

Developmental Constraints on Genome Evolution in Four Bilaterian Model Species

Jialin Liu^{1,2} and Marc Robinson-Rechavi^{1,2,*}

¹Department of Ecology and Evolution, University of Lausanne, Switzerland

²Swiss Institute of Bioinformatics, Lausanne, Switzerland

*Corresponding author: E-mail: marc.robinson-rechavi@unil.ch.

Accepted: August 17, 2018

Abstract

Developmental constraints on genome evolution have been suggested to follow either an early conservation model or an “hourglass” model. Both models agree that late development strongly diverges between species, but debate on which developmental period is the most conserved. Here, based on a modified “Transcriptome Age Index” approach, that is, weighting trait measures by expression level, we analyzed the constraints acting on three evolutionary traits of protein coding genes (strength of purifying selection on protein sequences, phyletic age, and duplicability) in four species: Nematode worm *Caenorhabditis elegans*, fly *Drosophila melanogaster*, zebrafish *Danio rerio*, and mouse *Mus musculus*. In general, we found that both models can be supported by different genomic properties. Sequence evolution follows an hourglass model, but the evolution of phyletic age and of duplicability follow an early conservation model. Further analyses indicate that stronger purifying selection on sequences in the middle development are driven by temporal pleiotropy of these genes. In addition, we report evidence that expression in late development is enriched with retrogenes, which usually lack efficient regulatory elements. This implies that expression in late development could facilitate transcription of new genes, and provide opportunities for acquisition of function. Finally, in *C. elegans*, we suggest that dosage imbalance could be one of the main factors that cause depleted expression of high duplicability genes in early development.

Key words: evo-devo, constraint, selection, transcriptome.

Introduction

Evolutionary changes in the genome can cause changes in development, which are subject to natural selection. This leads developmental processes to constrain genome evolution. More precisely, selection on the output of development affects evolution of the genomic elements active in development. Currently, based on morphological similarities during development, two popular models have been proposed to bridge developmental and evolutionary biology.

The early conservation model, modified from the “third law” of Von-Baer (1828) (as cited in Kalinka and Tomancak 2012), suggests that the highest morphological similarities among species from the same phylum occur in early development, followed by a progressive evolutionary divergence over ontogeny. It should be noted that Von-Baer in fact based his observations on postgastrulation embryos (Kalinka and Tomancak 2012; Abzhanov 2013). The “developmental burden” concept was proposed to explain this model.

It suggested that the development of later stages is dependent on earlier stages, so that higher conservation should be found in the earlier stages of development (Garstang 1922; Riedl 1978) (as discussed in Irie and Kuratani 2014).

On the basis of renewed observations in modern times, however, Duboule (1994) and Raff (1996) proposed the developmental “hourglass model”. This model suggested that a “phylotypic period” (Richardson 1995) in middle development has higher morphological similarities than early or late development. Several mechanisms have been proposed to explain this observation. Duboule (1994) proposed that it may be due to colinear Hox cluster gene expression in time and space. Raff (1996) suggested a high interdependence in signaling among developmental modules in middle development. Galis and Metz (2001) also highlighted the high number of interactions at this period, although Comte et al. (2010) did not find any molecular evidence for these interactions. It is worth noting that the hourglass model was not supported by

a comprehensive study of vertebrate embryonic morphology variation (Bininda-Emonds et al. 2003). A number of alternatives have been proposed, for example the “adaptive penetrance model” (Richardson et al. 1997) and the “ontogenetic adjacency model” (Poe and Wake 2004). Of note, the higher divergence of late development could also be due to stronger adaptive selection, the “Darwin hypothesis” (Artieri et al. 2009). This question is independent of the pattern of constraints (early or hourglass) which are the focus of this study, and we explore it in a companion study (Liu and Robinson-Rechavi 2017).

Both main models have been supported by recent genomic level studies based on different properties (such as expression divergence, sequence divergence, duplication, or phyletic age), different species, and different analysis methods. Concerning expression divergence, interestingly, all studies are consistent across different species and research groups (Kalinka et al. 2010; Irie and Kuratani 2011; Yanai et al. 2011; Levin et al. 2012; Wang et al. 2013; Gerstein et al. 2014; Ninova et al. 2014; Zalts and Yanai 2017). All of them suggested that middle development has the highest transcriptome conservation, that is, the hourglass pattern. On the other hand, when animals are compared between different phyla, middle development has been reported to have the highest divergence (Levin et al. 2016), although this conclusion has been criticized on methodological grounds (Dunn et al. 2018). From other properties, however, the results are inconclusive based on different methods (Castillo-Davis and Hartl 2002; Cutter and Ward 2005; Davis et al. 2005; Hazkani-Covo et al. 2005; Hanada et al. 2007; Irie and Sehara-Fujisawa 2007; Cruickshank and Wade 2008; Roux and Robinson-Rechavi 2008; Artieri et al. 2009; Domazet-Loso and Tautz 2010; Quint et al. 2012; Piasecka et al. 2013; Cheng et al. 2015; Drost et al. 2015).

Generally, the methods used to measure developmental constraints at the genomic level can be divided into three categories: Proportion based analysis, module analysis, and transcriptome index analysis. Proportion based analysis consists in testing the proportion of genes with a given property within all expressed genes (Roux and Robinson-Rechavi 2008). The method is less used following the emergence of accurate transcriptome-scale data, since it does not take into account the contributions of expression abundance. Module analysis consists in studying evolutionary properties of distinct sets of genes (modules) which are specifically expressed in groups of developmental stages (Piasecka et al. 2013). This method can avoid problems caused by genes expressed over all or a large part of development. For example, trends might be diluted by highly expressed housekeeping genes, which contribute to the average expression at all developmental stages. However, this approach can only measure the developmental constraints for a specific subset of genes, instead of considering the composition of the whole transcriptome. Transcriptome index analysis is a weighted mean: The mean

value of an evolutionary parameter is weighted by each gene’s expression level (Domazet-Loso and Tautz 2010). This method has the benefit of detecting evolutionary constraints on the whole transcriptome, but patterns can be driven by a subset of very highly expressed genes, or even by a few outliers, because the difference between highly and lowly expressed genes can span several orders of magnitude. For instance, Domazet-Loso and Tautz (2010) reported that transcriptomes of middle development stages of *Danio rerio* have a higher proportion of old genes than transcriptomes of early and late development stages, using the Transcriptome Age Index (TAI). However, Piasecka et al. (2013) re-analyzed the same data and reported that the highest proportion of old genes was in transcriptomes of early development stages, once a standard log-transformation of microarray signal intensities was done, a result confirmed by module analysis and proportion based analysis.

Several statistical methods have been proposed to distinguish the hourglass model from the early conservation model. The parabolic test is based on fitting both first degree and second degree polynomial models (Roux and Robinson-Rechavi 2008). The hourglass model is supported if the parabolic function provides a significantly better fit and its minimum corresponds to middle development. This method has been criticized for being too specific and insensitive to other nonparabolic hourglass patterns (Drost et al. 2015). The flat line test simply tests whether variance of transcriptome indexes across development is significantly higher than variance from random samples (Domazet-Loso and Tautz 2010; Quint et al. 2012). But a significant difference does not necessarily imply the existence of an hourglass pattern (Drost et al. 2015). Since these two methods are either too strict or without power to distinguish the hourglass model, Drost et al. (2015) proposed a “reductive hourglass test” which focuses on testing the presence of an hourglass pattern of divergence: High-low-high. For this, development can be divided into three periods (early, phylotypic, and late), based on the known phylotypic period from morphological studies. Then, a permutation method is used to test whether the mean value in the phylotypic period is significantly lower than in early and late periods.

