

Research



Cite this article: Li XC, Fuqua T, van Breugel ME, Crocker J. 2023 Mutational scans reveal differential evolvability of *Drosophila* promoters and enhancers. *Phil. Trans. R. Soc. B* **378**: 20220054.
<https://doi.org/10.1098/rstb.2022.0054>

Received: 17 October 2022
Accepted: 26 January 2023

One contribution of 13 to a theme issue 'Interdisciplinary approaches to predicting evolutionary biology'.

Subject Areas:

genetics, evolution, developmental biology, molecular biology

Keywords:

promoter, enhancer, *cis*-regulatory evolution, *Drosophila*, *shavenbaby*, transcription

Authors for correspondence:

Xueying C. Li
e-mail: xueying.li@embl.de
Justin Crocker
e-mail: justin.crocker@embl.de

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6432526>.

Mutational scans reveal differential evolvability of *Drosophila* promoters and enhancers

Xueying C. Li, Timothy Fuqua, Maria Elize van Breugel and Justin Crocker

European Molecular Biology Laboratory, Heidelberg, Baden-Württemberg 69117, Germany

id XCL, 0000-0002-1100-4694; TF, 0000-0003-4005-3329; MEv, 0000-0001-8442-6516; JC, 0000-0002-5113-0476

Rapid enhancer and slow promoter evolution have been demonstrated through comparative genomics. However, it is not clear how this information is encoded genetically and if this can be used to place evolution in a predictive context. Part of the challenge is that our understanding of the potential for regulatory evolution is biased primarily toward natural variation or limited experimental perturbations. Here, to explore the evolutionary capacity of promoter variation, we surveyed an unbiased mutation library for three promoters in *Drosophila melanogaster*. We found that mutations in promoters had limited to no effect on spatial patterns of gene expression. Compared to developmental enhancers, promoters are more robust to mutations and have more access to mutations that can increase gene expression, suggesting that their low activity might be a result of selection. Consistent with these observations, increasing the promoter activity at the endogenous locus of *shavenbaby* led to increased transcription yet limited phenotypic changes. Taken together, developmental promoters may encode robust transcriptional outputs allowing evolvability through the integration of diverse developmental enhancers.

This article is part of the theme issue 'Interdisciplinary approaches to predicting evolutionary biology'.

'Regulators, mount up [at transcriptional promoters].' - Warren G & Nate Dogg, 1994 [1]

1. Introduction

Mutations may be largely random, but the loci of evolution are not. Through analysing causal variants underlying natural variation, previous studies have found that some specific genes or nucleotide substitutions are more often used in evolution than others [2–4]. For example, *cis*-regulatory changes are shown to be favoured by long-term evolution and morphological traits [2], although their relative contribution to evolution compared to coding changes is still under debate [5,6]. Therefore, evolution may be predictable if we gain a more comprehensive understanding of the roles of different kinds of molecular changes and their possible contributions to evolution. However, the ability to predict evolution requires a full construction of the genotype-to-fitness map, which is difficult to achieve by analysing naturally occurring variations that are limited in number and shaped by selection [7].

Mutational scans of regulatory and coding sequences in microorganisms and cell lines have begun to map genotype-to-phenotype relationships in broad sequence spaces [8–11] and to reveal principles of regulatory grammar [12–14] and adaptation [15,16]. However, such data have been largely lacking for developmental systems, where the additional challenge is to understand how mutations impact the spatial and temporal patterns of gene expression across development

and populations. Recently, mutational scans have been applied to developmental enhancers in fruit flies, where it was found that almost all mutations altered gene expression [17]. This study was further extended to additional elements [18], suggesting that developmental enhancers are often sensitive to perturbation, and may be highly constrained.

Metazoan promoters are traditionally thought to be functionally separated from enhancers, with the former primarily interacting with the transcription machinery (e.g. Pol II) and the latter interacting with transcription factors carrying spatial and temporal information. However, recent studies suggest that the boundary between promoters and enhancers can be blurry: enhancers can initiate certain levels of transcription, just like promoters, and many known promoters can influence transcription initiation of other genes, which is the classical definition of enhancers [19–21]. From an evolutionary standpoint, it has been found that rapid evolution of enhancers is a general feature of mammalian genomes [22]. By contrast, the genomic enrichment of key histone marks H3K27 acetylation and H3K4 trimethylation associated with promoters was partially or fully conserved across these species, suggesting that there is slow evolution of promoters in animal genomes. However, it is not known how this information is encoded genetically at developmental promoters, which have been distinguished from ‘housekeeping’ promoters by their distinct properties in the epigenetic and sequence signatures [23], the level of PolII stalling [24] and enhancer preferences [25].

Here, to understand if mutations in developmental promoters have a different distribution of effects on gene expression from those in enhancers, we examined random mutation libraries of three *Drosophila* promoters and compared them to a previously surveyed *Drosophila* *E3N* enhancer. By contrast with the previous findings that enhancers may be highly sensitive to mutations [17], we found that mutations in these promoters sometimes change the level of gene expression, but never the spatial pattern of expression. Together, these findings suggest that developmental promoters may encode robust transcriptional outputs allowing evolvability through the integration of diverse developmental enhancers.

2. Results

We focused our analyses on the regulatory sequences of *shavenbaby* (*svb*), a gene that encodes an essential regulator of trichome development in *Drosophila*. The evolution of the *svb* regulatory regions has been extensively studied due to contributions to phenotypic evolution across many *Drosophila* species [26–29]. Through these works, seven transcriptional enhancers have been characterized for *svb*; each integrates information from multiple patterning networks giving rise to the overall expression of *svb* across the embryo (figure 1a).

To explore how the native *svb* promoter (*svbp*) integrates these diverse activities, we tested the activity of the *svbp* using integrated reporter gene assays. The *svbp* shows high regulatory activity based on ReMap density [31] and is conserved among *Drosophila* species [32] (figure 1a). It does not contain TATA-box or other strong transcription motifs, consistent with signatures of developmentally regulated promoters in *Drosophila* [23]. Deletion of *svbp* resulted in severe depletion of ventral trichomes (figure 1b,c), recapitulating *svb* mutant phenotypes [33], although these lines were homozygous viable. In

order to understand how different promoters control levels and patterns of transcription activities driven by developmental enhancers, we generated reporter constructs of *svb* promoter and *Drosophila* synthetic core promoter (DSCP), an artificially engineered promoter known to drive high levels of expression [34]. Both promoters were placed downstream of the *E3N* enhancer of *svb*, which drives expression in a pattern of eight stripes in the abdominal region (A1 to A8) in stage 15 embryos [17]. The design with the *E3N* enhancer was necessary because, in the absence of an enhancer, the construct (with *hsp70* promoter) only drove a low level of background expression in stage 15 embryos [18]. We found that the two promoters drove different levels of reporter gene expression in the stripes, using the second abdominal stripe (A2) as a focal region for quantification (figure 1d, electronic supplementary material, figure S1). The nuclei intensity from DSCP was on average 2.4-fold higher than that of *svbp*. However, we found no differences in the overall gene expression patterns in the stripes (figure 1e,f). Additionally, we tested constructs without a promoter and with a random sequence as the promoter. We found that the reporter expression was below the level of detection in both lines (electronic supplementary material, figure S2), ruling out any promoter activity contributed by *E3N* in the constructs.

