



The hourglass model of evolutionary conservation during embryogenesis extends to developmental enhancers with signatures of positive selection

Jialin Liu, Rebecca R. Viales, Pierre Khoueiry, et al.

Genome Res. 2021 31: 1573-1581 originally published online July 15, 2021
Access the most recent version at doi:[10.1101/gr.275212.121](https://doi.org/10.1101/gr.275212.121)

References This article cites 46 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/31/9/1573.full.html#ref-list-1>

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Research

The hourglass model of evolutionary conservation during embryogenesis extends to developmental enhancers with signatures of positive selection

Jialin Liu,^{1,2,4} Rebecca R. Viales,³ Pierre Khoueiry,^{3,5} James P. Reddington,³ Charles Girardot,³ Eileen E.M. Furlong,³ and Marc Robinson-Rechavi^{1,2}

¹Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; ²Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; ³European Molecular Biology Laboratory, Genome Biology Unit, 69117 Heidelberg, Germany

Inter-species comparisons of both morphology and gene expression within a phylum have revealed a period in the middle of embryogenesis with more similarity between species compared with earlier and later time points. This “developmental hourglass” pattern has been observed in many phyla, yet the evolutionary constraints on gene expression, as well as the underlying mechanisms of how this is regulated, remain elusive. Moreover, the role of positive selection on gene regulation in the more diverged earlier and later stages of embryogenesis remains unknown. Here, using DNase-seq to identify regulatory regions in two distant *Drosophila* species (*D. melanogaster* and *D. virilis*), we assessed the evolutionary conservation and adaptive evolution of enhancers throughout multiple stages of embryogenesis. This revealed a higher proportion of conserved enhancers at the phylotypic period, providing a regulatory basis for the hourglass expression pattern. Using an in silico mutagenesis approach, we detect signatures of positive selection on developmental enhancers at early and late stages of embryogenesis, with a depletion at the phylotypic period, suggesting positive selection as one evolutionary mechanism underlying the hourglass pattern of animal evolution.

[Supplemental material is available for this article.]

Embryological development has long been characterized by deep conservation, from morphology to mechanisms. Animals that belong to the same phylum share a group of structural and developmental characteristics, the so-called basic body plan or *Bauplan* (Wallace 2000; Valentine 2004; Irie and Kuratani 2014). For example, arthropods share a set of anatomic structures such as jointed legs, an exoskeleton made of chitin, and segmented bodies (Zrzavý and Štys 1997). Based on morphological conservation, Duboule (1994) and Raff (1996) proposed the hourglass model. Under this model, within a phylum, embryos at mid-embryonic stages (the phylotypic period) (Richardson 1995) are morphologically more conserved than embryos in both early and late development. However, this model was not supported by later morphological studies (Richardson et al. 1997; Bininda-Emonds et al. 2003). To overcome difficulties of comparing morphological features across species, more recent studies used comparative transcriptomics (Yanai 2018), as changes in gene expression play a central role in the morphological differences between species (King and Wilson 1975; Carroll 2008).

Transcriptome comparisons in different phyla (Kalinka et al. 2010; Irie and Kuratani 2011; Levin et al. 2012; Schep and Adryan 2013; Hu et al. 2017) indicate that expression divergence is lower in the phylotypic period compared with the early and late stages of embryogenesis, supporting the hourglass model. One of the pi-

oneer studies was conducted in six *Drosophila* species by Kalinka et al. (2010), quantifying expression divergence at different stages during embryogenesis. They found that expression divergence follows an hourglass pattern with the minimum divergence at the extended germband stage (8–10 h after laying egg) generally regarded as part of the arthropod phylotypic period (Sander 1976). Notably, this expression hourglass pattern also extends to the population level (Zalts and Yanai 2017) and even to the level of variation between isogenic individuals (Liu et al. 2020). Based on a single-embryo transcriptome time series of *Drosophila* embryonic development, with a high number of isogenic replicates, we found that the phylotypic period also has lower nongenetic expression variability (Liu et al. 2020).

Despite many transcriptomic comparisons, the role of the underlying regulatory regions (e.g., developmental enhancers) on the evolution of expression during embryogenesis remains to be elucidated. What is more, although purifying selection and mutational robustness can explain the hourglass expression divergence pattern (Zalts and Yanai 2017; Liu et al. 2020), the contribution of positive selection to the hourglass model remains unknown. For example, this pattern may also result from enhanced positive selection at both the early and late development stages. Moreover, the two evolutionary mechanisms (purifying vs. positive selection) may not be mutually exclusive. In terms of protein sequence, for example, the lower sequence evolution in the phylotypic period appears to be caused by both strong purifying selection and weak positive selection (Liu and Robinson-Rechavi 2018a,b; Coronado-Zamora et al. 2019). To investigate the underlying regulatory mechanisms, as well as the contribution of positive

Present addresses: ⁴Biozentrum, University of Basel, CH-4056 Basel, Switzerland; ⁵American University of Beirut (AUB), Department of Biochemistry and Molecular Genetics, Beirut 1107 2020, Lebanon
Corresponding authors: furlong@embl.de, marc.robison-rechavi@unil.ch

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.275212.121>. Freely available online through the *Genome Research* Open Access option.

© 2021 Liu et al. This article, published in *Genome Research*, is available under a Creative Commons license (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

selection at regulatory elements, to the hourglass model, we performed DNase-seq to identify active regulatory elements across multiple matched embryonic developmental stages in two distant *Drosophila* species: *Drosophila melanogaster* and *Drosophila virilis*.

Results

DNase-seq across five stages of embryogenesis in two species

To study the evolution of enhancers in the context of embryonic development, we extended our previously published DNase I hypersensitive site sequencing (DNase-seq) data on three embryonic stages of *D. virilis* and *D. melanogaster* (Peng et al. 2019) to five equivalent embryonic stages in both species (TP1 to TP5) (Fig. 1A). TP3 is part of the phylotypic period (Fig. 1A), with the two new time points extending to later stages beyond the phylotypic period. Regulatory regions were identified using DNase-seq in tightly staged whole embryos. Every stage had two biological replicates from each species, and high-confidence peaks (bound regulatory regions) were called at a 5% irreproducible discovery rate (IDR; a measure ensuring equivalent reproducibility between replicates) (Li et al. 2011). On average, we identified 15,831 peaks in each stage in *D. virilis* and 14,995 peaks in *D. melanogaster* (Supplemental Fig. S1). Replicates are highly concordant both for raw reads (median Spearman's correlation coefficient 0.96 for *D. virilis*, 0.92 for *D. melanogaster*) and for significant peaks (median Spearman's correlation coefficient = 0.94 for *D. virilis*, 0.90 for *D. melanogaster*). Distal regulatory elements, including putative enhancers, were defined as peaks >500 bp from an annotated transcriptional start site (TSS). Although these regions may also include other regulatory elements such as insulators, for the sake of simplicity, we refer to these putative enhancers simply as en-

hancers in the rest of the study. We detected more distal elements (enhancers) in *D. virilis* than in *D. melanogaster* (Fig. 1B), similar to our previous findings using ChIP-seq against transcription factors (Khoueiry et al. 2017), which likely reflects the larger size of the *D. virilis* noncoding genome. In addition, in both species, we found that TP3 has more enhancers than other stages (Fig. 1B). The general trend in the relative number of enhancers across development is consistent between the two species.