Overall, the transcriptome index analysis should be the best method to measure developmental constraints on the whole transcriptome, if care is taken to properly treat the expression values. Moreover, the reductive hourglass test should be used to objectively test the hourglass model, alone or in combination with other methods.

Because previous studies used different methodologies, and few studies adopted a transformed transcriptome index analysis, their conclusions cannot be compared consistently, making a biological conclusion concerning developmental constraints across species and features difficult. What’s more, while many studies focus on distinguishing between early conservation model

and hourglass conservation model, we still know very little of the factors driving these patterns.

To measure developmental constraints on genome evolution, we calculated transcriptome indexes over the development of four species (*Caenorhabditis elegans*, *Drosophila melanogaster*, *D. rerio*, and *Mus musculus*), for three evolutionary parameters (strength of purifying selection on coding sequences [ω_0], phyletic age, and duplicability [paralog number]), with three transformations of expression values (nontransformed, \log_2 transformed, and square root transformed). For *C. elegans*, the strength of purifying selection on coding sequences was not reliably estimated, with no data in the Selectome database (Moretti et al. 2014) and very high values of estimated synonymous distances (dS) from Ensembl Metazoa (Kersey et al. 2016) (supplementary fig. S1, Supplementary Material online); thus, we did not include this parameter in the study of *C. elegans*. In general, we found results consistent with the hourglass model for sequence evolution, but with early conservation for phyletic age and paralog number, in the four species. In addition, \log_2 transformed transcriptome indexes are always consistent with square root transformed transcriptome indexes but not with nontransformed transcriptome indexes.

Materials and Methods

Data files and analysis scripts are available on our GitHub repository: https://github.com/ljijoling1010/developmental_constraints_genome_evolution; last accessed May 24, 2018.

Expression Data Sets

Main Data Sets

For *C. elegans* and *D. melanogaster*, we downloaded processed (nontransformed but normalized) RNA-seq data from <http://www.stat.ucla.edu/~jingyi.li/software-and-data.html> (Li et al. 2014), which originally comes from (Gerstein et al. 2010; Graveley et al. 2011).

For *D. rerio*, we used the processed (log-transformed and normalized) microarray data from our previous study (Piasecka et al. 2013). This data originally comes from Domazet-Lošo and Tautz (2010).

For *M. musculus*, we obtained processed (nontransformed but normalized) RNA-seq data from Hu et al. (2017).

Supplementary Data Sets

For *C. elegans*, *D. melanogaster*, and *D. rerio*, we downloaded the processed (nontransformed and nonnormalized) RNA-seq data from Levin et al. (2016). This data set was generated by CEL-Seq (Hashimshony et al. 2012), a technique for multiplexed single cell RNA-seq. Because CEL-Seq retains only the 3' end of the transcript, we performed sample normalization based on transcripts

per million, but without transcript length normalization. Since only one embryo was sequenced in each stage, there could be larger technical error among lowly expressed genes, like in single cell RNA-seq, especially in early development. So, after normalization, we removed genes with mean expression across samples <1 .

For *M. musculus*, the processed (log-transformed and normalized) microarray data were retrieved from Bgee (release 13.1, July 2015; Bastian et al. 2008), a database for gene expression evolution. This data originally comes from (Irie and Kuratani 2011).

The detail information of the two data sets listed in supplementary table S1, Supplementary Material online.

Omega0 (ω_0)

The ω_0 values were downloaded from Selectome (Moretti et al. 2014), a database of positive selection based on the branch-site model (Zhang et al. 2005). Selectome excludes ambiguously aligned regions before model fitting, using Guidance bootstrapping and M-Coffee consistency. Omega0 is the dN/dS ratio (dN is the rate of nonsynonymous substitutions, dS is the rate of synonymous substitutions) of the subset of codons which have evolved under purifying selection according to the branch-site model. We used ω_0 from the Clupeocephala branch, the Murinae branch, and the Melanogaster group branch for *D. rerio*, *M. musculus*, and *D. melanogaster*, respectively. One gene could have two ω_0 values in the focal branch because of duplication events. In this case, we keep the value of the branch following the duplication and exclude the value of the branch preceding the duplication.

Phyletic Age Data

Phyletic ages were retrieved from Ensembl version 84 (Yates et al. 2016) using the Perl API. For each gene, we browsed its gene tree from the root and dated it by the first appearance. We assigned the oldest genes with phyletic age value of 1 and the youngest genes with the highest phyletic age value. So, genes can be split into discrete “phylostrata” by phyletic age. We classified 3 phylostrata, 4 phylostrata, 9 phylostrata, and 18 phylostrata respectively for *C. elegans*, *D. melanogaster*, *D. rerio*, and *M. musculus*. The definition of phylostrata is dependent on the available genome sequences in related lineages, hence the differences between species.

Number of Paralogs

We retrieved the number of paralogs from Ensembl release 84 (Yates et al. 2016) using BioMart (Kinsella et al. 2011).

Retrogene Data

For *C. elegans*, we retrieved 33 retrogenes from Zou et al. (2012). For *D. melanogaster*, we retrieved 72 retrogenes from

retrogeneDB (Kabza et al. 2014). For *D. rerio* we retrieved 113 retrogenes from Fu et al. (2010). For *M. musculus* we retrieved 134 retrogenes from Potrzebowski et al. (2008).

Connectivity Data

We retrieved connectivity (protein–protein interactions) data from the OGEE database (Chen et al. 2012).

Testis Specific Genes

We first retrieved processed (normalized and log-transformed) RNA-seq data of 22 *M. musculus* tissues and 6 *D. melanogaster* tissues from Kryuchkova-Mostacci and Robinson-Rechavi (2016a).

Then, we calculated tissue specificity based on Tau (Yanai et al. 2005; Kryuchkova-Mostacci and Robinson-Rechavi 2016b):

$$\text{Tau} = \frac{\sum_{i=1}^n (1 - \hat{x}_i)}{n - 1}; \hat{x}_i = \frac{x_i}{\max_{1 \leq i \leq n} (x_i)},$$

where n is the number of tissues, and x_i is the expression of the gene in tissue i . This index ranges from zero (broadly expressed genes) to one (genes specific to one tissue). All genes that were not expressed in at least one tissue were removed from the analysis.

Finally, we defined genes with highest expression in testis and with tissue specificity value ≥ 0.8 as testis specific genes.

Transcriptome Index Analysis for Different Evolutionary Parameters

The TEI (transcriptome evolutionary index) is calculated as:

$$\text{TEI}_s = \frac{\sum_{i=1}^n E_i e_{is}}{\sum_{i=1}^n e_{is}},$$

where s is the developmental stage, E_i is the relevant evolutionary parameter (ω_0 , paralog number, phyletic age, stage specificity, or protein connectivity) of gene i , n is the total number of genes, and e_{is} is the expression level of gene i in developmental stage s ; by default we use log-transformed expression levels for e_{is} .

Confidence Interval Analysis

Firstly, we randomly sampled gene IDs from each original data set 10,000 times with replacement. Then, we computed transcriptome indexes for the 10,000 samples. Finally, the 95% confidence interval is defined as the range from quantile 2.5% to quantile 97.5% of the 10,000 transcriptome indexes. This approach was integrated into myTAI (Drost et al. 2018), a R package for evolutionary transcriptome index analysis.

