Full title: Pleiotropy in enhancer function is encoded through diverse genetic architectures.

Short title: Enhancer pleiotropy and *cis*-regulatory architecture

Ella Preger-Ben Noon[1¶], Gonzalo Sabarís[2¶], Daniela Ortiz[2], Jonathan Sager[1], Anna Liebowitz[3], David L. Stern[1*] and Nicolas Frankel[2*]

[1] HHMI, Janelia Research Campus, Ashburn, Virginia, USA

[2] Departamento de Ecología, Genética y Evolución, IEGEBA-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[3] Department of Molecular Biology, Princeton University, USA.

¶These authors contributed equally to this work (E.P.B.N. and G.S. are listed in alphabetical order)

*corresponding authors: David L. Stern <sternd@janelia.hhmi.org> and Nicolas Frankel nfrankel@ege.fcen.uba.ar

Abstract

Developmental genes can have complex *cis*-regulatory regions, with multiple enhancers scattered across stretches of DNA spanning tens or hundreds of kilobases. Early work revealed remarkable modularity of enhancers, where distinct regions of DNA, bound by combinations of transcription factors, drive gene expression in defined spatio-temporal domains. Nevertheless, a few reports have shown that enhancer function may be required in multiple developmental stages, implying that regulatory elements can be pleiotropic. In these cases, it is not clear whether the pleiotropic enhancers employ the same transcription factor binding sites to drive expression at multiple developmental stages or whether enhancers function as chromatin scaffolds, where independent sets of transcription factor binding sites act at different stages. In this work we have studied the activity of the enhancers of the *shavenbaby* gene throughout *D. melanogaster* development. We found that all seven *shavenbaby* enhancers drive gene expression in multiple tissues and developmental stages at varying levels of redundancy. We have explored how this pleiotropy is encoded in two of these enhancers. In one enhancer, the same transcription factor binding sites contribute to embryonic and pupal expression, whereas for a second enhancer, these roles are largely encoded by distinct transcription factor binding sites. Our data suggest that enhancer pleiotropy might be a common feature of *cis*-regulatory regions of developmental genes and that this pleiotropy can be encoded through multiple genetic architectures.

Introduction

Developmental genes can have complex *cis*-regulatory regions, with multiple enhancers scattered across stretches of DNA spanning tens or hundreds of kilobases [1-4]. Over many years, numerous studies have revealed a remarkable modularity of enhancer function, where distinct regions of DNA, bound by combinations of transcription factors, drive gene expression in defined spatio-temporal domains [5]. It has long been hypothesized that enhancer modularity facilitates evolution, because mutations in one enhancer can alter gene function without affecting the activity of other enhancers, thereby minimizing pleiotropic effects [6-8]. It is not clear, however, if the apparent modularity of enhancers reflects ascertainment bias, since few studies have looked explicitly for enhancer pleiotropy.

Many of the genes that regulate development have a pleiotropic role and their function is required in multiple developmental stages. A paradigmatic case of pleiotropy is that of *Hox* genes, a family of master transcription factors that specify the identity of body parts [9]. Recently, it has been uncovered that the same mechanism activates *Hox* genes in different organs of the mouse: the same enhancers activate *Hox* genes in both digits and genitalia [10]. Clearly, this implies that enhancers can have pleiotropic functions. It is not evident, however, whether these enhancers employ the same transcription factor binding sites to drive expression at multiple developmental stages or whether enhancers function as chromatin scaffolds, where independent sets of binding sites act at different stages.

shavenbaby (*svb*) encodes a transcription factor that orchestrates the differentiation of non-sensory cuticular projections (hereafter called trichomes) in *Drosophila melanogaster*

[11,12]. *Svb* expression has been studied in detail mainly in the late embryonic stages, where it directs development of the epidermis and, concomitantly, the first-instar larval cuticle ([13], Fig 1B). S*vb* is also expressed in the pupal epidermis where it is required for trichome development in part of the wing, notum and abdomen [14,15] and for proper development of leg joints [16].

The *cis*-regulatory region of the *svb* gene has been experimentally dissected in *D. melanogaster* [17-21]. We have shown that the embryonic expression of *svb* is generated by seven enhancers that are located in a ~80 kb region upstream of the transcription start site of the gene ([11], Fig 1A). These seven enhancers drive partially overlapping expression patterns in the late embryo, and these overlapping patterns are required for robust gene expression [20]. Evolutionary changes in five of these enhancers led to reduced *svb* expression in the dorsum of the *D. sechellia* embryo, resulting in differentiation of naked cuticle, rather than trichomes, in *D. sechellia* [11,20,21].

In this work we show that all seven *svb* enhancers drive gene expression in multiple tissues and developmental stages at varying levels of redundancy. We have explored how this pleiotropy is encoded in two of these enhancers. In one enhancer, the same transcription factor binding sites contribute to embryonic and pupal expression, whereas for a second enhancer these roles are largely encoded by distinct sites. Our data suggest that enhancer pleiotropy might be a common feature of *cis*-regulatory regions of developmental genes, and that this pleiotropy can be encoded through multiple genetic architectures.

Results

Shavenbaby is expressed in the larval and pupal epidermis

To characterize the expression of *svb* in larval and pupal tissues, we engineered a BAC carrying the complete *cis*-regulatory region of *svb* by placing the coding sequence of a nuclear GFP upstream of *svb* ATG (Fig 1A). We stably integrated this BAC, named *svb*BAC-GFP, in the fly genome through attP/attB recombination. We confirmed that *svb*BAC-GFP recapitulates expression of the native gene in embryos (Fig 1B). This epidermal expression prefigures the location of trichomes in the first-instar larva cuticle (Fig 1C). We then examined *svb* expression in later stages. We observed GFP expression in the epidermis of third-instar larvae (data not shown). This may reflect persistence of the GFP reporter from second-instar larvae, when *svb* expression is probably required to cause differentiation of trichomes that will decorate the cuticle of third-instar larvae. We also detected GFP expression in larval non-epidermal structures of ectodermal origin that do not produce trichomes. Specifically, we observed GFP in the central nervous system (Fig 1D), the foregut (Figs 1E-F), the imaginal discs (Fig 1G), and the trachea (data not shown). We also found that *svb*BAC-GFP pupae display GFP expression in all epidermal tissue (Fig 1H), which is consistent with the fact that the adult exoskeleton is almost completely covered with trichomes (Fig 1I).

Shavenbaby is required for the formation of many, but not all, adult trichomes

Next we asked whether *svb* function is required for trichome development in the pupal epidermis, as it is in the embryo. Flies carrying *svb* null mutations normally die before they eclose, but we identified a few male escapers (carrying a null *svb* allele on their single X chromosome) that allowed us to assess the requirement of *svb* for trichome development in the adult cuticle. Escapers had fewer trichomes in the wing, the legs, and the dorsal abdomen than control flies, but they still retained trichomes over much of the exoskeleton (Fig 2). However, in most regions, trichomes were smaller than normal or were misshapen.

We observed that no trichomes developed in male *svb* escapers on the dorsal abdominal segment 5 (compare Figs 2H and 2J). This observation stimulated a detailed inspection of the wild type trichome pattern and we found a sexual dimorphism in the shape and size of trichomes in the dorsal abdomen: females produce trichomes of similar density and stoutness on abdominal segments A1 through A5 (S1A Fig), whereas males produce qualitatively different trichomes on abdominal segments A1-A4 vs A5 (S1A Fig). We observed that *svb* expression is lower in abdominal segment 5 versus more anterior segments in both sexes and this difference may contribute to the sexual dimorphism in trichome patterning (S1B Fig).

We confirmed the results observed with the male escapers by generating *svb-/-* clones in the adult. We observed loss of trichomes in *svb-/-* clones in the same regions where trichomes were lost in male escapers (see S2A Fig for an example), confirming that *svb* function is required for the production of some adult trichomes. We also observed loss of trichomes, change in trichome morphology and altered trichome distribution in the head cuticle (S2B Fig). Loss of *svb* function also modified the anatomy of the antennal arista (S2

Fig). In summary, although *svb* is expressed throughout the pupal epidermis (Fig 1H), it is required for the normal development of many, but not all, adult trichomes.

The embryonic enhancers of svb drive expression in larva and pupa

Since the BAC containing the complete *svb* regulatory region drives expression in embryonic, larval and pupal stages, we wondered whether the previously characterized embryonic enhancers also drove expression in later stages. In the third-instar larvae, we found that of the seven embryonic *svb* enhancers, five drove expression in the epidermis, six drove expression in the foregut and four drove expression in the central nervous system of L3 (Fig 3A and S3 Fig). These results are consistent with the expression of the *svb*BAC-GFP (Fig 1). In the pupa at 90 h after puparium formation (APF), all embryonic *svb* enhancers drove widespread epidermal expression (Fig 3A-B and S3 Fig.). Most notably, all seven enhancers drive expression throughout the dorsal abdomen (Fig 3B). This level of overlapping expression far exceeds patterns of overlapping embryonic expression that we reported previously [20]. Hence, the seven embryonic *svb* enhancers are both pervasively pleiotropic and redundant.

