simE6B::LacZ*, as determined by β -gal staining in control (B) and a representative *ab* RNAi line (C).

(D and E) *Ab* expression in stage 11 (D) and stage 14 (E) embryos.

(F–I) Expression of the *sec_site2-simE6B::LacZ* reporter construct in wild-type (F) and *ab* mutant (G) embryos, and plots (H and I) of average expression in the region outlined in (F) ($n = 10$ for each genotype). Black and magenta lines denote expression driven by *simE6B* and *sec_site2-simE6B*, respectively. Shaded areas indicate ± 1 SD. AU, arbitrary units of fluorescence intensity. In *ab* background, *sec_site2-simE6B* activity is derepressed.

(J) EMSAs with increasing amounts of the *Ab* zinc finger domain and *mel_site2* or *sec_site2* oligonucleotides. *Ab* bound specifically to the *sec_site2* probe. In the competition lane, a 10-fold excess of unlabeled *sec_site2* probe was used (lane 11) and the intensity of the shifted band was reduced. An *Ab* mutant, carrying a single amino acid change in the zinc finger domain, did not bind to *sec_site2* (lanes 12–15).

See also Figure S7 and Table S3.

2010). Second, the ventral *svb* enhancers *E3N* and *7H* contain homotypic clusters of Hox binding sites that provide an additional layer of regulatory robustness (Crocker et al., 2015). Similarly, we have found that the dorsal *svb* enhancer *E6* encodes multiple Awh binding sites that are required for robust expression. Given all these layers of robustness, it is not clear how a few nucleotide changes can lead to dramatic changes in gene expression. Our finding that the activity of multiple Awh binding sites can be overcome by gain of a single repressor binding site suggests that gain of repressor binding sites is a key mechanism to overcome enhancer robustness.

Clusters of binding sites for transcriptional activators have been documented in many other developmental enhancers

(Crocker et al., 2015; He et al., 2012; Lifanov et al., 2003; Nachman et al., 2015; Ochoa-Espinosa et al., 2005; Uhl et al., 2016). Loss of single activator binding sites rarely has much effect on enhancer function. Evolution entirely through loss of activator binding sites requires changes at many sites or deletions that remove many sites. The gain of a small number of repressor binding sites may in such cases provide an evolutionarily shorter route to reduction in gene expression than loss of multiple activator binding sites. This would provide an evolutionary advantage to evolving by gain of repressor binding sites.

Evolution by gain of repressor binding sites may also be favored for enhancers with pleiotropic roles. Some enhancers are active in multiple tissues or at multiple stages during

Drosophila melanogaster

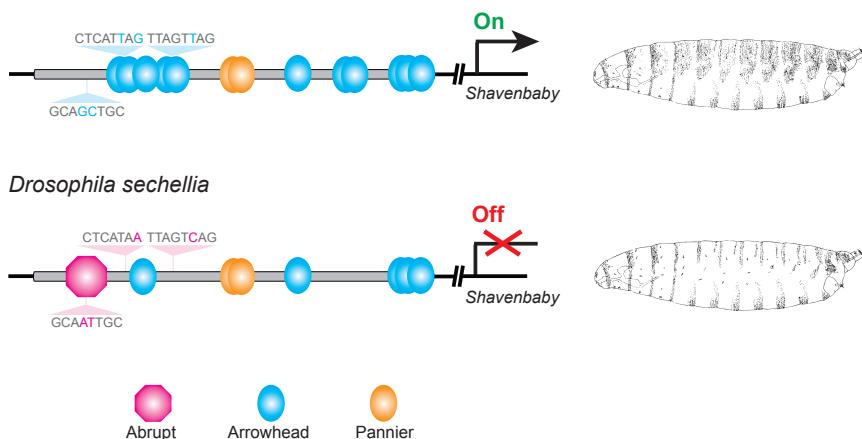


Figure 7. A Model Summarizing Transcription Factor Occupancy at the Evolved *svb* E6 Enhancer

D. melanogaster *E6* encodes multiple binding sites for the transcriptional activators Awh and Pnr that direct *svb* expression in quaternary cells, resulting in trichome formation. Four of the Awh binding sites were lost in the evolutionary lineage leading to *D. sechellia*, resulting in a partial loss of *E6* function. The acquisition of a binding site for the strong repressor Abrupt caused complete suppression of *E6* activity, contributing to the transformation of trichomes to naked cuticle in *D. sechellia*. Sequences of evolved sites are shown.