Enhancer conservation over embryonic development

To compare evolutionary conservation of enhancers between stages, we first identified stage-specific enhancers for each time point in each species. For example, *D. melanogaster* TP3-stage-specific enhancers were defined as regions with a DNase peak in this stage and no significant peak at other stages in *D. melanogaster* (Fig. 2A; Table 1). TP3 has the highest proportion of stage-specific enhancers in *D. melanogaster* and the second-highest proportion of stage-specific enhancers in *D. virilis* (Supplemental Fig. S2A,B). For all stage-specific enhancers in one species, we identified their corresponding orthologous regions in the other species with psMap (see Methods) (Zhu et al., 2007), restricting to one-to-one orthologous regions. In both species, we found that TP3, within the phylotypic period, has a higher proportion of enhancers with orthologous regions (Supplemental Fig. S3), indicating stronger sequence conservation for phylotypic period-specific enhancers. Next, we identified conserved stage-specific enhancers. For example, if a *D. melanogaster* TP3-specific enhancer has both an orthologous region and an overlap (by orthologous translation) with a *D. virilis* TP3-specific enhancer, we defined this as a conserved TP3-specific enhancer (Fig. 2A). Finally, to quantify the overall conservation, at each stage, we calculated the Jaccard index (Fig. 2B), which ranges from zero (none of the stage-specific enhancers are conserved) to one (all stage-specific enhancers are conserved). We found that TP3 has a significantly higher proportion of conserved enhancers than do other stages (Fig. 2C). Given the larger size of non-coding genome in *D. virilis*, the genome coordinates of orthologous enhancers can be shifted by insertions and deletions. To account for this, we repeated the analysis with a relaxed definition of conserved enhancer: The stage-specific enhancers in the two species had to be within 1 kb of each other, but not necessarily overlapping. We found a very similar pattern with this more relaxed definition, with a higher proportion of conserved enhancers at TP3 (Fig. 2D). In addition, similar patterns were observed when we used all enhancers, not restricting to stage-specific enhancers (Supplemental Fig. S4).

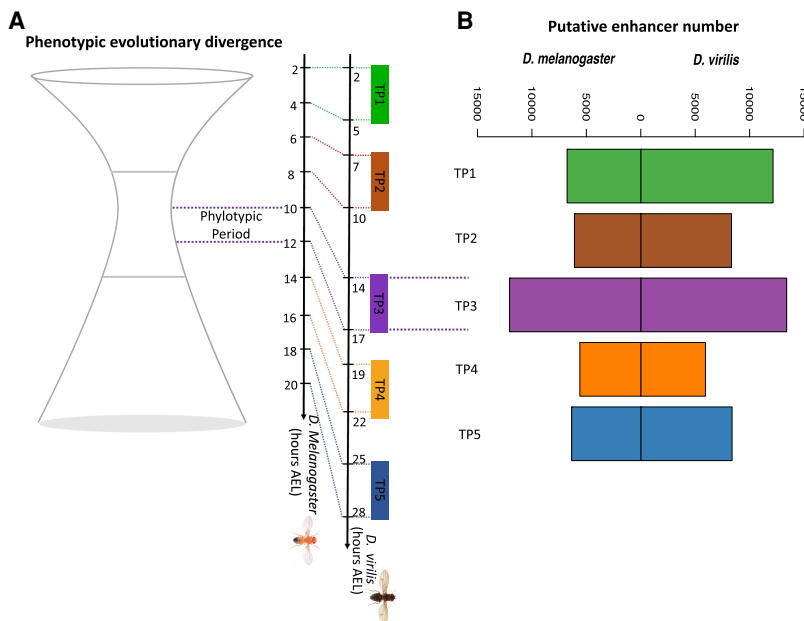


Figure 1. Studying regulatory element evolution throughout embryogenesis. (A) We performed DNase-seq at two matched late embryonic development stages in *D. melanogaster* and *D. virilis* (TP4 and TP5) and combined this with our previous DNase I hypersensitive site (DHS) data (TP1–3) in both species. The corresponding time points (TP1, TP2, TP3, TP4, and TP5) for the five embryonic stages are shown on the developmental axis. Different color bars represent different time points sampled. The developmental axes are scaled in hours after egg laying (AEL). (B) Number of putative enhancers in the five embryonic development stages in each species.

Detecting positive selection on enhancers

Higher conservation can be explained either by stronger purifying selection or by weaker positive selection. To test for the latter, we scanned for signatures of