Stages before the Start of Maternal to Zygote Transition (MZT)

The stages before the start of the MZT were defined from Tadros and Lipshitz (2009). For *C. elegans* and *D. melanogaster*, it's the first 1 h of embryo development; for *D. rerio*, it's the first 2 h of embryo development; for *M. musculus*, it's the first day of embryo development.

Phylogenic Period

From both morphological and genomic studies, we defined the phylogenic period of each species as follows: for *C. elegans*, the phylogenic period is defined as the ventral enclosure stage (Levin et al. 2012); for *D. melanogaster*, the phylogenic period is defined as an extended germband stage (Sander 1983; Kalinka et al. 2010); for *D. rerio*, the phylogenic period is defined as the segmentation and pharyngula stages (Ballard 1981; Wolpert 1991; Slack et al. 1993; Domazet-Loaso and Tautz 2010); for *M. musculus*, the phylogenic period is defined as Theiler stages 13–20 (Ballard 1981; Wolpert 1991; Slack et al. 1993; Irie and Kuratani 2011).

Permutation Test

We first assigned all development stages to three broad development periods (early: after the start of MZT and before the phylogenic period, middle: the phylogenic period, and late: after the phylogenic period). Next, we calculated the difference of mean transcriptome indexes between the early module and the middle module ($\Delta e-m$). Then, we permuted the values of the relevant parameter (ω_0 , paralog number, phyletic age, stage specificity or protein connectivity) 10,000 times. Finally, we approximated a normal distribution for $\Delta e-m$ based on 10,000 $\Delta e-m$ values computed from the permuted samples. The P -value of the hourglass model versus the early conservation model for each parameter is the probability of a randomly sampled $\Delta e-m$ exceeding the observed $\Delta e-m$. For protein connectivity, the P -value of the hourglass model is the probability that a randomly sampled $\Delta e-m$ lower than the observed $\Delta e-m$.

Results and Discussion

Effect of Expression Value Transformation on Transcriptome Indexes

As mentioned in the Introduction, the pattern from a transcriptome index analysis may not reflect the global behavior of the transcriptome, but that of a small fraction of very highly expressed genes, or even of a few outliers. In order to systematically test this issue, we calculated 95% confidence intervals of transcriptome indexes based on \log_2 transformed, square root transformed, and nontransformed expression values (see Materials and Methods). Then, for the purpose of comparing the range of confidence intervals in the same

scale, we plotted the ratio of upper to lower confidence interval boundary across development. Clearly, at a given confidence level (95% here), we can see that the ratio of nontransformed transcriptome indexes is much higher and more variable than transformed transcriptome indexes (supplementary fig. S2, Supplementary Material online), indicating that the transcriptome indexes estimated from transformed expression are more stable. The most stable pattern comes from \log_2 transformed transcriptome indexes, although it is quite similar with square root transformation. We note that while log-transformation is routine in most application of transcriptomics, many analyses of “hourglass” patterns use nontransformed expression data.

In summary, although a subset of genes with dramatically different expression values in different stages could be interesting in some sense, when the goal is to investigate the general tendency of the transcriptome, log- or square-root-transformation for expression value is necessary and efficient to reach a stable estimation.

Variation of Evolutionary Transcriptome Indexes across Development

Here, based on \log_2 transformed expression values, we calculated transcriptome indexes for strength of purifying selection on coding sequences (ω_0), phyletic age, and duplicability (paralog number). In order to objectively distinguish the hourglass model from the early conservation model, we used a permutation test method similar to that of Drost et al. (2015) (see Materials and Methods). For all parameters considered the highest divergence is observed in late development, and there are many more stages sampled from late development, so we only compared the difference between early and middle development. Thus, a significant *P*-value for lower divergence in middle versus early development supports the hourglass model, whereas a lack of significance supports the early conservation model. We consider early conservation to cover both stronger conservation in early than middle development, and similar strong conservation over early and middle development, and hence we use a one-sided test. Notably, for early development, we did not consider the stages before the start of MZT (see Materials and Methods), because these stages are dominated by maternal transcripts.

For the transcriptome index of purifying selection on coding sequence (Transcriptome Divergence Index: TDI), we found that genes with stronger purifying selection tend to be more expressed at middle developmental stages, suggesting an hourglass pattern (fig. 1). However, for the transcriptome indexes of phyletic age (TAI) and of paralog number (Transcriptome Paralog Index: TPI), we observed that genes with higher duplicability and younger phyletic age trend to be expressed at later developmental stages, which corresponds to early conservation (fig. 1). In addition, we also repeated these analyses based on square root transformed expression

values (supplementary fig. S3, Supplementary Material online) and on nontransformed expression values (supplementary fig. S4, Supplementary Material online). In general, the results from square root transformation are highly consistent with those from \log_2 transformation, but not with those from nontransformation. For example, with nontransformed expression data, the TPI in *C. elegans* became very noisy; the TAI in *D. rerio* changed from early conservation to hourglass pattern; or the TDI in *M. musculus* changed into an unexpected early divergence pattern. Finally, we confirmed our observations with other data sets (supplementary fig. S5, Supplementary Material online). For *C. elegans*, *D. melanogaster*, and *D. rerio*, we used a high resolution time series single embryo RNA-seq data set (Levin et al. 2016). Since this data set is without replicates, and is generated from single embryo, the transcriptome indexes are noisy and present extreme values in some time points. However, generally, all the results from the new data set are consistent with the results from our previous data sets except the TDI in *D. rerio*. For the latter, we only observed the first two-thirds of an hourglass pattern in the new data set. This is because the new data set only covers embryo development, whereas the increased TDI in late development is driven by postembryonic development stages. For *M. musculus*, we also confirmed our results based on a microarray data set (Irie and Kuratani 2011) (supplementary fig. S5, Supplementary Material online).

In *D. melanogaster*, we did not confirm the results of Drost et al. (2015) for phyletic age. After \log_2 transformation of expression data, we found an early conservation pattern instead of the hourglass pattern which they reported (supplementary fig. S6B, Supplementary Material online). It appears that the hourglass pattern of phyletic age in their study is driven by a few highly expressed genes, consistently with our previous observations in *D. rerio* (Piasecka et al. 2013). This is verified by excluding the top 10% most expressed genes and analyzing without transformation (supplementary fig. S6C, Supplementary Material online). Of note, that fly phyletic age hourglass with uncorrected expression was also reported earlier (Domazet-Loso and Tautz 2010).

Overall, these results suggest that genes under strong purifying selection on their protein sequence tend to be expressed in middle development; it remains to be seen how much these observations extend to more arthropods or chordates. They also extend our previous observations that genes expressed earlier have a lower duplicability and an older age (Roux and Robinson-Rechavi 2008; Piasecka et al. 2013). In addition, it poses the question whether a pattern driven by the minority of very highly expressed genes is relevant to understanding Evo-Devo, which is generally driven by regulatory genes (Carroll 2008), such as transcription factors, with typically not very high and rather tissue-specific expression.