development (Glassford et al., 2015). Conservation of activator binding sites might be required for the redeployment of these enhancers in different places or at different times in development. Conversely, gain of a binding site for a spatially and/or temporally limited repressor may provide a mechanism for reducing enhancer activity in a subset of these spatiotemporal domains. For example, in Dipterans, some wing-patterning genes, such as *knot* and *splat*, are expressed in wing discs and repressed specifically in the halteres by Ultrabithorax (*Ubx*) through multiple *Ubx* binding sites (Galant et al., 2002; Hersh and Carroll, 2005). It has been proposed that these *Ubx* binding sites were gained in the lineage leading to Dipterans and promoted transformation of hindwings to halteres (Prud'homme et al., 2007), although there is not yet direct evidence for this transition. Interestingly, we observed that the *E6* enhancers from both *D. melanogaster* and *D. sechellia* drive expression in multiple larval tissues (E.P.B.N., unpublished data). It is possible that Ab is not expressed in these tissues, which would make the presence of the Ab binding site in the *D. sechellia* *E6* enhancer irrelevant to these later functions, much as it is irrelevant to enhancer expression in the amnioserosa (Figure S2). The potential requirement for *E6* function during late developmental stages might also explain why this enhancer evolved through multiple mutations with specific effects and not through a large deletion, as has been observed in other cases of enhancer evolution (Chan et al., 2010; Indjeian et al., 2016; Jeong et al., 2008; McLean et al., 2011). For example, human-specific deletion of a hindlimb enhancer for the *Gdf6* gene may have contributed to the evolution of the human foot (Indjeian et al., 2016). In addition, enhancer deletions have repeatedly contributed to morphological evolution in recently evolved freshwater populations of three-spined sticklebacks (Chan et al., 2010). It is possible that such enhancers have fewer pleiotropic roles than enhancers that evolve by subtle nucleotide changes (Frankel et al., 2011; O'Brown et al., 2015; Rebeiz et al., 2009). In addition, human and freshwater stickleback populations have probably experienced stronger selection and smaller population sizes than most *Drosophila* species, both of which can contribute to evolution through mutations that would be less likely to be favored in large, stable populations (Stern, 2011).

It is not clear how often gain of repression has contributed to enhancer evolution because there are currently few other studies with which to compare our results. Rebeiz et al. (2009) observed that a reduction in *ebony* expression in Ugandan *D. melanogaster* strains evolved in part by strengthening of an orphan repressor binding site located outside of the minimal abdominal enhancer of *ebony*. Similar to our results, this change in repressor activity had the largest effect on enhancer function of all mutations tested, although it is not known whether the other changes encode loss of activation or gain of repression. While gain of repression can reduce enhancer activity, evolutionary loss of repression can increase expression. For example, Johnson et al. (2015) showed that evolutionary loss of the activity of a large silencer element caused increased expression of *ebony*, although the specific binding site alterations have not yet been identified. On a genome-wide scale, Prescott et al. (2015) observed that during human evolution enhancer functions were sometimes gained apparently by loss of putative repressor sites, although these observations have not yet been supported by functional experimental data. Much future work remains—requiring the more complete elucidation of the architecture of evolving enhancers—to determine whether gain of repression is a common mode by which evolution breaks robust regulatory linkages.

EXPERIMENTAL PROCEDURES

Transgenic Constructs and Fly Strains

DNA fragments were synthesized by Genscript and cloned into the reporter constructs *placZattB* and *pS3AG* (Table S4). Plasmids were integrated into the *attP2* landing site by Rainbow Transgenic Flies. For RNAi experiments, the reporter constructs were integrated into *attP40*. Additional strains used were *Oregon*^R, *w;P{f+7.7} w[+mC]=GMR43B11-GAL4}attP2* (Jenett et al., 2012), *Awh¹⁶*, *Awh¹¹* (Curtiss and Heilig, 1995), *ab^{1D}* (Hu et al., 1995), *Df(2L)Exel6028*, and *Df(3L)Exel6098* (Parks et al., 2004).