In order to understand the evolutionary potential of promoters in regulating the level and pattern of gene expression in a developmental context, we generated random mutation libraries of *svbp* and DSCP at a mutation rate of 1–2%, in a similar manner to our previous mutational scan on the *E3N* enhancer [17]. In the previous study, which examined the mutational profile of *E3N* in combination with the *hsp70* promoter, single-point mutations in *E3N* almost always decreased the gene expression level (18/18 variants). They often changed the state or locations of gene expression (11/18, 61%). In the follow-up studies that examined *E3N* variants with 1–10 mutations, it was found that the majority of variants caused decreases in the level of gene expression (83/91 based on median intensity, 91%) [18] and reduction in the number of nuclei in A1 to A8 stripes (81/91, 89%) [35]. We independently recapitulated these results by analysing ten randomly selected lines with 2–3 mutations from the *E3N* library. We found that all ten lines had a reduced number of nuclei expressing *lacZ* (figure 2a–c, electronic supplementary material, figure S3) and that seven showed reduced levels of expression (false discovery rate (FDR)-adjusted $p < 0.05$, Wilcoxon test, figure 2a). Together, these results are consistent with the previous finding that enhancers encode dense spatial information [17,18,36].

By contrast, we analysed 28 *svb* promoter variants (figure 2d–h), together covering 58 base pairs, and did not find any variants changing the pattern of gene expression (representative images in figure 2f–h). Unlike the enhancer, only two variants showed significantly lower expression than the wild-type *svb* promoter (figure 2d). Furthermore, one line showed higher expression levels than the wild-type promoter. The three lines with a changed expression level (89–6, 88–13 and 9–4) together contained seven unique mutations, and three were at highly conserved positions at the proximal end of the promoter (electronic supplementary material, figure S4A), suggesting potential functional relevance. However, we did not find a significant correlation between the presence of phenotypic effect and the level of conservation (phyloP score from 124 insects [32]), possibly

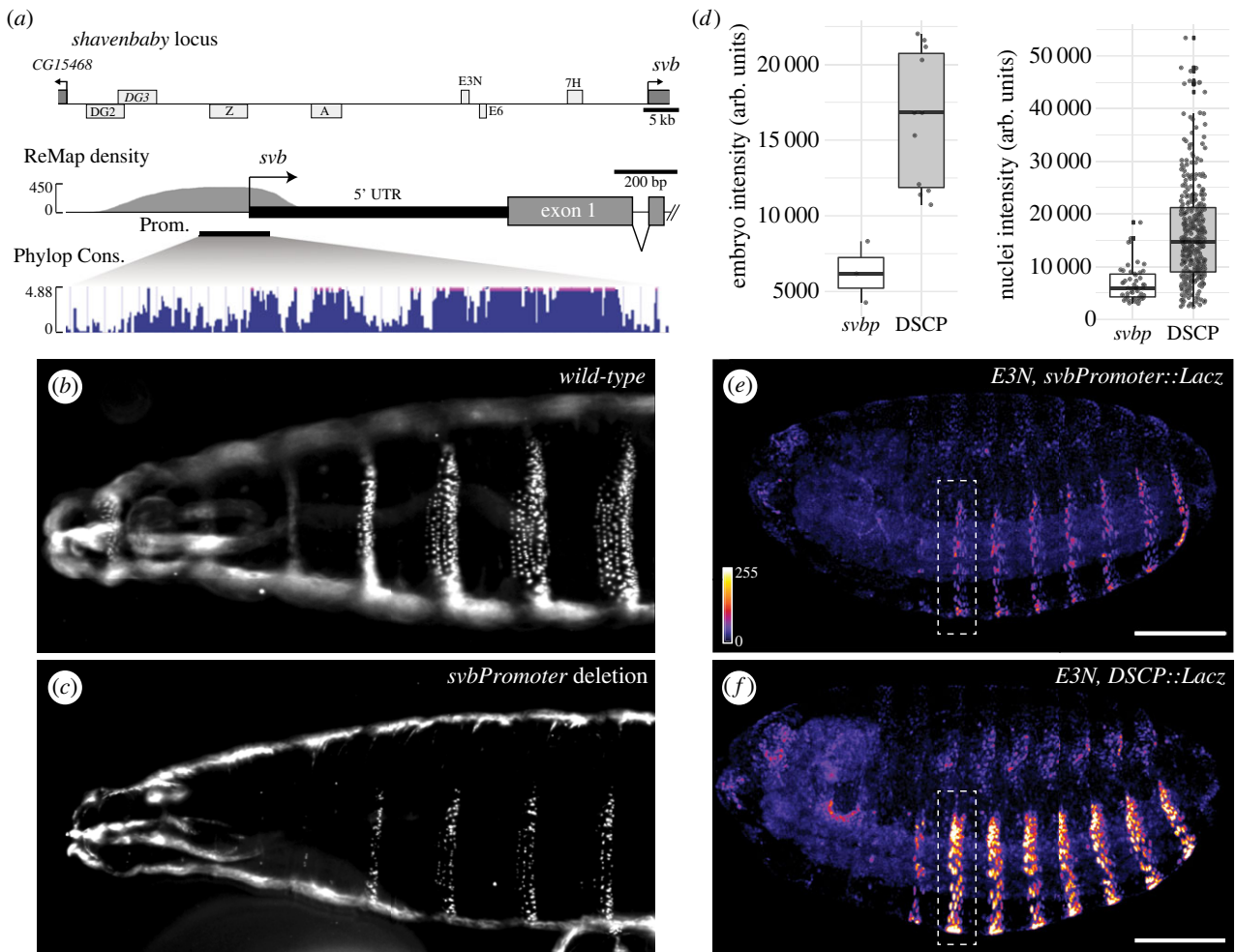


Figure 1. *LacZ* expression driven by *svb* promoter and *Drosophila* synthetic core promoter (DSCP). (a) *shavenbaby* locus. The regulatory region of *svb* spans approximately 87 kb, consisting of seven enhancer regions (top) [30]. The promoter of *svb* is 226bp-long, with high regulatory activity (ReMap score, middle) and conservation level (PhyloP score of 124 insects, bottom). Prom., promoter; Cons., conservation. (b,c) Cuticles of wild-type and *svb* promoter (*svbp*) deletion flies, respectively. (d) Levels of *lacZ* expression driven by *svbp* and DSCP under the control of *E3N*. Nuclei intensity was quantified by extracting average intensity around local maxima of A2 stripe (shown by white dashed boxes in e and f). Left, mean intensity per embryo, $p < 0.01$. Right, nuclei intensity across embryos, $p < 0.001$; arb. units, arbitrary units. p values were from Wilcoxon tests. (e,f) Representative images of stage 15 embryos, showing the pattern of *LacZ* expression in abdominal stripes driven by *E3N-svbp* and *E3N-DSCP* respectively, detected by anti-beta-Gal staining. Scale bars, 100 μm .

due to the small sample size (Wilcoxon test, $p > 0.05$). Interestingly, line 9-4 harvested a single T-to-A mutation at the first nucleotide position of the *svb* 5' UTR, which showed the most severe reduction in the level of gene expression in the library (figure 2d), suggesting a potentially critical role of transcription start sites. Taken together, given that the promoter library had a comparable mutation rate to the *E3N* enhancer library (on average 1%), our results suggest that developmental promoters are more robust than enhancers when subjected to the same mutation load.