Pupal expression of shavenbaby is conserved in D.sechellia

We have previously shown that five of the *svb* enhancers evolved reduced embryonic activity in *D. sechellia*, a species closely related to *D. melanogaster* [11,20,21]. This loss of

enhancer function reduces *svb* expression in *D. sechellia* embryos, causing the loss of many first-instar trichomes [11,17]. In contrast, *D. sechellia* adults, like *D. melanogaster* adults, are completely covered with trichomes (data not shown). To test whether pupal *svb* expression was conserved in *D. sechellia*, we generated a *D. sechellia svb*BAC-GFP with the same genomic boundaries as the *D. melanogaster svb*BAC-GFP (S4A Fig). The *D. sechellia svb*BAC-GFP recapitulated the embryonic expression pattern of *D. sechellia svb*, and no expression was detected in quaternary cells of the dorsal and lateral epidermis (S4B Fig). In contrast, the *D. sechellia svb*BAC-GFP drove GFP expression throughout the dorsal and ventral pupal epidermis, just like the *D. melanogaster svb*BAC-GFP (S4C Fig). Therefore, it is likely that at least some of the *D. sechellia svb* enhancers that lost embryonic expression still drive expression in pupa. To explore this problem further, we examined the embryonic and pupal functions of two evolved *svb* enhancers in more detail.

The same transcription factor binding sites within E6 are used in both embryo and pupa

We showed previously that the *D. melanogaster E6* enhancer (*melE6*) encodes multiple transcription factor binding sites for the transcriptional activators *Arrowhead* (*Awh*) and *Pannier* (*Pnr)* (Fig 4A-B, [17]). *D. sechellia E6* (*secE6*) lost four *Awh* sites and acquired a transcription factor binding site for the strong repressor *abrupt*, causing complete loss of its embryonic function (Fig 4C, [17]). We exploited this evolutionary transition to explore pleiotropic roles of *E6*.

176 To compare the activities of *melE6* and *secE6* in pupa, we generated new transgenes
177 with fluorescent reporters that were integrated in the same genomic location. We reasoned
178 that if transcription factor binding sites required for embryonic function are required also for
179 pupal expression, then *secE6* should not drive pupal expression. Consistent with our previous
180 observations (Fig 3), the new *melE6* transgene drove strong expression in the wing and dorsal
181 abdomen of the developing pupa (Fig 4D and 4F). In contrast, *secE6* did not drive expression
182 in most of the pupal domains (Fig 4E and 4G). Thus, the evolutionary changes in transcription
183 factor binding sites of *secE6* that led to reduced embryonic expression also reduced pupal
184 expression.

185 Next, we tested the contributions of individual classes of transcription factor binding
186 sites. We examined a mutated version of the minimal enhancer, *E6B* (Fig 4A), which lacks six
187 *Awh* sites. Like the full length *E6, E6B* drove both embryonic and pupal expression (Figs 4H-J,
188 [17]. Disruption of the *Awh* sites from *E6B* eliminated most pupal abdominal expression (Fig
189 4K). Additional mutations in the *Pnr* sites led to an even stronger reduction in pupal
190 expression (Fig 4L). All together, these experiments demonstrate that transcription factor
191 binding sites within *E6* are 'reused' at multiple developmental stages.

192

193 *Different regulatory information generates the various expression patterns of enhancer Z*

194

195 Next, we analyzed the *Z* (*melZ*) enhancer, another *svb* enhancer whose embryonic
196 activity was lost in *D. sechellia* [20]. We dissected the ~4.4 Kb of *melZ* and identified a ~1.3 Kb
197 minimal enhancer we named *melZ1.3* (S5 Fig). As expected, the orthologous sequence from

198 *D. sechellia* (*secZ1.3*) did not drive embryonic expression (Fig 5B, [20]). *melZ1.3* recapitulated
199 the full *Z* expression pattern in pupae (Fig 3), including expression in the wings (Fig 5C), legs
200 (Fig 5E), notum and abdomen (Fig 5G). In contrast to our observations for the *E6* enhancer, we
201 found that the *secZ1.3* enhancer drove expression in all tissues where *melZ1.3* is active (Figs
202 5D, 5F and 5H). Therefore, the evolutionary changes that led to the loss of *Z* expression in *D.*
203 *sechellia* embryos did not alter *Z* function in pupae.

204 To test whether the *Z* enhancer is divided into discrete modules that drive expression
205 at different developmental stages, we dissected *melZ1.3* into smaller fragments (Fig 5I) and
206 tested their ability to drive expression in embryos and pupae. We found that embryonic
207 expression is encoded mostly in *melZ1.3L* (Figs 5J-K), while pupal expression is encoded
208 mainly by *melZ1.3R* (Figs 5L-O and S5 Fig). In agreement with the results for *secZ1.3*, we found
209 that *secZ1.3R* drove expression in pupal epidermis (data not shown). Thus, the embryonic and
210 pupal expression patterns are encoded by adjacent sequences in the *Z* enhancer region.

211 The embryonic and pupal enhancer domains are not strictly separated, however. We
212 identified a 300 bp fragment named *melZ0.3* that drives strong embryonic expression (Fig 5P
213 and S5 Fig). Interestingly, *melZ0.3* drove stronger expression than the longer *melZ1.3* (Fig 5Q),
214 suggesting that regions outside *melZ0.3* contain binding sites for transcriptional repressors,
215 including sites within the pupal enhancer region *melZ1.3R* (Fig 5R). Nonetheless, the
216 transcription factor binding sites that activate gene expression appear to be present in non-
217 overlapping adjacent domains. Thus, the *Z* enhancer generates its many functions with
218 different binding sites, which are located in mostly non-overlapping regions.

219

220 *Deletion of individual enhancers has contrasting outcomes in embryo and pupa*

221

222 In recent years it has become evident that the expression of many developmental
223 genes is controlled by multiple enhancers with redundant functions [22]. We have previously
224 demonstrated that the *svb* enhancers drive partially redundant expression patterns in the
225 embryo (Fig 3A), providing phenotypic robustness for larval trichome patterns in the face of
226 environmental and genetic variation [20]. Remarkably, pupae display even greater
227 redundancy of *svb* enhancer expression than embryos (Fig. 3B). We therefore decided to
228 explore the functional consequences of this redundancy.

229 We used BAC recombineering to generate deletions of individual enhancers (*Z1.3, E6*
230 and *7H*) in the *svb*BAC-GFP, and integrated these BACS in a specific *attP* site of the *D.*
231 *melanogaster* genome (Fig 6A). As an internal control we used a wild-type *svb*BAC with a
232 DsRed reporter (*svb*BAC-DsRed) that was integrated into a different *attP* site (to avoid
233 transvection effects between BACs). We then quantified expression patterns of the BACs
234 carrying deletions (expressing GFP) relative to the control BAC (expressing DsRed) in the same
235 animal (Fig 6, see Materials and Methods for details).

236 Removing *Z1.3*, *E6*, or *7H* resulted in a decrease of the mean GFP expression in
237 embryos of 28%, 46%, and 38% respectively (Fig 6B). However, none of the BACs with
238 enhancer deletions drove reduced reporter expression in the pupal epidermis of abdominal
239 segment A4 (Fig 6C). On the contrary, we determined that the deletion of single enhancers
240 slightly increases reporter expression. This fact suggests that the function of *svb* enhancers
241 changes during development.

242

243 *Deletion of multiple enhancers in the svb locus does not affect the adult trichome pattern*

244

245 To examine the importance of this redundancy on the phenotypic output of *svb*
246 function we examined the effects of several deficiencies that remove part of the *svb*
247 regulatory region in the native locus. We used *Df(X)svb*^{108}, a deletion of the three most distal
248 enhancers (*DG2*, *DG3* and *Z*, [20]) and a larger deletion, *Df(X)svb*^{106}, that removes the four most
249 distal enhancers (*DG2*, *DG3*, *Z* and *A*, Fig 7A). We showed previously that the *Df(X)svb*^{108} line
250 produces a normal number of first-instar larval trichomes when embryos develop at their
251 optimal temperature of 25°C [20]. However, when embryos are grown at a stressful
252 temperature, 32°C, they develop with significantly fewer larval trichomes [20]. We found that
253 *Df(X)svb*^{106} produces similar results (data not shown).

254 We could not find any gross changes in the trichome pattern of females (data not
255 shown) when pupae are grown at 25°C. We did notice, however, a small but consistent
256 trichome defect in males carrying either deficiency grown at 25°C: the dorsum of abdominal
257 segment A5 had fewer trichomes than wild type males (Fig 7B). This result is consistent with
258 the phenotype of male *svb* escapers (Fig 2) and the observation that *svb* is expressed in this
259 abdominal segments at lower levels (S1B Fig). Growing *Df(X)svb*^{108} or *Df(X)svb*^{106} pupae at 32°C
260 did not alter the adult trichome pattern (data not shown). Hence, the adult trichome pattern
261 is largely robust to removal of up to four of the seven *svb* enhancers. In contrast to the effect
262 of these deficiencies on first-instar larvae, stressful growth temperatures in pupa do not
263 significantly alter adult trichome development. Finally, adult *D. sechellia* males displayed an

264 A5 trichome pattern identical to that of *D. melanogaster* (Fig 7B) despite the fact that *secE6*
265 drives reduced expression in the pupal epidermis.
266

267 **Discussion**

268 The *svb* gene encodes a master transcription factor that determines the fate of
269 epidermal cells in *Drosophila melanogaster* [14]. It has been known for some time that the
270 embryonic activity of SVB is necessary to pattern the first-instar larva cuticle [13]. In this work
271 we show that *svb* is expressed in several structures of third-instar larvae which have an
272 ectodermal origin (foregut, central nervous system, imaginal discs and epidermis).
273 Furthermore, guided by previous research [14,15], we demonstrate that *svb* is expressed all
274 over the pupal epidermis, and that this expression is required for the development of most of
275 the adult trichomes.