Embryo Staining and Image Analysis

Stage 15 embryos were collected, fixed, and stained using standard protocols with mouse anti-β-galactosidase (β-gal; 1:500, Promega), chicken anti-GFP (1:500, Aves Labs) or rabbit anti-Ab (1:200, Hu et al., 1995), and anti-mouse, anti-chicken, or anti-rabbit Alexa Fluor (1:500, Invitrogen) antibodies. Embryos carrying reporter constructs were imaged on a Leica SPE confocal microscope. Image analysis and fluorescence intensity quantification procedures were described previously (Crocker et al., 2015). In brief, sum projections of confocal stacks were assembled, images were scaled, background was subtracted using a 50-pixel rolling-ball radius, and plot profiles of fluorescence

intensity in abdominal segments A2–A4 or A2–A5 were analyzed using ImageJ software (<http://rsb.info.nih.gov/ij>). Data from the plot profiles were analyzed further in MATLAB (<http://www.mathworks.com>). For representative embryo images, staining is shown in inverted mode. Some images have been rotated and adjacent embryos cropped out, with the background filled in.

RNAi Screen

Males from each *UAS::RNAi* carrying line (Dietzl et al., 2007; Ni et al., 2011; Table S3) were crossed with virgin females of stock *w;sec_site2-simE6B;GMR43B11-GAL4* and placed in 24-chamber Fly Condos (Flystuff). Eggs were collected for 16 hr at 28°C, dechorionated, and fixed in fixative (2% formaldehyde, 0.2% glutaraldehyde in PBS) and heptane by vigorous shaking for 20 min. The fixed embryos were washed three times with PBS and β-gal staining was performed using the β-gal staining kit (Life Technologies). Embryos were screened for expression of *simE6B* in abdominal segments under a Leica M80 stereo microscope.

Cell Sorting and RNA Isolation

Approximately 50 mL of young wild-type flies or flies carrying the *melE10::EGFP* transgene were reared in large population cages at 25°C. Following 1 hr of egg collection, embryos were left to age for approximately 11 hr at 25°C. Cells from early stage 14 embryos were dissociated using the method described by Salmand et al. (2011) and EGFP-positive cells were sorted on a Becton-Dickinson FACSVantage SE w/Diva flow cytometer as described in *Supplemental Experimental Procedures*. The collected cells were centrifuged at 850 × g for 10 min at 4°C and resuspended in 350 μL of RLT buffer (Qiagen), and RNA was purified using a Qiagen RNeasy Plus Micro kit according to the manufacturer's instructions.

cDNA Amplification, Library Construction, and Data Analysis

Five nanograms of total RNA was used for cDNA synthesis and amplification using the Ovation RNA-Seq System V2 (NuGen). To quantify the RNA-seq results, we added ERCC RNA Spike-In Control Mix 2 (1 μL of 1:10,000 dilution, Ambion) to the samples prior to amplification. One microgram of each amplified sample was used for library construction using the Encore NGS Library System I (NuGen). RNA-seq data were analyzed using the Tuxedo protocol (Trapnell et al., 2012) as described in *Supplemental Experimental Procedures*.

Electromobility Shift Assays

Protein purification and EMSAs were performed as described by Uhl et al. (2010). See *Supplemental Experimental Procedures* for a detailed protocol.

ACCESSION NUMBERS

All sequencing datasets were deposited in the NCBI GEO repository under accession number GEO: GSE80790.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.devcel.2016.10.010>.

AUTHORS CONTRIBUTIONS

E.P.B.N. and D.L.S. conceived the experimental plan. E.P.B.N. performed all the experiments and analyzed the data. F.P.D. and E.P.B.N. analyzed the RNA-seq data. E.P.B.N. and D.L.S. wrote the manuscript with input from F.P.D.

ACKNOWLEDGMENTS

We thank Jessica Treisman, Steve Crews, VDRC, and the Bloomington Drosophila Stock Center for fly stocks and reagents. We thank Aishwarya Koragaonkar for help with protein purification. We thank the Janelia Fly Core Facility for help with fly work and the Flow Cytometry Resource Facility at Princeton University for sorting experiments. E.P.B.N is grateful for the generous hospi-

tality of Paul Schedl and his laboratory members, especially Tsutomu Aoki and Girish Deshpande, and of Adi Salzberg and her laboratory members. We thank Nicolas Frankel, Anastasios Pavlopoulos, and Stern and Schedl laboratory members for helpful discussions. The manuscript was improved by the critical comments of Richard Mann, François Payre, and Nicolas Frankel. E.P.B.N was supported by post-doctoral fellowships from the Human Frontier Science Program (LT000528/2011) and EMBO (ALTF 1364-2010).