We next extended our analysis to the DSCP. The DSCP was created by adding initiator (Inr) motif, motif ten element (MTE) and downstream promoter element (DPE) to a TATA-containing promoter of the developmental gene *even skipped* (*eve*)—creating one of the strongest promoters available in fruit flies [34]. We quantitatively analysed 45 variants of DSCP, with the number of mutations ranging from 1–8 across the 255bp-long sequence and an average mutation rate of 1.5% (figure 3). There were 117 nucleotide positions mutated in total, and 9 of them fell in the four functional motifs mentioned above (shaded regions in the left panel of figure 3a). We found that mutations in DSCP changed the expression level of the reporter gene more often than those

in *svbp*, with 13 mutant lines showing significant changes, suggesting that the endogenous *svb* promoter might be more robust than the synthetic promoter. Among the variants showing changes in expression, 7–2 had a mutation in TATA, 17–2 had a mutation in Inr, and lines 13–11 and 49–1 both had mutations in MTE. However, we did not find a statistically significant enrichment of mutations in these transcriptional motifs in the lines that showed changes in expression versus ones that did not (Fisher's exact test, $p > 0.05$). Furthermore, there was not a correlation between the presence of an effect on gene expression and the level of conservation (for regions from the *eve* promoter, see electronic supplementary material, figure S4B, Wilcoxon test, $p > 0.05$).

Interestingly, although DSCP drove a high level of expression, mutations in DSCP increased its activity even further in 10 mutant lines. This suggests that developmental promoters such as those of *svb* and *eve* might have the evolutionary potential to drive higher expression. However, it remains to be tested if strong transcriptional motifs such as those artificially engineered into DSCP are prerequisites of such evolvability. Due to the multiple mutational paths that led to higher levels of expression (either through transcriptional motifs or point mutations), developmental promoters

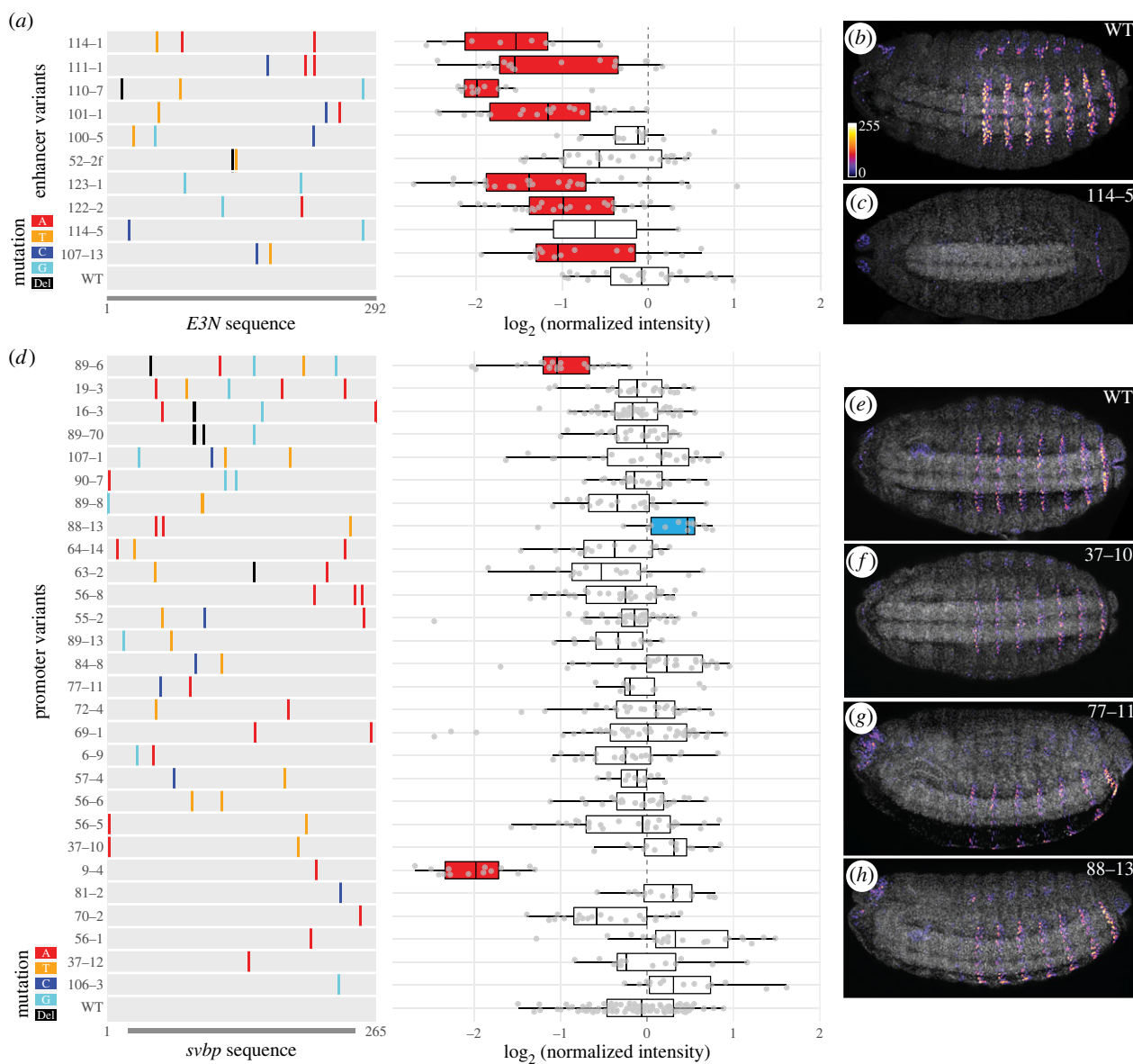


Figure 2. Different mutational profiles of *svb* enhancer and promoter. (a) Mutations and expression level of *E3N* variants, with representative images in (b,c). (d) Mutations and expression level of *svb* promoter variants, with representative images in (e–h). The mutant lines were ordered by the number of mutations, from low (bottom) to high (top). Coloured lines show the position and identity of mutations. Del, deletion. The level of expression was represented by mean intensity of nuclei in the A2 stripe in each embryo, normalized to internal wild-type controls within each batch. The colour of boxplots indicates significant difference from wild-type, tested by Wilcoxon test within batches (FDR-adjusted $p < 0.05$): red, reduced expression; blue, increased expression. The grey channel in (b,c) and (e–h) shows DAPI staining.

might have been selected to maintain low transcriptional activity during evolution.

Consistent with our findings from the *svbp*, mutations in the DSCP did not change gene expression patterns, supported by the 45 lines quantified above (e.g. figure 3b–d) and 21 additional DSCP variants examined quantitatively (electronic supplementary material, figure S5A–C). To further validate these results, we generated a mutation library of the *hsp70* promoter (*hsp70p*), a promoter commonly used to drive constitutive expression in *Drosophila* and used in the *E3N* library. Similarly, we did not find any variants causing a change in the expression pattern in the 31 variants examined (covering 74 out of 268 bp; electronic supplementary material, figure S5D–F). Together, these results are consistent with the traditional view of promoters encoding little spatial information [37].

