

Adaptation and Conservation throughout the *Drosophila melanogaster* Life-Cycle

Marta Coronado-Zamora^{1,†}, Irepan Salvador-Martínez^{2,3,†}, David Castellano⁴, Antonio Barbadilla^{1,*}, and Isaac Salazar-Ciudad^{1,2,5,*}

¹Genomics, Bioinformatics and Evolution, Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

²Evo-Devo Helsinki Community, Centre of Excellence in Experimental and Computational Developmental Biology, Institute of Biotechnology, University of Helsinki, Finland

³Department of Genetics, Evolution and Environment, University College London, United Kingdom

⁴Bioinformatics Research Center, Aarhus University, Denmark

⁵Centre de Recerca Matemàtica, Cerdanyola del Vallès, Spain

[†]These authors contributed equally to this work.

*Corresponding authors: E-mails: antonio.barbadilla@uab.cat; isaac.salazar@helsinki.fi.

Accepted: April 16, 2019

Abstract

Previous studies of the evolution of genes expressed at different life-cycle stages of *Drosophila melanogaster* have not been able to disentangle adaptive from nonadaptive substitutions when using nonsynonymous sites. Here, we overcome this limitation by combining whole-genome polymorphism data from *D. melanogaster* and divergence data between *D. melanogaster* and *Drosophila yakuba*. For the set of genes expressed at different life-cycle stages of *D. melanogaster*, as reported in modENCODE, we estimate the ratio of substitutions relative to polymorphism between nonsynonymous and synonymous sites (α) and then α is decomposed into the ratio of adaptive (ω_a) and nonadaptive (ω_{na}) substitutions to synonymous substitutions.

We find that the genes expressed in mid- and late-embryonic development are the most conserved, whereas those expressed in early development and postembryonic stages are the least conserved. Importantly, we found that low conservation in early development is due to high rates of nonadaptive substitutions (high ω_{na}), whereas in postembryonic stages it is due, instead, to high rates of adaptive substitutions (high ω_a).

By using estimates of different genomic features (codon bias, average intron length, exon number, recombination rate, among others), we also find that genes expressed in mid- and late-embryonic development show the most complex architecture: they are larger, have more exons, more transcripts, and longer introns. In addition, these genes are broadly expressed among all stages. We suggest that all these genomic features are related to the conservation of mid- and late-embryonic development. Globally, our study supports the hourglass pattern of conservation and adaptation over the life-cycle.

Key words: adaptation, conservation, natural selection, evo-devo, DFE-alpha, hourglass hypothesis.

Introduction

The relationship between phylogeny and ontogeny has been a hotly debated topic in evolutionary biology since its origins (Darwin 1872; Gould 1977; Raff 1996). Any change in an organism's morphology is first a change in the developmental process leading to that morphology (Alberch 1980) and, thus, morphological divergence between species in evolution implies divergence in their development. It seems, thus, intuitive, that

there should be a relationship between divergence in species in a phylogeny and different stages along development. Several theories and hypotheses about this relationship have been proposed over the years (Gould 1977; Irie and Kuratani 2014). According to von Baer's hypotheses or laws (1828), early development stages are the most similar between species within a phylogenetic group (e.g., vertebrates), whereas late development stages are the most divergent.

© The Author(s) 2019. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Multiple theoretical justifications for this “law” have been proposed (Irie and Kuratani 2011). The most immediate is that changes in early development can have consequences in later development, because late developmental processes are causally dependent on the correct functioning of earlier developmental processes, whereas late developmental processes do not retroactively affect early developmental processes (Arthur 1977). As a result, early development should be more constrained.

The advent of developmental genetics in mouse and *Drosophila* and some observations in comparative embryology (Sander 1983) led to a different hypothesis: the hourglass hypothesis of embryonic development evolution (Medawar 1954; Slack et al. 1993; Duboule 1994). According to this hypothesis, early and late development would be more divergent between species than intermediate developmental stages (middevelopment). There is no consensus about which would be, exactly, these middevelopment conserved stages, but many researchers suggest that it should be some short time after gastrulation (Wilt and Hake 2004). The most conserved stage within a phylum is called the phylotypic stage (Sander 1983). There is no consensus either on what is actually conserved: embryonic morphology (von Baer 1828), developmental mechanisms (Raff 1996), or expression patterns of specific genes (Slack et al. 1993; Duboule 1994).

There is an extensive literature about the processes that may lead to an hourglass pattern. Some propose that many whole-body scale interactions take place during middevelopment, whereas during early and late development interactions are at a much more restricted spatial scale (Raff 1996), both at the level of mechanical interactions and molecular signaling between tissues. Accordingly, changes in middevelopment would be much more likely to affect the whole embryo and changes at other stages would have much more spatially restricted effects. Other authors argue that the hourglass pattern arises from different selection pressures acting in early and late development (Slack et al. 1993; Kalinka and Tomancak 2012).

There is also a long ongoing discussion about whether the hourglass hypothesis or the von Baer’s law (or something else) fit the observed patterns of divergence among developmental stages in phylogenies (Richardson et al. 1997; Poe and Wake 2004; Salazar-Ciudad 2010; Kalinka and Tomancak 2012; Hu et al. 2017; Yang et al. 2018). In principle, the von Baer and hourglass hypotheses could equally apply at the morphological, gene expression, or genomic level and there are indeed studies at these three levels. In here, we briefly explain their main conclusions to clarify why the questions, approaches, and results in this article significantly add to the understanding of the relationship between phylogeny and development. We focus, however, on studies combining the transcriptomic and genomic levels.

Many studies measure conservation at the DNA sequence level for genes expressed at different developmental stages to

explore whether some stages are indeed more conserved than others. The d_N/d_S ratio, also known as K_a/K_s or ω (ω from now on), is widely used as a proxy for conservation at the sequence level (Yang and Nielsen 1998; Hurst 2002). ω quantifies the level of constraint by comparing the rate of nonsynonymous substitutions per nonsynonymous site (d_N), which are presumably under selection, to the rate of synonymous substitutions per synonymous site (d_S), which are presumably neutral. Departures of neutrality are detected when ω is different of 1. Davis et al. (2005) used this statistic on 4,028 genes that were orthologous between *Drosophila melanogaster* and *Drosophila pseudoobscura*. By combining these analyses with *D. melanogaster* microarray expression data through development, they found that genes with the highest rates of nonsynonymous substitutions were expressed at low levels in late embryonic development and at high levels in the larva, pupa, and adult. The genes with the lowest rates of nonsynonymous substitution (the most conserved genes) were expressed at high levels in late embryonic development and at low levels before and after late embryonic development. This suggests, according to these authors, an hourglass pattern where embryonic stages spanning from 12 to 22 h are highly conserved between *D. melanogaster* and *Drosophila pseudoobscura*.

In another study by Kalinka et al. (2010), the number of nonsynonymous substitutions, d_N , was calculated for 3,019 genes in six different *Drosophila* species and their expression was measured by microarrays in each species in eight 2-h intervals during development. Consistently with the hourglass model, the study found that middevelopment, around the 10-h stage, is the period in which gene expression is the most conserved among the six species. The set of genes expressed in those stages has also the lowest average d_N values across the six species. A similar study by Mensch et al. (2013) estimated ω for more than 2,000 genes across six different *Drosophila* species for three categories of genes: maternal genes, genes expressed in early development, and genes expressed in late development. Maternal genes (whose mRNA is left by the mother in the egg) and late embryonic genes show higher ω than early expressed genes. In contrast, another *D. melanogaster* study by Artieri et al. (2009) found that genes expressed in the adult have higher ω than genes expressed in the pupa and those of the pupa have higher ω than those expressed in the embryo (for 7,180 analyzed genes), thus favoring von Baer’s law. In another study (Drost et al. 2015), the ω of the genes expressed in each stage are weighted by their expression to provide a stage-conservation measure. This measure supports the hourglass hypothesis in the zebrafish, *D. melanogaster*, and in the plant *Arabidopsis thaliana*. Similar studies exist for other species of animals (Roux and Robinson-Rechavi 2008; Piasecka et al. 2013; Liu and Robinson-Rechavi 2018b). The latter of such studies found that the genes expressed in middevelopment tend to be expressed in more developmental stages and that,

thus, they can be seen as more pleiotropic. This higher degree of pleiotropy would then explain the higher degree of conservation in middevelopment, at least in *Caenorhabditis elegans*, *D. melanogaster*, *Danio rerio*, the mouse *Mus musculus* (Liu and Robinson-Rechavi 2018b), and in other eight species of chordates (Hu et al. 2017).

A limitation of the aforementioned studies is that ω is not able to disentangle adaptive from nonadaptive substitutions. In fact, nonsynonymous substitutions can turn out to be adaptive, neutral, or slightly deleterious (strongly deleterious mutations are purged from the population and so they do not contribute to differences between species). In other words, a high ω can be the result of relaxed selection (low conservation), a high adaptive substitution rate, or a combination of both. Because adaptive substitutions tend to be fixed quickly (McDonald and Kreitman 1991), they will rarely be detected as polymorphic variants but only as divergent ones between species. Thus, adaptive substitutions can be inferred when there is an excess of nonsynonymous substitutions between species relative to nonsynonymous substitutions within a population, this is an excess of divergence in respect to polymorphism. Thus, in the McDonald and Kreitman test (MKT) (McDonald and Kreitman 1991), the divergence ratio (D_r/D_s) is divided by the polymorphism ratio (P_r/P_s) and adaptive substitutions are inferred if the ratio is larger than 1 [$(D_r/D_s)/(P_r/P_s) > 1$]. From that MKT, the proportion of adaptive to nonadaptive substitutions that have been fixed, α , can be estimated (Smith and Eyre-Walker 2002). The divergence and polymorphism in synonymous sites, which are assumed to be neutral, are used to estimate the underlying mutation rate and the expected polymorphism and divergence under a neutral scenario.

To try to circumvent known biases in the MKT, several modifications have been proposed over the years (Bustamante et al. 2002, 2005; Smith and Eyre-Walker 2002; Sawyer et al. 2003; Bierne and Eyre-Walker 2004; Mackay et al. 2012; Messer and Petrov 2013). Under stationary population size, slightly deleterious mutations lead to an underestimation of α because they tend to contribute more to polymorphism than to divergence (Eyre-Walker 2002). Because slightly deleterious substitutions tend to segregate at low frequency, its effect can be partially controlled by removing low-frequency polymorphisms from the analysis (Fay et al. 2001). However, Charlesworth and Eyre-Walker (2008) showed that even removing low-frequency variants, α estimates are always downwardly biased. Mackay et al. (2012) proposed an extension to the MKT, which we will call here extended MKT (eMKT), in which the count of segregating sites in nonsynonymous sites is partitioned into the number of neutral variants and the number of weakly deleterious variants. This increases the power for detecting selection and allows the independent estimation of both adaptive and weakly deleterious selection (see Materials and Methods).

Other methods correcting for the potential biases of the MKT estimates are based on the estimation of the distribution of fitness effects (DFE of mutations from DNA sequence polymorphism data) at functional sites (Bustamante et al. 2002, 2005; Eyre-Walker 2006; Eyre-Walker and Keightley 2007, 2009; Keightley and Eyre-Walker 2007; Boyko et al. 2008). One of the most popular DFE-based methods is the DFE-alpha (Eyre-Walker and Keightley 2009). The DFE-alpha method corrects for the segregation of slightly deleterious substitutions by first estimating the DFE at selected sites by a gamma distribution and then calculating how many nonadaptive substitutions are expected to become fixed given the inferred DFE from polymorphism data. Additionally, this method attempts to correct for possible effects of demography. Although this method is relatively recent, it has been used to estimate the rate of adaptive nonsynonymous substitutions in a number of studies (Strasburg et al. 2011; Carneiro et al. 2012; Kousathanas et al. 2014; Ávila et al. 2015; Santpere et al. 2015; Cagan et al. 2016; Galtier 2016; James et al. 2016; Murray et al. 2017; Steige et al. 2017).

There is no reason to expect that the patterns of adaptation and conservation in the genes expressed over developmental stages should be directly related to each other. It could be, for example, that middevelopment would be the most conserved stage but that it would still accumulate more adaptive nonsynonymous substitutions than other stages of development because nonsynonymous substitutions in the latter would be mostly neutral. We expect, however, that most adaptive substitutions occur in the adult, the larva, or the late stages of embryonic or pupal development. The reason for that would be that the larva and the adult have simply more functional aspects of the phenotype to select from (more cell types, more tissues, more organs), than developmental stages where most of the phenotype is still in the making. In addition, because *Drosophila* is a holometabolous genus, the larva and the adult effectively live in different habitats and, thus, one may exhibit larger rates of adaptive nonsynonymous substitutions than the other.