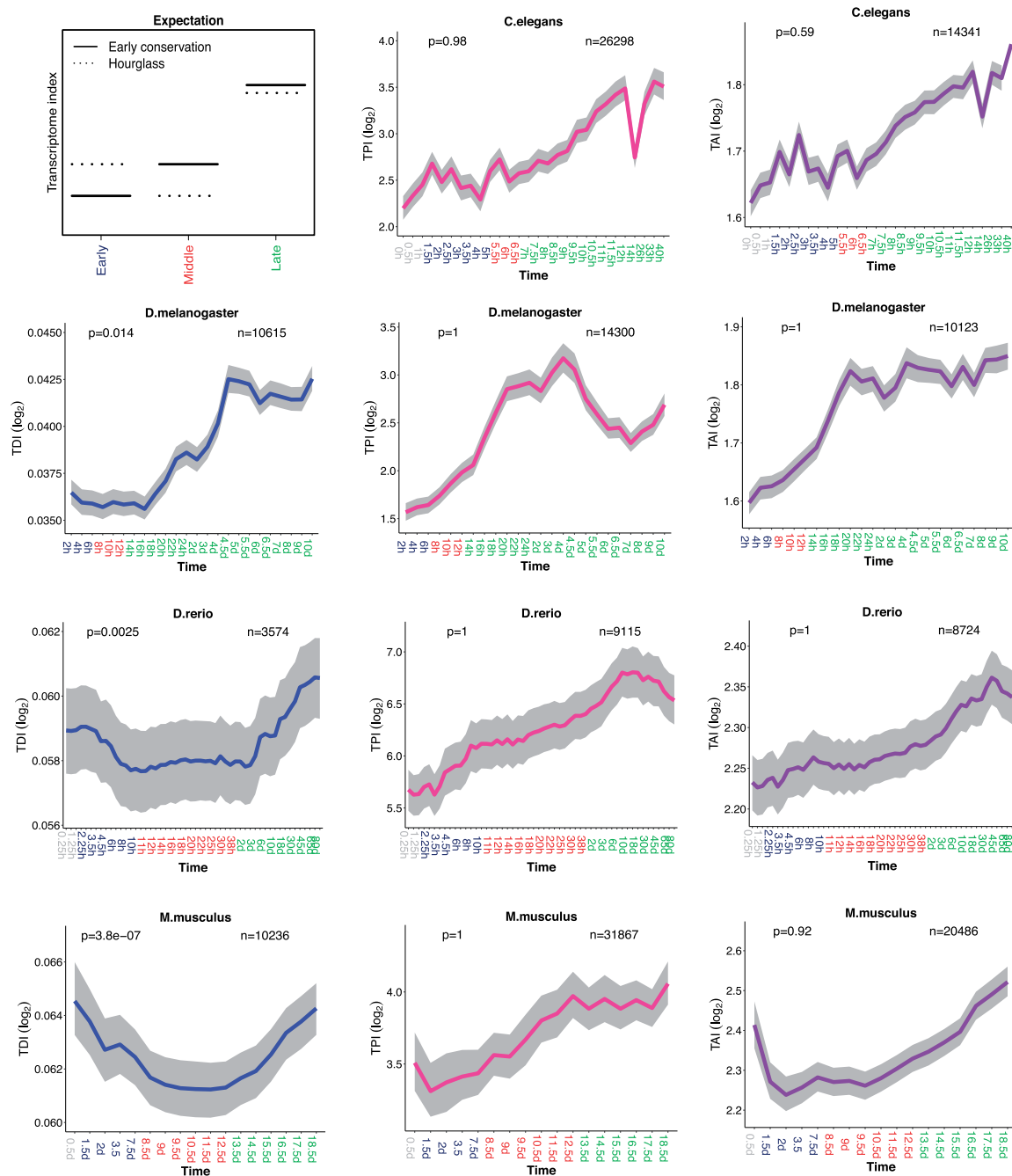


Fig. 1.—Evolutionary transcriptome indexes based on \log_2 transformed expression values. The top-left plot schematically shows the expected transcriptome index patterns of the early conservation and hourglass models. For the other plots, grey, dark blue, red, and green marked time points in the x-axis represent stages before the start of MZT, early developmental stages, middle developmental stages, and late developmental stages, respectively. Transcriptome index of divergence (TDI): Blue line; transcriptome index of paralog number (TPI): Pink line; transcriptome index of phyletic age (TAI): Purple line. The grey area indicates 95% confidence interval estimated from bootstrap analysis. The *P*-values for supporting the hourglass model (permutation test, early vs. middle development) are indicated in the top-left corner of each plot. The numbers of genes analyzed are noted in the top-right corner of each plot.