276 In this work we demonstrate that the seven enhancers of *svb* that drive expression in
277 the embryo also generate expression in third-instar larva and pupa. Thus, the seven
278 enhancers of *svb* have a pleiotropic role during development. Recently, studies of chromatin
279 conformation showed that the same genomic regions are active in the regulation of gene
280 expression in both developing limb and developing genitalia of mouse [10]. These data
281 strongly suggested that enhancers might be used in entirely different developmental
282 contexts [10]. Later studies reported that the HLEB enhancer of the *Tbx4* gene functions
283 during both limb and genitalia formation in mice, corroborating the idea that the same
284 regulatory element can be active in two dissimilar contexts [23]. Altogether, our data and
285 previous reports [10,23,24] suggest that pleiotropy of enhancer function might be a common

feature of *cis*-regulatory regions. A possible explanation for the pleiotropic role of enhancers is that these regions of the genome are structurally or topologically special, and that their qualities facilitate the interaction with basal promoters. This idea is supported by the fact that there is conservation in the position of enhancers in distant lineages [25,26]. Alternatively, it can be hypothesized that new expression patterns are easier to evolve within pre-existent regulatory landscapes and, thus, there is nothing special about the position of enhancers in the genome.

There are two fundamentally distinct models by which pleiotropic enhancers could encode expression in different spatio-temporal domains. First, the same transcription factor binding sites could be used to drive expression in different domains or, second, distinct transcription factor binding sites within the same enhancer region could drive different patterns. We studied two pleiotropic *svb* enhancers in detail and found one example that supports each model. For example, we find that Awh and Pnr binding sites, which activate *E6* in the embryo, are also needed to activate *E6* in the pupa. Similarly, it has been shown that Abd-B and STAT binding sites are required for the function of a *Poxn* enhancer in two developmental contexts [27]. In contrast, in the *svb Z1.3* enhancer, we find that transcription factor binding sites that function in the embryo in pupa are mostly independent. This type of architecture might explain how the *Tbx4* enhancer lost its hind-limb function in snakes without losing its activity in genitalia [23]. In summary, we found that there are multiple genetic architectures through which single enhancers can drive pleiotropic expression patterns.

Our results suggest that the classical view of enhancers as strongly modular genomic elements should be reevaluated: transcription factor binding sites can be 'reused' during development and enhancer pleiotropy seems to be a common phenomenon. These facts allow us to challenge the notion that conceives enhancers as elements that are active in a single developmental context and evolutionarily unconstrained. Often, enhancer function is schematized with a univocal relationship between a DNA fragment and an expression pattern (one enhancer gives only one expression pattern; for example see [28-30]). This schematization, though generally made to illustrate a concept, conveys the wrong idea of enhancers as always being active in a single context. On the other hand, the 'reuse' of transcription factor binding sites may impose a limit for enhancers to evolve new functions. Thus, pleiotropic enhancers may sometimes constrain and sometimes facilitate evolution, depending on precisely how pleiotropy is encoded.

We observed extensive redundancy of *svb* enhancer activity in pupal stages. This redundancy far exceeds the redundancy we had characterized previously for the embryonic expression pattern, which is required for phenotypic robustness [17,20]. In fact, we observed that deleting individual enhancers has contrasting outcomes in embryo and pupa. Whereas in embryos the loss of one enhancer diminishes gene expression, in pupa the lack of a single enhancer generates a slight increase in gene expression. Furthermore, flies carrying only three of the total seven *svb* enhancers still produce largely normal adult trichome patterns, even when pupae are grown under stressful conditions. In summary, the function of *svb cis*-regulatory region in pupa appears to result in strong robustness of the adult trichome pattern.

We have shown that *svb* enhancers are pleiotropic and that their expression is highly redundant. Indeed, in *D. sechellia* these enhancers drive enough pupal *svb* expression through stage-specific transcription factor binding sites that the embryonic expression pattern was free to evolve without altering the adult expression pattern. However, it is also possible to imagine a scenario with less redundancy and where pleiotropy is encoded in enhancers through the same transcription factor binding sites (as in the case of enhancer *E6*), which would strongly constrain the evolution of expression patterns. At present, it is unclear how many enhancers in the genome are pleiotropic, and how their pleiotropy tends to be encoded. Further studies should help determine the extent and encoding of enhancer pleiotropy, clarifying the potential role of enhancer pleiotropy in evolution.

Materials and Methods

Genetic constructs and transgenesis

P[acman] CH321-64E24 (https://bacpacresources.org) contains a 91,307 bp. insert that includes the *cis*-regulatory region, the first exon and part of the first intron of *shavenbaby* (https://www.ncbi.nlm.nih.gov/clone/33521512). We used BAC recombineering [31] to insert a GFP-NLS or a DsRed-NLS in the initiation codon of *svb* to generate *svb*BAC-GFP and *svb*BAC-DsRed and to delete specific enhancers in the context of *svb*BAC-GFP. All primers and constructs that were used for BAC recombineering are summarized in S1 Table.

The *D. sechellia svb* gene, including the *cis*-regulatory region, the first exon and part of the first intron (droSec1: super_4:1,797,878-1,880,229) was subcloned from the BAC DSE1-

007L13 (RIKEN BioResource Center DNA Bank) into P[acman]. Subsequently, BAC recombineering was used to insert a GFP-NLS in the initiation codon of *svb* to generate *sec-svb*-BAC-GFP.

The GFP-NLS in pS3AG (a gift from Thomas Williams, addgene plasmid # 31171) was replaced with a DsRed-NLS to generate pS3AR. The DsRed-NLS was released from pRed H-Stinger with enzymes XhoI and SpeI. GFP-NLS was removed from pS3AG by cutting with enzymes XhoI and SpeI. The pS3AG backbone (without GFP-NLS) was then ligated to the DsRed-NLS. All other transgenes generated in this study were constructed by Genscript (summarized in S2 Table). These constructs were integrated into the fly genome through attP/attB recombination (Rainbow Transgenic flies).

Fly strains

Enhancer-reporter lines are summarized in S2 Table. In order to generate *Df(X)svb*^{106}, pBacPtp4E[f02952] and pBac[f06356] were recombined onto the same X chromosome and a homozygous stock was generated. This stock was crossed to a line containing a hs::flipase and larvae were heat shocked at 37°C for 1 h each day during larval development. After crossing these adults to *white* flies, we selected adults that had lost one copy of the white+ transgene (originating on one of the pBac transgenes), which is expected if the two FRT sites recombined to generate a deletion. The deletion was confirmed by PCR, with primers located right outside the deletion (5'- CGTACCGCCTGTTTGCCATA-'3 and 5'-TCCAGACGGATTTTATGGCC-'3), which amplified the expected 7.3 kb fragment containing a

373 pBac transposon. We then generated a stock homozygous for the deletion. *Df(X)svb*^{108} was
374 previously described [20]. *D.sechellia* 14021-0248.28 was obtained from the Drosophila
375 Species Stock Center at the University of California.
376
377 We generated large clonal territories of *svb*$^{-}$ tissue by employing the Minute technique
378 (Morata & Ripoll, 1975). With this technique clones that contain two wild-type alleles of
379 *Minute*$^{+}$ over-proliferate relative to neighboring cells that are heterozygous for a *Minute* null
380 mutation. To mark *svb*$^{-}$ tissue, we recombined three visible markers (y^{1}, w^{1}, and f^{36a}) onto a
381 chromosome together with a null mutation for *svb* (*svb*^{1}), to generate y^{1} w^{1} *svb*^{1} f^{36a}. We then
382 crossed this strain to flies carrying a dominant Minute allele on the X chromosome. We
383 exposed larvae carrying the genotype y^{1} w^{1} *svb*^{1} f^{36a} / M to X-Rays (1000 Rad) between 24-72
384 hours after egg laying. We screened females for clones homozygous for *svb*^{1} by searching for
385 cuticle containing bristles that were both yellow and forked. We compared trichome patterns
386 in these clones with trichome patterns on flies homozygous for the f^{36a} allele.
387
388 *X-gal staining*
389
390 Third instar larvae were dissected in PBS and fixed in PBS with 4% formaldehyde for 10
391 min. Staged pupae were removed from the pupal case and then fixed in PBS with 4%
392 formaldehyde for 15 min. After washing in PBT (1X PBS + 0.1% Triton X-100), samples were
393 incubated with X-Gal solution (5 mM $K_4[Fe^{+2}(CN)_6]$, 5 mM $K_3[Fe^{+2}(CN)_6]$, 1 mg/ml X-Gal in PBT)
394 at 37°C for 1 hour. The samples were mounted and imaged with bright-field microscopy.