In this article, we study 1) whether *D. melanogaster* development follows the hourglass model or the von Baer's law, 2) whether there are differences not just in conservation but also in the rates of adaptive substitutions between stages, and 3) how these rates are related to specific genomic features such as gene size, codon usage bias, average intron length, intergenic distance, Guanine–Cytosine (GC) content, number of protein–protein interactions (PPI), number of exons and transcripts, expression bias, or recombination rates.

Question (2) has been recently approached through a completely different method (Liu and Robinson-Rechavi 2018a): the branch-site likelihood test (Zhang et al. 2005). This test is an extension of the ω ratio (Yang and Nielsen 1998; Hurst 2002), for the detection of positive selection for at least some codons in a phylogenetic branch of interest. Using this method, Liu and Robinson-Rechavi study three

phylogenetic branches, the Clupeocephala branch, the Murinae branch, and the Melanogaster group branch for focal species *Danio rerio*, *Mus musculus*, and *D. melanogaster*, respectively. The test contrasts two hypotheses: no positive selection occurred (H_0) in the phylogenetic branch of interest versus at least some codons experienced positive selection (H_1). One major advantage of the test is that it allows positive selection to vary both among codon sites and among phylogenetic branches. In our study, we use tests (MKT, DFE-alpha) that use both polymorphic and divergence data. Polymorphic data allow taking into account purifying selection in the divergent ω ratio, strongly increasing the power of detecting positive selection. These tests covers a shorter time scale, from the divergence of the outgroup species to the present, lowering the contingencies associated with longer evolutionary processes in the branch-site likelihood test. These include, for example, the poor sequence alignments due to increased divergence, which in turn can result in false positives.

In our study, we used the expression data in the modENCODE project (Graveley et al. 2011) from FlyBase. This is the most complete gene expression database through *D. melanogaster* life-cycle (it includes 17,788 genes over most developmental and life-cycle stages). We used divergence and polymorphism data for the genes expressed in each developmental stage to estimate four selection statistics α , ω , ω_a , and ω_{na} . ω_a is $\omega \times \alpha$, or the rate of adaptive substitutions per nonsynonymous site divided by the synonymous substitution rate (d_s) (Gossmann et al. 2012), whereas ω_{na} is $\omega \times (1 - \alpha)$, the rate of nonadaptive substitutions divided by d_s (Galtier 2016). In other words, ω_a is a measure of the rate of adaptive nonsynonymous substitutions, ω_{na} is the same measure for nonadaptive substitutions. Thus, ω_a informs about the overall rate of adaptive substitutions occurred in respect to the neutral mutation rate, the overall level of adaptation, whereas α simply measures how much common are adaptive substitutions in respect to nonadaptive ones (but does not tell if there has been many of them or not, as ω_a does). Polymorphism data come from the *Drosophila* Genetic Reference Panel (DGRP) project (Mackay et al. 2012) and divergence comes from comparing the genomes of *D. melanogaster* and *Drosophila yakuba*. Choosing *D. yakuba* as outgroup has been reported to provide a more reliable estimation of the selection statistics than the closest *Drosophila simulans* (Keightley and Eyre-Walker 2012) (see Materials and Methods).

Materials and Methods

Gene Expression

Gene Expression through the Developmental and Life-Cycle Stages

Gene expression data of 17,875 genes comes from RNA-seq experiments in the modENCODE project (Graveley et al.

2011). The data set contains the expression data for 30 stages of the whole life-cycle of *D. melanogaster*, including 12 embryonic samples collected at 2-h intervals for 24 h, six larval, six pupal, and three sexed adult stages at 1, 5, and 30 days after eclosion (Graveley et al. 2011). These data were downloaded from FlyBase (release 6.06, July 2015, file: "gene_rpkms_report_fb_2015_03," last accessed December 2015). The downloaded file reports gene expression values as reads per kilobase per million reads (RPKM). RPKM values are provided only for exonic regions of the gene (excluding segments that overlap with other genes), except for genes derived from dicistronic/polycistronic transcripts, where all exon regions were used for the estimation of RPKM expression. See Gelbart and Emmert (2013) for details. All RPKM values were \log_2 -transformed.

We tried different criteria to consider whether a gene is expressed in a stage or not. The following five criteria were applied: A "low stringent criterion," in which a gene is considered expressed in a stage if its RPKM is larger than 0 in that stage. This criterion is used in many other publications (e.g., Hebenstreit et al. 2011; Wagner et al. 2013; Guillén et al. 2019). A "low stringent criterion with 2-fold differential expression," in which a gene is considered expressed in a stage if its RPKM is larger than 0 in that stage and if, in addition, its maximal gene expression (over all the stages) is at least twice that of its minimal gene expression (also over all the stages). A "low stringent criterion with a 4-fold differential expression" that is as the 2-fold criterion but with a 4-fold differential expression criterion. A "medium stringent criterion," in which a gene is considered expressed if its RPKM is equal or higher than 2 and a "high stringent criterion" in which a gene is considered expressed if its RPKM is equal or higher than 10. This criterion is also used as a stringent criterion in RNA-seq analysis by other authors (e.g., Dezso et al. 2008). [Supplementary figure S1, Supplementary Material](#) online, shows the number of genes for each criterion and stage. As a result of applying each of these five criteria on the same modENCODE RNA-seq data, we obtain five different lists of genes for each developmental and life-cycle stages.

Gene Expression Profile Clustering

To identify shared temporal expression patterns among the genes of the modENCODE RNA-seq experiments, we applied a soft clustering method to the \log_2 -transformed expression values. We used a fuzzy c-means algorithm with the *mfuzz*() function of the R package *Mfuzz* (Futschik 2015). The *Mfuzz* soft clustering algorithm uses Euclidean distance as distance statistic and requires two main parameters (c = number of clusters and m = fuzzification parameter). For the clustering, we z-standardized the \log_2 -transformed expression values, so that the average expression value for each gene is 0 and the standard deviation is 1. The fuzzy soft clustering method is

different from hard clustering (like hierarchical clustering), in that genes are not uniquely assigned to one cluster. Instead of this, a gene i has gradual degrees of membership μ_{ij} to a cluster j . High membership values indicates a high correlation between gene i with the cluster centroid c_j (Futschik 2015). The *mfuzz()* function uses the fuzzy c -means algorithm, based on minimization of a weighted square error function with which the clusters centroids c_j result from the weighted sum of all clusters members. The membership value (μ_{ij}) indicates how well the gene i is represented by cluster j . Soft clustering is especially useful when clusters are not well defined, as in gene expression time-course data (Futschik and Carlisle 2005). We discarded genes that were constitutively expressed in all stages. Genes having a cluster membership lower than 0.8 were excluded. This resulted in 3,819 genes in eight clusters (one of the clusters was discarded as it was mostly noise) for the embryonic development out of the 5,514 genes that are expressed during the embryo development under the low stringent criteria, and 8,167 genes in nine clusters for the life-cycle of the 9,241 genes expressed in the whole life-cycle (discarding females). The values of c and m were optimized using the procedure described in Futschik and Carlisle (2005) and Futschik (2015), resulting in $c = 9$ for both data sets and $m = 1.23$ and 1.08 for the embryonic development and life-cycle, respectively. [Supplementary tables S1 and S2, Supplementary Material](#) online, show the number of genes expressed in each cluster for the two analyzes (for the embryo development and the whole life-cycle, respectively).

Testis and Immune Genes

For determining testis- and immune-related genes, we implemented the developed pipeline in Castellano et al. (2016). We downloaded the Gene Ontology (GO) terms for our gene set through the R package *biomaRt* (Durinck et al. 2005) using the *D. melanogaster* ENSEMBL database. When a gene was associated to any term related to the testis or the immune system (see [supplementary table S3, Supplementary Material](#) online, for the related GO terms) we removed it from the data set (a total of 171 out of 2,869 genes were removed). Those 171 exhibit higher rates of adaptation as it would be expected (permutation test, ω_a , $P = 0.028$; ω , $P < 0.001$) when compared against our whole data set (6,690 genes).

Sex-Biased Genes and Sex-Linked Genes

To control for the effect of genes with a sex-biased expression, we estimated the gene sex-bias ratio with the expression:

$$\text{sex bias ratio} = \log_2 \left(\frac{\text{RPKM}_{\text{♀}}}{\text{RPKM}_{\text{♂}}} \right).$$

And we discarded genes with a sex-biased above a $>|2|$ threshold in any of the three sampled adult points.

Additionally, we also removed genes that were exclusively expressed in one sex but no in the other. See [supplementary table S4, Supplementary Material](#) online, for the genes analyzed.

To control for the effect of the sex chromosome, we discarded X-linked genes from the analyses. See [supplementary table S5, Supplementary Material](#) online, for the genes analyzed.

Maternal, Maternal–Zygotic, and Zygotic Genes

A list of maternal, maternal–zygotic, and zygotic genes was obtained from Thomsen et al. (2010) data using egg and early development microarray analyses. Maternal genes are those genes whose mRNA is laid in the egg and are never transcribed by the embryo, maternal–zygotic those genes whose mRNA is laid in the egg by the mother but that are also transcribed by the embryo and zygotic genes are genes whose mRNA is not laid in the egg by the mother. The maternal gene list was obtained joining the original Thomsen's categories for not transcribed genes: “maternal decay,” “mixed decay,” “stable,” and “zygotic decay” categories (4,942 genes). The maternal–zygotic gene list was created by joining the categories of genes that are both present in the egg and transcribed later (1,332 genes analyzed): “maternal decay + transcription” and “stable transcription” categories. Finally, the zygotic gene list equals the “purely zygotic” category (850 genes). In the end, we get three lists of genes, one for the maternal genes, one for the maternal–zygotic, and one for the zygotic genes. See [supplementary table S6, Supplementary Material](#) online, for the genes analyzed.

Population Genomics

Selection Statistics

We used three different tests to estimate the selection statistics on the lists of genes expressed in each life-cycle stage: MKT standard (McDonald and Kreitman 1991), eMKT (Mackay et al. 2012), and DFE-alpha (Eyre-Walker and Keightley 2009). For the gene expression profiles and for maternal, maternal–zygotic, and zygotic gene categories, we used only the DFE-alpha method.

These three methods rely on polymorphism and divergence data to estimate adaptation. For obtaining this data, coding exon annotations from *D. melanogaster* were retrieved from FlyBase (release 5.50, www.flybase.org, last accessed March 2013). Genes 1:1 orthologs across *D. yakuba*–*D. melanogaster* were obtained from FlyBase (www.flybase.org, last accessed March 2013). We used *D. yakuba* as outgroup species because, aside from its high coverage (9.1×, Clark et al. 2007), the time elapsed because its divergence from *D. melanogaster* (7.4 Ma, Tamura et al. 2004) is more suitable for estimating adaptation (Keightley and Eyre-Walker 2012). In closely related species (as is the case of *D. melanogaster* and

D. simulans, which diverged 4.3 Ma [Cutter 2008]), the estimated rate of adaptive substitution can be biased due to 1) an erroneous attribution of polymorphism to divergence, 2) ancestral polymorphism contributing to divergence, and 3) differences in the rate of fixation of neutral and adaptive mutations (Keightley and Eyre-Walker 2012). The authors (Keightley and Eyre-Walker 2012) find that the adaptive rate estimated from closely related species (as in the case of *D. melanogaster* and *D. simulans*) may be underestimated by 10% or more. Choosing *D. yakuba* as outgroup gives a more reliable estimation of the adaptive rate than the closest *D. simulans* (Keightley and Eyre-Walker 2012).

The population genomic data comes from 168 inbred lines of *D. melanogaster* sequenced in the Freeze 1.0 of the DGRP project (Mackay et al. 2012). The DGRP population was created collecting gravid females from a single population of Raleigh, NC, followed by the full-sibling inbreeding approach during 20 generations to obtain full homozygous individuals. After this, the residual heterozygosity in the samples is expected to be 1.4% (inbreeding coefficient $F = 0.986$). The expected 1.4% of residual heterozygosity was true for 90% of the sequenced chromosome lines. DGRP lines showing high values of residual heterozygosity (>9%) were observed to be associated with large polymorphic inversions and, they were not included in our analyses (Huang et al. 2014). The inbreeding approach can alter the frequency spectrum of the strongly deleterious recessive mutations. However, alternative resources such as the *Drosophila* Population Genomics Project (DPGP) (Langley et al. 2012) would encounter the same problem. In fact, previous works comparing adaptation and DFE estimates between DGRP and DPGP data sets have shown no differences between those data sets (Castellano et al. 2016, 2018). And given that the DPGP sample size is smaller than the DGRP sample size, it is likely that these mutations contribute very marginally to the estimations of polymorphisms, DFE, and adaptation in both databases.