Although reporter constructs allowed us to examine the promoter variants in a controlled setting, it remains unknown whether the changes in transcription can lead to phenotypic outcomes at the endogenous locus, where complex promoter–enhancer interactions are involved. Therefore, we next

tested if a change in the promoter activity at the endogenous locus could lead to phenotypic outcomes. We knocked out the *svb* promoter at its endogenous locus and replaced it with DSCP using CRISPR/Cas9. We found that the stronger DSCP promoter produced higher levels of transcription based on the local levels of nascent *svb* transcription compared to the endogenous promoter (figure 4a–c), consistent with the finding from reporter constructs. However, the changes in transcription levels did not directly translate into morphological changes, i.e. the pattern of ventral trichomes in larval cuticles (figure 4d,e): the DSCP knock-in rescued the knock-out phenotype (severe depletion of trichomes, figure 1) to the wild-type level, but there were no apparent differences in the trichome patterns from the wild-type.

3. Discussion

Although it remains debatable whether enhancers and promoters are functionally different elements from a transcription

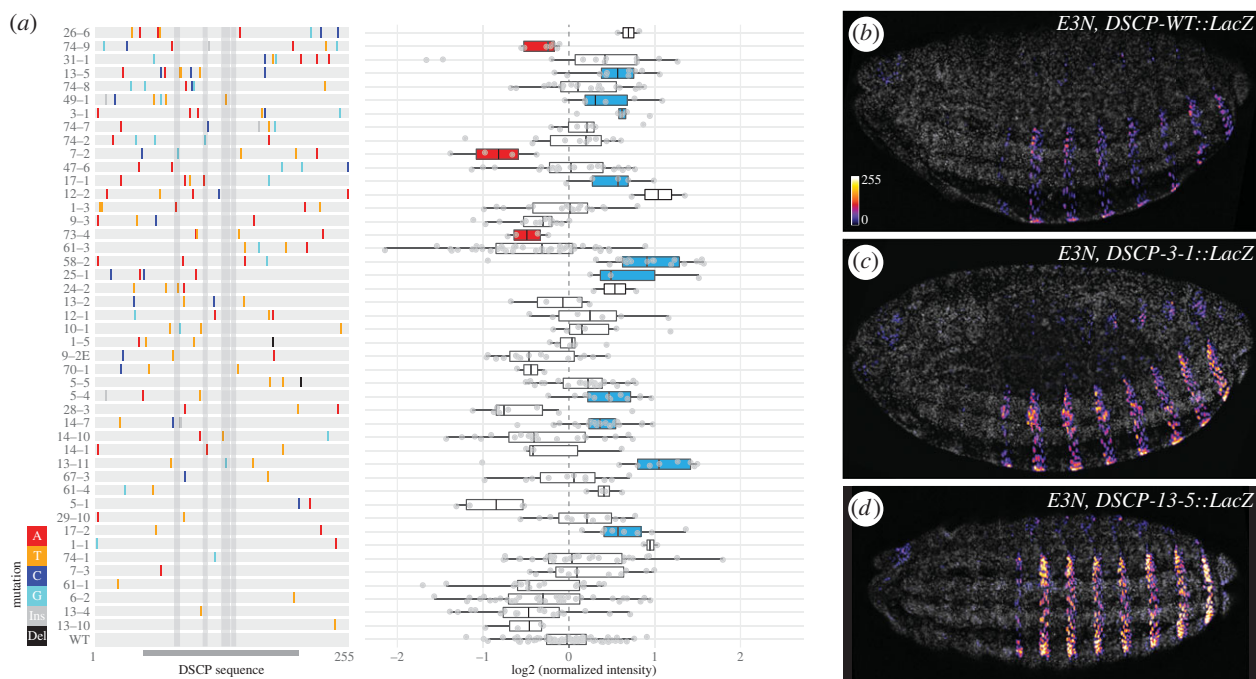


Figure 3. Mutational profiles of *Drosophila* synthetic core promoter (DSCP). (a) Mutations and expression level of DSCP variants, with representative images in (b–d). The variants were ordered by the number of mutations, from low (bottom) to high (top). Grey shades show regions of TATA, Inr, and MTE-DPE motifs, respectively. Coloured lines show the position and identity of mutations. Ins, insertion; Del, deletion. The level of expression was represented by mean intensity of nuclei in A2 stripes in each embryo, normalized to internal wild-type controls within each batch. The colour of boxplots indicates significant difference from the wild-type promoter, tested by Wilcoxon test within batches on lines with a minimum sample size of three embryos (FDR-adjusted $p < 0.05$): red, reduced expression; blue, increased expression. The grey channel in (b–d) shows DAPI staining. The images were background-subtracted and displayed in the same intensity range.

perspective [19,20], there is evidence that they are under different selective pressures and possible evolutionary constraints. For example, comparative studies have shown that enhancer sequences undergo rapid sequence divergence while maintaining their regulatory functions via binding site turnover, consistent with stabilizing selection [38,39]. Gains and losses of enhancers were also found to be frequent in different lineages [22,39]. Promoters have been shown to exhibit higher levels of sequence divergence than surrounding regions in insects, possibly associated with an increased mutation rate [40]. Changes in promoters tend to be neutral [41], consistent with our findings. They have also been shown to evolve more slowly than enhancers in mammals [22]. Still, the level of constraint on promoters can differ among different types of promoters [42], with highly constrained promoters associated with developmental functions [43].

Empirical characterization of the mutational space of enhancers and promoters in a developmental context was only made possible recently through mutational scans [17] and automation of embryo handling [44]. Recent mutational scans have found that developmental enhancers encode dense regulatory information and are strongly constrained [17,18,36]. In this study, we found that *Drosophila* promoters have different mutational profiles from enhancers. At a comparable mutation rate to the previously published E3N enhancer library [17], variants in our promoter libraries did not show any changes in the pattern of gene expression (figures 2 and 3, electronic supplementary material, figure S5). By contrast, almost all mutant lines of E3N changed the pattern (figure 2, electronic supplementary material, figure S3). Mutations in promoters can change the level of expression in either direction (figures 2 and 3), whereas mutations in enhancers tended to reduce expression (figure 2) [17,18].

Together, these findings suggest that *Drosophila* promoters might be more robust to mutations than enhancers.

Interestingly, this difference seems to exist in yeast promoters, if one considers a yeast promoter to be a mixture of enhancer (binding transcription factors) and promoter (initiating transcription) sequences: in the study of *TDH3* promoter, mutations in transcription factor binding sites (enhancer) greatly reduced transcription whereas other mutations only fine-tuned the level of expression [15]. Additionally, our results indicated that promoters might have little potential to evolve new spatial patterns of expression, consistent with a previous finding that promoters were less likely to be repurposed as enhancers than the other way around in mammalian evolution [45]. However, this observation remains to be tested with more promoters and beyond the context of reporter constructs. Furthermore, the effects of promoter variants on the level of gene expression did not correlate with the number or the location (e.g. in TATA or other motifs) of mutations (figures 2 and 3), suggesting that regulatory information might be randomly distributed in these promoters and a saturated mutational scan might be required to fully decode the regulatory potential of promoter sequences.