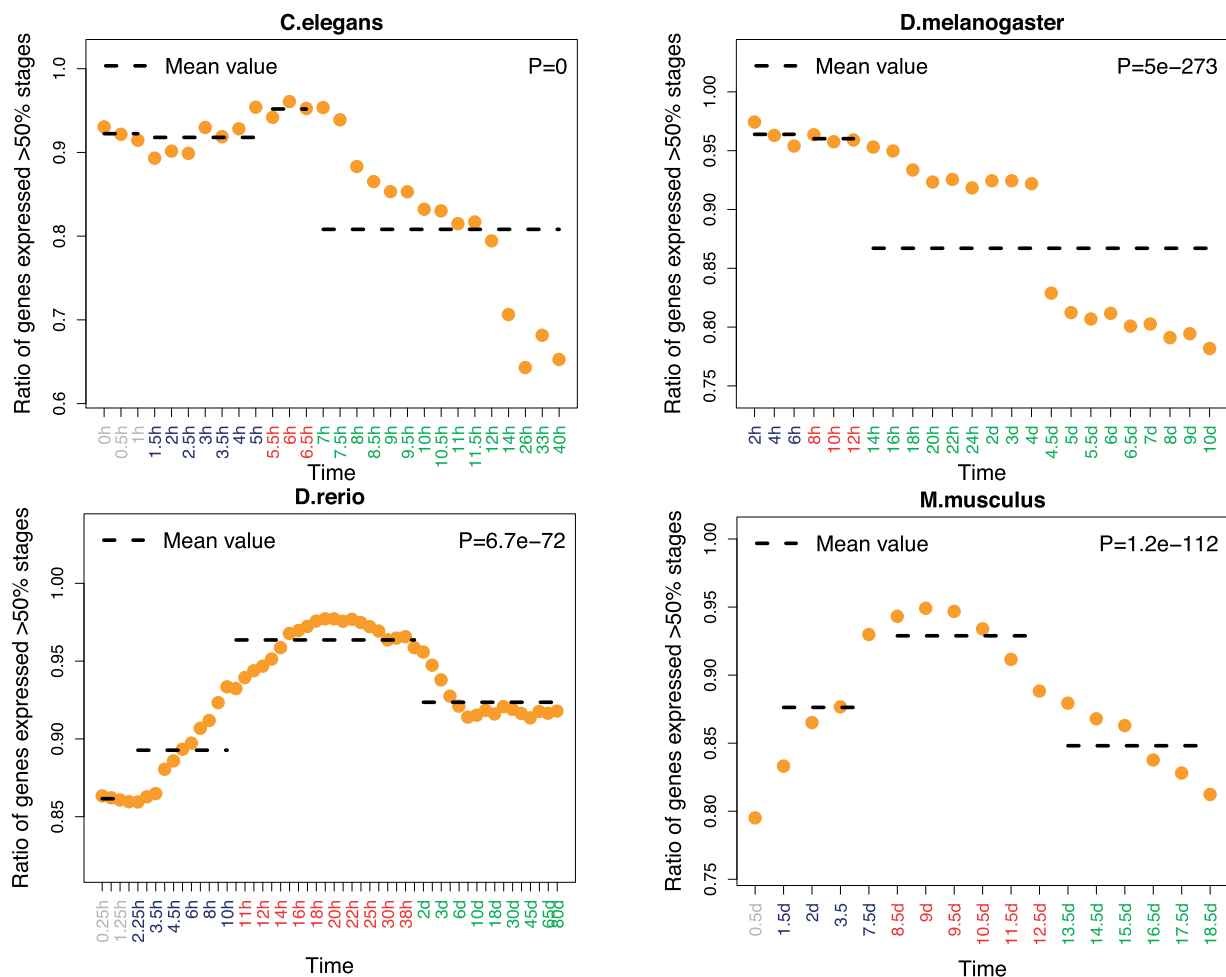


FIG. 2.—Proportion of temporal pleiotropic genes across development. Grey, dark blue, red, and green marked time points in the x-axis represent stages before the start of MZT, early developmental stages, middle developmental stages, and late developmental stages respectively. The proportion of temporal pleiotropic genes is plotted as orange circles. The P -values from chi-square goodness of fit test are indicated in the top-right corner of each graph. Pleiotropic genes are defined as expressed in $>50\%$ of stages sampled. The proportion of pleiotropic genes is defined as the number of pleiotropic genes divided by the number of all genes expressed in the corresponding stage.

Expression of Temporal Pleiotropy Genes across Development

Several models have been proposed to explain why some developmental stages are more conserved than others, as presented in the Introduction. In all models, a common point is that high conservation is caused by selection against deleterious pleiotropic effects of mutations. This implies that higher sequence conservation in middle developmental stages is caused by higher pleiotropy of genes expressed in these stages, pleiotropy being one of the major factors that constrain sequence evolution (Fraser et al. 2002).

In order to test this hypothesis, we used one type of development related pleiotropic effect: temporal pleiotropy (Artieri et al. 2009) (expression breadth across development). This is similar to spatial pleiotropy (Larracuente et al. 2008; Kryuchkova-Mostacci and Robinson-Rechavi 2015) (expression breadth across tissues) or connective pleiotropy

(Fraser et al. 2002) (protein–protein connectivity). The more stages a gene is expressed in, the more traits it could affect, so it is expected to be under stronger evolutionary constraints (Wagner and Zhang 2011). For *C. elegans*, *D. melanogaster*, and *M. musculus*, we defined FPKM > 1 as expressed. For *D. rerio* we set genes with microarray signal rank in top 70% as expressed.

We calculated the proportion of potentially pleiotropic genes as expressed in $>50\%$ of development stages. In all the species, interestingly, we found pleiotropic genes enriched in middle development (fig. 2). In *D. melanogaster*, the evidence is weaker, because of the low sampling of early and middle development stages in the main data set; but the pattern was clear in the high resolution single embryo RNA-seq data set (supplementary fig. S7, Supplementary Material online). We also found similar patterns when we define pleiotropic genes as expressed in $>70\%$ of development stages

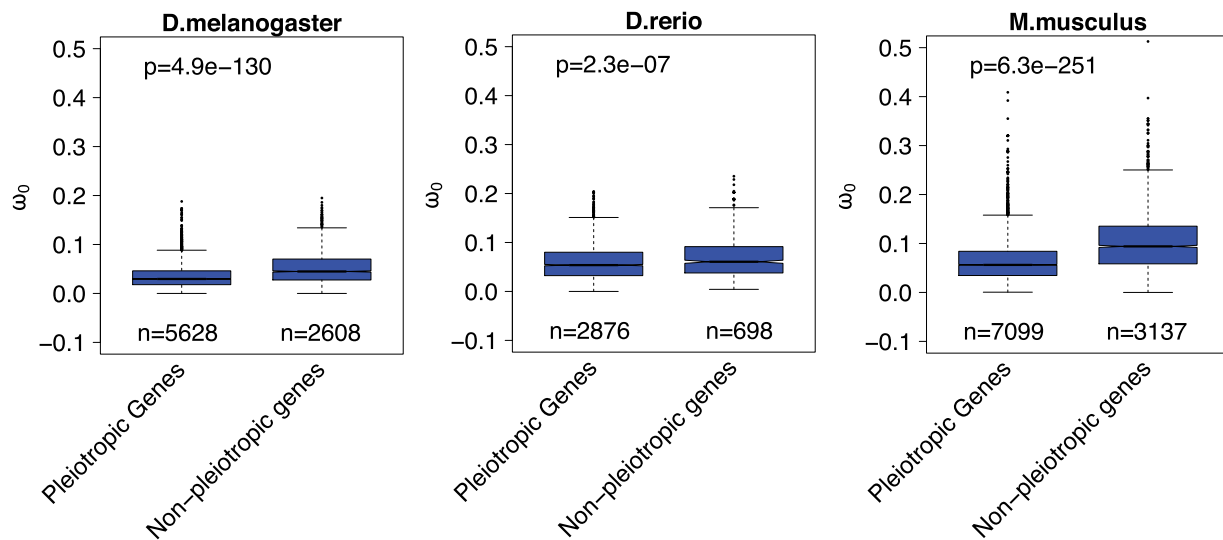


Fig. 3.—Comparison of ω_0 between temporal pleiotropic genes and nonpleiotropic genes. The number of genes in each category is indicated below each box. The P -values from a Wilcoxon test comparing categories are reported above boxes. The lower and upper intervals indicated by the dashed lines (“whiskers”) represent 1.5 times the interquartile range, or the maximum (respectively minimum) if no points are beyond 1.5 IQR (default behavior of the R function boxplot).

(supplementary fig. S8, Supplementary Material online). Since the late development of *D. melanogaster* can clearly be divided into two periods with distinct patterns from the pleiotropy analysis, we removed the second period of late development, and found the same overall trend (supplementary fig. S9, Supplementary Material online). For *D. rerio*, in addition, we observed consistent results based on setting expressed genes as microarray signal rank in the top 90% or 50% (supplementary fig. S10, Supplementary Material online). Similar observations of higher temporal pleiotropy for genes in middle development in vertebrates were recently reported by Hu et al. (2017).

On the basis of these observations, we further checked whether higher temporal pleiotropic constraint could explain stronger purifying selection on sequence evolution. As expected, we found that pleiotropic genes have lower ω_0 than nonpleiotropic genes (fig. 3).

In summary, we found that middle development stages with a higher proportion of broadly expressed genes are under stronger pleiotropic constraint on sequence evolution.

Higher Expression of Retrogenes in Later Development Stages

In adult anatomy, young genes are mainly enriched for expression in testis (Kaessmann 2010). Two main factors have been proposed to explain this pattern. Firstly, permissive chromatin in testis facilitates the transcription of most genes, including new genes (Soumillon et al. 2013). The widespread expression in testis appears related to regulation of gene evolution rates based on transcription coupled repair (Xia et al. 2018). Secondly, as the most rapidly evolving organ at

genomic level, there is least purifying selection acting on new genes expressed in testis (Kaessmann 2010). Is there a similar explanation for the ontogenic pattern of young genes tending to be expressed in late development stages? As testis constitutes the most rapidly evolving organ transcriptome, late development represents the most rapidly evolving stage transcriptome, owing to both relaxed purifying selection (Artieri et al. 2009) and to increased positive selection (Liu and Robinson-Rechavi 2017). Thus, we suggest that expression in late development might, like in testis, promote the fixation and functional evolution of new genes.