A multiple genome alignment between the DGRP isogenic lines (Mackay et al. 2012) and the *D. yakuba* genome using the BDGP 5 coordinates (Berkeley *Drosophila* Genome Project 5) was obtained from the publicly available database at <http://popdrowser.uab.cat> (Ràmia et al. 2012) (last accessed May 2010). For each gene, we took all nonoverlapping coding exons, independently of their inclusion levels. When two exons overlapped, the largest was chosen for subsequent analyses. Only exons without frameshifts, gaps, or early stop codons were retained. In this way, we tried to avoid potential alignment errors that would inflate our mutation rate estimates and create an artifactual positive correlation between them. Exonic sequences were trimmed in order to contain only full codons. We calculated the number of substitutions and the folded site frequency spectrum (SFS, Ronen et al. 2013) of the minor allele frequency (MAF) for short introns, 0-fold and 4-fold degenerate sites, using an ad hoc Perl script. The SFS was folded to avoid difficulties with misidentification

of the ancestral state (Hernandez et al. 2007) and because it performs well for inferring deleterious DFE (Eyre-Walker and Keightley 2007; Boyko et al. 2008; Tataru et al. 2017). We used a custom-made Perl script to estimate the number of short-intron substitutions and to compute the folded SFS. Jukes and Cantor correction for multiple hits was applied (Jukes and Cantor 1969). For computational reasons, one of the programs used for estimating adaptation (DFE-alpha, Eyre-Walker and Keightley 2009, see below) needs that all sites are sampled in the same number of individuals (Eyre-Walker and Keightley 2009). Hence, the original data of 168 lines set have been reduced to 128 isogenic lines by randomly sampling the polymorphisms at each site without replacement. Finally, residual heterozygous sites and sites with no quality value were excluded from the analysis.

MKT, Standard McDonald and Kreitman Test

The MKT (McDonald and Kreitman, 1991) was developed to be applied to protein-coding sequences, combining both divergent (D) and polymorphic (P) sites, and categorizing mutations as synonymous (P_S , D_S) and nonsynonymous (P_N , D_N). If all mutations are either strongly deleterious or neutral, then D_N/D_S is expected to roughly equal P_N/P_S . In contrast, adaptive mutations rapidly reach fixation and thus contribute relatively more to divergence than to polymorphism when compared with neutral mutations, and then $D_N/D_S > P_N/P_S$. Assuming that adaptive mutations contribute little to polymorphism but substantially to divergence, the proportion of adaptive nonsynonymous substitutions that have been fixed by selection can be inferred as $\alpha = 1 - (P_N/P_S)/(D_N/D_S)$ (Smith and Eyre-Walker 2002). The significance of effect can be easily quantified using a simple 2×2 contingency table (see table 1).

eMKT or Extended MKT

The null hypothesis of neutrality is rejected in a MKT not only when $D_N/D_S > P_N/P_S$ inferring adaptation but also when $P_N/P_S > D_N/D_S$. In this later case, there is an excess of polymorphism relative to divergence for the nonsynonymous class N , due to 1) slightly deleterious variants segregating at low frequency in the population, which contribute to polymorphism but not to divergence or 2) relaxation of selection where sites previously under strong or weak purifying selection have become neutral, causing an increased level of polymorphism relative to divergence.

Adaptive mutations and weakly deleterious selection act in opposite directions on the MKT, so α will be underestimated when the two selection regimes occur. Because slightly deleterious mutations tend to segregate at lower frequencies than do neutral mutations, they can be partially controlled for by removing low-frequency polymorphisms from the analysis, generally the 5% (Fay et al. 2001). However, this method is still expected to lead to biased estimates (Charlesworth and Eyre-Walker 2008). To take adaptive and slightly deleterious

Table 1

Standard MKT and eMKT Tables

Standard MKT Table			Number of Segregating Sites by MAF Category		
Site Class	Polymorphism	Divergence	Site Class	MAF < 5%	MAF ≥ 5%
S (neutral class)	P_S	D_S	S (neutral class)	P_S MAF < 5%	P_S MAF ≥ 5%
N (selected class)	P_N	D_N	N (selected class)	P_N MAF < 5%	P_N MAF ≥ 5%

mutation into account, P_N , the count of segregating sites in the nonsynonymous class, should be separated into the number of neutral variants and the number of weakly deleterious variants, $P_N = P_{N \text{ neutral}} + P_{N \text{ weakly del}}$. If both numbers are estimated, adaptive and weakly deleterious selection can be evaluated independently. Consider the following pair of 2×2 contingency tables (table 1).

The table on the left represents the standard MKT table with the theoretical counts of segregating sites and divergent sites for each cell. The table on the right contains the count of P_N and P_S for two-frequency categories. Mackay et al. (2012) showed using DGRP data that the ideal threshold to separate neutral and weakly deleterious mutations is a 5%, and that similar results were obtained using a 10% threshold, suggesting that most of the weakly deleterious variants are below a 5%, whereas the mutations above can be considered as nearly neutral variants. The estimate of the fraction of sites segregating neutrally within the MAF < 5% is $f_{\text{neutral MAF}<5\%} = P_{S \text{ MAF}<5\%}/P_S$. The expected number of segregating sites in the nonsynonymous class which are neutral within the MAF < 5% is $P_{N \text{ neutral MAF}<5\%} = P_N \times f_{\text{neutral MAF}<5\%}$. The expected number of neutral segregating sites in the nonsynonymous class is $P_{N \text{ neutral}} = P_{N \text{ neutral MAF}<5\%} + P_{N \text{ MAF}\geq 5\%}$.

To estimate α from the standard MKT table correcting by the segregation of weakly deleterious variants, we have to substitute the P_N by the expected number of neutral segregating sites, $P_{N \text{ neutral}}$. The correct estimate of α is then $\alpha = 1 - (P_{N \text{ neutral}}/P_S)/(D_N/D_S)$.

DFE-alpha

DFE-alpha (Eyre-Walker and Keightley 2009), an extension of the MKT, also corrects for the segregation of slightly deleterious alleles, providing a more accurate estimation than the MKT and other methods that do not take polymorphism data into account. Briefly, this software uses a maximum-likelihood method based on polymorphism data to infer the DFE of new mutations. It assumes two classes of sites in the genome: neutral sites (synonymous) and selected sites (nonsynonymous) and contrasts the SFS at these two classes. As a neutral reference, we used positions 8–30 of short introns (≤ 65 bp) following Halligan and Keightley (2006) (and 4-fold degenerated sites for some cross-validating analysis). As selected sites, we used 0-fold degenerate sites. Provided the SFS at both neutral and selected sites together with divergence data, the DFE-alpha method allows calculating of the proportion

of fixed substitutions that are adaptive (α) and the rate of adaptive substitutions relative to the neutral rate (ω_a , estimated as $\omega \times \alpha$, Gossmann et al. 2012). Furthermore, in our analysis, we include another statistic, ω_{na} (estimated as $\omega - \omega_a$), which represents the proportion of nonadaptive substitutions (slightly deleterious and neutral) relative to the neutral rate (Galtier 2016). Thus, we are able to decompose the classical ω ratio, into these two statistics: ω_a and ω_{na} , and differentiate whether high rates of ω are due to adaptive or nonadaptive nonsynonymous substitutions. We estimated natural selection on coding regions under a two-epoch demographic model.

Finally, to estimate these selection statistics with DFE-alpha, it is necessary to concatenate data from several genes because estimates from a single gene cannot be obtained due to the lack of segregating (or divergent) sites for some site classes. Thus, we calculated the selection statistics for the genes expressed in each stage and calculated the confidence intervals by randomly resampling the genes expressed in each stage, cluster or gene category (bootstrapping) 100 times with replacement (see *Bootstrapping* section).

Genomic Features

For the genes in each list and gene expression profile, we measured a number of genomic features. Coding exons and short-intron annotations from *D. melanogaster* were obtained from FlyBase (release 5.50, www.flybase.org, last accessed March 2013).

Average intron length. It is the average distance, in base pairs, between the successive exons of a gene. The effect of average intron length and total intron length into ω has been reported to be very similar (Marais et al. 2005).

Intergenic distance. It is the average distance, in base pairs, between the two closest genes to a given gene.

Gene size. Length of the coding DNA of a gene sequence (CDS).

Number of exons and transcripts. Number of different exons and transcripts of a gene, respectively. For the latter, we count all the transcripts reported for a gene in Fly Base. This is not the same as the number of different transcripts expressed in a stage.

Codon bias. Measured as the frequency of optimal codons, *Fop*. We used the software CodonW for the estimation of *Fop* (Peden 1999, www.codonw.sourceforge.net, last accessed June 2012). The index is estimated as the ratio of optimal

codons to synonymous codons. Values range between 0 (no optimal codons are used) and 1 (only optimal codons are used).

Expression bias. Proportion of development stages in which a gene is expressed according to modENCODE. Following Yanai et al. (2005) and Larracuente et al. (2008), we estimated the expression bias, τ , as (eq. 1):

$$\tau = \frac{\sum_{j=1}^n 1 - \frac{\log(S_j)}{\log(S_{\max})}}{n - 1}, \quad (1)$$

where $\log(S)$ is the logarithm of the RPKM and n is the number of developmental stages. τ ranges from 0 to 1. Values close to 0 indicate broadly expressed genes and values close to 1 indicate genes with highly biased expression. $\tau = 1$ means that expression is only detectable in a single sample.

Recombination rates. Recombination rates estimates at 100-kb nonoverlapping windows and microscopically observed crossing-over (CD) events were obtained from Comeron et al. (2012).

CG content. We used the software CodonW (Peden 1999, www.codonw.sourceforge.net, last accessed June 2012) to calculate the GC content of each protein-coding gene.

PPI. We downloaded PPIs for 10,631 protein-coding genes available at the Droid database, version Droid_v2018_08 (Murali et al. 2011, <http://www.droidb.org/Downloads.jsp>; last accessed January 2019). We integrated the information of six data sets available in Droid: PPI Curagen yeast two-hybrid, DPIM co-AP/MS, Finley Lab yeast two-hybrid, FlyBase curated, from other databases (BioGRID, IntAct, MINT, and BIND) and Hybrigenics yeast two-hybrid and counted the total unique protein interactions of each protein-coding gene. The number of genes included in FlyBase curated database was too small to base our analysis only on it.

Statistical Analysis

Bootstrapping

As explained before, to estimate the rate of adaptive evolution with the DFE-alpha method (Eyre-Walker and Keightley 2009), it is necessary to pool several genes together. The sampling distribution of parameters estimated by the DFE-alpha method at each stage, cluster or gene category was estimated by randomly sampling 100 times with replacement the genes at each temporal stage (or cluster or category). This concatenation was not necessary for the MKT and eMKT methods and, in fact, the same results are obtained with and without concatenation.

Permutation Test

To assess whether stages, gene clusters, or gene categories undergo differential selection compared with the genes not expressed in such stage, gene cluster or gene category, a

permutation test was applied. For obtaining a null distribution for the differences between gene groups, we shuffled without replacement 1,000 times the complete list of genes by means of ad hoc bash and Perl scripts. We estimated the rates of adaptive nonsynonymous substitutions and nonadaptive synonymous substitutions in each randomized list, obtaining an expected null distribution. The two-tailed P value was obtained by counting the number of replicates below and above the observed difference divided by the total number of replicates (i.e., 1,000), thus obtaining a two-tail P value. Multiple comparisons for each analysis were corrected by the false recovery rate approach.

See in the following [supplementary table S7, Supplementary Material](#) online, a summary of the permutation analysis performed in this study and [supplementary figure S2, Supplementary Material](#) online, for a graphical summary of the permutation procedure.

Spearman's Rank Correlations

Correlations between temporal profiles were carried out by Spearman's rank correlations, calculated by using the `cor.test()` function of the R program (R Core Team 2015).

Spearman's Partial Correlations

In order to test for the joint effect of ten genomic features on the selective statistics, we performed three independent multiple regression linear models, one for each selection statistic (ω , ω_a , and ω_{na}).