When comparing *svb* promoter and DSCP, it is clear that the endogenous *svb* promoter had low activity (figure 1), consistent with previous views [19]. The fact that both *svbp* and DSCP had access to mutations that can increase the expression level (figures 2 and 3) suggested that the low activity of endogenous promoters might be a result of selection. Furthermore, the high-activity, artificially engineered promoter was more ‘evolvable’ (or ‘breakable’) in the sense that many mutations led to changes in the level of gene expression, whereas the low-activity, endogenous *svb* promoter was relatively robust to mutations, suggesting that developmental promoters might have evolved to encode robust transcriptional outputs. This robustness may facilitate evolvability through the rapid integration of developmental

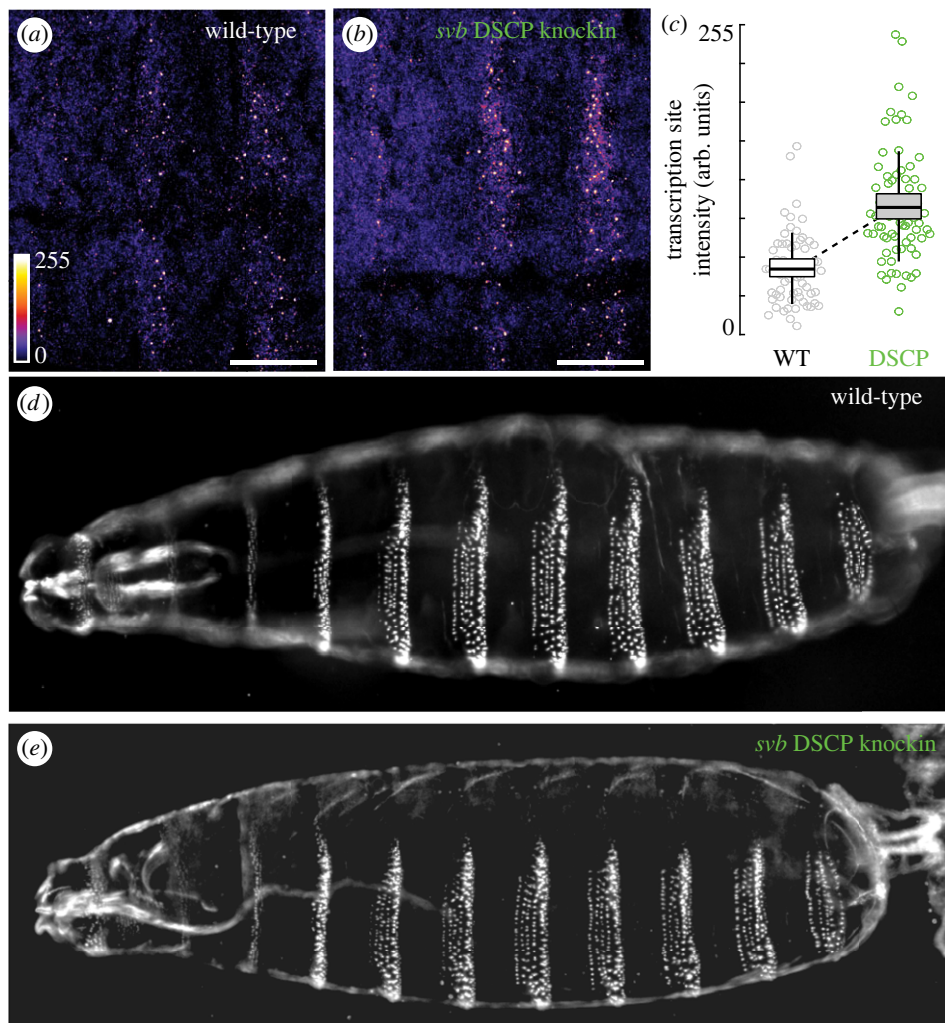


Figure 4. Replacing *svb* promoter with DSCP at its endogenous locus. (a,b) Transcription sites of *svb* in stage 15 embryos, detected by fluorescent *in situ* hybridization. Scale bar = 20 μ m. (c) Transcription site intensity, quantified across 3 embryos for each genotype. Arb. units, arbitrary units. (d,e) Cuticles of wild-type and *svbp Δ ::DSCP* larvae, showing ventral trichomes.

enhancers that drive cell type-specific expression patterns—including novel or co-opted elements (figure 5).

The relationship between the effects of *cis*-regulatory mutations on transcription and on fitness is often nonlinear [15,46]. In our study, we found that deletion of *svb* promoter led to a severe reduction of trichomes, similar to *svb* knockout phenotypes [47]. However, changes in promoter activity at the endogenous locus of *svb* by knocking-in DSCP did not cause a change in larval trichome patterns (figure 4), suggesting that changes in transcription level could be buffered by the downstream network [2], consistent with developmental traits being highly robust systems [48,49]. Alternatively, the nonlinear relationship could be explained by a threshold model, where the downstream patterning is elicited when *svb* transcription is above a certain threshold and a higher level of transcription does not change the patterning outcome [47].

Although promoters seem to be more robust to mutations than enhancers, the *svb* promoter shows a high level of sequence conservation, suggesting a certain degree of constraint. There could be a few explanations for this seemingly conflicted observation. First, promoters might be more ‘essential’ to transcription than enhancers because transcription is usually closely associated with one promoter but possibly multiple enhancers with redundant roles. Perturbation of promoters at their endogenous loci often has large phenotypic

effects [50,51], whereas perturbation of redundant enhancers may only manifest its effects under challenging conditions [52–54]. Future mutational scans of promoters and enhancers at the endogenous locus that focus on fitness effects are expected to provide insights in this direction. Second, different enhancers might interact with different sequence motifs in the promoter [55] that constrain promoter sequences, but that were not explored in the current study. Examination of a combinatorial promoter–enhancer mutation library may be required to address this possibility.

Together, ours and previous studies [17,18] highlight the power of mutational scans in providing insights for developmental evolution. This approach allows us to fully explore ‘the possible and the actual’ [56] of *cis*-regulatory evolution, which is currently lacking, especially in a developmental context. The differential constraints observed in different *cis*-regulatory elements can help us predict where evolutionarily relevant substitutions could occur within a locus. They also support the previous findings that the evolution of *svb* consists of multiple small-effect substitutions throughout the locus in different *Drosophila* species [27,29]. In the future, mutational scans by allele replacement at the endogenous loci will provide further insights into the fitness landscape of regulatory elements in a developmental context, paralleling those in microorganisms [10,15] and cell lines [57].

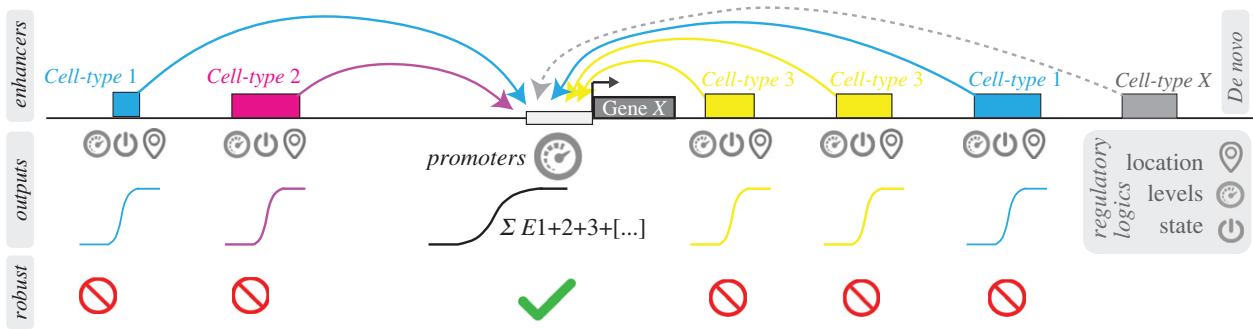


Figure 5. Model of enhancer and promoter evolution. Cell type-specific enhancers encode information for the location, levels and states of gene expression, whereas promoters encode information for the level of gene expression and integrate the transcriptional outputs from multiple enhancers. Promoters are relatively robust to mutations, allowing evolutionary changes through enhancers, including novel or co-opted changes.