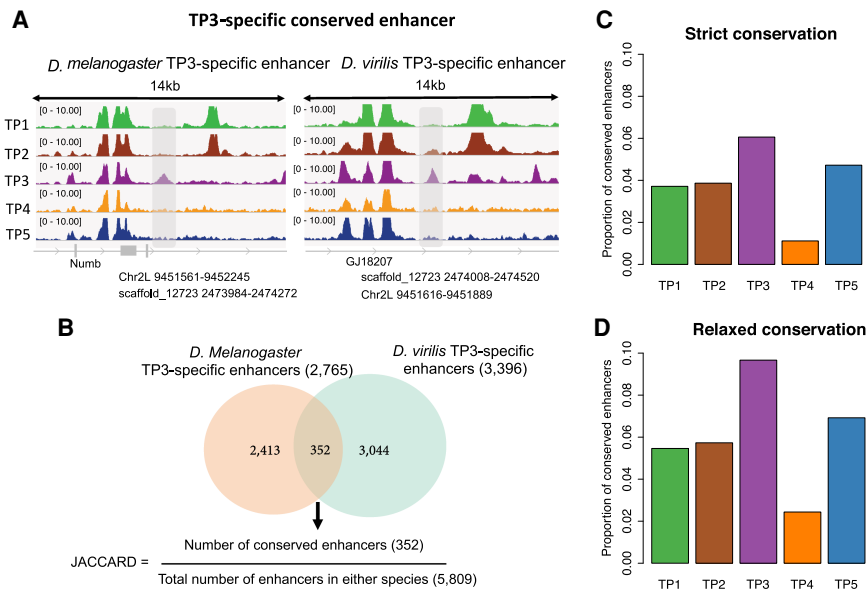


Figure 2. The phylotypic period has a higher proportion of conserved regulatory elements. (A) Illustration of a TP3-specific conserved enhancer. The left panel is the DNase-seq signal in different development stages (TP1, TP2, TP3, TP4, and TP5) for a *D. melanogaster* TP3-specific enhancer, covered by the gray area. The coordinates of this enhancer in *D. melanogaster*, as well as the orthologous coordinates in *D. virilis*, are indicated below the gray arrow. The right panel is the DNase-seq signals in different development stages (TP1, TP2, TP3, TP4, and TP5) for a *D. virilis* TP3-specific enhancer, covered by the gray area. The coordinates of this enhancer in *D. virilis*, as well as the orthologous coordinates in *D. melanogaster*, are illustrated below the gray arrow. Because there is overlap between the *D. melanogaster* TP3-specific enhancer and the *D. virilis* TP3-specific enhancer based on orthologous position, we define the two enhancers as a TP3 conserved enhancer. (B) Venn diagram of orthologous TP3-specific enhancers (only one-to-one orthologs) conserved between both species. (C) Proportion of conserved stage-specific enhancers at each development stage. Here the conservation means there is at least 1-bp overlap between stage-specific enhancers in the two species. The *P*-values from pairwise Fisher's exact tests between TP3 and TP1, TP2, TP4, and TP5 are 6.51×10^{-7} , 3.69×10^{-3} , 1.44×10^{-14} , and 2.22×10^{-2} , respectively. (D) Proportion of conserved stage-specific enhancers at each development stage. Here the conservation means the distance between stage-specific enhancers in the two species must be <1 kb, not necessarily overlapping. The *P*-values from pairwise Fisher's exact tests between TP3 and TP1, TP2, TP4, and TP5 are 1.89×10^{-12} , 4.28×10^{-5} , 2.7×10^{-17} , and 2.25×10^{-4} , respectively.

positive selection in all *D. melanogaster* stage-specific enhancers. Our approach considers the effects of substitutions on enhancer accessibility and is derived from a new method to detect positive selection in transcription factor binding site evolution (Liu and Robinson-Rechavi 2020). Briefly, a gapped *k*-mer support vector machine (gkmSVM) classifier is trained on these stage-specific enhancers. This gkmSVM identifies sequence features that determine chromatin accessibility and thus enhancer occupancy. This allows computing SVM weights of all possible 10-mers, which are predictions of their contribution to enhancer accessibility. We can then predict the accessibility impact of substitutions by calculating deltaSVM, the difference of sum weights between two homologous sequences. The significance of the observed deltaSVM was evaluated by comparing it with a null distribution of deltaSVM, constructed by scoring the same number of random substitutions 10,000 times. The *P*-value can be interpreted as the probability that the observed deltaSVM could arise by chance under the assumptions of the randomization.

Adaptive evolution on enhancer chromatin accessibility is expected to push them from a suboptimal accessibility toward an optimal accessibility or from an old optimum to a new one. Thus, chromatin-accessible sites evolving adaptively are expected to accumulate substitutions that consistently change the phenotype toward stronger or weaker accessibility, whereas sites evolving un-

der relaxed purifying selection are expected to accumulate substitutions that increase or diminish accessibility randomly around a constant optimum.

As this positive selection scanning method needs to be applied to sequences with relatively low divergence, we tested for substitutions on the *D. melanogaster* branch after divergence from *Drosophila simulans* (Fig. 3A), rather than the much more distant *D. virilis* branch. To check whether there were sequence features that differ between enhancers of different stages, we first separately trained a gkmSVM for each stage (see Methods). Not only can the gkmSVM trained in the corresponding stage accurately distinguish enhancers from random sequences, it also has higher performance than the gkmSVM trained from other stages (see Methods) (Fig. 3B). In addition, a gkmSVM trained from an adjacent stage has higher performance than a gkmSVM trained from a distant stage. For example, the gkmSVM model trained from TP2 has higher power to distinguish TP1-specific enhancers from random sequences than the gkmSVM model trained from TP5. These results suggest that the gkmSVM's predictions are not only informative but also specific to developmental stage.

We focused on enhancers with at least two substitutions to their sequence between species (Table 1). In each stage, we calculated the deltaSVM score of every stage-specific enhancer based on its corresponding gkmSVM. We associated a *P*-value for each enhancer by in silico mutagenesis (see Methods). Thus, we can identify enhancers whose substitution pattern on the *D. melanogaster* branch has effects on chromatin accessibility that are inconsistent with neutrality and therefore imply positive selection on enhancers. In each stage, the distributions of *P*-values for all stage-specific enhancers show a strong skew toward low *P*-values (Fig. 4A), indicating evidence for positive selection. For all downstream analyses, we use $q < 0.05$ (i.e., 5% false positives) as a threshold to define an enhancer as having evolved under positive selection (hereafter "positive selection enhancer"). Because mutations under positive selection will spread through a population rapidly, they are expected to

Table 1. Number of enhancers for each sampled stage

	TP1	TP2	TP3	TP4	TP5
Stage-specific enhancers					
<i>D. melanogaster</i>	1582	549	3333	677	1238
<i>D. virilis</i>	4765	1022	3862	893	2115
Stage-specific enhancers with two or more substitutions on the <i>D. melanogaster</i> phylogenetic branch	1467	521	3198	657	1216