In order to test this, we analyzed the expression of retrogenes across development. Since retrogenes usually lack regulatory elements, most of them fail to acquire transcription and achieve function (Kaessmann et al. 2009). So, if late development, like testis, can facilitate the transcription of new genes, promoting their fixation, we should observe higher expression of retrogenes in later developmental stages. Because retrogenes have higher expression in testis, and testis is already differentiated after middle development, we excluded testis genes in our analyses for *D. melanogaster* and *M. musculus*, where the information of testis gene expression was available. As expected, the median expression of retrogenes is higher in late development (fig. 4), with a significant positive correlation. Generally, in *C. elegans*, *D. rerio*, and *M. musculus*, the median expression progressively increases; in *D. melanogaster*, all the median values are 0 until stage 4 days, and then it progressively increases.

These results confirm that late development could allow more transcription of new gene copies, which usually lack efficient regulatory elements and transcriptional activity.

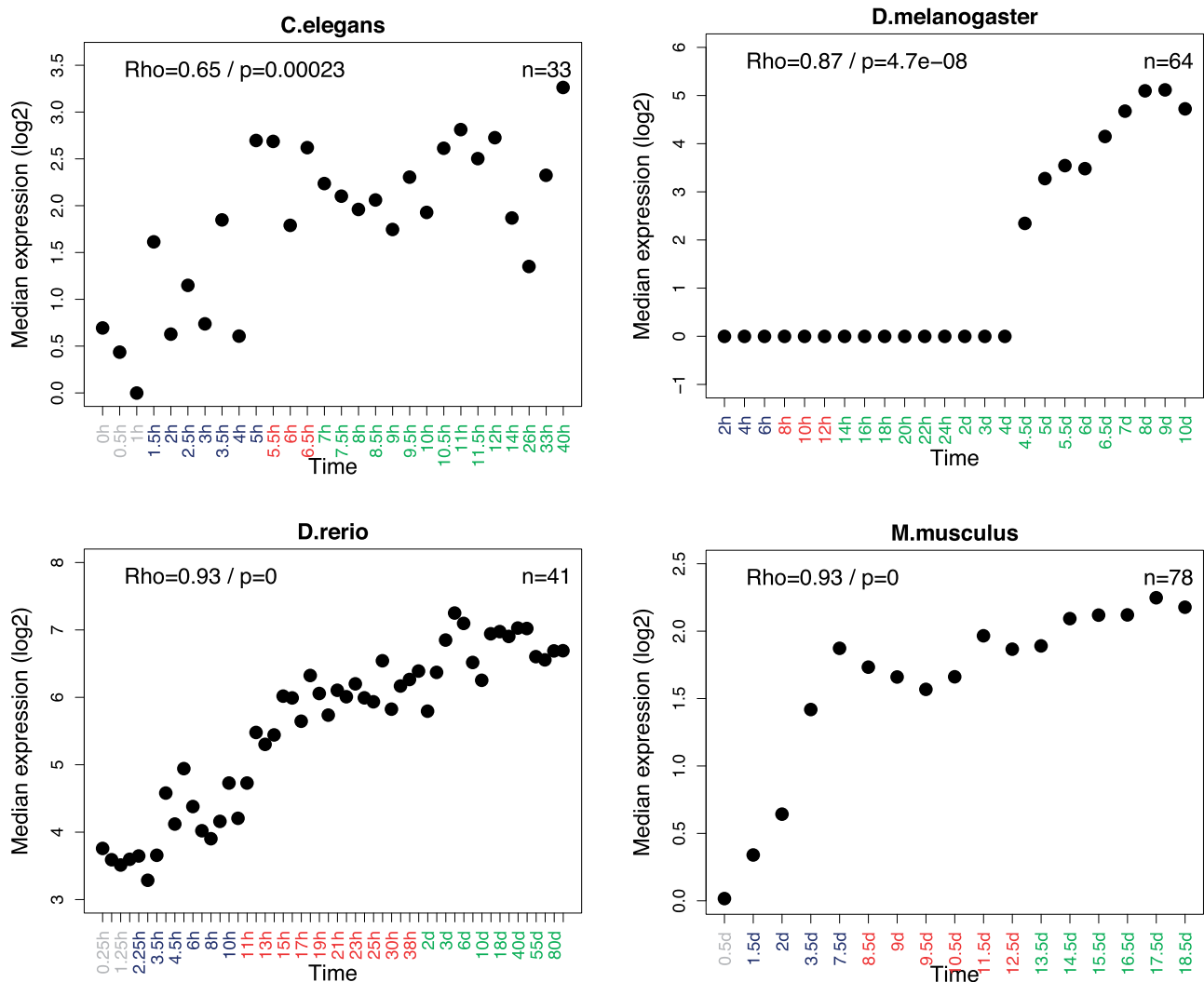


FIG. 4.—Expression of retrogenes in development. Grey, dark blue, red, and green marked time points in the x-axis represent stages before the start of MZT, early developmental stages, middle developmental stages and late developmental stages respectively. A Spearman correlation was computed between development and median expression. The correlation coefficient (Rho) and *P*-value are indicated in the top-left corner of each plot. The numbers of genes analyzed are noted in the top-right corner of each plot.

Since the first step to functionality is acquiring transcription, we suggest that the functional acquisition and survival at the beginning of life history for new genes could be promoted by expression in late development. When beneficial mutations come, a subset of these new gene candidates could subsequently obtain adaptive functions in late development, evolve efficient regulatory elements, and finally be retained long term in the genome. Thus, the higher proportion of young genes expressed in later development stages can be in part explained by these stages favoring the fixation of new genes.

Connectivity and Dosage Imbalance

It has previously been found that, in both *Saccharomyces cerevisiae* and *C. elegans*, gene duplicability is negatively

correlated with protein connectivity (Hughes and Friedman 2005; Prachumwat and Li 2006) which might be explained by dosage balance (Veitia 2002; Papp et al. 2003). Firstly, we checked the relationship of connectivity and duplicability in our data sets. We found, indeed, a negative relationship in *C. elegans* (supplementary fig. S11, Supplementary Material online). In *D. melanogaster* and in *D. rerio*, there is a nonmonotonous pattern (increasing first, and then decreasing), but the overall trend is more connectivity with less duplicability. In *M. musculus*, however, we did not observe a significant relationship between connectivity and duplicability. Secondly, we calculated a transcriptome index of connectivity (Transcriptome Connectivity Index: TCI). In *C. elegans* and *M. musculus*, earlier developmental stages have