Immunofluorescence

Stage 15 embryos were collected, fixed, and stained using standard protocols with chicken anti-GFP (1:300, Aves Labs), rabbit anti-RFP (1:150, MBL), mouse anti-βGal (1:500, Promega), anti-chicken AlexaFluor 488 (1:250, ThermoFisher), anti-rabbit AlexaFluor 647 (1:150, ThermoFisher) and anti-mouse AlexaFluor 546 (1:400, ThermoFisher). Pupal tissues were dissected, fixed and stained with mouse anti-βGal and anti-mouse AlexaFluor 546 as described [32].

Microscopy and image analysis

Embryos were prepared using standard protocols and immunostained with the antibodies described above. Pupae of the desired stages were removed from the pupal case and placed in a microscope slide for imaging. To analyze the effect of enhancer deletions in *svb*BACs we measured GFP and DsRed levels in embryos and pupae carrying *svb*BAC-GFP (WT and deletions) and *svb*BAC-DsRed (WT). GFP and DsRed signals were measured sequentially over a z-stack in a confocal microscope. Images were analyzed using ImageJ software (http://rsb.info.nih.gov/ij/). Firstly, background was subtracted using a 50-pixel rolling-ball radius in each slice of the confocal z-stack. Then, we calculated the Sum projection of the z-stacks for each channel in order to compare GFP versus DsRed levels. Max projections were obtained in order to analyze GFP levels between abdominal segments A4 and A5. For embryos, we applied the segmentation masks using the Sum projection of the DsRed channel with the ImageJ autothreshold tool ("IJ_IsoData dark"). For pupal abdomens, segmentation

masks were applied with Ilastik 1.2.0 software (http://ilastik.org) to Sum projections of the GFP channel (GFP versus DsRed levels in A4) and Max projections (GFP levels in A4 and A5). We measured the fluorescence mean intensities of each nucleus with the 'Analyze particles' tool in ImageJ. Then, we calculated the average of the fluorescence mean intensity of all segmented nuclei. Last, we calculated the ratio GFP/DsRed in each nucleus and calculated the average ratio for all segmented nuclei.

Cuticle preparation

Adults were collected and frozen until used. Adult cuticles were dissected in PBS and mounted in a microscope slide with a drop of 1:1 Hoyer's:lactic acid mixture. After overnight drying, the cuticles of adults were imaged with bright-field microscopy. The images were processed using Adobe Photoshop.

Acknowledgements

We thank François Payre for providing the f^{36a} / FM6 stock. We thank Xiaorong Zhang of the Janelia Molecular Biology Shared Resource for help with the *sec-svb*-BAC-GFP recombineering and the Janelia Fly core facility for help with fly work. We thank the Bioimaging Core Facility at the Technion Rappaport Faculty of Medicine for help with imaging. E.P.B.N is grateful for the generous hospitality of Adi Salzberg and her lab members. E.P.B.N was supported by post-doctoral fellowships from the Human Frontier Science

Program. This work was supported in part by Fundación Bunge y Born and Agencia Nacional de Promoción Científica y Tecnológica (PICT 2013-2138) grants to N.F.

Figure captions

Fig 1. *svb* expression throughout *Drosophila melanogaster* development.

(A) Schematic representation of *svb*BAC-GFP. Gray boxes represent the seven embryonic enhancers. The site of insertion of the GFP-NLS is indicated in the scheme. (B) GFP expression recapitulates the expression pattern of *svb* in the embryo. (C) Trichome pattern of the first-instar larva. (D-G) GFP expression in non-epidermal structures of the third-instar larva: central nervous system (D), pharynx and salivary glands (E), esophagus and proventriculus (F), and wing imaginal disc (G) DAPI stain in magenta. (H) GFP expression in pupal epidermis. (I) Representation of the trichome pattern in the dorsum of an adult fly.

Fig 2. *svb* is required for the production of trichomes in the adult cuticle.

(A-J) The cuticle of control f[36a] adult wing (A-B), leg (E) and abdomen (G-H) is covered with trichomes. In *svb* null male escapers (f[36a], svb^{1}/Y, C-D, F and I-J) many trichomes, but not all, are replaced by naked cuticle. A complete loss of trichomes is observed in legs (F) and abdominal segment A5 (J). Blue boxes within the cartoon demarcate the imaged area.

Fig 3. Pleiotropy and redundancy in the activity of the seven *svb* embryonic enhancers.

(A) Schematization of the expression pattern driven by each enhancer (blue) in embryo (top), third-instar larva (middle) and pupa (bottom) (see S2 Fig for details). E: Epidermis, F: Foregut, ph: pharynx, e: esophagus, pv: proventriculus, C.N.S: Central Nervous System. (B) Expression pattern generated by each enhancer in the dorsal epidermis of the head, thorax and abdomen (90 hours APF).

Fig 4. Reuse of TFBSs in the pleiotropic enhancer *E6*.

(A) Scheme of *D. melanogaster E6* and *E6B* enhancers. Cyan and orange ovals represent TFBSs for *Awh* and *Pnr*, respectively. (B-C) Expression driven by *D. melanogaster E6* (B, *melE6*) and *D. sechellia E6* (C, *secE6*) in *D. melanogaster* stage 15 embryos. (D-G) Expression driven by *melE6* (D and F) and *secE6* (E and G) in pupal wings (D-E, 74 h APF) and dorsal abdomen (F-G, 84 h APF). (H-L) Expression driven by *melE6* (H), *D. simulans E6* (I, *simE6*), *D. simulans E6B* (B, J, *simE6B*) and *simE6B* mutants (K and L) in pupa (74 h APF). Red crosses in the schemes below images indicate mutated TFBSs (K, L).

Fig 5. The regulatory information in enhancer *Z* is partially separated in two modules.

(A-B) Expression driven by *D. melanogaster Z1.3* (A, *melZ1.3*) and *D. sechellia Z1.3* (B, *secZ1.3*) in *D. melanogaster* stage 15 embryos. (C-H) Expression of *melZ1.3* (C, E, G) and *secZ1.3* (D, F, H) in *D. melanogaster* pupal wings (C-D, 36 h AFP), legs (E-F, 36 h APF) and dorsal epidermis of the pupa (G-H, 74 h AFP). (I) Scheme of a subset of the *melZ1.3* enhancer fragments tested with transgenic reporter constructs. (J-P) Expression driven by *D.melanogaster Z1.3L* (*melZ1.3L*), *D.melanogaster Z1.3R* (*melZ1.3R*) *and D.melanogaster Z0.3* (*melZ0.3*) in stage 15 embryos (J, K,

P), pupal wings (L-M, 36 h AFP) and legs (N-O, 36 h AFP). (Q) Quantification of reporter activity in embryos (n=10) carrying *melZ1.3L, melZ1.3R and melZ0.3* reporter constructs. Mean intensity is shown with a black cross. (R) Genetic architecture of *melZ1.3*. *Z0.3* (dashed rectangle) carries most of the regulatory information used for embryonic expression, while Z1.3R (solid rectangle) contains the regulatory information used for pupal expression. Putative binding sites for a transcriptional repressor, which acts in the embryo, are indicated with red hexagons.

Fig 6. *Cis*-regulation of *svb* varies between embryo and pupa.

(A) Wild type and mutated versions of the *svb*BAC. The green and red triangles depict the coding sequence of GFP and DsRed, respectively. (B) Effect of enhancer deletions in embryonic expression; Δ*Z1.3* (top), Δ*E6* (middle) and Δ*7H* (bottom). Open circles indicate the average ratio (GFP/DsRed) for each individual. Closed black circles and vertical lines indicate mean and one standard deviation, respectively. P values were calculated with two-tailed unpaired t-tests (*p<0.05, **p<0.005, ****p<0.0001). Boxes within embryo cartoons specify analyzed regions. (C) Effect of enhancer deletions on pupal expression. The GFP/DsRed ratio was measured in part of abdominal segment A4 (rectangle) of pupae 90 hours APF. Open circles indicate the average ratio (GFP/DsRed) for each individual. Closed black circles and vertical lines indicate mean and one standard deviation, respectively. P values were calculated with two-tailed unpaired t-tests (*p<0.05, **p<0.005, ****p<0.0001).

Fig 7. Enhancer deletions in the native *svb* locus alter the A5 trichome pattern only in males.

(A) Diagram of the *svb* locus showing the genomic deletions on the X chromosome of lines *Df(X)svb^{108}* and *Df(X)svb^{106}*. Deletion in *Df(X)svb^{108}* removes enhancers *DG2*, *DG3* and *Z*, while *Df(X)svb^{106}* deletion removes enhancers *DG2*, *DG3*, *Z* and *A*. (B) A4 and A5 trichome pattern in adult males of *D. melanogaster*, *Df(X)svb^{108}*, *Df(X)svb^{106}* and *D. sechellia*. Blue boxes within the cartoon demarcate the imaged area.