4. Methods

(a) Promoter libraries

Random mutation libraries of *Drosophila* synthetic core promoter (DSCP), *hsp70* and *svb* promoters were synthesized at Genscript with a mutation rate of 10–20 point mutations per kb. In particular, the DSCP sequence (155 bp) was flanked by 50 bp-long sequences from *hsp70p* at each end, and the *svb* sequence (226 bp) was flanked by 19 and 20 bp-long sequences from *hsp70p* at each end, respectively. The flanking sequences were also subjected to mutagenesis. The variants were cloned into E3N-placZattB [17] to replace the wild-type *hsp70* promoter, which was positioned downstream of an E3N enhancer and upstream of a *lacZ* reporter [17]. The libraries were integrated into the fly genome at the attP2 site, with the injection service provided by GenetiVision. G0 transformants were crossed to w1118, and their offspring (F1) were screened for the presence of the construct by eye colour. The red-eye F1 flies were individually crossed to w1118 to establish isogenic lines, which were subsequently homozygosed by sibling crosses. The mutant lines were then sequenced to identify mutations in the promoters, with primer 5'-CCAAGTTGGTGGAGTTCATAATTCC-3' or 5'-AGGCATTGGGTGTGAGTTCCTC-3'. The sequences are listed in electronic supplementary material, table S1.

Additionally, we used two negative controls: 1) a construct with the E3N enhancer and *LacZ* but without any promoter; and 2) a construct with the E3N enhancer, a 200-bp-long 'inert' spacer sequence that was computationally screened for lack of transcription factor motifs in the place of the promoter (GAAGTTTCGACTAGTCTGAACTTCTACACAGACCGTATTAGAACTAT TACTAGCTACAAGCTCCTAGTGTCTTGAAAGCTATAACCTTAAGATGC TGTAGTATCTCAACCGACTTACTGCGAGACTTGACGAATTCGAA GTTCAGAACTAGTCTCTGAGTTGCGAGGTACATTTAGCAATGTAAGAA CCTCGGCT), and *LacZ*. The control constructs were integrated at attP2 sites as described above.

(b) Embryo collection and immunostaining

Embryos were collected from an overnight laying period at 25°C, using a standard fixation protocol [18]. During fixation and staining, a wild-type promoter control was always included in each batch, to account for batch effects.

Expression of *lacZ* was detected with a chicken anti-βGal antibody (1 : 500, abcam ab9361). ELAV was stained with mouse anti-ELAV supernatant (1 : 25, Developmental Studies Hybridoma Bank Elav-9F8A9) as a fiducial marker for the automated imaging pipeline to rotate the images, as well as for the experimenter to visually stage the embryos. For DSCP, E3N and *hsp70p* libraries, as well as for comparing DSCP and *svb* activity (data in figures 1 and 3, electronic supplementary material, figure S3 and figure S5), Alexa-Fluor 488 and 633 (1 : 500) were used as secondary antibodies for

βGal and ELAV, respectively. Due to the extremely weak signal of *svb* lines, we used extra staining steps for the *svb* mutation library to enhance the signal (data in figure 2). After a secondary incubation of AlexaFluor 555/488 (goat anti chicken, 1 : 500), Biotin conjugate was used as tertiary antibody (donkey anti sheep, 1 : 500, 1 h incubation) and NeutrAvidin 550 was used for quaternary staining (1 : 500, 30min to 1 h incubation). AlexaFluor 488/647 (1 : 500) was used as the secondary antibody for ELAV in this case.

The stained embryos of DSCP and *svb* libraries were mounted in ProLong Gold with DAPI. A subset of DSCP lines and all of the *hsp70p* lines were mounted in benzyl alcohol/benzyl benzoate (BABB) [44] and were analysed qualitatively due to lower imaging quality. The mutation libraries were imaged on a Zeiss LSM 880 confocal microscope with an automated pipeline under a 20× objective (air, 0.8 NA) as previously described [44] or manually under the same setting. Embryos used for comparison between DSCP and *svb* promoter in figure 1 were imaged manually under a 25× (oil, 0.8 NA) objective.

(c) Quantification of *lacZ* expression

We focused on cells in the second abdominal stripe (A2) in stage 15 embryos for analysing the pattern and intensity of *lacZ* expression. In each embryo, the A2 region was manually selected, max-projected, and background-subtracted with a rolling ball radius of 50 pixels. To select for *LacZ*-expressing cells in the region, we first performed a Gaussian blur with a radius of 2 pixels to remove noise, and then identified regions of interest (ROIs) by automatically thresholding the image with the Otsu method in ImageJ (electronic supplementary material, figure S1A). The ROIs were applied to the background-subtracted image and analysed with the Analyze Particles function to extract mean intensity of each ROI. Mean intensity per embryo was calculated by $I_e = \sum i_R \times A_R / \sum A_R$ where i_R was the mean intensity and A_R was the area of the ROI, respectively.

The embryo mean intensity of mutant lines was compared to the wild-type in the same batch with a Wilcoxon test. In the case where two biological replicates from different batches showed inconsistency in expression changes (i.e. one was different from wild-type and other was not), we took a conservative approach and removed both of them. A few *svb* lines (57–4, 19–3, 64–14, 69–1, 89–13, 89–70, 90–7) were imaged along with two wild-type controls that were different from the control of other batches, and their intensities were scaled to the other control by linear conversion to eliminate differences caused by background variation in the different controls. The data were normalized to the control line in each batch when combined in one plot. Data from biological replicates were merged.

To compare expression level between DSCP and *svb*, we extracted nucleus intensity by identifying local maxima with a prominence of 2000 and selecting a circular region with a radius of 0.55 μm around each local maximum. Mean intensity

of the circular regions was extracted from the background-subtracted images (electronic supplementary material, figure S1B).

(d) Allele replacement with CRISPR

We performed allele replacement following a two-step process, using a 3XP3-RFP marker as an intermediate step to easily select for integration events [58]. A 182 bp-long sequence of *svb* promoter immediately upstream of *svb* 5'UTR was targeted with two gRNAs, 5'-cgagatattcgccgttgctc-3' and 5'-gaatcacgtaagttgcgagc-3', which were cloned into pCFD4. A repair template containing the 3XP3-RFP sequence (1.86 kb) [58] and a 983bp-long homology arm at each end was synthesized and cloned into pUC57. The gRNA (75 ng μl^{-1}) and the repair template (225 ng μl^{-1}) were mixed and injected into a fly stock expressing Cas9 in the germline (BDSC#51324: w[1118]; PBac[y1+mDint2] GFP[E.3xP3]=vas-Cas9)VK00027). Flies from the injection were crossed to an FM6 balancer and subsequently screened for RFP expression in the eyes, which indicates successful replacement of *svb* promoter by the 3XP3-RFP cassette. The RFP-positive transformants were then homozygosed for both RFP and GFP markers to establish a fly line for the second round of allele replacement.