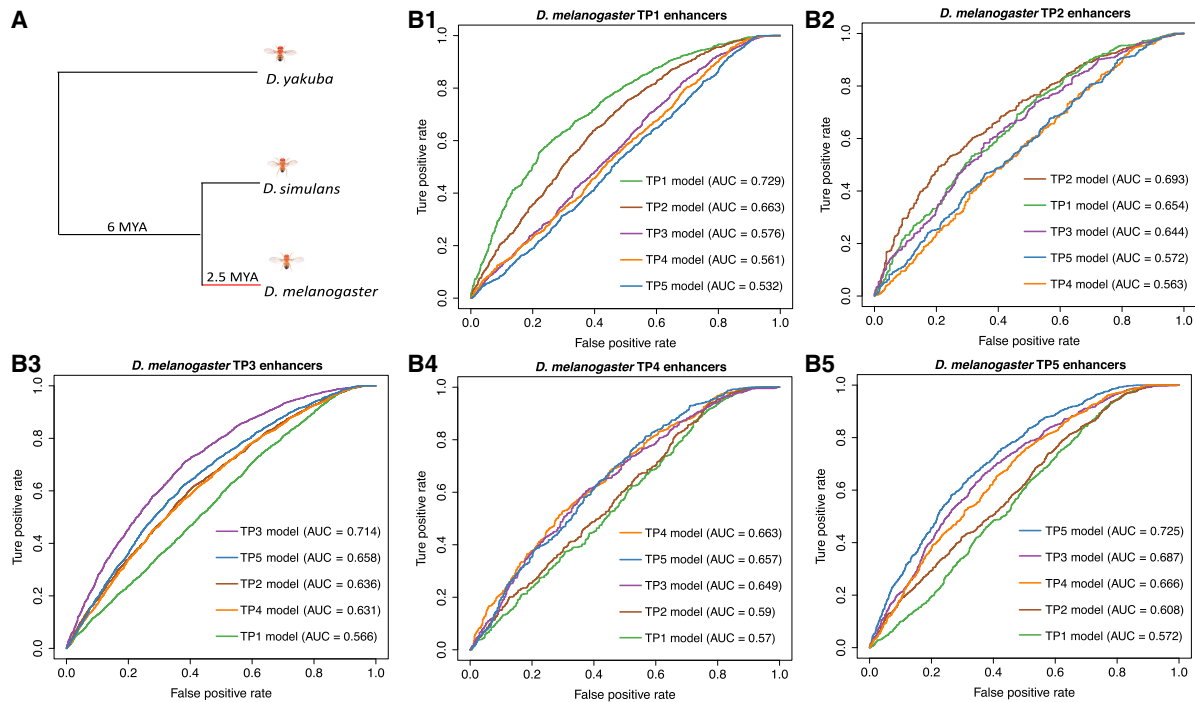


Figure 3. Gapped *k*-mer support vector machine (gkmSVM) can predict stage-specific enhancers. (A) Topological illustration of the phylogenetic relationships between the three *Drosophila* species used to detect positive selection on enhancers. We want to detect positive selection that occurred on the lineage of *D. melanogaster* after divergence from *D. simulans*, as indicated by the red branch. *Drosophila yakuba* is the outgroup used to infer enhancer sequence in the ancestor of *D. melanogaster* and *D. simulans*. (B1–B5) Receiver operating characteristic (ROC) curves for gkmSVM classification performance on stage-specific enhancers. AUC values represent areas under the ROC curve and provide an overall measure of predictive power.

decrease polymorphisms (intra-species variation) while increasing substitutions (inter-species variation) (McDonald and Kreitman 1991). Thus, we expect that positive selection enhancers should have higher substitutions to polymorphisms ratios than nonpositive selection enhancers. To test this, we counted the number of substitutions between *D. melanogaster* and *D. simulans*, as well as the number of polymorphisms among 205 *D. melanogaster* inbred lines from wild isolates, separately for positive selection and nonpositive selection enhancers (see Methods). As predicted, in all stages, positive selection enhancers have a significant excess of fixed nucleotide changes (Fig. 4B), confirming that we are indeed detecting positive selection in enhancer evolution.

Having identified enhancers that evolved under positive selection, we investigated whether their distribution varies across development. We previously found that species-specific gains in transcription factor binding sites have a higher proportion of positive selection than conserved ones (Liu and Robinson-Rechavi 2020). Thus, we separated *D. melanogaster* enhancers into conserved and nonconserved enhancers, based on conservation with *D. virilis* enhancers. As expected, the nonconserved enhancers generally have a higher proportion of positive selection than do the conserved ones (Fig. 5A,B). Moreover, over development, the phylotypic period has a much lower proportion of enhancers with evidence of positive selection than the other stages (Fig. 5A, B). This suggests that positive selection contributes to the evolution of enhancers and that the phylotypic period is characterized by less positive selection.

To check whether our test could be biased by mutation patterns, we analyzed the rates of all possible substitutions and their corresponding deltaSVM (Liu and Robinson-Rechavi 2020) in

each development stage. We split all substitutions into two categories: substitutions on CpG, and substitutions not on CpG. Overall, as expected, we found that the transition rate is much higher than the transversion rate, but we did not find any trend for specific substitution types to strengthen or weaken deltaSVM (Supplemental Figs. S5–S9). We also checked whether neighboring substitutions (dinucleotide substitutions) have a general tendency to change deltaSVM in the same direction. Indeed, this is the case (Supplemental Fig. S10), suggesting that our test could be liberal or conservative for dinucleotide substitutions, depending on the direction of deltaSVM. Finally, to check whether dinucleotide substitutions and mutation bias affect the pattern we found, we excluded dinucleotide substitution sequences from all binding sites, and we integrated the transition and transversion rate (3:1) (estimated from Supplemental Figs. S5–S9) into our null model. With these controls, we found a very consistent pattern to the original analysis (Supplemental Fig. S11). Thus, our results are robust to the main known mutational biases.

It has been suggested that evolutionary pressure on enhancers varies in a tissue-specific manner across development (Nord et al. 2013). To check whether lower positive selection in the phylotypic period holds true at a tissue level, we assigned each enhancer from TP2 (6–8 h) and TP3 (10–12 h) into three categories (muscle, neuron, and others) based on a recent lineage-resolved chromatin accessibility data set in *D. melanogaster* (see Methods) (Reddington et al. 2020). This data set contains muscle- and neuron-specific distal regulatory elements (putative enhancers) across four development stages, including two overlapping with TP2 and TP3 of our study. The pattern is consistent in each category (Supplemental Fig. S12), with lower positive selection at the

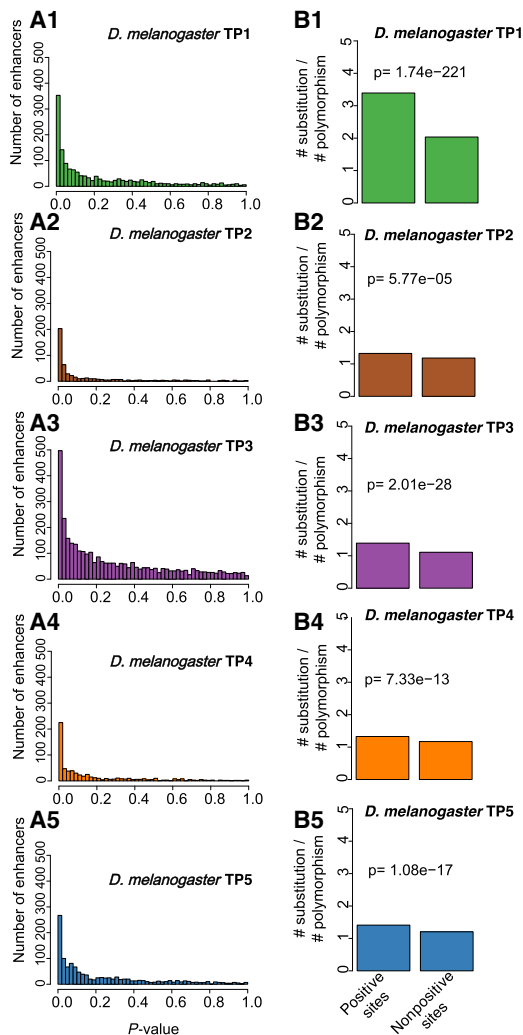


Figure 4. Evidence of positive selection throughout embryogenesis. (A1–A5) The distribution of deltaSVM P -values (test for positive selection) for each stage-specific enhancer. (B1–B5) The ratio between the number of substitutions and the number of polymorphisms (SNPs) for each stage-specific enhancer. Positive sites are enhancers with evidence of positive selection (deltaSVM q -value < 0.05); nonpositive sites are enhancers without evidence of positive selection. The P -value from the Fisher's exact test is reported *above* the bars.

phylogenetic period and no strong difference between neural and muscle lineages. Thus, the result is not driven by a specific pattern in one tissue, at least at this resolution.