The selection statistics were considered response variables in each model and the ten genomic features as main effects (no interaction terms were included). Some of the ten genomic features showed significant pairwise correlations ([supplementary table S16, Supplementary Material](#) online) as determined using Spearman's partial correlations (Kim 2015). The analysis using multiple regression models were intended to disentangle their effects on the selection statistics. To assess the relative importance of each analyzed genomic feature in each linear regression model, we used the *pmvd* metric introduced by Feldman (2005) included in the R package *relaimpo* (Grömping 2006), that averages the R^2 of each regressor over all possible orderings using weighted averages.

Results

Overall Temporal Pattern of Adaptation and Conservation

We calculated four selection statistics (three based on polymorphism and divergence data ω_a , ω_{na} , and α and one based on divergence data alone, ω) for the set of genes expressed in each developmental and life-cycle stage.

We estimated these statistics using three different methods, DFE-alpha (Eyre-Walker and Keightley 2009) (fig. 1), the standard MKT (McDonald and Kreitman 1991)

(supplementary fig. S3, Supplementary Material online), and the eMKT (Mackay et al. 2012) (supplementary fig. S3, Supplementary Material online), and using different criteria to consider whether a gene is expressed in a stage or not (supplementary figs. S4–S6, Supplementary Material online). In these calculations, the mutation rate was estimated using short introns (fig. 1) and 4-fold degenerated sites (supplementary fig. S7, Supplementary Material online). Finally, we also explored whether the number of genes in a stage has an effect on the values of the estimated metrics. For that, we repeated the analyses by sampling 350 genes per stage (with replacement) 100 times and calculating the mean values for the selection metrics (supplementary fig. S8, Supplementary Material online). For each combination of methods and expression criteria, we obtain very similar patterns of change of ω_a , ω_{na} , ω , and α over stages (what we call the temporal pattern of these selection statistics).

Figure 1 shows the temporal pattern found when using the DFE-alpha method, short-intron sites as a proxy for the mutation rate, and considering all the genes with nonzero expression (excluding the 6,655 genes that were constitutively expressed throughout all stages, see Materials and Methods and supplementary table S8, Supplementary Material online, for the genes analyzed). A total of 2,869 genes are considered with this criterion. The rate of adaptive (ω_a) and, especially, nonadaptive (ω_{na}) nonsynonymous substitutions on the set of genes expressed in each developmental stage (fig. 1) was the highest in the first embryonic stage. Both rates gradually decrease until the 10-h embryo stage. The next developmental stages (across mid- and late-embryonic development) show, on the contrary, the lowest rates of fixation of substitutions (either adaptive or not). At the third larval stage (L3), the rates of adaptive substitution (ω_a), and to a lesser extent of nonadaptive substitution (ω), increase and remain high through all the pupal stages. Finally, in the male adult stage, ω_a and ω values are very similar to those of the pupa, whereas female adults exhibit lower values.

Overall, all the selection statistics show a similar temporal trend except for nonadaptive substitutions being seemingly more abundant in the genes expressed in the very early stages as compared with later stages (a difference we further analyze in a coming section). To analyze whether these differences between stages were statistically significant, we merged stages into eight developmental periods: embryo 0–2 h, embryo 2–6 h, embryo 6–24 h, larva 1–3, larva 4–6, pupa, female, and male adults (see supplementary table S9, Supplementary Material online, for the genes analyzed). We calculated by a permutation test the chances that the genes expressed in a period undergo differential selection compared with the genes not expressed in that period. First, we calculated the difference in selection statistics between the two groups, that is, the genes of a period and the genes not expressed in that period. To obtain

the null distribution, we shuffled without replacement 1,000 times the complete list of genes expressed during the whole life-cycle (2,869 genes) and randomly assigned the genes into one of the two groups. We estimated the selection statistics in each randomized list, obtaining an expected null distribution that we then compared with the observed difference. Multiple comparisons were corrected by the false recovery rate approach. This analysis shows that mid- and late-embryonic development, the beginning of the larva and genes expressed in female adults show significantly low rates of nonsynonymous substitutions (supplementary fig. S9, Supplementary Material online). The relatively high rates of substitutions in early development and in the larva, pupa, and males are not significant in this analysis (P values can be consulted on supplementary table S10, Supplementary Material online).

However, if the same permutation test is done using all genes (expressed in all stages or not) the test shows that, in addition, early development, late larva, pupa, and male adult exhibit significantly high ω_a , ω , and α values (supplementary fig. S10, Supplementary Material online). See the P values of all comparisons in supplementary table S11, Supplementary Material online.

Immune system and testis-related genes have been reported to be under faster rates of adaptation than other genes (Pröschel et al. 2006; Obbard et al. 2009). Immune system genes are expressed mostly in adults, and testis genes are expressed only in adult males. It could then be that the low rates of conservation in the genes expressed in the male adult stages would be explained by the low rates of conservation of immune- and testis-related genes. To explore this possibility, we repeated our analysis excluding immune system and testis genes. The exclusion of these genes does not modify our results much (supplementary fig. S11, Supplementary Material online, see supplementary table S3, Supplementary Material online, for the GO terms excluded), so the high rates of substitution we observed in adult males are not exclusively due to these immune- and testis-related genes. We additionally controlled for the effect of genes with a sex-biased expression (see Materials and Methods) by removing genes with a sex-biased ratio above a ratio of 2 (supplementary fig. S12, Supplementary Material online). As expected, sex-biased genes are responsible of the observed differences between males and females adaptation rates. When removing sex-biased genes the female stages exhibit selection statistics values similar to those of males. This indicates that the genes that are overexpressed in females tend to be more constrained.

X-linked genes are known to be under faster rates of evolution compared with autosomal genes (faster-X effect) (Meisel and Connallon 2013). When we excluded X-linked genes from the analyses (supplementary fig. S13, Supplementary Material online) the selection pattern

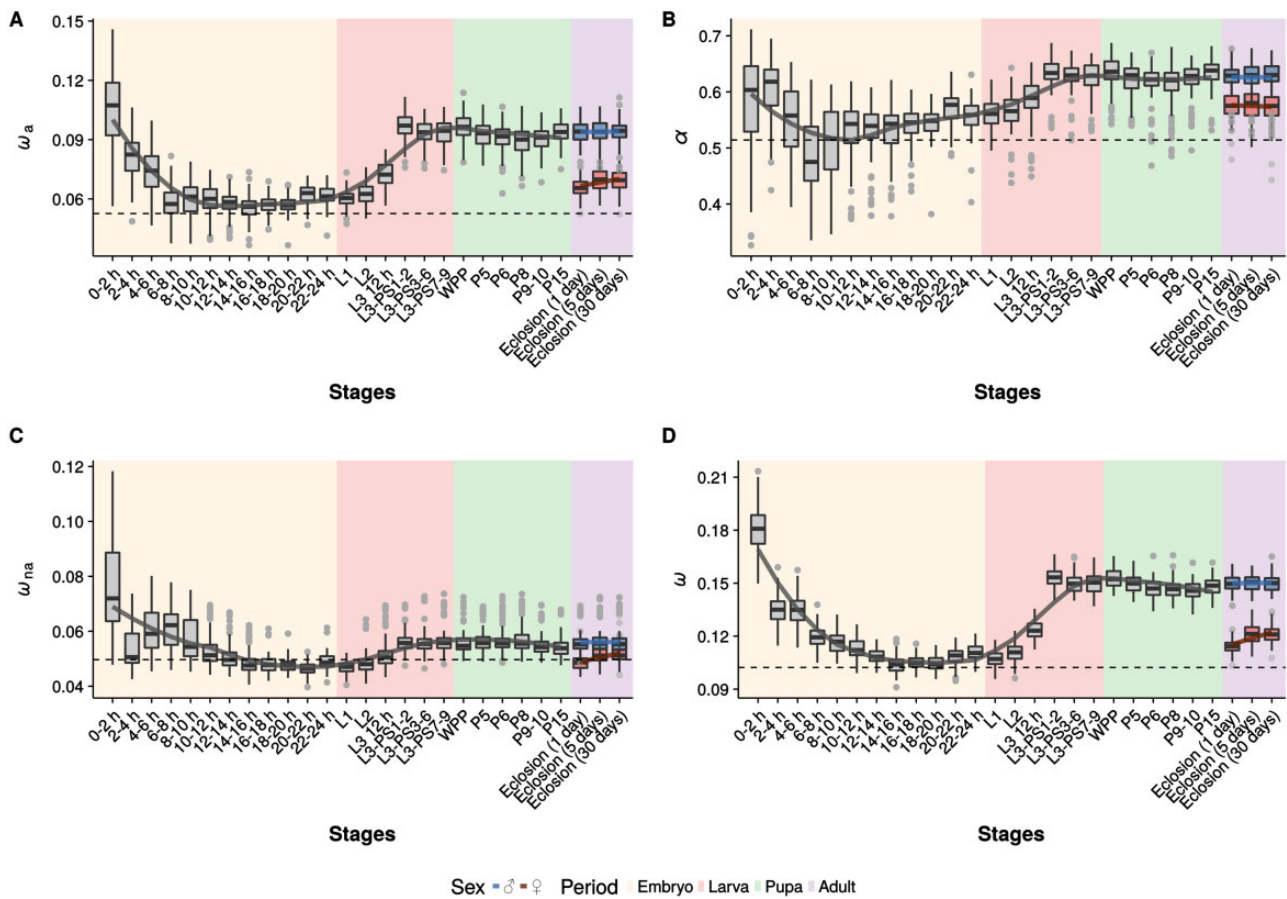


FIG. 1.—Temporal pattern of the four selection statistics estimated with DFE-alpha (ω_a , α , ω_{na} , and ω). (A) ω_a , the rate of adaptive nonsynonymous substitutions relative to the mutation rate. (B) α , the proportion of base substitutions fixed by natural selection. (C) ω_{na} , the rate of nonadaptive nonsynonymous substitutions relative to the mutation rate. (D) ω , the rate of nonsynonymous substitutions relative to the mutation rate. Each boxplot (A–E, 100 bootstrap replicates per stage) in a plot is calculated for a randomly drawn sample of the set of genes expressed in a stage with replacement. The solid line going through the boxplot is inferred by LOESS. For the male and female stages, the line is simply a linear regression. The dashed line shows the mean value of each statistic for the genes that are expressed in all stages (again with 100 bootstrap replicates). The embryonic stages are named by the hour’s intervals (from 0 to 24 h), the larval stages are the first instar (L1), second instar (L2), and third instar (L3). The L3 stages are subdivided into the first 12 h (L3-12h) and several puff stages (L3-PS1 to L3-PS7). WPP is the white prepupae stage. The pupal stages with RNA-seq are phanerocephalic pupa, 15 h (P5), 25.6 h pupa (P6), yellow pharate, 50.4 h (P8), amber eye-pharate, 74.6 h (P9–10), and green meconium pharate, 96 h (P15). Adult stages are 1, 5, and 30 days after eclosion (1, 5, and 30 days). Number of genes analyzed is in [supplementary table S8, Supplementary Material](#) online. The earliest stages show more variation in the selection statistics because they have less genes.

remained similar, so the rates of substitutions observed are not exclusively due to X-linked genes.

Gene Expression Profiles Clustering

There are at least three different scenarios that could explain the observed temporal pattern of change in the selection statistics. In the case of ω_a , for example, it could be that a subset of genes with high ω_a is expressed just with the observed temporal pattern (everywhere but in mid- and late-development and in females). Alternatively, it could be that each of the time periods with high ω_a would express a distinct group of genes that have high levels of adaptive substitutions. It could also be that no simple correspondence exists between

the high ω_a in a time period and the expression of a specific subset of genes in it.