Received: March 30, 2016

Revised: July 26, 2016

Accepted: October 14, 2016

Published: November 10, 2016

REFERENCES

- Abouchar, L., Petkova, M.D., Steinhardt, C.R., Gregor, T., Thompson, D., Cross, M., Greenside, H., Wigglesworth, V., Smith, J.M., Lawrence, P., et al. (2014). Fly wing vein patterns have spatial reproducibility of a single cell. *J. R. Soc. Interf.* 11, 20140443.
- Arif, S., Kittelmann, S., and McGregor, A.P. (2015). From shavenbaby to the naked valley: trichome formation as a model for evolutionary developmental biology. *Evol. Dev.* 17, 120–126.
- Cannavò, E., Khoury, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, J.O., and Furlong, E.E.M. (2016). Shadow enhancers are pervasive features of developmental regulatory networks. *Curr. Biol.* 26, 38–51.
- Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327, 302–305.
- Chanut-Delalande, H., Fernandes, I., Roch, F., Payre, F., and Plaza, S. (2006). Shavenbaby couples patterning to epidermal cell shape control. *PLoS Biol.* 4, e290.
- Cheng, Y., Ma, Z., Kim, B.-H., Wu, W., Cayting, P., Boyle, A.P., Sundaram, V., Xing, X., Dogan, N., Li, J., et al. (2014). Principles of regulatory information conservation between mouse and human. *Nature* 515, 371–375.
- Crocker, J., Abe, N., Rinaldi, L., McGregor, A.P., Frankel, N., Wang, S., Alsawadi, A., Valenti, P., Plaza, S., Payre, F., et al. (2015). Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* 160, 191–203.
- Curtiss, J., and Heilig, J.S. (1995). Establishment of Drosophila imaginal precursor cells is controlled by the Arrowhead gene. *Development* 121, 3819–3828.
- Debat, V., Debelle, A., and Dworkin, I. (2009). Plasticity, canalization, and developmental stability of the *Drosophila* wing: joint effects of mutations and developmental temperature. *Evolution* 63, 2864–2876.
- Degenhardt, K.R., Milewski, R.C., Padmanabhan, A., Miller, M., Singh, M.K., Lang, D., Engleka, K.A., Wu, M., Li, J., Zhou, D., et al. (2010). Distinct enhancers at the Pax3 locus can function redundantly to regulate neural tube and neural crest expressions. *Dev. Biol.* 339, 519–527.
- Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151–156.
- Driever, W., Thoma, G., and Nüsslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the bicoid morphogen. *Nature* 340, 363–367.
- Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., and Levine, M.S. (2015). Suboptimization of developmental enhancers. *Science* 350, 325–328.
- Flores, G.V., Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., Noll, M., and Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. *Cell* 103, 75–85.

- Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., and Stern, D.L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466, 490–493.
- Frankel, N., Ereyilmaz, D.F., McGregor, A.P., Wang, S., Payre, F., and Stern, D.L. (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474, 598–603.
- Frankel, N., Wang, S., and Stern, D.L. (2012). Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl. Acad. Sci. USA* 109, 20975–20979.
- Galant, R., Walsh, C.M., and Carroll, S.B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* 129, 3115–3126.
- Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M.L., Natoli, G., et al. (2010). Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Mol. Cell* 37, 418–428.
- Glassford, W.J., Johnson, W.C., Dall, N.R., Smith, S.J., Liu, Y., Boll, W., Noll, M., and Rebeiz, M. (2015). Co-option of an ancestral hox-regulated network underlies a recently evolved morphological novelty. *Dev. Cell* 34, 520–531.
- Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A., and Carroll, S.B. (2005). Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433, 481–487.
- Hammonds, A.S., Bristow, C.A., Fisher, W.W., Weiszmann, R., Wu, S., Hartenstein, V., Kellis, M., Yu, B., Frise, E., and Celtniker, S.E. (2013). Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biol.* 14, R140.
- He, X., Duque, T.S.P.C., and Sinha, S. (2012). Evolutionary origins of transcription factor binding site clusters. *Mol. Biol. Evol.* 29, 1059–1070.
- Hersh, B.M., and Carroll, S.B. (2005). Direct regulation of knot gene expression by Ultrabithorax and the evolution of cis-regulatory elements in *Drosophila*. *Development* 132, 1567–1577.
- Hiller, M., Schaar, B.T., and Bejerano, G. (2012). Hundreds of conserved non-coding genomic regions are independently lost in mammals. *Nucleic Acids Res.* 40, 11463–11476.
- Hu, S., Fambrough, D., Atashi, J.R., Goodman, C.S., and Crews, S.T. (1995). The *Drosophila* abrupt gene encodes a BTB-zinc finger regulatory protein that controls the specificity of neuromuscular connections. *Genes Dev.* 9, 2936–2948.
- Indjeian, V.B., Kingman, G.A., Jones, F.C., Guenther, C.A., Grimwood, J., Schmutz, J., Myers, R.M., and Kingsley, D.M. (2016). Evolving new skeletal traits by cis-regulatory changes in bone morphogenetic proteins. *Cell* 164, 45–56.
- Jackman, W.R., and Stock, D.W. (2006). Transgenic analysis of Dlx regulation in fish tooth development reveals evolutionary retention of enhancer function despite organ loss. *Proc. Natl. Acad. Sci. USA* 103, 19390–19395.
- Jenett, A., Rubin, G.M., Ngo, T.T., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., et al. (2012). A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* 2, 991–1001.
- Jeong, S., Rokas, A., and Carroll, S.B. (2006). Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* 125, 1387–1399.
- Jeong, S., Rebeiz, M., Andolfatto, P., Werner, T., True, J., and Carroll, S.B. (2008). The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132, 783–793.
- Johnson, W.C., Ordway, A.J., Watada, M., Pruitt, J.N., Williams, T.M., and Rebeiz, M. (2015). Genetic changes to a transcriptional silencer element confers phenotypic diversity within and between *Drosophila* species. *PLoS Genet.* 11, e1005279.
- Lifanov, A.P., Makeev, V.J., Nazina, A.G., and Papatsenko, D.A. (2003). Homotypic regulatory clusters in *Drosophila*. *Genome Res.* 13, 579–588.
- Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.-Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., et al. (2015). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 44, D110–D115.
- McGregor, A.P., Orgogozo, V., Delon, I., Zanet, J., Srinivasan, D.G., Payre, F., and Stern, D.L. (2007). Morphological evolution through multiple cis-regulatory mutations at a single gene. *Nature* 448, 587–590.
- McLean, C.Y., Reno, P.L., Pollen, A.A., Bassan, A.I., Capellini, T.D., Guenther, C., Indjeian, V.B., Lim, X., Menke, D.B., Schaar, B.T., et al. (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* 471, 216–219.
- Menoret, D., Santolini, M., Fernandes, I., Spokony, R., Zanet, J., Gonzalez, I., Latapie, Y., Ferrer, P., Rouault, H., White, K.P., et al. (2013). Genome-wide analyses of Shavenbaby target genes reveals distinct features of enhancer organization. *Genome Biol.* 14, R86.
- Nachman, A., Halachmi, N., Matia, N., Manzur, D., and Salzberg, A. (2015). Deconstructing the complexity of regulating common properties in different cell types: lessons from the delilah gene. *Dev. Biol.* 403, 180–191.
- Ni, J.-Q., Zhou, R., Czech, B., Liu, L.-P., Holderbaum, L., Yang-Zhou, D., Shim, H.-S., Tao, R., Handler, D., Karpowicz, P., et al. (2011). A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat. Methods* 8, 405–407.
- O'Brown, N.M., Summers, B.R., Jones, F.C., Brady, S.D., and Kingsley, D.M. (2015). A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. *Elife* 4, e05290.
- Ochoa-Espinosa, A., Yucel, G., Kaplan, L., Pare, A., Pura, N., Oberstein, A., Papatsenko, D., and Small, S. (2005). The role of binding site cluster strength in Bicoid-dependent patterning in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 102, 4960–4965.
- Parks, A.L., Cook, K.R., Belvin, M., Dompe, N.A., Fawcett, R., Huppert, K., Tan, L.R., Winter, C.G., Bogart, K.P., Deal, J.E., et al. (2004). Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nat. Genet.* 36, 288–292.
- Payne, J.L., and Wagner, A. (2015). Mechanisms of mutational robustness in transcriptional regulation. *Front. Genet.* 6, 322.
- Payre, F., Vincent, A., and Carreno, S. (1999). ovo/svb integrates Wingless and DER pathways to control epidermis differentiation. *Nature* 400, 271–275.
- Perry, M.W., Boettiger, A.N., Bothma, J.P., and Levine, M. (2010). Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* 20, 1562–1567.
- Prescott, S.L., Srinivasan, R., Marchetto, M.C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F.H., Swigut, T., and Wysocka, J. (2015). Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* 163, 68–83.
- Prud'homme, B., Gompel, N., and Carroll, S.B. (2007). Emerging principles of regulatory evolution. *Proc. Natl. Acad. Sci. USA* 104, 8605–8612.
- Rebeiz, M., Pool, J.E., Kassner, V.A., Aquadro, C.F., and Carroll, S.B. (2009). Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326, 1663–1667.
- Rockman, M.V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66, 1–17.
- Rowan, S., Siggers, T., Lachke, S.A., Yue, Y., Bulyk, M.L., and Maas, R.L. (2010). Precise temporal control of the eye regulatory gene Pax6 via enhancer-binding site affinity. *Genes Dev.* 24, 980–985.
- Salmandi, P.-A., Iché-Torres, M., and Perrin, L. (2011). Tissue-specific cell sorting from *Drosophila* embryos: application to gene expression analysis. *Fly (Austin)* 5, 261–265.
- Shim, S., Kwan, K.Y., Li, M., Lefebvre, V., and Sestan, N. (2012). Cis-regulatory control of corticospinal system development and evolution. *Nature* 486, 74–79.
- Shirangi, T.R., Dufour, H.D., Williams, T.M., and Carroll, S.B. (2009). Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. *PLoS Biol.* 7, e1000168.
- Small, S., Kraut, R., Hoey, T., Warrior, R., and Levine, M. (1991). Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes Dev.* 5, 827–839.