Discussion

Based on DHSs in two distant *Drosophila* species across multiple matched embryonic stages, we identified a set of highly conserved stage-specific developmental enhancers. There is a higher proportion of conserved enhancers at the phylotypic period than at other embryonic stages, suggesting that conserved expression in the phylotypic period can be at least partly explained by conservation in gene regulation (Fig. 6). This provides a regulatory basis for the hourglass expression pattern. It was suggested that pleiotropic constraints might play important roles in the conservation of phylotypic stage (Duboule 1994; Raff 1996; Hu et al. 2017; Liu and

Robinson-Rechavi 2018a). Because we found the highest conservation at the phylotypic stage for both stage-specific enhancers and all enhancers (Fig. 2C; Supplemental Fig. S4) and because we found higher stage specificity at the phylotypic stage (Supplemental Fig. S2A,B), it is less likely that the evolution of enhancers in the phylotypic stage was constrained by ontogenetic pleiotropic effects (e.g., enhancers used in multiple stages). However, we cannot rule out that our estimate of stage specificity might lack power, as it is based on a sampling of five embryonic stages. If data were available for more stages with a finer temporal resolution, we might find that TP3 (phylotypic-stage)-specific enhancers are also used in neighboring developmental stages. In addition, the pleiotropic effects might be manifested at an anatomical level. For example, TP3-specific enhancers might be active in more cell types or tissues and hence have higher pleiotropy not measured here. Although enhancer conservation can be owing to stronger purifying selection, it can also result from weaker positive selection. Here, for the first time, we provide evidence that the higher enhancer conservation at the phylotypic period can be explained in part by the latter (Fig. 6). This is consistent with similar results at the protein sequence level (Liu and Robinson-Rechavi 2018b; Coronado-Zamora et al. 2019). Although both gene regulation and protein might be constrained at mid-development by lower embryo modularity, the new pattern we find for enhancer evolution can be more directly linked to previous observations of conservation of gene expression and of embryo morphology. Overall, we found that regulatory elements are more conserved in the phylotypic period and less accessible to adaptive evolution. Thus, the phylotypic period can be regarded as an evolutionary regulatory lockdown.

Our results indicate that widespread positive selection shaped the evolution of developmental enhancers, especially of early and late embryonic enhancers. This signature does not appear to be driven by a specific tissue, at least at the level of the two major lineages examined (muscle and neuron). Higher adaptation in late embryonic stages could be owing to the greater diversity of challenges to which natural selection needs to respond in the next stages of larva and juvenile development, compared with early and mid-embryogenesis. This fits with Darwin's "selection opportunity" (Liu and Robinson-Rechavi 2018b). Adaptation in the early embryo is less expected. Whereas early developmental proteins are influenced by the maternal contribution, which could be impacted by selection on reproduction, early embryonic enhancers are specific to the zygote. Although there is evidence that the evolution of *cis*-regulatory elements is the main driver of morphologic diversity (Gompel et al. 2005; McGregor et al. 2007; Wray 2007; Jeong et al. 2008), for most changes in early embryogenesis, there are no clear consequences on adult morphology (Kalinka and Tomancak 2012). This raises the question of why there are so many adaptive enhancer changes in early embryogenesis. One explanation, proposed by Kalinka and Tomancak (2012), is that much of the variation in early embryogenesis results from adaptation to diverse ecological circumstances. For example, the evolution of long germ segmentation in some insects (Liu and Kaufman 2005) might play a role in shortening the embryonic development time, which is likely an adaptive strategy to a particular ecological niche (Kalinka and Tomancak 2012). Another explanation is that adaptive evolution of maternally contributed *trans* factors might drive rapid compensatory evolution of early zygotic enhancers. It should also be noted that the maxima of positive selection are at the second and fourth time points, that is, not at the first and last. An interesting speculation is that the stages with the

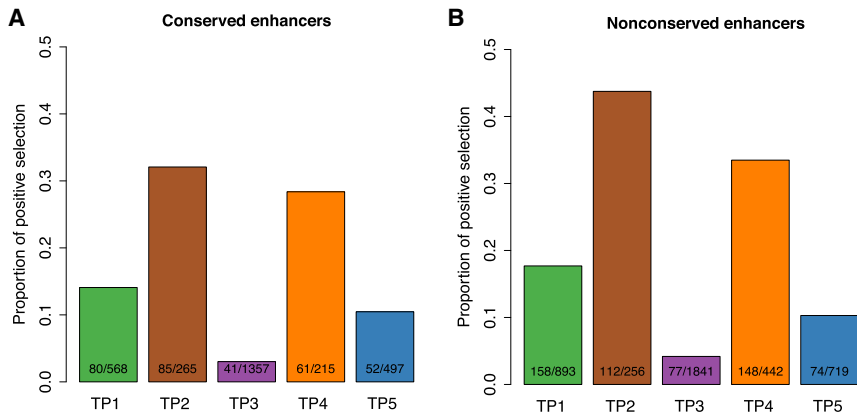


Figure 5. The phylotypic period has a lower proportion of enhancers with evidence of positive selection. The proportion of enhancers with evidence of positive selection in the five stages. Positive sites are enhancers with evidence of positive selection (deltaSVM q -value < 0.05). The number of stage-specific enhancers and the number of stage-specific enhancers with evidence of positive selection in each development stage are indicated inside each bar. Only enhancers with at least two substitutions were used for this analysis. (A) The P -values from pairwise Fisher's exact tests between TP3 and TP1, TP2, TP4, and TP5 are 1.08×10^{-15} , 1.01×10^{-33} , 6.75×10^{-25} , and 9.94×10^{-9} , respectively. (B) The P -values from pairwise Fisher's exact tests between TP3 and TP1, TP2, TP4, and TP5 are 5.04×10^{-25} , 1.47×10^{-46} , 4.57×10^{-46} , and 1.64×10^{-7} , respectively.

most positive selection on enhancers are those in which the most cell differentiation is occurring, but this is highly speculative in the present state of our data and our knowledge, and although it fits to our later time points (TP4 and TP5), it does not fit to the early ones (TP1, TP2).