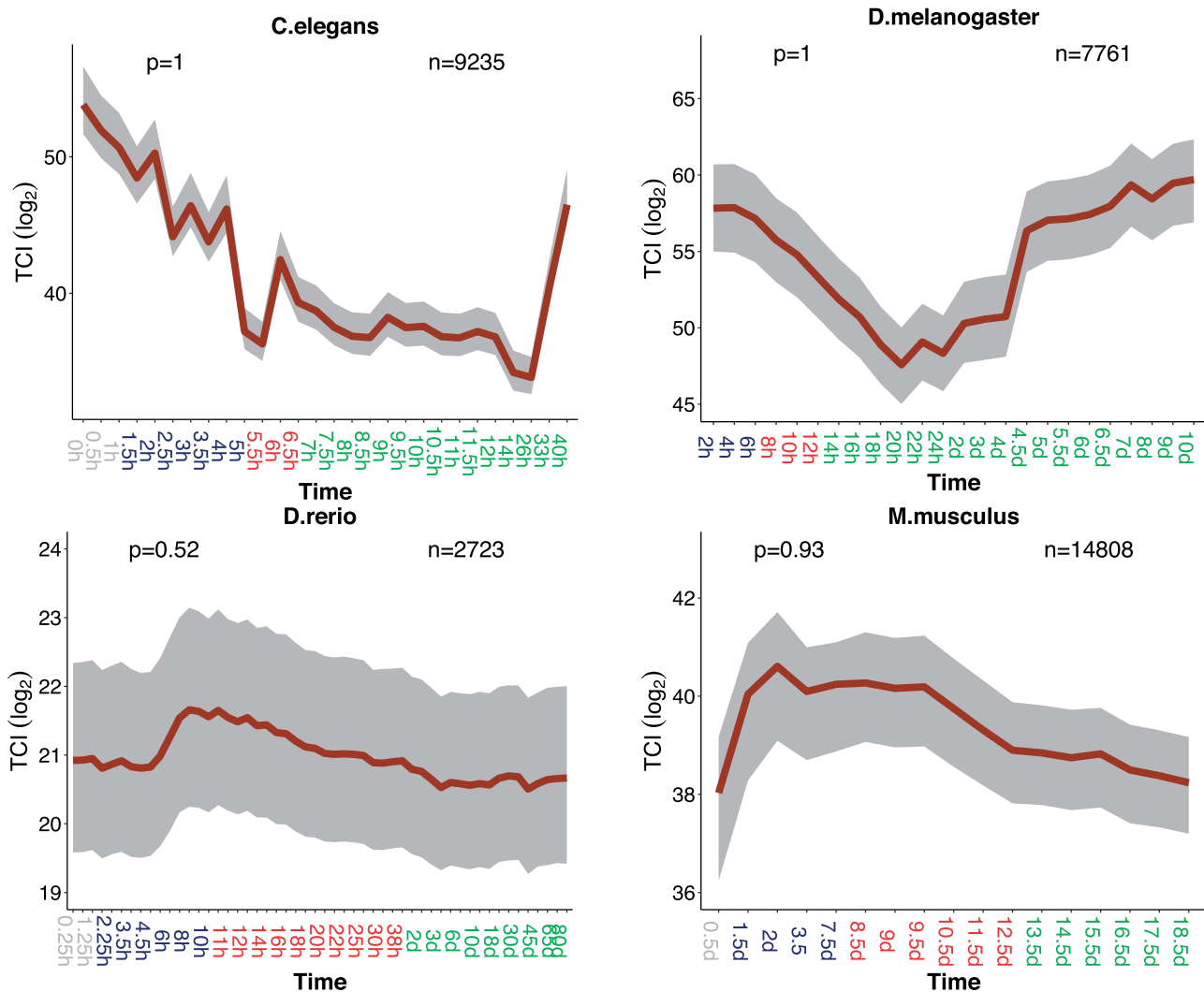


FIG. 5.—Transcriptome index of connectivity (TCI) across development. Grey, dark blue, red, and green marked time points in the x-axis represent stages before the start of MZT, early developmental stages, middle developmental stages, and late developmental stages respectively. TCI is plotted in dark red line. The grey area indicates 95% confidence interval estimated from bootstrap analysis. The P -values for supporting the hourglass model (permutation test, early vs. middle development) are indicated in the top-left corner of each graph. The numbers of genes analyzed are noted in the top-right corner of each plot.

higher TCI, which means that these stages trend to have higher expression of more connected genes (fig. 5). In *D. melanogaster* and in *D. rerio*, there is no clear pattern based on individual stages, but the mean TCI of each developmental period (early, middle and late development) also gradually decreases.

These results indicate that, at least in *C. elegans*, earlier stages trend to express higher connectivity genes, which are less duplicable because more sensitive to dosage imbalance, but that this cannot be generalized to other animals. Of course, this is not exclusive with an adaptive scenario that early stages lack opportunities for neo or subfunctionalization, because of simpler anatomical structures, which could also diminish fixation of duplicates in early development.

Conclusion

Our results concern both patterns and processes of evolution over development. For patterns, we tested the early conservation and hourglass models by using three evolutionary properties: strength of purifying selection, phyletic age and duplicability. The strength of purifying selection on protein sequence supports the hourglass model. Genes under stronger purifying selection are more expressed at middle developmental stages. Both duplicability and phyletic age support the early conservation model. Less duplicated genes and phyletically older genes are more expressed at earlier stages.

For processes, we investigated the potential causes of the observed patterns. The hourglass pattern of sequence

evolution appears to be driven by temporal pleiotropy of gene expression. Genes expressed at middle development evolve under stronger temporal pleiotropic constraints. The enrichment in young phyletic age genes in late development might be related to a testis-like role of late development that facilitates the expression of retrogenes. Finally, in *C. elegans*, connectivity appears to be the main force explaining higher duplicability of genes expressed in later development.

Supplementary Material

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

Acknowledgments

We thank Marie Zufferey and Nadezda Kryuchkova for help with preliminary work. We thank the Bgee team, Marie Zufferey, Nadezda Kryuchkova, and Barbara Piasecka for help with data retrieval, Julien Roux for help with the Ensembl API, Kamil Jaron and Andrea Komljenovic for help with programming, all members of the Robinson-Rechavi lab and Hajk-Georg Drost for helpful discussions. Part of the computations was performed at the Vital-IT (<http://www.vital-it.ch>) centre for high-performance computing of the SIB Swiss Institute of Bioinformatics. This work was supported by Swiss National Science Foundation grant 31003A_153341/1.

Author Contributions

J.L. and M.R.R. designed the work. J.L. performed the data gathering and analysis. J.L. and M.R.R. interpreted the results. J.L. wrote the first draft of the paper. J.L. and M.R.R. finalized the paper.