To explore these possibilities, we categorized all the analyzed genes into classes based on their temporal profiles of expression. To do that we use an unsupervised soft clustering algorithm (Futschik and Carlisle 2005) as explained in the Materials and Methods. Genes within each temporal expression class show relatively similar changes in gene expression levels over time. We consider eight such classes for embryonic development (fig. 2A, number of genes analyzed in each cluster in [supplementary table S1, Supplementary Material](#) online) and nine classes for the whole life-cycle (see [supplementary fig. S14, Supplementary Material](#) online, for the temporal profiles, number of genes analyzed in each cluster in

supplementary table S2, Supplementary Material online). For embryonic development, clusters 1 and 2 are the ones showing significantly the highest ω_a and ω compared with the other clusters using a permutation test (cluster 1, ω : $P < 0.001$; ω_a : $P = 0.008$; cluster 2, ω : $P < 0.001$; ω_a : $P = 0.059$) (fig. 2B–E). These clusters correspond to the genes that are expressed at high levels in the earliest development and that rapidly decrease their expression to very low levels (fig. 2A). ω_a , ω , and α values in clusters 1 and 2 are larger than those in the first three developmental stages and, thus, it is likely that the genes in these clusters, but not the other genes that are expressed in these three stages, are responsible for the high ω_a , ω , ω_{na} , and α values in the earliest development. The decline in the values of these selection statistics over early development would then just be a simple reflection of the decrease in expression of the genes in those clusters over time. Cluster 8 also shows larger ω than that observed in the other clusters (ω : $P < 0.001$). This cluster is composed of genes whose expression increases only in the last stages of embryonic development. This high ω_a cannot be detected when directly analyzing the genes in each stage because the other genes expressed in these late stages have lower ω values, as it can be seen for cluster 5, which expresses genes from the 10-h onward. Thus, the temporal pattern of change in the selection statistics seems to come from the temporal dynamics of expression of three different sets of genes (those of clusters 1, 2, and 8). P values can be consulted on supplementary table S12, Supplementary Material online. Similar results were found when the clustering was done over the whole life-cycle (supplementary fig. S15 and table S13, Supplementary Material online).

Genomic Features

To search for sequence properties affecting the temporal pattern of the selection statistics, we measured a number of genomic features for each of the genes expressed in development (see Materials and Methods). These are gene size (the length of the coding sequence of a gene), number of exons (total number of exons of a gene; not to confuse with the number of exons a gene expresses in a specific developmental stage), codon usage bias (measured as frequency of optimum codons, *Fop*), number of transcripts per gene (this is not the number of alternative transcripts of a gene in a stage but the total number of alternative transcripts of a gene according to FlyBase annotations), average intron length (measured as the average distance in base pairs between the exons of a given gene), intergenic distance (average distance in base pairs between two adjacent genes), expression bias (a measure of how evenly distributed the expression of a gene is over time), recombination rate (based in observed cross-overs in 100-kb intervals, from Comeron et al. [2012]), the GC

content of each gene, and the number of PPIs (Murali et al. 2011).

First of all, we assessed the relationship between the genomic features and selection statistics, clumping all expressed genes together irrespectively of the stage in which they are expressed. For that, we categorized the genes in our data set in five different categories based on each genomic feature (see supplementary table S14, Supplementary Material online, for the number of genes considered in each category) and resampled with replacement 100 times the genes in each category and estimated the selection statistics in each such category. We found clear negative correlations between ω_a and gene size, number of exons, number of transcripts, average intron length, codon bias, GC content, and PPIs (similar correlations were found for ω_{na} [supplementary fig. S16, Supplementary Material online], ω , and α [data not shown]) and positive correlations between ω_a and expression bias, recombination rate, and intergenic distance (supplementary fig. S17, Supplementary Material online). Thus, in general, less complex genes (shorter, with fewer exons and transcripts and short introns), expressed in only a small number of stages, with lower GC content and/or less PPIs are more likely to accumulate adaptive substitutions (high ω_a).

We performed linear multiple regression models to test for the dependence of the selective statistics (ω , ω_a , and ω_{na}) estimated on each protein-coding gene using the eMKT on the ten genomic features (supplementary table S15 and fig. S18, Supplementary Material online). Supplementary table S16, Supplementary Material online, show the matrix of correlation coefficients from the pairwise correlations between the ten genomic features included in the linear models. The coefficient of determination (multiple R^2) was the highest for ω , explaining a 21.21% of the total variance ($P = 2.2 \times 10^{-16}$). For ω_a , it was 3.34% ($P = 2.2 \times 10^{-16}$) and for ω_{na} it was 9.36% ($P = 2.2 \times 10^{-16}$). Average intron length, size, and PPI are the only genomic features that do not have a significant effect on any of the selective regimes. Expression bias is the genomic feature with the highest effect on ω , as observed in Guillén et al. (2019).

When considered over developmental stages we found that these genomic features exhibit a temporal pattern that is either very similar to that of ω_a or its opposite (see fig. 3). To analyze that we calculated the correlation between each stage's average ω_a and the average of each genomic feature in each stage (see table 2). Gene size, number of exons, codon usage bias, and number of transcripts per gene follow a temporal pattern that is the reverse of that of ω_a . Average intron length also shows a temporal pattern contrary to that of ω_a except that no clear differences between stages are found after embryonic development. The intergenic distance shows a temporal pattern similar to that of ω_a , except that this distance is low in the earliest stages in which ω_a is high. The average expression bias shows a temporal pattern similar to that of ω_a except for an overall increase over time. The same

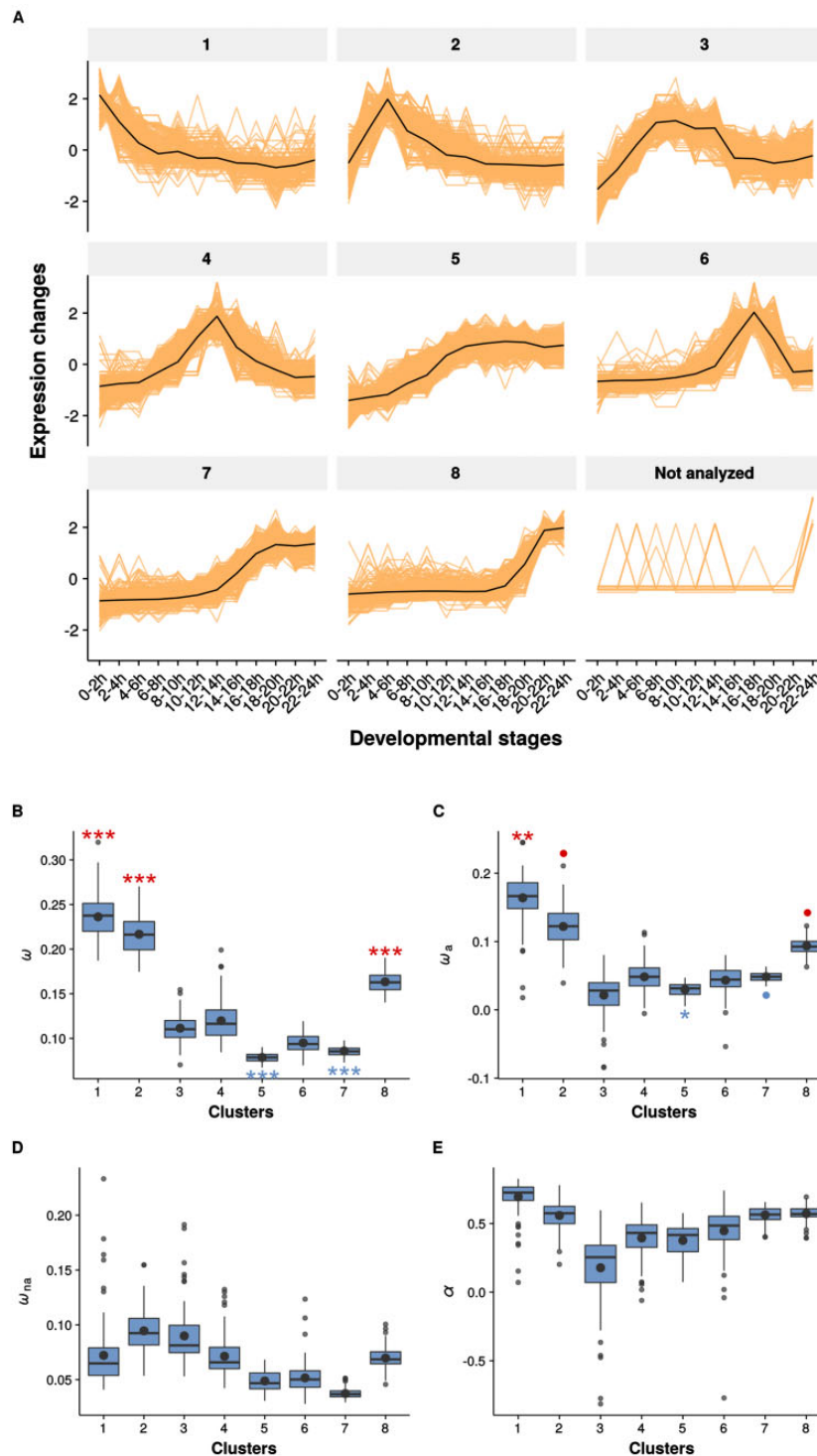


FIG. 2.—Clusters of temporal profiles of expression of embryonic development genes and their ω_{ar} , α , ω_{na} , and ω estimated using DFE-alpha. (A) Temporal expression profile for all the genes belonging to each cluster. (B) ω for each cluster. (C) ω_a sampling for each cluster. (D) ω_{na} for each cluster. (E) α for each cluster. Each point in the plots in (B)–(E) is calculated for a randomly drawn sample of the set of genes in each cluster with replacement (100 bootstrap replicates per cluster). Asterisks represent the significance by a permutation test, the color indicates whether the value was higher (red) or lower (blue) than expected (\bullet , $0.1-0.05$, $* < 0.05$, $** < 0.01$, and $*** < 0.001$). Number of genes analyzed in [supplementary table S1, Supplementary Material online](#). Permutation P values are shown in [supplementary table S12, Supplementary Material online](#).

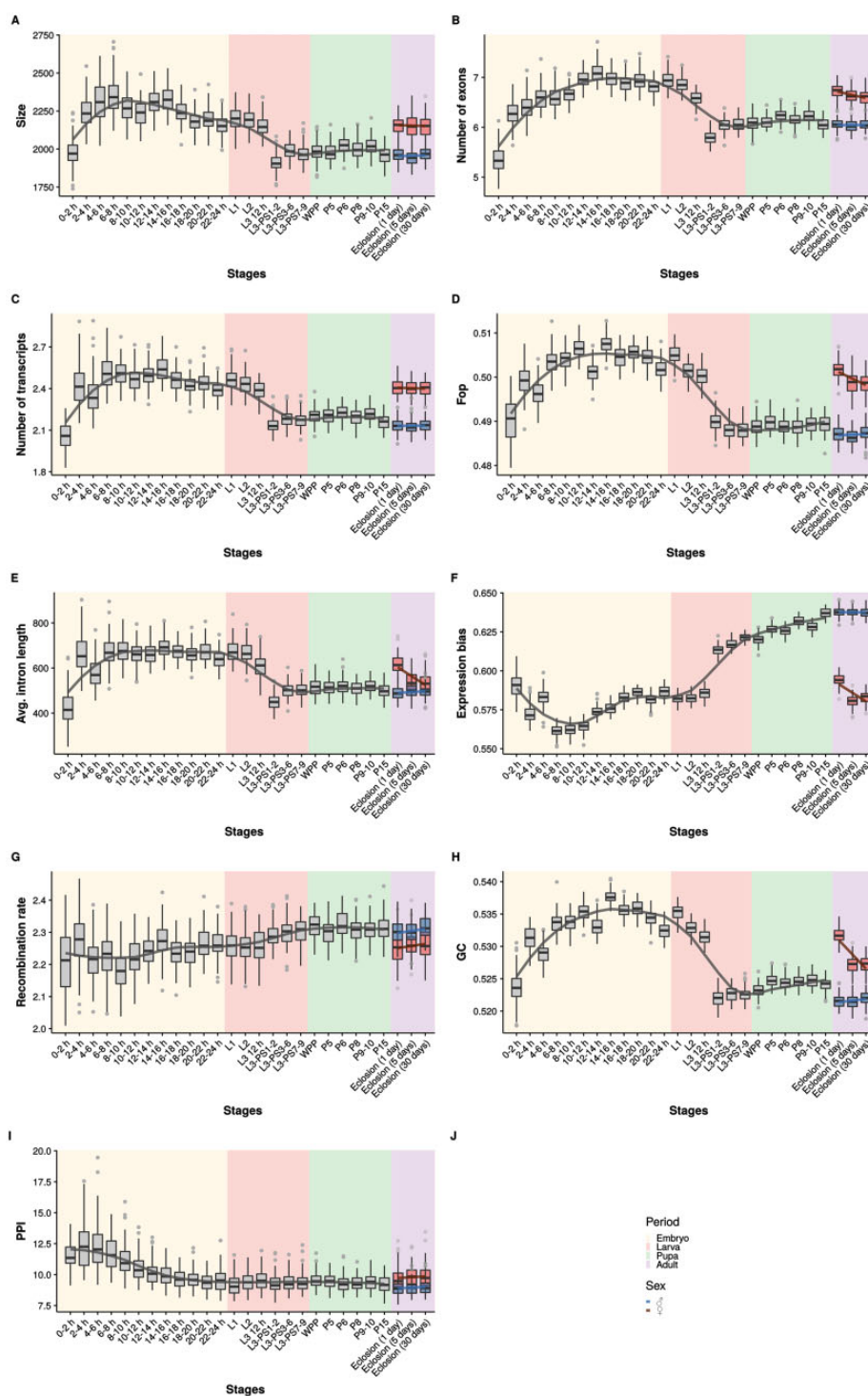


FIG. 3.—Nine genomic features over developmental stages. Lines and stages as in figure 1. (A) Size is the coding sequences length of a gene in base pairs. (B) Number of exons is the number of exons for the genes expressed in a stage. (C) Number of transcripts is the number of different isoforms of each gene expressed in a stage. (D) *Fop* is a measure of codon usage bias: the ratio of optimal codons to synonymous codons. (E) Average intron length is the average distance, in bases, between the exons of a gene. (F) The expression bias is a measure of how much the expression of a gene is restricted to one or few stages estimated as equation (1) (see Materials and Methods). (G) Recombination rate is estimated in 100-kb nonoverlapping windows. (H) CG content of each gene. (I) PPIs are estimated as the number of PPIs of each gene. Mean sampling distribution was obtained by resampling 100 times with replacement the genes from each stage. See [supplementary table S8, Supplementary Material](#) online, for the genes analyzed. The same patterns are found when using 4-fold data, see [supplementary figure S19, Supplementary Material](#) online.