Although we show a role for lower adaptive evolution in the regulation of more conserved expression at the phylotypic period, several recent studies also found evidence of developmental constraints, both purifying selection and mutational robustness. For example, in order to mostly eliminate the influence of positive selection on gene expression evolution, Zalts and Yanai (2017) quantified expression variability of 20 *Caenorhabditis elegans* mutation accumulation strains throughout embryogenesis. They found that the nematode phylotypic period has lower expression variability, indicating that purifying selection can contribute to the hourglass model of expression evolution. These results are also compatible with a role of mutational robustness. In *D. melanogaster*, we have recently compared the expression variability of isogenic single-embryo transcriptomes across development and found lower variability at the phylotypic period, suggesting that genes expressed at this stage are intrinsically less sensitive to perturbations on gene expression (Liu et al. 2020). Here, we found a much higher number of enhancers at the phylotypic period, which suggests more redundancy and thus higher regulatory robustness in gene expression. This could also contribute to mutational robustness (Cannavò et al. 2016) and thus lower expression divergence. Overall, the low expression divergence at the phylotypic period seems to have been shaped by the interplay of purifying selection, positive selection, and mutational robustness.

Methods

DNase-seq on *D. melanogaster* and *D. virilis* embryos

The first three time points were from our published study (Peng et al. 2019) with EMBL-EBI ArrayExpress accession number E-MTAB-3797. Here we supplemented those by two additional time points in each species to extend our time course beyond

the phylotypic period (Fig. 1A). Stage-matched *D. melanogaster* and *D. virilis* embryos (stages defined by Khoueiry et al. 2017) were collected and DNase-seq performed as described previously (Khoueiry et al. 2017). Two biological replicates were generated for each species at every stage.

DNase-seq sample processing

All analysis were performed in the Galaxy platform (Afgan et al. 2018). Raw paired-end reads were first trimmed with Trim Galore! (Galaxy Tool version 0.4.3.1; <https://github.com/FelixKrueger/TrimGalore>), and reads were clipped to maximum of 94 bases using Trimmomatic (Galaxy Tool version 0.36.6) (Bolger et al. 2014). Then the trimmed reads were mapped to the *D. melanogaster* genome (FlyBase Assembly 6 version: dm6) and to the *D. virilis* genome (FlyBase-R1.2 assembly version: droVir3), respectively, by using Bowtie 2 (Langmead and Salzberg 2012) with standard param-

eters (Galaxy Tool version 2.3.4.2). Multimapping reads were discarded, and duplicates were removed using MarkDuplicates (Galaxy Tool version 2.7.1.1) from the Picard suite (<https://github.com/broadinstitute/picard>). For peak calling, we used MACS2 (Zhang et al. 2008) with standard parameters (MACS2 Galaxy Tool version 2.1.1.20160309.5). We derived peaks using a 5% IDR (Galaxy Tool version 2.0.3) threshold for biological replicates.

Enhancer temporal pleiotropy analysis

Enhancers identified in different stages were integrated by intersecting all enhancers across stages, with at least 1-bp overlap

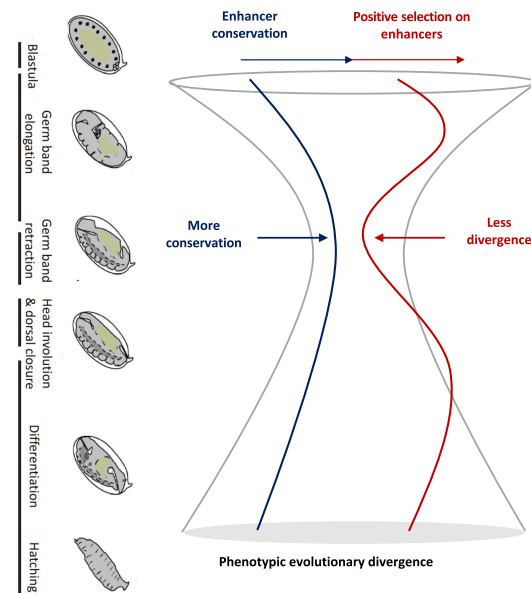


Figure 6. A simple model of the evolutionary forces on gene regulation for the hourglass pattern. Embryo images of *D. melanogaster* adapted from Levin et al. (2016).

used as the merge criteria. Then, in each stage, we calculated the proportion of enhancers used in one stage, two stages, three stages, four stages, and five stages (Supplemental Fig. S2A). In addition, for stages with two adjacent stages sampled (TP2, TP3, TP4), we calculated the proportion of enhancers used in both adjacent stages, in only one adjacent stage, or in neither (Supplemental Fig. S2B).

Genome coordinate translation

Because *D. melanogaster* and *D. virilis* are highly divergent, the *D. virilis* (*D. melanogaster*) peak coordinates were translated to *D. melanogaster* (*D. virilis*) genome coordinates by using the pslMap (Zhu et al. 2007), as suggested by Khoueiry et al. (2017).

Sequence alignment files

The pairwise whole-genome alignments between *D. melanogaster* and *D. simulans* or *D. yakuba* were downloaded from Haeussler et al. (2019; <https://hgdownload.soe.ucsc.edu/downloads.html>, accessed in December 2018).

Single-nucleotide polymorphism data

Over 4.8 million single-nucleotide polymorphisms (SNPs) for 205 *D. melanogaster* inbred lines were downloaded from the *Drosophila melanogaster* Genetic Reference Panel (DGRP) (Huang et al. 2014; <http://dgrp2.gnets.ncsu.edu/>, accessed in December 2018).

In silico mutagenesis for detecting positive selection on enhancers

Training the gkmSVM

gkmSVM is a method for predicting regulatory DNA sequence by using *k*-mer frequencies (Ghandi et al. 2014). For the gkmSVM training, we followed the same approaches as Lee et al. (2015). First, we defined a positive training set and its corresponding negative training set. The positive training set is stage-specific enhancers. The negative training set is an equal number of sequences randomly sampled from the genome with matched length, GC content, and repeat fraction as in the positive training set. This negative training set was generated by using “genNullSeqs,” a function of the gkmSVM R package (Ghandi et al. 2016). Then, we trained a gkmSVM with default parameters except $-l=10$ (meaning we use 10-mers as feature to distinguish positive and negative training sets). The classification performance of the trained gkmSVM was measured by using receiver operating characteristic (ROC) curves with fivefold cross-validation. The gkmSVM training and cross-validation were achieved by using the “gkmtrain” function of “LS-GKM” (Lee 2016; also see <https://github.com/Dongwon-Lee/lsgkm>).

Testing the stage specificity of gkmSVM

To test the performance of gkmSVM trained in stages other than the focal stage (e.g., trained TP2, TP3, TP4, TP5 to predict in TP1), we first scored both positive and negative training sets in the focal stage by using the gkmSVM from other stages. We used the “gkmpredict” function of “LS-GKM”. Then the ROC curve was used to evaluate the prediction performance.