Supporting information

S1 Fig. *svb*BAC-GFP expression in pupa and adult trichome patterns

(A) Adult trichome pattern in abdominal segments A3 to A6 in female (left) and male (right). (B) *svb*BAC-GFP expression in the dorsum of abdominal segments A2 to A5 of pupae 50 hours APF (left). GFP fluorescence quantification in segments A4 and A5 (right). Each point represents the average mean intensity of all segmented nuclei in A4 and A5 for one individual (n=7). Significance was calculated with a two-tailed paired t-test, *** $p<0.0005$.

S2 Fig. *svb* is required for the formation of adult cuticular structures.

(A-B) *Svb$^-$* clones on T1 leg (A) and head (B) are outlined in cyan. (C-D) Wild type (C) and *svb-* clone modified (D) antennal arista.

S3 F. Expression driven by the seven *svb* embryonic enhancers in larva and pupa.

(A-B) X-Gal staining of dissected tissues for each enhancer. (A) Enhancer expression in foregut, central nervous system and epidermis of third-instar larva. (B) Enhancer expression in leg 1 and pupal wing. N.E.: No expression.

S4 Fig. *shavenbaby* (*svb*) expression in different developmental stages of *Drosophila sechellia*. (A) Schematic representation of *D. sechellia svbBAC-GFP* (*sec-svbBAC-GFP*). Gray boxes represent the seven embryonic enhancers. The site of insertion of the GFP-NLS is indicated in the scheme. (B) *sec-svb-BAC-GFP* expression in stage 15 embryo. Quaternary cells on abdominal segment A2 are outlined. (C) sec-*svbBAC-GFP* expression in dorsal (left) and ventral (right) epidermis of a 90 hour APF pupa.

S5 Fig. Dissection of the embryonic and pupal functions of *svb Z* enhancer.

(A) Scheme of fragments tested for epidermal enhancer activity with reporter constructs (upper panel). Yellow boxes show fragments with enhancer activity (representative stage 15 embryos are shown in the bottom). Gray boxes indicate no expression. (B) Schematic of a subset of the *svb Z1.3* fragments tested for enhancer activity with reporter constructs. Stage 15 embryos carrying enhancer::*lacZ* reporters stained with an antibody against β-Gal (C) Expression patterns of reporter constructs in pupal dorsal abdomen (74 hours APF) as determined by X-Gal staining.

S6 Fig. Effect of enhancer deletions on *svb*BAC-GFP expression.

(A-B) Representative images of *svb*BAC-GFP and *svb*BAC-DsRed expression in abdominal segments of embryo (A) and pupa (B). Black boxes in embryos demarcate analyzed regions for each deletion. The white box in segment A4 of the pupa demarcates the analyzed region.

S1 Table. List of primers used in this study.

S2 Table. List of transgenic lines used in this sudy.

References

1. Kim AR, Martinez C, Ionides J, Ramos AF, Ludwig MZ, et al. (2013) Rearrangements of 2.5 kilobases of noncoding DNA from the Drosophila even-skipped locus define predictive rules of genomic cis-regulatory logic. PLoS Genet 9: e1003243.

2. Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H, 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.

3. Williamson I, Lettice LA, Hill RE, Bickmore WA (2016) Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity. Development 143: 2994-3001.

4. Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, et al. (2011) A regulatory archipelago controls Hox genes transcription in digits. Cell 147: 1132-1145.

5. Davidson EH (2010) The Regulatory Genome: Gene Regulatory Networks In Development And Evolution: Elsevier Science.

563 6. Stern DL (2000) Evolutionary developmental biology and the problem of variation. Evolution 54: 1079-
564 1091.
565 7. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of
566 morphological evolution. Cell 134: 25-36.
567 8. Wittkopp PJ, Kalay G (2011) Cis-regulatory elements: molecular mechanisms and evolutionary
568 processes underlying divergence. Nat Rev Genet 13: 59-69.
569 9. Akam M (1998) Hox genes: from master genes to micromanagers. Curr Biol 8: R676-678.
570 10. Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D (2014) Convergent evolution of complex
571 regulatory landscapes and pleiotropy at Hox loci. Science 346: 1004-1006.
572 11. Stern DL, Frankel N (2013) The structure and evolution of cis-regulatory regions: the shavenbaby
573 story. Philos Trans R Soc Lond B Biol Sci 368: 20130028.
574 12. Arif S, Kittelmann S, McGregor AP (2015) From shavenbaby to the naked valley: trichome formation
575 as a model for evolutionary developmental biology. Evol Dev 17: 120-126.
576 13. Payre F, Vincent A, Carreno S (1999) ovo/svb integrates Wingless and DER pathways to control
577 epidermis differentiation. Nature 400: 271-275.
578 14. Delon I, Chanut-Delalande H, Payre F (2003) The Ovo/Shavenbaby transcription factor specifies actin
579 remodelling during epidermal differentiation in Drosophila. Mech Dev 120: 747-758.
580 15. Chanut-Delalande H, Hashimoto Y, Pelissier-Monier A, Spokony R, Dib A, et al. (2014) Pri peptides are
581 mediators of ecdysone for the temporal control of development. Nat Cell Biol 16: 1035-1044.
582 16. Pueyo JI, Couso JP (2011) Tarsal-less peptides control Notch signalling through the Shavenbaby
583 transcription factor. Dev Biol 355: 183-193.
584 17. Preger-Ben Noon E, Davis FP, Stern DL (2016) Evolved Repression Overcomes Enhancer Robustness.
585 Dev Cell 39: 572-584.

18. Crocker J, Abe N, Rinaldi L, McGregor AP, Frankel N, et al. (2015) Low affinity binding site clusters confer hox specificity and regulatory robustness. Cell 160: 191-203.

19. Frankel N, Erezyilmaz DF, McGregor AP, Wang S, Payre F, et al. (2011) Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. Nature 474: 598-603.

20. Frankel N, Davis GK, Vargas D, Wang S, Payre F, et al. (2010) Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466: 490-493.

21. McGregor AP, Orgogozo V, Delon I, Zanet J, Srinivasan DG, et al. (2007) Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature 448: 587-590.

22. Frankel N (2012) Multiple layers of complexity in cis-regulatory regions of developmental genes. Dev Dyn 241: 1857-1866.

23. Infante CR, Mihala AG, Park S, Wang JS, Johnson KK, et al. (2015) Shared Enhancer Activity in the Limbs and Phallus and Functional Divergence of a Limb-Genital cis-Regulatory Element in Snakes. Dev Cell 35: 107-119.

24. Schep R, Necsulea A, Rodriguez-Carballo E, Guerreiro I, Andrey G, et al. (2016) Control of Hoxd gene transcription in the mammary bud by hijacking a preexisting regulatory landscape. Proc Natl Acad Sci U S A 113: E7720-E7729.

25. Frankel N, Wang S, Stern DL (2012) Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. Proc Natl Acad Sci U S A 109: 20975-20979.

26. Cande J, Goltsev Y, Levine MS (2009) Conservation of enhancer location in divergent insects. Proc Natl Acad Sci U S A 106: 14414-14419.

27. Glassford WJ, Johnson WC, Dall NR, Smith SJ, Liu Y, et al. (2015) Co-option of an Ancestral Hox-Regulated Network Underlies a Recently Evolved Morphological Novelty. Dev Cell 34: 520-531.

28. Smith E, Shilatifard A (2014) Enhancer biology and enhanceropathies. Nat Struct Mol Biol 21: 210-219.

610 29. Krijger PH, de Laat W (2016) Regulation of disease-associated gene expression in the 3D genome.
611 Nat Rev Mol Cell Biol 17: 771-782.
612 30. Gilbert SF (2010) Developmental Biology: Sinauer Associates.
613 31. Wang S, Zhao Y, Leiby M, Zhu J (2009) A new positive/negative selection scheme for precise BAC
614 recombineering. Mol Biotechnol 42: 110-116.
615 32. Halachmi N, Nachman A, Salzberg A (2012) Visualization of proprioceptors in Drosophila larvae and
616 pupae. J Vis Exp: e3846.