Literature Cited

- Abzhanov A. 2013. von Baer's law for the ages: lost and found principles of developmental evolution. *Trends Genet.* 29(12):712–722.
- 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.* 7:42.
- Ballard WW. 1981. Morphogenetic movements and fate maps of vertebrates. *Am Zool.* 21(2):391–399.
- Bastian F, et al. 2008. Bgee: integrating and comparing heterogeneous transcriptome data among species. *Data Integr Life Sci.* 5109:124–131.
- Bininda-Emonds ORP, Jonathan JE, Richardson MK. 2003. Inverting the hourglass: quantitative evidence against the phylotypic stage in vertebrate development. *Proc Biol Sci.* 270(1513):341–346.
- Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25–36.
- Castillo-Davis CI, Hartl DL. 2002. Genome evolution and developmental constraint in *Caenorhabditis elegans*. *Mol Biol Evol.* 19(5):728–735.
- Chen W-H, Minguez P, Lercher MJ, Bork P. 2012. OGEE: an online gene essentiality database. *Nucleic Acids Res.* 40(Database issue):D901–D906.
- Cheng X, Hui JHL, Lee YY, Wan Law PT, Kwan HS. 2015. A “developmental hourglass” in fungi. *Mol Biol Evol.* 32(6):1556–1566.
- Comte A, Roux J, Robinson-Rechavi M. 2010. Molecular signaling in zebrafish development and the vertebrate phylotypic period. *Evol Dev.* 12(2):144–156.
- 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, Ward S. 2005. Sexual and temporal dynamics of molecular evolution in *C. elegans* development. *Mol Biol Evol.* 22(1):178–188.
- Davis JC, Brandman O, Petrov DA. 2005. Protein evolution in the context of *Drosophila* development. *J Mol Evol.* 60(6):774–785.
- Domazet-Loso T, Tautz D. 2010. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* 468(7325):815–818.
- 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.
- Drost H-G, Gabel A, Liu J, Quint M, Grosse I. 2018. myTAI: evolutionary transcriptomics with R. *Bioinformatics* 34(9):1589–1590.
- 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.
- Dunn CV, Zapata F, Munro C, Siebert S, Hejnol A. 2018. Pairwise comparisons across species are problematic when analyzing functional genomic data. *Proc Natl Acad Sci U S A.* 115(3):E409–E417.
- Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW. 2002. Evolutionary rate in the protein interaction network. *Science* (80-) 296(5568):750–752.
- Fu B, Chen M, Zou M, Long M, He S. 2010. The rapid generation of chimerical genes expanding protein diversity in zebrafish. *BMC Genomics.* 11:657.
- Galis F, Metz JA. 2001. Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation. *J Exp Zool.* 291(2):195–204.
- Garstang W. 1922. The theory of recapitulation: a critical re-statement of the biogenetic law. *J Linn Soc Lond Zool.* 35(232):81–101.
- Gerstein MB, et al. 2010. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* (80-) 330(6012):1775–1787.
- Gerstein MB, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature* 512(7515):445–448.
- Graveley BR, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471(7339):473–479.
- Hanada K, Shiu S-H, Li W-H. 2007. The nonsynonymous/synonymous substitution rate ratio versus the radical/conservative replacement rate ratio in the evolution of mammalian genes. *Mol Biol Evol.* 24(10):2235–2241.
- Hashimshony T, Wagner F, Sher N, Yanai I. 2012. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* 2(3):666–673.
- Hazkani-Covo E, Wool D, Graur D. 2005. In search of the vertebrate phylotypic stage: a molecular examination of the developmental hourglass model and von Baer's third law. *J Exp Zool B Mol Dev Evol.* 304B(2):150–158.
- Hu H, et al. 2017. Constrained vertebrate evolution by pleiotropic genes. *Nat Ecol Evol.* 1(11):1722–1730.
- Hughes AL, Friedman R. 2005. Gene duplication and the properties of biological networks. *J Mol Evol.* 61(6):758–764.
- 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.

- Irie N, Sehara-Fujisawa A. 2007. The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. *BMC Biol.* 5:1.
- Kabza M, Ciombarowska J, Makaowska I. 2014. RetrogeneDB – a database of animal retrogenes. *Mol Biol Evol.* 31(7):1646–1648.
- Kaessmann H. 2010. Origins, evolution, and phenotypic impact of new genes. *Genome Res.* 20(10):1313–1326.
- Kaessmann H, Vinckenbosch N, Long M. 2009. RNA-based gene duplication: mechanistic and evolutionary insights. *Nat Rev Genet.* 10(1):19–31.
- 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.
- Kersey PJ, et al. 2016. Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Res.* 44(D1):D574–D580.
- Kinsella RJ, et al. 2011. Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database* 2011:bar030.
- Kryuchkova-Mostacci N, Robinson-Rechavi M. 2015. Tissue-specific evolution of protein coding genes in human and mouse. *PLoS ONE.* 10(6):e0131673–e0131615.
- Kryuchkova-Mostacci N, Robinson-Rechavi M. 2016a. Tissue-specificity of gene expression diverges slowly between orthologs, and rapidly between paralogs. *PLOS Comput Biol.* 12:e1005274.
- Kryuchkova-Mostacci N, Robinson-Rechavi M. 2016b. A benchmark of gene expression tissue-specificity metrics. *Brief Bioinform.* 18:205–214.
- 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.
- Levin M, Hashimshony T, Wagner F, Yanai I. 2012. Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Dev Cell.* 22(5):1101–1108.
- Li JJ, Huang H, Bickel PJ, Brenner SE. 2014. Comparison of *D. melanogaster* and *C. elegans* developmental stages, tissues, and cells by modENCODE RNA-seq data. *Genome Res.* 24(7):1086–1101.
- Liu J, Robinson-Rechavi M. 2017. Adaptive evolution of proteins expressed in late and post-embryonic development in animals. *bioRxiv* doi:10.1101/161711.
- Moretti S, et al. 2014. Selectome update: quality control and computational improvements to a database of positive selection. *Nucleic Acids Res.* 42(D1):D917–D921.
- Ninova M, Ronshaugen M, Griffiths-Jones S. 2014. Conserved temporal patterns of microRNA expression in *Drosophila* support a developmental hourglass model. *Genome Biol Evol.* 6(9):2459–2467.
- Papp B, Pál C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. *Nature* 424(6945):194–197.
- 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.
- Przybylowski L, et al. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol.* 6(4):e80.
- Prachumwat A, Li W-H. 2006. Protein function, connectivity, and duplicability in yeast. *Mol Biol Evol.* 23(1):30–39.
- Quint M, et al. 2012. A transcriptomic hourglass in plant embryogenesis. *Nature* 490(7418):98–101.
- Raff RA. 1996. *The shape of life: genes, development, and the evolution of animal form.* Chicago: University of Chicago Press.
- Richardson MK. 1995. Heterochrony and the phylotypic period. *Dev Biol.* 172(2):412–421.
- 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 (Berl).* 196(2):91–106.
- Riedl R. 1978. *Order in living organisms.* West Sussex, UK: Wiley-Interscience.
- Roux J, Robinson-Rechavi M. 2008. Developmental constraints on vertebrate genome evolution. *PLoS Genet.* 4(12):e1000311.
- Sander K. 1983. *Development and evolution: the sixth symposium of the british society for developmental biology.* Cambridge University Press.
- Slack JMW, Holland PWH, Graham CF. 1993. The zootype and the phylotypic stage. *Nature* 361(6412):490–492.
- Soumillon M, et al. 2013. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep.* 3(6):2179–2190.
- Tadros W, Lipshitz HD. 2009. The maternal-to-zygotic transition: a play in two acts. *Development* 136(18):3033–3042.
- Veitia RA. 2002. Exploring the etiology of haploinsufficiency. *BioEssays* 24(2):175–184.
- Von-Baer KE. 1828. *Über entwicklungsgeschichte der tiere: beobachtung und reflexion.* Königsberg: Gebrüder Bornträger.
- Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat Rev Genet.* 12(3):204–213.
- Wang Z, et al. 2013. The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat Genet.* 45(6):701–706.
- Wolpert L. 1991. *The triumph of the embryo.* Oxford: Oxford University Press.
- Xia B, et al. 2018. Widespread transcriptional scanning in testes modulates gene evolution rates. *bioRxiv* doi:10.1101/282129.
- Yanai I, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21(5):650–659.
- Yanai I, Peshkin L, Jorgensen P, Kirschner MW. 2011. Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Dev Cell.* 20(4):483–496.
- Yates A, et al. 2016. Ensembl 2016. *Nucleic Acids Res.* 44(D1):D710–D716.
- Zalts H, Yanai I. 2017. Developmental constraints shape the evolution of the nematode mid-developmental transition. *Nat Ecol Evol.* 1(5):113.
- 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.
- Zou M, Wang G, He S. 2012. Evolutionary patterns of RNA-based gene duplicates in *Caenorhabditis* nematodes coincide with their genomic features. *BMC Res Notes.* 5:398.

Associate editor: Gunter Wagner