Generating SVM weights of all possible 10-mers

The SVM weights of all possible 10-mers were generated with the “gkmpredict” function of “LS-GKM.” A positive value means increasing chromatin accessibility, a negative value means decreasing accessibility, and value close to zero means no impact on chromatin accessibility (the function measured in this case).

Inferring ancestor sequence

The ancestor sequence was inferred from sequence alignment between *D. melanogaster* and *D. simulans* by using *D. yakuba* as an outgroup.

Calculating deltaSVM

We calculated the sum of weights of all 10-mers for ancestor sequence and focal sequence, respectively. The deltaSVM is the sum of weights of the focal sequence minus the sum of weights of the ancestor sequence. A positive deltaSVM indicates substitutions increasing the chromatin accessibility in the focal sequence and vice versa.

Generating empirical null distribution of deltaSVM

First, we counted the number of substitutions between each ancestor sequence and focal sequence. Then we generated a random pseudofocal sequence by randomly introducing the same number of substitutions to the ancestor sequence. Finally, we calculated the deltaSVM between the pseudofocal sequence and the ancestor sequence. We repeated this process 10,000 times to get 10,000 expected deltaSVMs.

Calculating P-value of deltaSVM

The *P*-value was calculated as the probability that the expected deltaSVM is higher than the observed deltaSVM. The *P*-value can be interpreted as the probability that the observed deltaSVM could arise by chance under the assumptions of the randomization.

Estimating substitution rate

The substitution rate, for example, C → T, was estimated as the number of C → T divided by the number of nucleotide C in the ancestor sequence.

Definition of conserved and nonconserved enhancers

We split stage-specific enhancers into two categories: conserved and nonconserved. A *D. melanogaster* TP_{*i*}-specific enhancer whose orthologous region in *D. virilis* overlaps at least 1 bp with a *D. virilis* enhancer is defined as conserved. All other *D. melanogaster* TP_{*i*} enhancers are defined as nonconserved.

Muscle- and neuron-specific enhancer assignment

We first downloaded muscle- and neuron-specific enhancers (DNase-seq identified peaks >500 bp from an annotated TSS) from Reddington et al. (2020). This study included DNase-seq in isolated muscle and neural lineages across four stages (4–6 h, 6–8 h, 8–10 h, 10–12 h) of *D. melanogaster* embryogenesis. Then, based on these tissue-specific enhancers, we split our whole-embryo-identified TP2 (6–8 h)-specific and TP3 (10–12 h)-specific enhancers into three categories: muscle-specific enhancers, neuron-specific enhancers, and remaining enhancers. A TP2-specific enhancer that overlaps at least 1 bp with a muscle (resp. neuron)-specific enhancer is defined as a TP2 muscle (resp. neuron)-specific enhancer. All other enhancers are defined as “remaining” enhancers. The same was applied to TP3-specific enhancers.

Data access

The DNase-seq data generated in this study have been submitted to the EMBL-EBI ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>) under accession number E-MTAB-9480. Data files and analysis scripts are available as Supplemental Code and on

GitHub (<https://github.com/ljljolinq1010/Chromatin-accessibility-evolution-during-Drosophila-embryogenesis>).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank the EMBL Genomics Core Facilities for support and David Garfield and Gunter Wagner for critical reading of an earlier version of the manuscript. We thank members of the Furlong and Robinson-Rechavi laboratories for helpful discussions. We thank Irepan Salvador-Martínez, Axel Visel, and Matt Benton for comments on a previous version in preprint. Part of the computations were performed at the Vital-IT (<http://www.vital-it.ch>) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. E.E.M.F. is supported by the Marie Curie EvoNet ITN and the Deutsche Forschungsgemeinschaft (DFG) grant FU 750; J.L. and M.R.-R. are supported by Swiss National Science Foundation grant 31003A_173048.

Author contributions: E.E.M.F. initiated the project and supervised all experiments. R.R.V., P.K., and J.P.R. performed all experiments. C.G. performed raw data analysis and DNase-seq peak calling. J.L. designed the detailed study with input from M.R.-R. J.L. performed all downstream bioinformatics analyses. J.L. and M.R.-R. interpreted the results with input from all other authors. J.L. wrote the first draft of the paper. J.L., E.E.M.F., and M.R.-R. finalized the paper with input from all authors.