correlations are found when we used 4-fold degenerated sites as a proxy for the mutation rate (see [supplementary table S17](#) and [fig. S19, Supplementary Material](#) online). A similar pattern is found when we analyze the genomic features in the gene expression clusters for the embryo development and life-cycle (see [supplementary fig. S20, Supplementary Material](#) online, and P values in [supplementary table S18](#) and [fig. S21, Supplementary Material](#) online, and P values in [supplementary table S19, Supplementary Material](#) online, respectively).

Analysis of Maternal, Maternal–Zygotic, and Zygotic Genes

To further explore why ω_a and ω_{na} were high in the earliest developmental stages, we analyzed separately maternal, maternal–zygotic, and zygotic genes. For that purpose, we used a microarray study (Thomsen et al. 2010) that categorized developmental genes as maternal, zygotic, and maternal–zygotic by determining which gene transcripts are already present in the egg and which ones are not. Maternal genes are defined as genes whose mRNA is laid within the egg by the mother and are never transcribed by the embryo. The embryo contains, thus, mRNAs coming from two different genomes, that of the mother and that of the embryo. Maternal–zygotic genes are genes whose mRNA is laid in the egg by the mother but that are also transcribed by the embryo. Zygotic genes are genes whose mRNA is not laid in the egg by the mother. We analyzed the genes reported in Thomsen et al. (2010) for each category with a permutation test, testing if the genes in each category undergo differential selection compared with the genes in the other two categories (see Materials and Methods for details). We found that ω_a was not significantly different between categories when assessed with a permutation test ([fig. 4](#)). This implies that the large ω_a of the earliest stages is not due to any specific gene category. Consistent with the hypothesis of lower efficiency of natural selection on maternal genes both ω ($P=0.024$) and ω_{na} ($P=0.003$) were higher for maternal genes than for zygotic genes (and intermediate for the maternal–zygotic genes). Zygotic genes show lower values of ω_a and ω_{na} than expected from the permutation test ($P=0.035$ and $P=0.036$, respectively). [Supplementary table S20, Supplementary Material](#) online, contains permutation P values. Finally, this analysis was repeated but using the gene that are in common with the genes expressed in the first hours according to modENCODE to check whether maternal genes account for the high ω_{na} in the first stages of the development ([fig. 1](#)). Very similar results are obtained, thus, indicating that the high ω_{na} values in the earliest stages are due to the maternal genes in these stages ([supplementary fig. S22, Supplementary Material](#) online). Even so, the ω_a in these earliest stages is still larger than in mid-and late-development.

Discussion

Three main conclusions can be derived from our analyses. First, the rate of adaptive substitution (ω_a) measured along the life-cycle of *D. melanogaster* reveals two peak periods: one encompassing the four initial hours of the embryonic development and one encompassing from the L3 larval stage onward. *Drosophila melanogaster*, as all holometabolous insects, has an indirect development with two active free-roaming phases, the larva and the adult, and two inactive sessile developmental phases, the embryo and the pupa. The larval and adult phenotypes, especially their morphology, arise primarily through the genetic, cellular and tissue interactions of embryonic and pupal (metamorphosis) development, respectively. Therefore, adaptation in the larva and the adult should be reflected, not only in the substitution rates of the genes expressed in the larva and adult but also in those expressed during embryonic development (for the larva) and pupal development (for the adult). The observation that genes expressed in mid-and late-embryonic development show relatively lower rates of nonsynonymous substitution than the genes expressed in the larval and pupal stages suggests that adaptation has occurred preferentially in the adult rather than in the larva. In contrast to a previous report using a smaller gene set (Artieri et al. 2009), we do not find that adults have higher rates of nonadaptive substitutions, neither in ω nor in ω_{na} , than the pupa. In a previous study, it was found that the 150 genes with the highest number of nonsynonymous substitutions (this is the highest ω) are more intensively expressed in the larva and pupa than in the embryo and that their highest level of expression is in the male adults (Davis et al. 2005).

An important novelty of our analysis is the incorporation of polymorphism data, which allows us to more precisely distinguish between the rates of nonsynonymous substitutions (with conservation indicated by low ω), adaptive nonsynonymous substitution (measured by ω_a), nonadaptive nonsynonymous substitution (measured by ω_{na}), and the proportion of fixed adaptive substitutions (α). In a previous study (Salvador-Martínez et al. 2018), we have estimated the same selection statistics over the embryonic anatomy of *D. melanogaster*, in this study, we do it over developmental time. This has allowed us to infer that the relatively lower level of conservation in early pupal and male stages is due to adaptive nonsynonymous substitutions and not to nonadaptive nonsynonymous substitutions. In the earliest developmental stages, instead, we can infer that the lower sequence conservation is due, mostly, to nonadaptive substitutions. Our results are, thus, compatible with those of Liu and Robinson-Rechavi (2018a) but the methods we use and the inclusion of polymorphism data, however, allow us to infer the nonadaptive nature of the nonsynonymous substitutions in the earliest stages of development.

Our analysis shows that the latter is due, mostly, to maternal genes. Previous studies have already pointed out that

Table 2Spearman's Correlations between ω_a and Genetic Features

Genomic Feature	Relation with ω_a	Correlation (r^2) for Females (P Value)	Correlation (r^2) for Males (P Value)
Average intron length	Negative	0.802 (1.12×10^{-6})	0.808 (1.08×10^{-6})
Size	Negative	0.731 (1.56×10^{-6})	0.764 (1.39×10^{-6})
Number of exons	Negative	0.862 (6.53×10^{-7})	0.886 (4.82×10^{-7})
Number of transcripts	Negative	0.874 (5.70×10^{-7})	0.870 (5.94×10^{-7})
<i>Fop</i>	Negative	0.759 (1.42×10^{-6})	0.688 (1.71×10^{-6})
Expression bias	Positive	0.508 (4.89×10^{-5})	0.552 (1.58×10^{-5})
Recombination	Positive	0.330 (2.07×10^{-3})	0.334 (1.91×10^{-3})
Intergenic distance	N.S.	0.043 (0.299)	0.082 (0.148)
CG content	Negative	0.905 (3.647×10^{-7})	0.772 (1.34×10^{-6})
PPI	Negative	0.152 (0.045)	0.075 (0.165)

NOTE.—Spearman's correlations performed between each stage's average ω_a and the average of each genomic feature in each stage. Females and males are separated because their gene expression is measured separately in the last three stages in the modENCODE. Thus, the data considered for females and males are the same for all the stages but the last three. Methodological details about how each genomic feature was calculated are explained in the Materials and Methods. *Fop*: frequency of optimum codons.

selection in maternal genes is less efficient (although these studies do not relate that to the lower conservation of early development [Cruickshank and Wade 2008; Wade et al. 2009]). This is because, in females, the alleles in the loci of maternal genes can affect the fitness of the offspring but that is not the case for males (because males lay no eggs and thus no maternal mRNAs in their offspring). As a result, for maternal genes, selection cannot act as effectively on males as on females. Natural selection is then less effective and many non-adaptive variants cannot be eliminated from the population (leading to the higher ω_{na}).

The high rate of nonsynonymous nonadaptive substitutions in the earliest developmental stages parallels recent findings in developmental genetics: the genes expressed in the earliest developmental stages were found to be different within *Diptera*, whereas the zygotic genes expressed right after, the gap genes, were found to be the same in all the *Diptera* species analyzed so far (Wotton et al. 2015).

A second conclusion of our study is that the temporal pattern of the selection statistics mirrors that of the genomic features analyzed. Thus, mid- and late-embryonic development express genes that have, on average, more exons, more different transcripts, a more optimal codon usage, larger introns, and larger gene size. Some of these genomic features were previously found to correlate with ω , and in some few cases with ω_a . However, their distribution over developmental stages and their relationship with the von Baer's law or the hourglass model have not been analyzed before, except for intron length (Liu and Robinson-Rechavi 2018b) and gene length (Liu and Robinson-Rechavi 2018a; Yang et al. 2018). This latter study also reports how other measures of gene complexity (such as the number of protein domains) are higher among developmentally expressed genes.

The negative correlation between codon bias and adaptive nonsynonymous substitution at the protein level was already known (Hershberg and Petrov 2008; Presnyak et al. 2015). Such negative correlation should be expected by mere

probability because, for any given protein, the codon changes that improve a protein function would often be different from the codon changes associated with more efficient codon usage (Hershberg and Petrov 2008; Presnyak et al. 2015), especially for highly expressed genes (Pal et al. 2001). The relationship between codon bias and GC content we observe, that is, identical temporal trends over development (see fig. 3), is explained by many optimal codons ending in G or C in *D. melanogaster* (Bierne and Eyre-Walker 2006).

Similarly, it has previously been reported that genes with a large expression bias tend to be less conserved (low ω , Larracuenta et al. 2008) and we found, that in addition, they also show higher rates of adaptive substitutions. On average, if a gene is expressed in a very specific time window, it is likely involved in regulating a smaller number of developmental interactions than if it is expressed in through many stages. Changes in such gene are then less likely to interfere with many different developmental processes and, thus, these changes are less likely to be deleterious. Similarly, it has been shown that genes with a high connectivity, as measured with the total number of PPIs, tend to be under selective constraint because they are involved in complex functions (Valdar and Thornton 2001; Caffrey et al. 2004). We do indeed find that our selection statistics correlate with PPI, although the temporal pattern of PPI is not similar to the temporal pattern of any of the selection statistics.