617

618

619

Figure 1

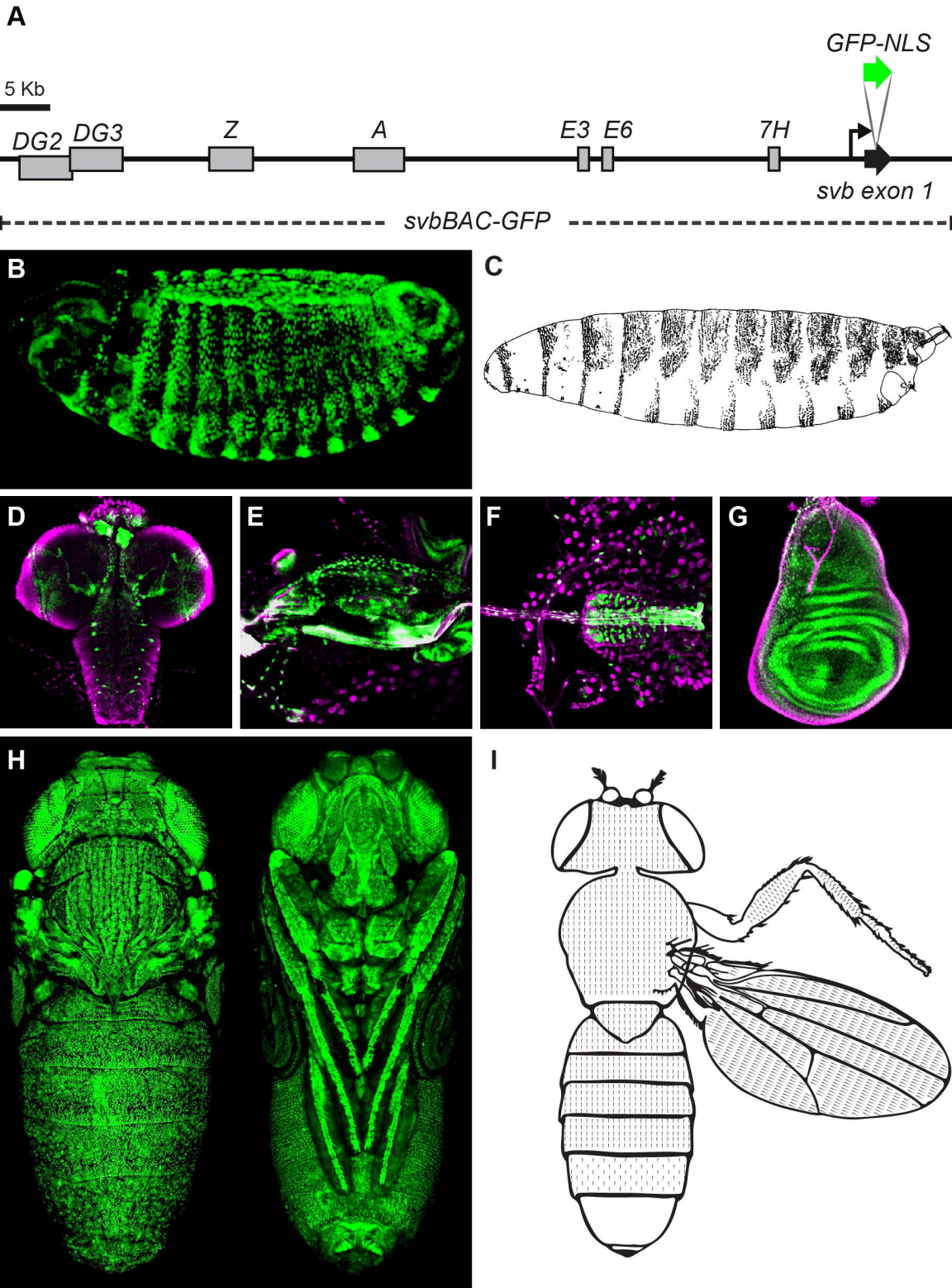


Figure 2

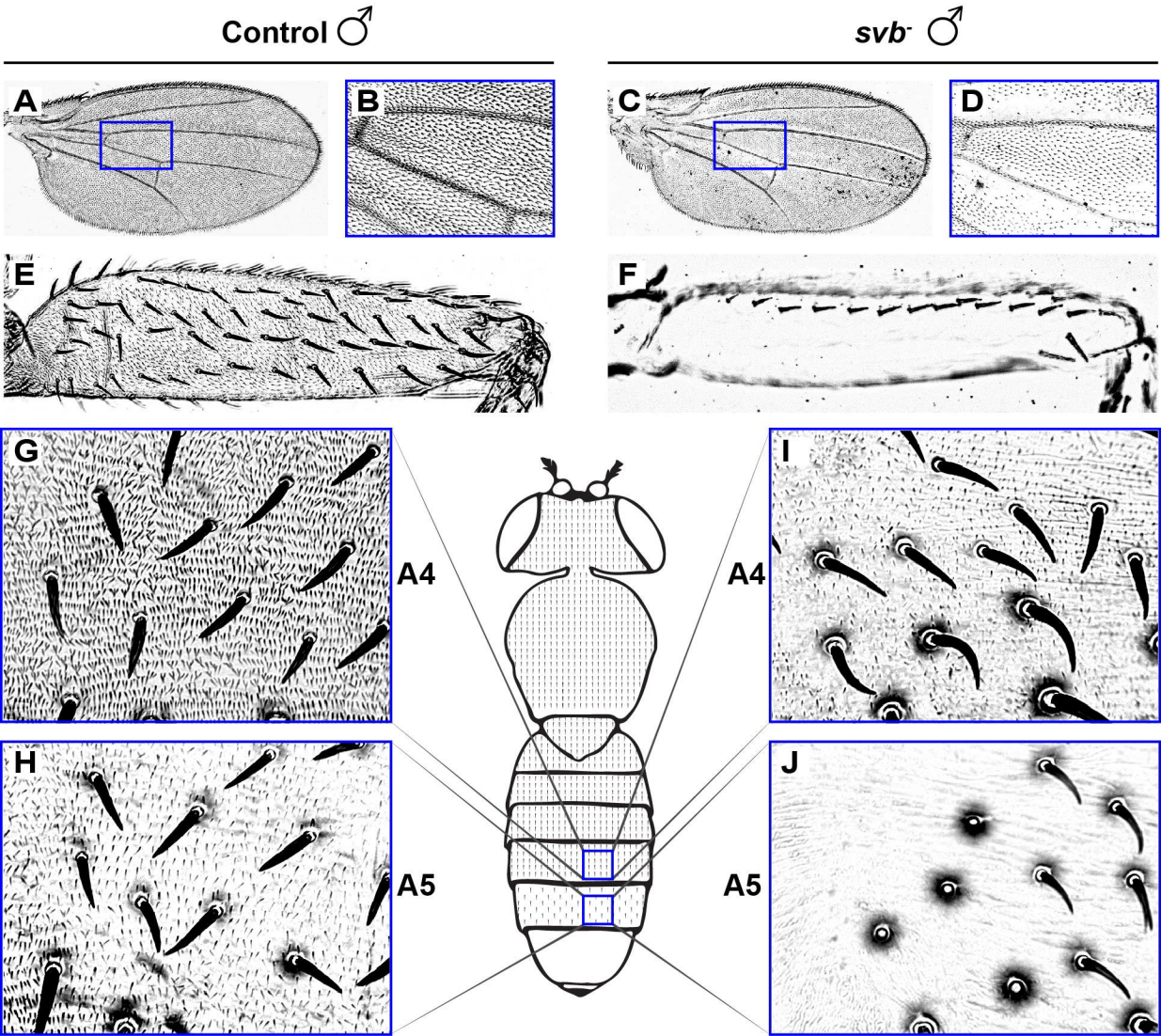


Figure 3

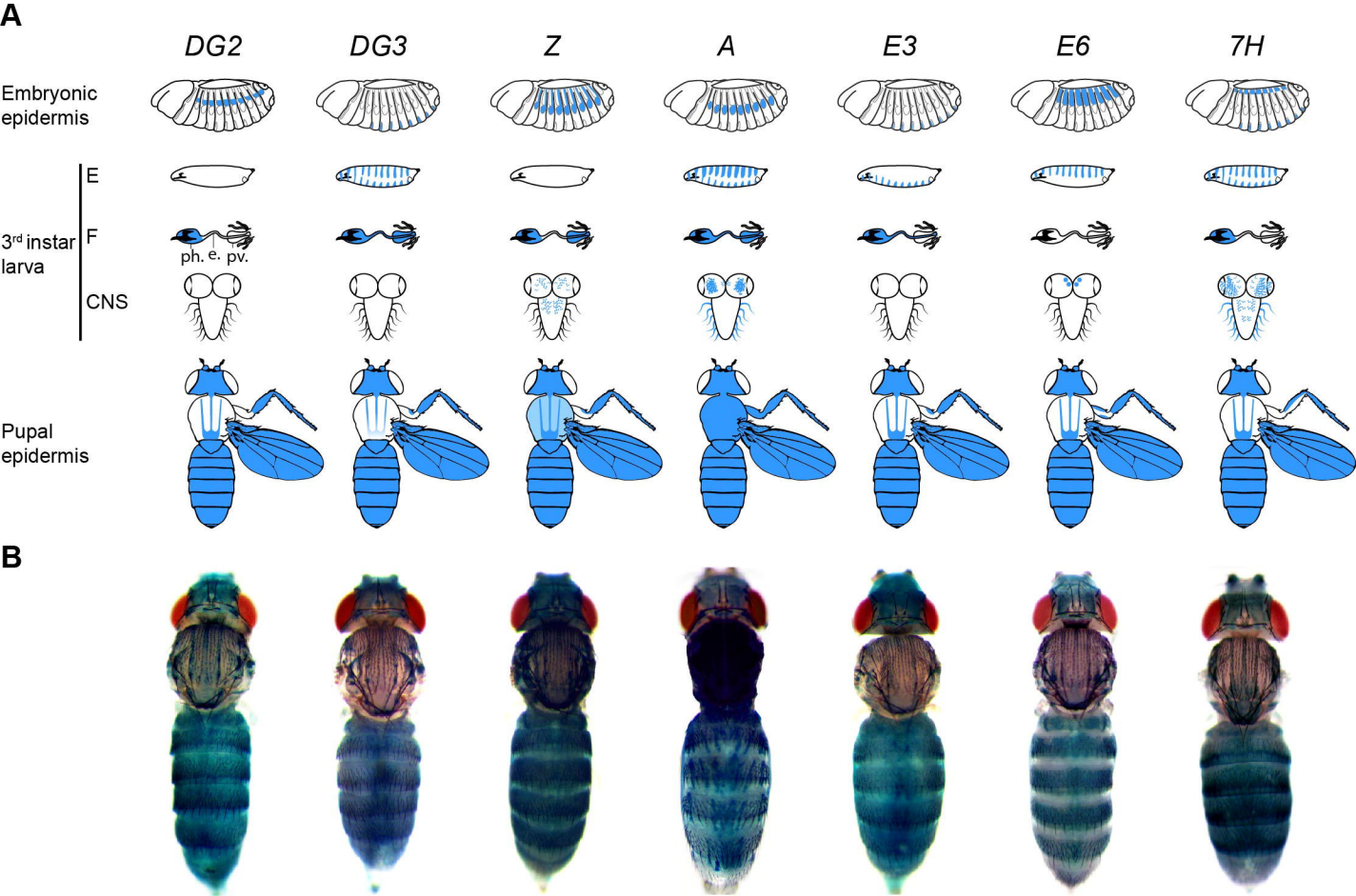

A
DG2
DG3
Z
A
E3
E6
7H
Embryonic epidermis
E
3rd instar larva
F
ph.
e.
pv.
CNS
Pupal epidermis
B