References

- Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning BA, et al. 2018. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* **46**: W537–W544. doi:10.1093/nar/gky379
- Bininda-Emonds ORP, Jeffery JE, Richardson MK. 2003. Inverting the hourglass: quantitative evidence against the phyletic stage in vertebrate development. *Proc Biol Sci* **270**: 341–346. doi:10.1098/rspb.2002.2242
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170
- Cannavò E, Khoueiry P, Garfield DA, Geeleher P, Zichner T, Gustafson EH, Ciglar L, Korbelt JO, Furlong EEM. 2016. Shadow enhancers are pervasive features of developmental regulatory networks. *Curr Biol* **26**: 38–51. doi:10.1016/j.cub.2015.11.034
- Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**: 25–36. doi:10.1016/j.cell.2008.06.030
- Coronado-Zamora M, Salvador-Martínez I, Castellano D, Barbadilla A, Salazar-Ciudad I. 2019. Adaptation and conservation throughout the *Drosophila melanogaster* life-cycle. *Genome Biol Evol* **11**: 1463–1482. doi:10.1093/gbe/evz086
- Duboule D. 1994. Temporal colinearity and the phyletic progression: a basis for the stability of a vertebrate bauplan and the evolution of morphologies through heterochrony. *Development* **1994**: 135–142. doi:10.1242/dev.1994.Supplement.135
- Ghandi M, Lee D, Mohammad-Noori M, Beer MA. 2014. Enhanced regulatory sequence prediction using gapped *k*-mer features. *PLoS Comput Biol* **10**: e1003711. doi:10.1371/journal.pcbi.1003711
- Ghandi M, Mohammad-Noori M, Ghareghani N, Lee D, Garaway L, Beer MA. 2016. gkmSVM: an R package for gapped-kmer SVM. *Bioinformatics* **32**: 2205–2207. doi:10.1093/bioinformatics/btw203
- Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. 2005. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**: 481–487. doi:10.1038/nature03235
- Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, Lee CM, Lee BT, Hinrichs AS, Gonzalez JN, et al. 2019. The UCSC genome browser database: 2019 update. *Nucleic Acids Res* **47**: D853–D858. doi:10.1093/nar/gky1095
- Hu H, Uesaka M, Guo S, Shimai K, Lu T-M, Li F, Fujimoto S, Ishikawa M, Liu S, Sasagawa Y, et al. 2017. Constrained vertebrate evolution by pleiotropic genes. *Nat Ecol Evol* **1**: 1722–1730. doi:10.1038/s41559-017-0318-0
- Huang W, Massouras A, Inoue Y, Peiffer J, Rämia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* genetic reference panel lines. *Genome Res* **24**: 1193–1208. doi:10.1101/gr.171546.113
- Irie N, Kuratani S. 2011. Comparative transcriptome analysis reveals vertebrate phyletic period during organogenesis. *Nat Commun* **2**: 248. doi:10.1038/ncomms1248
- Irie N, Kuratani S. 2014. The developmental hourglass model: a predictor of the basic body plan? *Development* **141**: 4649–4655. doi:10.1242/dev.107318
- Jeong S, Rebeiz M, Andolfatto P, Werner T, True J, Carroll SB. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* **132**: 783–793. doi:10.1016/j.cell.2008.01.014
- Kalinka AT, Tomancak P. 2012. The evolution of early animal embryos: conservation or divergence? *Trends Ecol Evol* **27**: 385–393. doi:10.1016/j.tree.2012.03.007
- Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, Ohler U, Bergman CM, Tomancak P. 2010. Gene expression divergence recapitulates the developmental hourglass model. *Nature* **468**: 811–814. doi:10.1038/nature09634
- Khouchou P, Girardot C, Ciglar L, Peng P-C, Gustafson EH, Sinha S, Furlong EE. 2017. Uncoupling evolutionary changes in DNA sequence, transcription factor occupancy and enhancer activity. *eLife* **6**: e28440. doi:10.7554/eLife.28440
- King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* **188**: 107–116. doi:10.1126/science.1090005
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359. doi:10.1038/nmeth.1923
- Lee D. 2016. LS-GKM: a new gkm-SVM for large-scale datasets. *Bioinformatics* **32**: 2196–2198. doi:10.1093/bioinformatics/btw142
- Lee D, Gorkin DU, Baker M, Strober BJ, Asoni AL, McCallion AS, Beer MA. 2015. A method to predict the impact of regulatory variants from DNA sequence. *Nat Genet* **47**: 955–961. doi:10.1038/ng.3331
- Levin M, Hashimshony T, Wagner F, Yanai I. 2012. Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Dev Cell* **22**: 1101–1108. doi:10.1016/j.devcel.2012.04.004
- Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, Senderovich N, Kovalev E, Silver DH, Feder M, et al. 2016. The mid-developmental transition and the evolution of animal body plans. *Nature* **531**: 637–641. doi:10.1038/nature16994
- Li Q, Brown JB, Huang H, Bickel PJ. 2011. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat* **5**: 1752–1779. doi:10.1214/11-AOAS466
- Liu PZ, Kaufman TC. 2005. Short and long germ segmentation: unanswered questions in the evolution of a developmental mode. *Evol Dev* **7**: 629–646. doi:10.1111/j.1525-142X.2005.05066.x
- Liu J, Robinson-Rechavi M. 2018a. Developmental constraints on genome evolution in four bilaterian model species. *Genome Biol Evol* **10**: 2266–2277. doi:10.1093/gbe/evy177
- Liu J, Robinson-Rechavi M. 2018b. Adaptive evolution of animal proteins over development: support for the Darwin selection opportunity hypothesis of Evo-Devo. *Mol Biol Evol* **35**: 2862–2872. doi:10.1093/molbev/msy175
- Liu J, Robinson-Rechavi M. 2020. Robust inference of positive selection on regulatory sequences in the human brain. *Sci Adv* **6**: eabc9863. doi:10.1126/sciadv.abc9863
- Liu J, Frochaux M, Gardeux V, Deplancke B, Robinson-Rechavi M. 2020. Inter-embryo gene expression variability recapitulates the hourglass pattern of evo-devo. *BMC Biol* **18**: 129. doi:10.1186/s12915-020-00842-z
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654. doi:10.1038/351652a0
- McGregor AP, Orgogozo V, Delon I, Zanet J, Srinivasan DG, Payre F, Stern DL. 2007. Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* **448**: 587–590. doi:10.1038/nature05988
- Nord AS, Blow MJ, Attanasio C, Akiyama JA, Holt A, Hosseini R, Phouanavong S, Plajzer-Frick I, Shoukyr M, Afzal V, et al. 2013. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**: 1521–1531. doi:10.1016/j.cell.2013.11.033
- Peng P-C, Khouchou P, Girardot C, Reddington JP, Garfield DA, Furlong EEM, Sinha S. 2019. The role of chromatin accessibility in *cis*-regulatory evolution. *Genome Biol Evol* **11**: 1813–1828. doi:10.1093/gbe/evz103
- Raff RA. 1996. *The shape of life: genes, development, and the evolution of animal form*. University of Chicago Press, Chicago.
- Reddington JP, Garfield DA, Sigalova OM, Karabacak Calviello A, Marco-Ferreres R, Girardot C, Viales RR, Degner JF, Ohler U, Furlong EEM.

2020. Lineage-resolved enhancer and promoter usage during a time course of embryogenesis. *Dev Cell* **55**: 648–664.e9. doi:10.1016/j.devcel.2020.10.009
- Richardson MK. 1995. Heterochrony and the phylotypic period. *Dev Biol* **172**: 412–421. doi:10.1006/dbio.1995.8041
- Richardson MK, Hanken J, Gooneratne ML, Pieau C, Raynaud A, Selwood L, Wright GM. 1997. There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development. *Anat Embryol (Berl)* **196**: 91–106. doi:10.1007/s004290050082
- Sander K. 1976. Specification of the basic body pattern in insect embryogenesis. *Adv Insect Phys* **12**: 125–238. doi:10.1016/S0065-2806(08)60255-6
- Schep AN, Adryan B. 2013. A comparative analysis of transcription factor expression during metazoan embryonic development. *PLoS One* **8**: e66826. doi:10.1371/journal.pone.0066826
- Valentine JW. 2004. *On the origin of phyla*. University of Chicago Press, Chicago.
- Wallace A. 2000. *The origin of animal body plans: a study in evolutionary developmental biology*. Cambridge University Press, Cambridge.
- Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet* **8**: 206–216. doi:10.1038/nrg2063
- Yanai I. 2018. Development and evolution through the lens of global gene regulation. *Trends Genet* **34**: 11–20. doi:10.1016/j.tig.2017.09.011
- Zalts H, Yanai I. 2017. Developmental constraints shape the evolution of the nematode mid-developmental transition. *Nat Ecol Evol* **1**: 0113. doi:10.1038/s41559-017-0113
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137. doi:10.1186/gb-2008-9-9-r137
- Zhu J, Sanborn JZ, Diekhans M, Lowe CB, Pringle TH, Haussler D. 2007. Comparative genomics search for losses of long-established genes on the human lineage. *PLoS Comput Biol* **3**: e247. doi:10.1371/journal.pcbi.0030247
- Zrzavý J, Štys P. 1997. The basic body plan of arthropods: insights from evolutionary morphology and developmental biology. *J Evol Biol* **10**: 353–367. doi:10.1007/s000360050029

Received January 4, 2021; accepted in revised form June 2, 2021.