Gene size, measured as the protein's primary structure length, has also been shown to negatively correlate with ω (Comeron et al. 1999; Duret and Mouchiroud 1999; Larracuenta et al. 2008) and even with a measure of the rate of adaptive substitutions (Liu and Robinson-Rechavi 2018a). A possible explanation for this observation could be the fact that for a given adaptive mutation, longer genes would have more adaptive segregating sites competing against each other in different haplotypes. This would produce a kind of intergenic or interexonic Hill–Robertson interference (Hill and Robertson 1966). Consistent with this

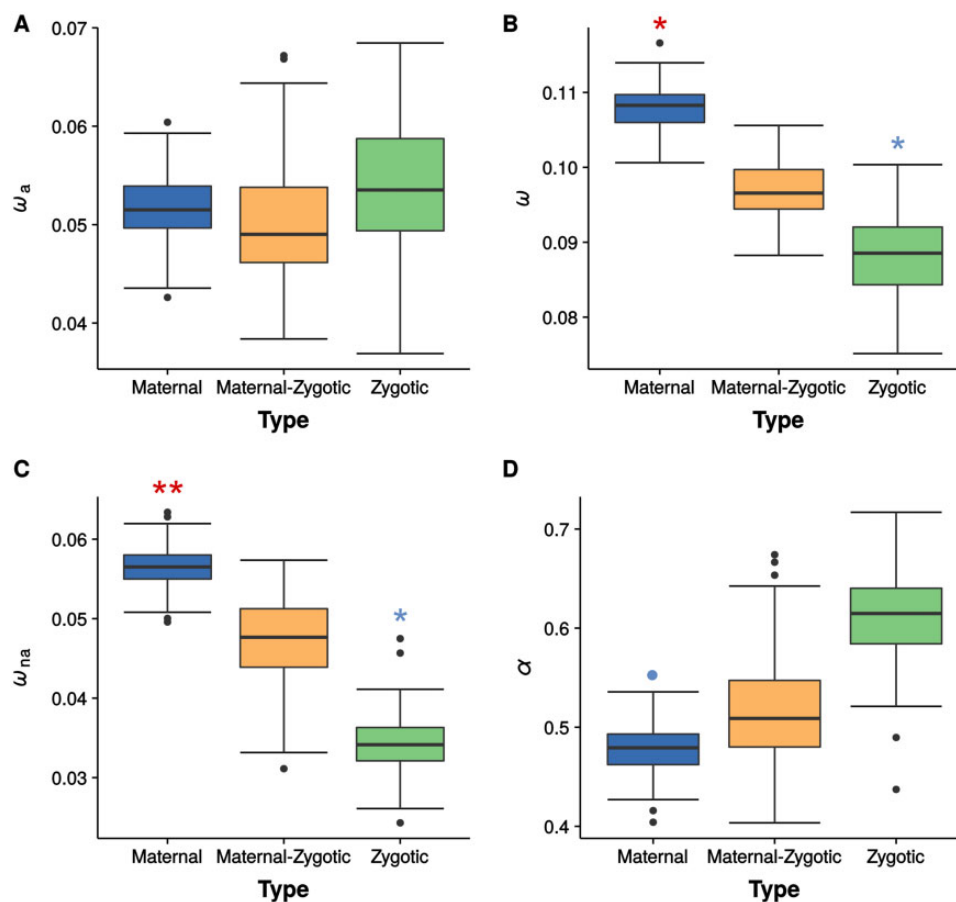


FIG. 4.—Selection statistics (ω_a , ω , ω_{na} , and α) for maternal, maternal-zygotic and zygotic genes. Maternal genes are those genes whose mRNA are laid by the mother in the egg and are never zygotically transcribed, maternal-zygotic are those genes whose mRNA is present in the egg but that are also transcribed by the zygote. Zygotic genes are genes whose mRNA is not laid in the egg by the mother. (A) ω_a is not statistically different between these gene categories. (B) ω is significantly higher in maternal genes than the other two gene categories ($P = 0.024$). (C) ω_{na} is significantly higher in maternal genes than in the other two gene categories (P value = 0.003). (D) α is marginally lower in maternal genes compared with the other two categories. Each point in a plot (100 bootstrap replicates per group) is calculated for a randomly drawn sample of the set of genes in each gene category. Asterisks represent the significance by a permutation test, the color indicates whether the value was higher (red) or lower (blue) than expected (\bullet , 0.1–0.05, $\ast < 0.05$, $\ast\ast < 0.01$, and $\ast\ast\ast < 0.001$). The number of genes analyzed is shown in [supplementary table S6, Supplementary Material online](#). P values in [supplementary table S20, Supplementary Material online](#).

previous work, we found that gene size correlates negatively with both ω and ω_a over developmental time.

To our knowledge, the number of exons and transcripts themselves have never been directly associated with neither ω nor ω_a , as we found in our study. As one would expect, however, the number of exons correlates with gene size (data not shown). In a similar way, the larger the number of exons, the larger the number of different transcripts that can be produced by reshuffling exons by alternative splicing. Then, the relationship between the number of exons and the number of transcripts with the selection statistics could be either direct or indirect through the correlation of these genomic features with each other and gene size.

It is not possible to establish whether the temporal pattern of adaptation is a consequence of differences in genomic

features over the life-cycle, or if, on the contrary, these genomic features are a consequence of differential adaptation over the life-cycle. To establish what leads to what, one would have to devise experiments, or at least some form of modeling, that is beyond the state of the art for fly's development and its underlying genetic bases. There are, however, some intrinsic characteristics of development that can be used to suggest that the first possibility is more likely.

A possible explanation for that would be that, as suggested from a more qualitative evo-devo perspective (Gellon and McGinnis 1998; Kennison 1993), embryonically expressed genes have a more complex regulation than postembryonically expressed genes. Although the former are expressed in wider areas of the embryo, their expression changes more in time and space than that of postembryonically expressed

genes (Salvador-Martínez and Salazar-Ciudad 2015). This more complex regulation may require a more complex genetic structure, such as manifested by those genes having more exons, more transcripts, larger genes and larger introns (Kennison 1993; Gellon and McGinnis 1998). The larger average intron length of developmental genes may also be a reflection of complex regulation, but at the level of *cis*-regulatory elements, because *cis*-regulatory elements can be located within introns too. Larger expression areas, less temporally restricted expression and more complex gene structure, may also reflect that mid- and late-development genes interact with more other genes than genes expressed later. This would make them, in rough terms, more pleiotropic and, thus, less likely to adapt. This explanation would be the reflection at the genomic level of the idea that there are wider and higher levels of interdependence among body parts in middevelopment than in late and early development (Duboule 1994; Raff 1996). From this perspective, genomic features would not be a consequence of differential adaptation over the life-cycle. Instead, the pattern of differential adaptation over the life-cycle would be a consequence of how genomic features have to be over development.

The above arguments would not apply to the earliest embryonic stages because of the low efficiency of natural selection on maternal genes. In addition, it has been suggested that the small intron and gene size of the genes expressed in the first hours of *D. melanogaster* development (Anderson 1973; Heyn et al. 2014) would be imposed by the very fast cell divisions occurring in those hours. Because cells divide very rapidly in early *D. melanogaster* development, there is no time to transcribe, splice, and translate long genes and genes with long introns. This would preclude the earliest expressed genes from having the complex regulatory gene structure (no long introns and no long genes) that in the later-expressed genes correlates with lower rates of nonsynonymous substitutions. This hypothesis is consistent with the reasoning of the previous paragraph because, in fact, there are not many genetic interactions, signaling or cell movement during these early fast cell-division states. In other words, no complex genetic regulation occurs until the early fast division stages have finished. This applies not only to the fly but also to many other animal groups, including vertebrates (O'Farrell et al. 2004; Heyn et al. 2014; Siefert et al. 2015). The hourglass model has also been suggested for vertebrates and it is then tempting to suggest that, in them, the relative lack of conservation of early development would also be related to the simpler gene structure required for short cell cycles and, possibly, to maternal genes. This simpler gene structure would be consistent with our results of shorter genes with fewer exons, shorter introns, and less transcripts in the genes expressed in the earliest stages of development.

A third main conclusion of our study is that our results are consistent, roughly, with the hourglass model but not with the von Baer's law. However, the fit to the hourglass model is

rather weak, because there are no major differences in ω between embryonic stages after the 8-h, except for genes in cluster 8. During the first 2 h, ω is significantly high (hours 0–2 h: $P=0.032$), but from 6–8 h to 22–24 h ω is lower than expected based on the permutation test ($P<0.001$). This is also the case for ω_a (supplementary table S10, Supplementary Material online). In contrast with some previous studies (Kalinka et al. 2010; Levin et al. 2016), we do not find that the latest stages of embryonic development are less conserved (this previous study measures only d_N and d_S). However, genes, whose expression is high only in late embryonic development (cluster 8, fig. 2), show a significant high ω and marginally significant high ω_a . These genes are only a small proportion of the genes expressed in the last embryonic developmental stages and, thus, have a minor effect on our calculations of ω_a , ω , and α of these stages (thus likely explaining the differences between our study and Kalinka et al. [2010]). The difference in d_N between mid- and late-development stages in Kalinka et al. (2010) is however rather subtle too. Overall, thus, our results are compatible with Kalinka et al. (2010). The hourglass model was proposed on the basis of what was understood of *D. melanogaster* and vertebrate's development (Slack et al. 1993; Duboule 1994; Raff 1996). The life-cycles of the fly and the mouse are quite different. Mice, as all amniotes, are direct developers, meaning that development gives rise to a juvenile and later, gradually, to an adult. Flies are indirect developers in which embryonic development gives rise to a free-roaming larva and that, by a rather abrupt process of metamorphosis, to an adult. If the hourglass model is understood for the whole of the life-cycle then our results are roughly consistent with it at the genetic level: genes expressed during embryonic development are highly conserved, except for the genes expressed in the earliest stages, whereas the genes expressed later, from the larval stage L3 onward, show less conservation and more adaptation. On the other hand, this temporal hourglass pattern can also be understood as development generally obeying von Baer's law, but departing from it in the earliest stages. We hypothesize that this departure would arise, in one hand, from the lower efficiency of selection on maternal genes and, in the other hand, as a consequence of the reduced gene structure complexity required for fast nuclei divisions in early development.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Roland Zimm, Miquel Marin-Riera, Pascal Hagolani, Rishi Das Roy, and Miguel Brun-Usan for comments. We thank Carla Giner for helping in developing R scripts. This

work was supported by the Finnish Academy to IS-C and the Spanish Ministerio de Economía y Empresa [Project MINECO CGL2017-89160-P awarded to AB], the Generalitat de Catalunya [grant 2017SGR 1379], and the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya [FI-DGR 2015 to MC-Z].

Literature Cited

- Alberch P. 1980. Ontogenesis and morphological diversification. *Am Zool.* 20(4):653–667.
- Anderson DT. 1973. Embryology and phylogeny in annelids and arthropods. Oxford, New York: Pergamon Press.
- Arthur W. 1977. The origin of animal body plans: a study in evolutionary developmental biology. Cambridge: Cambridge University Press.
- Artieri CG, Haerty W, Singh RS. 2009. Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of *Drosophila*. *BMC Biol.* 7(1):42.
- Ávila V, Campos JL, Charlesworth B. 2015. The effects of sex-biased gene expression and X-linkage on rates of adaptive protein sequence evolution in *Drosophila*. *Biol Lett.* 11:20150117.
- Bierne N, Eyre-Walker A. 2004. The genomic rate of adaptive amino acid substitution in *Drosophila*. *Mol Biol Evol.* 21(7):1350–1360.
- Bierne N, Eyre-Walker A. 2006. Variation in synonymous codon use and DNA polymorphism within the *Drosophila* genome. *J Evol Biol.* 19(1):1–11.
- Boyko AR, et al. 2008. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* 4(5):e1000083.
- Bustamante CD, et al. 2002. The cost of inbreeding in *Arabidopsis*. *Nature* 416(6880):531–534.
- Bustamante CD, et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* 437(7062):1153–1157.
- Caffrey DR, Somaroo S, Hughes JD, Mintseris J, Huang ES. 2004. Are protein–protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Sci.* 13(1):190–202.
- Cagan A, et al. 2016. Natural selection in the great apes. *Mol Biol Evol.* 33(12):3268–3283.
- Carneiro M, et al. 2012. Evidence for widespread positive and purifying selection across the European rabbit (*Oryctolagus cuniculus*) genome. *Mol Biol Evol.* 29(7):1837–1849.
- Castellano D, Coronado-Zamora M, Campos JL, Barbadilla A, Eyre-Walker A. 2016. Adaptive evolution is substantially impeded by Hill–Robertson interference in *Drosophila*. *Mol Biol Evol.* 33(2):442–455.
- Castellano D, James J, Eyre-Walker A. 2018. Nearly neutral evolution across the *Drosophila melanogaster* genome. *Mol Biol Evol.* 35:2685–2694.
- Charlesworth J, Eyre-Walker A. 2008. The McDonald–Kreitman test and slightly deleterious mutations. *Mol Biol Evol.* 25(6):1007–1015.
- Clark AG, et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450(7167):203–218.
- Comeron JM, Kreitman M, Aguadé M. 1999. Natural selection on synonymous sites is correlated with gene length and recombination in *Drosophila*. *Genetics* 151(1):239–249.
- Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genet.* 8(10):e1002905.
- Cruikshank T, Wade MJ. 2008. Microevolutionary support for a developmental hourglass: gene expression patterns shape sequence variation and divergence in *Drosophila*. *Evol Dev.* 10(5):583–590.
- Cutter AD. 2008. Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. *Mol Biol Evol.* 25(4):778–786.
- Darwin C. 1872. The origins of species by means of natural selection. Philadelphia (PA): John Wanamaker.
- Davis JC, Brandman O, Petrov DA. 2005. Protein evolution in the context of *Drosophila* development. *J Mol Evol.* 60(6):774–785.
- Dezso Z, et al. 2008. A comprehensive functional analysis of tissue specificity of human gene expression. *BMC Biol.* 6:49.
- Drost H-G, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. *Mol Biol Evol.* 32(5):1221–1231.
- Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development* 1994:135–142.
- Duret L, Mouchiroud D. 1999. Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. *Proc Natl Acad Sci U S A.* 96(8):4482–4487.
- Durinck S, et al. 2005. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 21(16):3439–3440.
- Eyre-Walker A. 2002. Changing effective population size and the McDonald–Kreitman test. *Genetics* 162(4):2017–2024.
- Eyre-Walker A. 2006. The genomic rate of adaptive evolution. *Trends Ecol Evol (Amst).* 21(10):569–575.
- Eyre-Walker A, Keightley PD. 2007. The distribution of fitness effects of new mutations. *Nat Rev Genet.* 8(8):610–618.
- Eyre-Walker A, Keightley PD. 2009. Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol Biol Evol.* 26(9):2097–2108.
- Fay JC, Wyckoff GJ, Wu C-I. 2001. Positive and negative selection on the human genome. *Genetics.* 158(3):1227–1234.
- Feldman B. 2005. Relative importance and value. Unpublished manuscript (Version 1.1, March 19, 2005). Available from: <http://www.gtcenter.org/Archive/Conf05/Downloads/Conf/Feldman60.pdf> (last accessed January 2019).
- Futschik M. 2015. Mfuzz: soft clustering of time series gene expression data. R package version 2.36.0. Available from: <http://www.sysbiolab.eu/software/R/Mfuzz/index.html> (last accessed December 2015).
- Futschik ME, Carlisle B. 2005. Noise-robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol.* 3(4):965–988.
- Galtier N. 2016. Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genet.* 12(1):e1005774.
- Gelbart WM, Emmert DB. 2013. FlyBase high throughput expression pattern data. FlyBase Analysis. Available from: <http://flybase.org/reports/FBrf0221009.html> (last accessed May 2014).
- Gellon G, McGinnis W. 1998. Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *BioEssays* 20(2):116–125.
- Gossmann TI, Keightley PD, Eyre-Walker A. 2012. The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. *Genome Biol Evol.* 4(5):658–667.
- Gould SJ. 1977. Ontogeny and phylogeny. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Graveley BR, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471(7339):473–479.
- Grömping U. 2006. Relative importance for linear regression in R: the package relaimpo. *J Stat Softw.* 17:1–27.
- Guillén Y, Casillas S, Ruiz A. 2019. Genome-wide patterns of sequence divergence of protein-coding genes between *Drosophila buzzatii* and *D. mojavensis*. *J Hered.* 110(1):92–101.
- Halligan DL, Keightley PD. 2006. Ubiquitous selective constraints in the *Drosophila* genome revealed by a genome-wide interspecies comparison. *Genome Res.* 16(7):875–884.
- Hebenstreit D, et al. 2011. RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Mol Syst Biol.* 7(1):497.