- Stanojevic, D., Small, S., and Levine, M. (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science* 254, 1385–1387.
- Stern, D.L. (2011). Evolution, Development, & the Predictable Genome (Roberts and Company Publishers).
- Stern, D.L., and Frankel, N. (2013). The structure and evolution of cis-regulatory regions: the shavenbaby story. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 368, 20130028.
- Struhl, G., Struhl, K., and Macdonald, P.M. (1989). The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* 57, 1259–1273.
- Sucena, E., and Stern, D.L. (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.
- Sucena, E., Delon, I., Jones, I., Payre, F., and Stern, D.L. (2003). Regulatory evolution of shavenbaby/ovo underlies multiple cases of morphological parallelism. *Nature* 424, 935–938.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578.
- Uhl, J.D., Cook, T.A., and Gebelein, B. (2010). Comparing anterior and posterior Hox complex formation reveals guidelines for predicting cis-regulatory elements. *Dev. Biol.* 343, 154–166.
- Uhl, J.D., Zandvakili, A., and Gebelein, B. (2016). A hox transcription factor collective binds a highly conserved distal-less cis-regulatory module to generate robust transcriptional outcomes. *PLoS Genet.* 12, e1005981.
- Van Moerkerke, A., Haring, M.A., and Schuurink, R.C. (2011). The transcription factor EMISSION OF BENZENOIDS II activates the MYB ODORANT1 promoter at a MYB binding site specific for fragrant petunias. *Plant J.* 67, 917–928.
- Visser, M., Kayser, M., and Palstra, R.-J. (2012). HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res.* 22, 446–455.
- Williams, T.M., Selegue, J.E., Werner, T., Gompel, N., Kopp, A., and Carroll, S.B. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134, 610–623.
- Yuh, C.H., and Davidson, E.H. (1996). Modular cis-regulatory organization of Endo16, a gut-specific gene of the sea urchin embryo. *Development* 122, 1069–1082.
- Zhu, L.J., Christensen, R.G., Kazemian, M., Hull, C.J., Enuameh, M.S., Basciotta, M.D., Brasefield, J.A., Zhu, C., Asriyan, Y., Lapointe, D.S., et al. (2011). FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res.* 39, D111–D117.