In the second round, we replaced the 3XP3-RFP sequence with the DSCP sequence. The gRNAs were designed based on the fused sequence of *svb* locus and 3XP3-RFP cassette: 5'-GGTACCGTACGAGATCTCTC-3' and 5'-GGCGCCTAAGGATC-GATAGC-3', cloned into pCFD4. The repair template contained a 255 bp-long DSCP sequence and the same homology arms as above. A mixture of plasmids carrying gRNAs and repair template was injected into the RFP-positive line mentioned above. Flies from the injection were crossed to an RFP/FM6 line and screened for loss of RFP. The resulting transformants were then homozygosed to establish a stable line of *svb*Promoter Δ ::DSCP genotype. The integration was confirmed by PCR and sequencing. A list of fly strains used in this study is provided in electronic supplementary material, table S2.

(e) Fluorescent *in situ* hybridization

svb transcripts were detected with DIG-labelled probes of *svb* as per [59]. Fixed *Drosophila* embryos were mounted in ProLong Gold + DAPI mounting media (Molecular Probes, Eugene, OR) and imaged on a Zeiss LSM 880 confocal microscope with FastAiryscan under a 63 \times objective (Carl Zeiss Microscopy, Jena,

Germany). Inside nuclei with *svb* transcription sites, the centre of the transcription site was identified using the find maximum function of Fiji/ImageJ. A circle with a diameter of 12 pixels [0.85 μm , region of interest (ROI)] centred on the transcription site was then created. The integrated fluorescent intensity inside the ROI was then reported. The intensity presented in the figures is the per-pixel average intensity with the maximum readout of the sensor normalized to 255.

(f) Cuticle preparation

Embryos from an overnight laying period were dechorionated with bleach and left in distilled water at room temperature for 24 h. After 24 h, the hatched larvae were transferred onto a glass slide and mounted in Hoyer's medium mixed with lactic acid (1 : 1). The slide was baked at 55°C for 2 days before being imaged with dark field microscopy.

Data accessibility. All processed data generated during this study are included in the manuscript and supporting files.

The data are provided in electronic supplementary material [60].

Authors' contributions. X.C.L.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, visualization, writing—original draft, writing—review and editing; T.F.: conceptualization, investigation, methodology, software, writing—review and editing; M.E.B.: conceptualization, investigation, writing—review and editing; J.C.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, software, supervision, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. X.C.L. is supported by a fellowship from the European Molecular Biology Laboratory Interdisciplinary Postdoc Programme (EIPOD) under Marie Skłodowska-Curie Actions COFUND (grant no. 664726). Research in the Crocker laboratory is supported by the European Molecular Biology Laboratory (EMBL).

Acknowledgement. We thank Alessandra Reversi and Matthew Benton for providing the injection service at EMBL. We thank Noa Otilie Borst for her help with *in situ* stainings, Natalia Misunou for reading the manuscript, Rafael Galupa for sharing reagents and other members of the Crocker group for their input on the project. We also thank members of Ella Preger-Ben Noon group, Alexander Stark group and Christa Buecker group for feedback on the preprint.

References

- Warren G (featuring Nate Dogg). 1994 *Regulate...* G Funk Era. New York, NY: Violator Records. (Distributed by Def Jam Recordings, Manhattan, NY.)
- Stern DL, Orgogozo V. 2008 The loci of evolution: how predictable is genetic evolution? *Evolution (NY)* **62**, 2155–2177. (doi:10.1111/j.1558-5646.2008.00450.x)
- Chan YF *et al.* 2010 Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**, 302–305. (doi:10.1126/science.1182213)
- Martin A, Orgogozo V. 2013 The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution (NY)* **67**, 1235–1250. (doi:10.1111/evo.12081)
- Hoekstra HE, Coyne JA. 2007 The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016. (doi:10.1111/j.1558-5646.2007.00105.x)
- Stern DL, Orgogozo V. 2009 Is genetic evolution predictable? *Science* **323**, 746–751. (doi:10.1126/science.1158997)
- Perkins ML, Gandara L, Crocker J. 2022 A synthetic synthesis to explore animal evolution and development. *Phil. Trans. R. Soc. B* **377**, 20200517. (doi:10.1098/rstb.2020.0517)
- Patwardhan RP *et al.* 2012 Massively parallel functional dissection of mammalian enhancers *in vivo*. *Nat. Biotechnol.* **30**, 265–270. (doi:10.1038/nbt.2136)
- Li C, Qian W, Madean CJ, Zhang J. 2016 The fitness landscape of a tRNA gene. *Science* **352**, 837–840. (doi:10.1126/science.aae0568)
- Venkataram S *et al.* 2016 Development of a comprehensive genotype-to-fitness map of adaptation-driving mutations in yeast. *Cell* **166**, 1585–1596.e22. (doi:10.1016/j.cell.2016.08.002)
- Kinsler G, Geiler-Samerotte K, Petrov D. 2020 Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation. *Elife* **9**, 1–52. (doi:10.7554/ELIFE.61271)
- Sharon E *et al.* 2012 Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* **30**, 521–530. (doi:10.1038/nbt.2205)
- Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA. 2012 Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl Acad. Sci. USA* **109**, 19 496–19 503. (doi:10.1073/pnas.1210678109)
- Qi Z *et al.* 2022 Large-scale analysis of *Drosophila* core promoter function using synthetic promoters. *Mol. Syst. Biol.* **18**, 1–26. (doi:10.15252/msb.20209816)