- Hernandez RD, Williamson SH, Bustamante CD. 2007. Context dependence, ancestral misidentification, and spurious signatures of natural selection. *Mol Biol Evol.* 24(8):1792–1800.
- Hershberg R, Petrov DA. 2008. Selection on codon bias. *Annu Rev Genet.* 42(1):287–299.
- Heyn P, et al. 2014. The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Rep.* 6(2):285–292.
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res.* 8(3):269–294.
- Hu H, et al. 2017. Constrained vertebrate evolution by pleiotropic genes. *Nat Ecol Evol.* 1(11):1722–1730.
- Huang W, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24(7):1193–1208.
- Hurst LD. 2002. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18(9):486–487.
- Irie N, Kuratani S. 2011. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat Commun.* 2:248.
- Irie N, Kuratani S. 2014. The developmental hourglass model: a predictor of the basic body plan? *Development* 141(24):4649–4655.
- James JE, Piganeau G, Eyre-Walker A. 2016. The rate of adaptive evolution in animal mitochondria. *Mol Ecol.* 25(1):67–78.
- Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro, HN, editor. *Mammalian protein metabolism*. New York: Academic Press. p. 21–32.
- Kalinka AT, Tomancak P. 2012. The evolution of early animal embryos: conservation or divergence? *Trends Ecol Evol.* 27(7):385–393.
- Kalinka AT, et al. 2010. Gene expression divergence recapitulates the developmental hourglass model. *Nature* 468(7325):811–814.
- Keightley PD, Eyre-Walker A. 2007. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. *Genetics* 177(4):2251–2261.
- Keightley PD, Eyre-Walker A. 2012. Estimating the rate of adaptive molecular evolution when the evolutionary divergence between species is small. *J Mol Evol.* 74(1-2):61–68.
- Kennison JA. 1993. Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. *Trends Genet.* 9(3):75–79.
- Kim S. 2015. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Commun Stat Appl Methods.* 22:665–674.
- Kousathanas A, Halligan DL, Keightley PD. 2014. Faster-X adaptive protein evolution in house mice. *Genetics* 196(4):1131–1143.
- Langley CH, et al. 2012. Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics* 192(2):533–598.
- Larracuent AM, et al. 2008. Evolution of protein-coding genes in *Drosophila*. *Trends Genet.* 24(3):114–123.
- Levin M, et al. 2016. The mid-developmental transition and the evolution of animal body plans. *Nature* 531(7596):637–641.
- Liu J, Robinson-Rechavi M. 2018a. Adaptive evolution of animal proteins over development: support for the Darwin selection opportunity hypothesis of evo-devo. *Mol Biol Evol.* 35:2862–2872.
- Liu J, Robinson-Rechavi M. 2018b. Developmental constraints on genome evolution in four bilaterian model species. *Genome Biol Evol.* 10:2266–2277.
- Mackay TFC, et al. 2012. The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482(7384):173–178.
- Marais G, Nouvellet P, Keightley PD, Charlesworth B. 2005. Intron size and exon evolution in *Drosophila*. *Genetics* 170(1):481–485.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351(6328):652–654.
- Medawar PB. 1954. The significance of inductive relationships in the development of vertebrates. *Br J Philos Sci.* 2:172–174.
- Meisel RP, Connallon T. 2013. The faster-X effect: integrating theory and data. *Trends Genet.* 29(9):537–544.
- Mensch J, Serra F, Lavagnino NJ, Dopazo H, Hasson E. 2013. Positive selection in nucleoporins challenges constraints on early expressed genes in *Drosophila* development. *Genome Biol Evol.* 5(11):2231–2241.
- Messer PW, Petrov DA. 2013. Frequent adaptation and the McDonald-Kreitman test. *Proc Natl Acad Sci U S A.* 110(21):8615–8620.
- Murali T, et al. 2011. DroiD 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for *Drosophila*. *Nucleic Acids Res.* 39(Suppl_1):D736–D743.
- Murray GGR, et al. 2017. Natural selection shaped the rise and fall of passenger pigeon genomic diversity. *Science* 358(6365):951–954.
- Obbard DJ, Welch JJ, Kim K-W, Jiggins FM. 2009. Quantifying adaptive evolution in the *Drosophila* immune system. *PLoS Genet.* 5(10):e1000698.
- O’Farrell PH, Stumpff J, Su TT. 2004. Embryonic cleavage cycles: how is a mouse like a fly? *Curr Biol.* 14:R35–R45.
- Pal C, et al. 2001. Highly expressed genes in yeast evolve slowly. *Genetics* 158:927–931.
- Peden JF. 1999. Analysis of codon usage. Nottingham (United Kingdom): University of Nottingham.
- Piasecka B, Lichocki P, Moretti S, Bergmann S, Robinson-Rechavi M. 2013. The hourglass and the early conservation models—co-existing patterns of developmental constraints in vertebrates. *PLoS Genet.* 9(4):e1003476.
- Poe S, Wake MH. 2004. Quantitative tests of general models for the evolution of development. *Am Nat.* 164(3):415–422.
- Presnyak V, et al. 2015. Codon optimality is a major determinant of mRNA stability. *Cell* 160(6):1111–1124.
- Pröschel M, Zhang Z, Parsch J. 2006. Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. *Genetics* 174(2):893–900.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: <http://www.R-project.org/> (last accessed December 2015).
- Raff RA. 1996. The shape of life: genes, development, and the evolution of animal form. Chicago (IL): University of Chicago Press.
- Ràmia M, Librado P, Casillas S, Rozas J, Barbadilla A. 2012. PopDrowser: the population *Drosophila* browser. *Bioinformatics* 28:595–596.
- Richardson MK, et al. 1997. There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development. *Anat Embryol.* 196(2):91–106.
- Ronen R, Udpa N, Halperin E, Bafna V. 2013. Learning natural selection from the site frequency spectrum. *Genetics* 195(1):181–193.
- Roux J, Robinson-Rechavi M. 2008. Developmental constraints on vertebrate genome evolution. *PLoS Genet.* 4(12):e1000311.
- Salazar-Ciudad I. 2010. Morphological evolution and embryonic developmental diversity in metazoa. *Development* 137(4):531–539.
- Salvador-Martínez I, Coronado-Zamora M, Castellano D, Barbadilla A, Salazar-Ciudad I. 2018. Mapping selection within *Drosophila melanogaster* embryo’s anatomy. *Mol Biol Evol.* 35(1):66–79.
- Salvador-Martínez I, Salazar-Ciudad I. 2015. How complexity increases in development: an analysis of the spatial-temporal dynamics of 1218 genes in *Drosophila melanogaster*. *Dev Biol.* 405(2):328–339.
- Sander K. 1983. The evolution of patterning mechanisms: gleanings from insect embryogenesis and spermatogenesis. In: Goodwin BB, Holder N, Wylie CC, editors. *Development and evolution*. New York: Cambridge University Press. p. 137–159.
- Santpere G, et al. 2015. Analysis of five gene sets in chimpanzees suggests decoupling between the action of selection on protein-coding and on noncoding elements. *Genome Biol Evol.* 7(6):1490–1505.
- Sawyer SA, Kulathinal RJ, Bustamante CD, Hartl DL. 2003. Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. *J Mol Evol.* 57(Suppl 1):S154–S164.

- Siefert JC, Clowdus EA, Sansam CL. 2015. Cell cycle control in the early embryonic development of aquatic animal species. *Comp Biochem Physiol C Toxicol Pharmacol.* 178:8–15.
- Slack JMW, Holland PWH, Graham CF. 1993. The zootype and the phenotypic stage. *Nature* 361(6412):490–492.
- Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in *Drosophila*. *Nature* 415(6875):1022–1024.
- Steige KA, Laenen B, Reimegård J, Scofield DG, Slotte T. 2017. Genomic analysis reveals major determinants of cis-regulatory variation in *Capsella grandiflora*. *Proc Natl Acad Sci U S A.* 114(5):1087–1092.
- Strasburg JL, et al. 2011. Effective population size is positively correlated with levels of adaptive divergence among annual sunflowers. *Mol Biol Evol.* 28(5):1569–1580.
- Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol.* 21(1):36–44.
- Tataru P, Mollion M, Glémin S, Bataillon T. 2017. Inference of distribution of fitness effects and proportion of adaptive substitutions from polymorphism data. *Genetics* 207(3):1103–1119.
- Thomsen S, Anders S, Janga SC, Huber W, Alonso CR. 2010. Genome-wide analysis of mRNA decay patterns during early *Drosophila* development. *Genome Biol.* 11(9):R93.
- Valdar WS, Thornton JM. 2001. Protein–protein interfaces: analysis of amino acid conservation in homodimers. *Proteins* 42(1):108–124.
- von Baer KE. 1828. Über Entwicklungsgeschichte der Thiere. Beobachtung und Reflexion. Königsberg (Russia): Bei den Gebrüdern Bornträger.
- Wade MJ, Priest NK, Cruickshank TE. 2009. A theoretical overview of genetic maternal effects. In: *Maternal effects in mammals*. Chicago (IL): University of Chicago Press. p. 38–63.
- Wagner GP, Kin K, Lynch VJ. 2013. A model based criterion for gene expression calls using RNA-seq data. *Theory Biosci.* 132(3):159–164.
- Wilt FH, Hake S. 2004. *Principles of developmental biology*. New York: W.W. Norton.
- Wotton KR, et al. 2015. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife* 4:e04785.
- Yanai I, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21(5):650–659.
- Yang D, et al. 2018. A two-level model for the role of complex and young genes in the formation of organism complexity and new insights into the relationship between evolution and development. *Evodevo* 9:22.
- Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J Mol Evol.* 46(4):409–418.
- Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22(12):2472–2479.

Associate editor: Maria Costantini