Figure 4

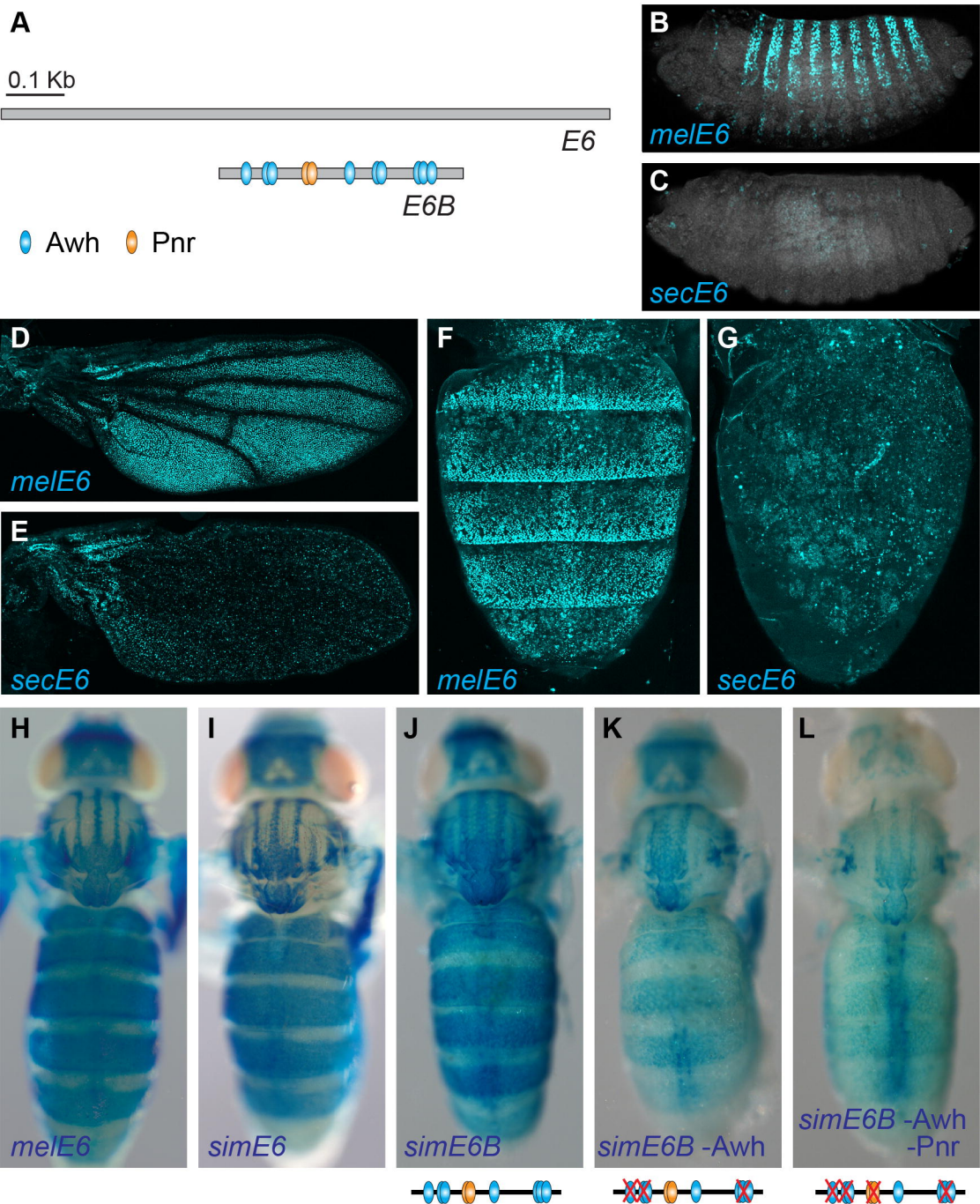

A
0.1 Kb
E6
E6B
Awh
Pnr
B
melE6
C
secE6
D
melE6
E
secE6
F
melE6
G
secE6
H
melE6
I
simE6
J
simE6B
K
simE6B -Awh
L
simE6B -Awh
-Pnr

Figure 5

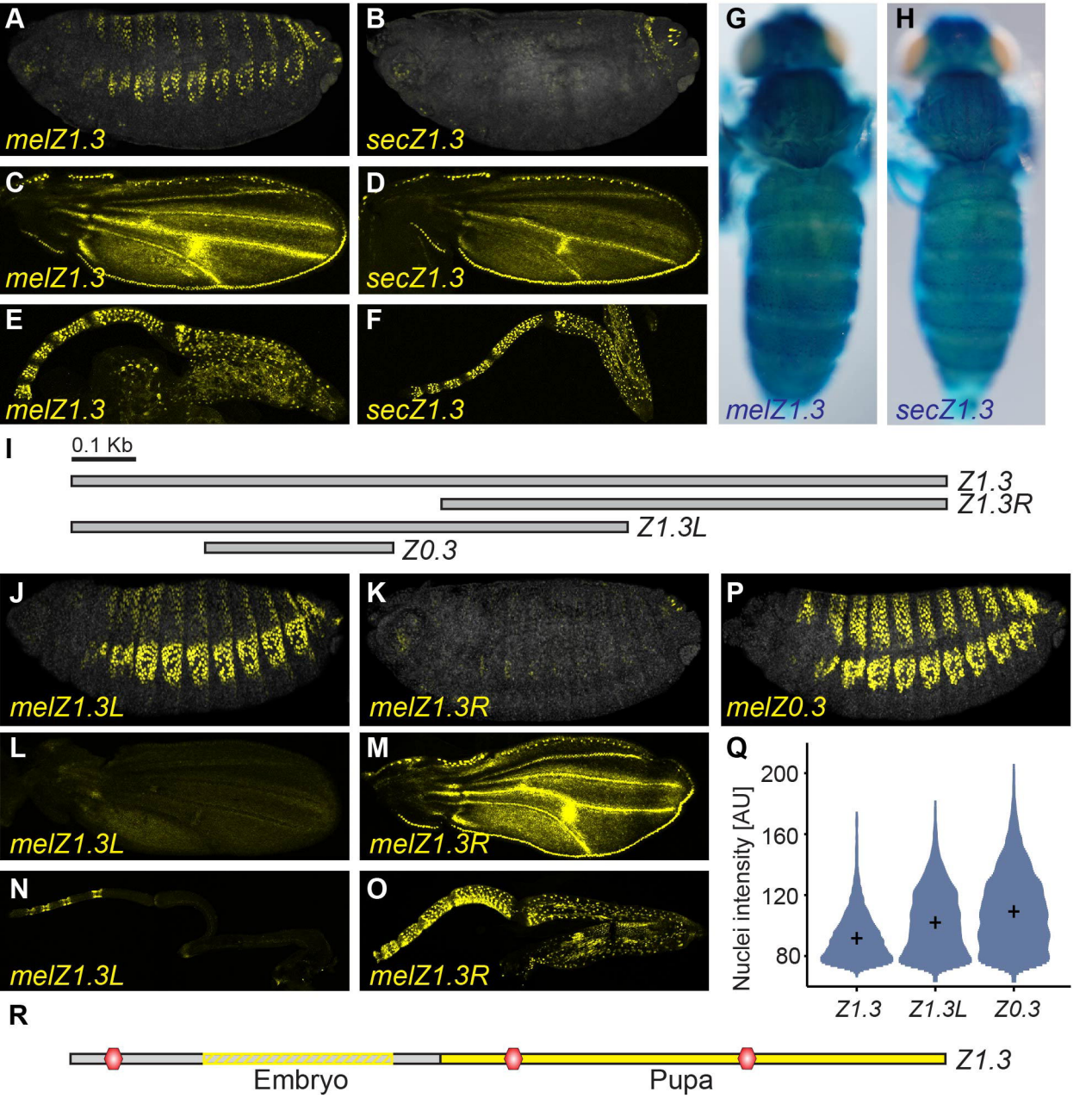
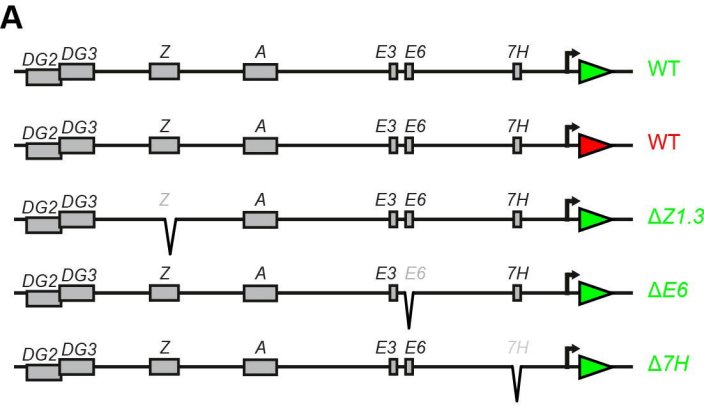