15. Metzger BPH, Yuan DC, Gruber JD, Duveau F, Wittkopp PJ. 2015 Selection on noise constrains variation in a eukaryotic promoter. *Nature* **521**, 344–347. (doi:10.1038/nature14244)
16. Park Y, Metzger BPH, Thornton JW. 2022 Epistatic drift causes gradual decay of predictability in protein evolution. *Science* **376**, 823–830. (doi:10.1126/science.abn6895)
17. Fuqua T *et al.* 2020 Dense and pleiotropic regulatory information in a developmental enhancer. *Nature* **587**, 235–239. (doi:10.1038/s41586-020-2816-5)
18. Galupa R *et al.* 2023 Enhancer architecture and chromatin accessibility constrain phenotypic space during *Drosophila* development. *Dev. Cell* **58**, 51–62.e4. (doi:10.1016/j.devcel.2022.12.003)
19. Haberle V, Stark A. 2018 Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **19**, 621–637. (doi:10.1038/s41580-018-0028-8)
20. Andersson R, Sandelin A. 2020 Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* **21**, 71–87. (doi:10.1038/s41576-019-0173-8)
21. Ramalingam V *et al.* 2022. Lola-I is a developmentally regulated promoter pioneer factor. *bioRxiv*:2022.04.25.489272. (doi:10.1101/2022.04.25.489272)
22. Villar D *et al.* 2015 Enhancer evolution across 20 mammalian species. *Cell* **160**, 554–566. (doi:10.1016/j.cell.2015.01.006)
23. Lenhard B, Sandelin A, Carninci P. 2012 Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* **13**, 233–245. (doi:10.1038/nrg3163)
24. Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. 2007 RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* **39**, 1512–1516. (doi:10.1038/ng.2007.26)
25. Zabidi MA, Arnold CD, Scherhuber K, Pagani M, Rath M, Frank O, Stark A. 2015 Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556–559. (doi:10.1038/nature13994)
26. Sucena E, Stern DL. 2000 Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven-baby*. *Proc. Natl Acad. Sci. USA* **97**, 4530–4534. (doi:10.1073/pnas.97.9.4530)
27. Frankel N, Erezyilmaz DF, McGregor AP, Wang S, Payre F, Stern DL. 2011 Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* **474**, 598–603. (doi:10.1038/nature10200)
28. Crocker J *et al.* 2015 Low affinity binding site clusters confer Hox specificity and regulatory robustness. *Cell* **160**, 191–203. (doi:10.1016/j.cell.2014.11.041)
29. Preger-Ben Noon E, Davis FP, Stern DL. 2016 Evolved repression overcomes enhancer robustness. *Dev. Cell* **39**, 572–584. (doi:10.1016/j.devcel.2016.10.010)
30. Stern DL, Frankel N. 2013 The structure and evolution of cis-regulatory regions: the *shavenbaby* story. *Phil. Trans. R. Soc. B* **368**, 20130028. (doi:10.1098/rstb.2013.0028)
31. Hammal F, de Langen P, Bergon A, Lopez F, Ballester B. 2022 ReMap 2022: a database of Human, Mouse, *Drosophila* and *Arabidopsis* regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* **50**, D316–D325. (doi:10.1093/nar/gkab996)
32. Siepel A *et al.* 2005 Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050. (doi:10.1101/gr.3715005)
33. Payre F, Vincent A, Carreno S. 1999 *ovo/svb* Integrates Wingless and DER pathways to control epidermis differentiation. *Nature* **400**, 271–275. (doi:10.1038/22330)
34. Pfeiffer BD *et al.* 2008 Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105**, 9715–9720. (doi:10.1073/pnas.0803697105)
35. Fuqua TJ. 2021 The evolutionary landscape of a developmental enhancer. PhD thesis, European Molecular Biology Laboratory, Heidelberg, Germany. See <https://archiv.ub.uni-heidelberg.de/volltextserver/30535/>.
36. Le Poul Y *et al.* 2020 Deciphering the regulatory logic of a *Drosophila* enhancer through systematic sequence mutagenesis and quantitative image analysis. (doi:10.1101/2020.06.24.169748)
37. Serfling E, Jasin M, Schaffner W. 1985 Enhancers and eukaryotic gene transcription. *Trends Genet.* **1**, 224–230. (doi:10.1016/0168-9525(85)90088-5)
38. Ludwig MZ, Bergman C, Patel NH, Kreitman M. 2000 Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**, 564–567. (doi:10.1038/35000615)
39. Arnold CD, Gerlach D, Spies D, Matts JA, Snytkova YA, Pagani M, Lau NC, Stark A. 2014 Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat. Genet.* **46**, 685–692. (doi:10.1038/ng.3009)
40. Main BJ, Smith AD, Jang H, Nuzhdin SV. 2013 Transcription start site evolution in *Drosophila*. *Mol. Biol. Evol.* **30**, 1966–1974. (doi:10.1093/molbev/mst085)
41. Hoffman MM, Birney E. 2010 An effective model for natural selection in promoters. *Genome Res.* **20**, 685–692. (doi:10.1101/gr.096719.109)
42. Carninci P *et al.* 2006 Genome-wide analysis of mammalian promoter architecture and evolution. *Nat. Genet.* **38**, 626–635. (doi:10.1038/ng1789)
43. Lindblad-Toh K *et al.* 2011 A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482. (doi:10.1038/nature10530)
44. Fuqua T, Jordan J, Halavatyi A, Tischer C, Richter K, Crocker J. 2021 An open-source semi-automated robotics pipeline for embryo immunohistochemistry. *Sci. Rep.* **11**, 1–16. (doi:10.1038/s41598-021-89676-5)
45. Carelli FN, Liechti A, Halbert J, Warnefors M, Kaessmann H. 2018 Repurposing of promoters and enhancers during mammalian evolution. *Nat. Commun.* **9**, 1–11. (doi:10.1038/s41467-018-06544-z)
46. Bergen AC, Olsen GM, Fay JC. 2016 Divergent MLS1 promoters lie on a fitness plateau for gene expression. *Mol. Biol. Evol.* **33**, 1270–1279. (doi:10.1093/molbev/msw010)
47. Delon I, Chanut-Delalande H, Payre F. 2003 The *Ovo/Shavenbaby* transcription factor specifies actin remodelling during epidermal differentiation in *Drosophila*. *Mech. Dev.* **120**, 747–758. (doi:10.1016/S0925-4773(03)00081-9)
48. Siegal ML, Bergman A. 2002 Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl Acad. Sci. USA* **99**, 10 528–10 532. (doi:10.1073/pnas.102303999)
49. Payne JL, Wagner A. 2015 Mechanisms of mutational robustness in transcriptional regulation. *Front. Genet.* **6**, 1–10. (doi:10.3389/fgene.2015.00322)
50. Lee AM, Wu CT. 2006 Enhancer–promoter communication at the *yellow* gene of *Drosophila melanogaster*: diverse promoters participate in and regulate *trans* interactions. *Genetics* **174**, 1867–1880. (doi:10.1534/genetics.106.064121)
51. Yokoshi M, Kawasaki K, Cambón M, Fukaya T. 2022 Dynamic modulation of enhancer responsiveness by core promoter elements in living *Drosophila* embryos. *Nucleic Acids Res.* **50**, 92–107. (doi:10.1093/nar/gkab1177)
52. Frankel N, Davis G, Vargas D, Wang S, Payre F, Stern DL. 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493. (doi:10.1038/nature09158.Phenotypic)
53. Perry MW, Boettiger AN, Bothma JP, Levine M. 2010 Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20**, 1562–1567. (doi:10.1016/j.cub.2010.07.043)
54. Osterwalder M *et al.* 2018 Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243. (doi:10.1038/nature25461)
55. Butler JEF, Kadonaga JT. 2001 Enhance–promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev.* **15**, 2515–2519. (doi:10.1101/gad.924301)
56. Jacob F. 1982 *The possible and the actual*. Seattle, WA: University of Washington Press.
57. Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, Cheng C, Regev A, Zhang F. 2016 High-resolution interrogation of functional elements in the noncoding genome. *Science* **353**, 1545–1549. (doi:10.1126/science.aaf7613)
58. Lamb AM, Walker EA, Wittkopp PJ. 2017 Tools and strategies for scarless allele replacement in *Drosophila* using CRISPR/Cas9. *Fly (Austin)* **11**, 53–64. (doi:10.1080/19336934.2016.1220463)
59. Tsai A, Alves MRP, Crocker J. 2019 Multi-enhancer transcriptional hubs confer phenotypic robustness. *Elife* **8**, 1–17. (doi:10.7554/eLife.45325)
60. Li XC, Fuqua T, van Breugel ME, Crocker J. 2023 Data from: Mutational scans reveal differential evolvability of *Drosophila* promoters and enhancers. Figshare. (doi:10.6084/m9.figshare.c.6432526)