A
melZ1.3
B
secZ1.3
G
H
C
melZ1.3
D
secZ1.3
E
melZ1.3
F
secZ1.3
melZ1.3
secZ1.3
I
0.1 Kb
Z1.3
Z1.3R
Z1.3L
Z0.3
J
melZ1.3L
K
melZ1.3R
P
melZ0.3
L
melZ1.3L
M
melZ1.3R
Q
Nuclei intensity [AU]
200
160
120
80
Z1.3
Z1.3L
Z0.3
N
melZ1.3L
O
melZ1.3R
R
Z1.3
Embryo
Pupa

Figure 6


A
DG2 DG3 Z A E3 E6 7H
WT
WT
ΔZ1.3
ΔE6
Δ7H

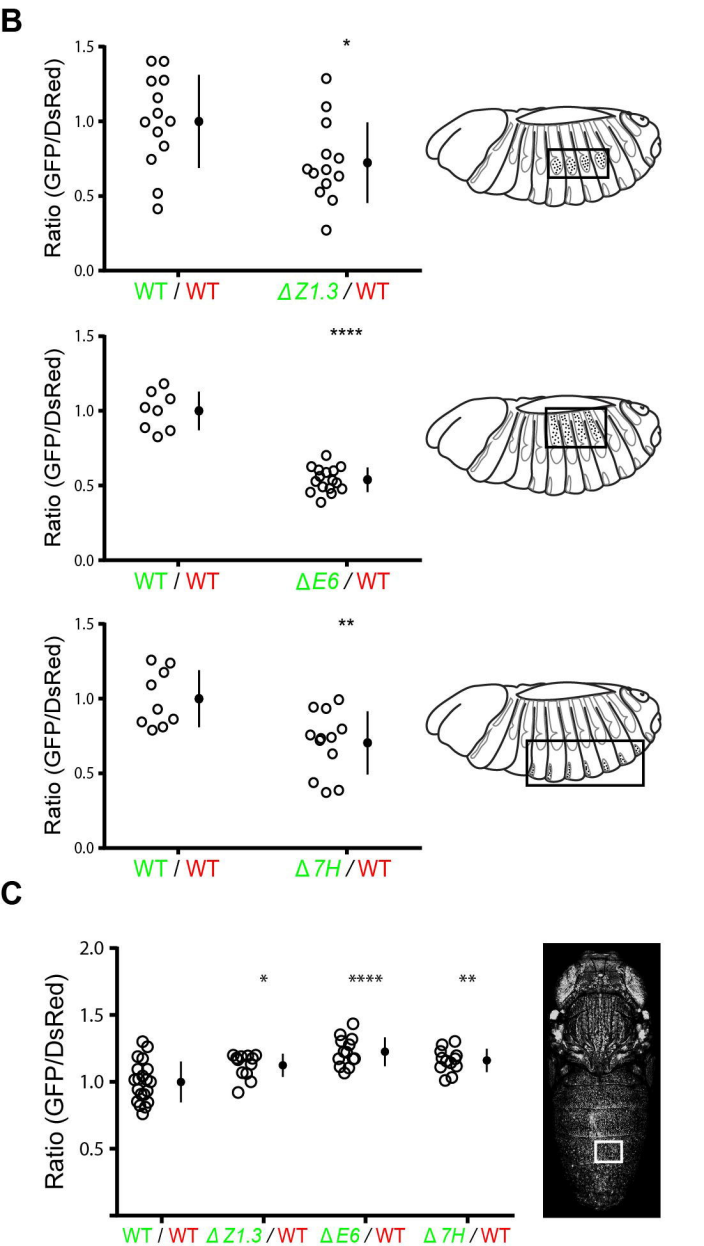
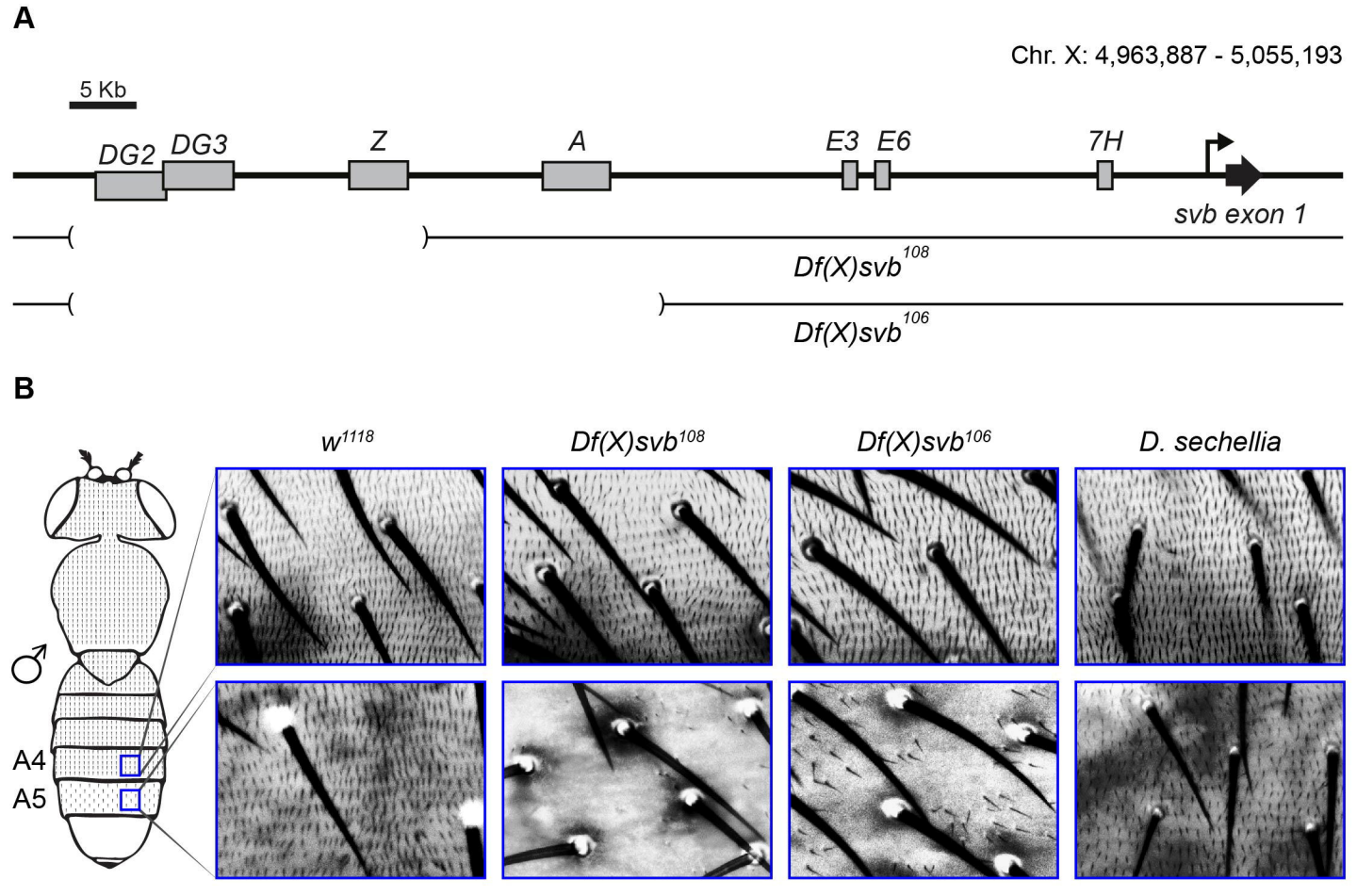

B
Ratio (GFP/DsRed)
1.5
1.0
0.5
0.0
*
WT / WT
ΔZ1.3 / WT
****
ΔE6 / WT
**
Δ7H / WT
C
2.0
1.5
1.0
0.5
*
****
**
WT / WT
ΔZ1.3 / WT
ΔE6 / WT
Δ7H / WT

Figure 7


A
Chr. X: 4,963,887 - 5,055,193
5 Kb
DG2
DG3
Z
A
E3
E6
7H
svb exon 1
Df(X)svb108
Df(X)svb106
B
w1118
Df(X)svb108
Df(X)svb106
D. sechellia
♂
A4
A5