Adaptive evolution of animal proteins over development: support for the Darwin selection opportunity hypothesis of Evo-Devo

Jialin Liu[1,2], Marc Robinson-Rechavi[1,2*]

1 Department of Ecology and Evolution, University of Lausanne, Switzerland
2 Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland
* corresponding author: marc.robinson-rechavi@unil.ch

Abstract

A driving hypothesis of Evo-Devo is that animal morphological diversity is shaped both by adaptation and by developmental constraints. Here we have tested Darwin's "selection opportunity" hypothesis, according to which high evolutionary divergence in late development is due to strong positive selection. We contrasted it to a "developmental constraint" hypothesis, according to which late development is under relaxed negative selection. Indeed, the highest divergence between species, both at the morphological and molecular levels, is observed late in embryogenesis and post-embryonically. To distinguish between adaptation and relaxation hypotheses, we investigated the evidence of positive selection on protein-coding genes in relation to their expression over development, in fly *Drosohila melanogaster*, zebrafish *Danio rerio*, and mouse *Mus musculus*. First, we found that genes specifically expressed in late development have stronger signals of positive selection. Second, over the full transcriptome, genes with evidence for positive selection trend to be expressed in late development. Finally, genes involved in pathways with cumulative evidence of positive selection have higher expression in late development. Overall, there is a consistent signal that positive selection mainly affects genes and pathways expressed in late embryonic development and in adult. Our results imply that the evolution of embryogenesis is mostly conservative, with most adaptive evolution affecting some stages of post-embryonic gene expression, and thus post-embryonic phenotypes. This is consistent with the diversity of environmental challenges to which juveniles and adults are exposed.

Introduction

There are two main models to explain the relationship of development and evolutionary divergence. The early conservation model suggests that embryonic morphology between different species within the same group progressively diverges across development (Von Baer 1828); such groups are usually understood to by phyla in a modern context. In contrast, the hourglass model proposes that middle development (the morphological 'phylotypic' period) has the highest morphological similarity (Duboule 1994; Raff 1996). Based on recent genomic studies, both models have some level of molecular support. Some studies support the early conservation model (Roux and Robinson-Rechavi 2008; Artieri et al. 2009), while most recent ones support the hourglass model (Kalinka et al. 2010; Irie and Kuratani 2011; Levin et al. 2012; Quint et al. 2012; Drost et al. 2015; Hu et al. 2017; Zalts and Yanai 2017). And in fact the two models may not be mutually exclusive (Piasecka et al. 2013; Liu and Robinson-Rechavi 2018).

Both the early conservation and hourglass models predict that late development has high evolutionary divergence. This high divergence of late development has been interpreted as a consequence of relaxed developmental constraints, i.e. weaker negative selection. For example, Garstang (1922) and Riedl (1978) suggested that the development of later stages is dependent on earlier stages, so higher divergence should be found in the later stages of development (cited in Irie and Kuratani 2014). Indeed, many studies have found evidence for relaxed purifying selection in late development (Castillo-Davis and Hartl 2002; Roux and Robinson-Rechavi 2008; Artieri et al. 2009; Kalinka et al. 2010; Liu and Robinson-Rechavi 2018). An alternative explanation, however, known as Darwin's "selection opportunity" hypothesis (Darwin 1871)(cited in Artieri et al. 2009), proposed that highly divergent late development could also be driven by adaptive evolution (positive selection), at least in part. This could be due to the greater diversity of challenges to which natural selection needs to respond in juvenile and adult life than in early and mid-development. Notably, weaker negative and stronger positive selection are not mutually exclusive. For example, Cai and Petrov (2010) found the accelerated sequence evolution rate of primate lineage specific genes driven by both relaxed purifying selection and enhanced positive selection. Necsulea and Kaessmann (2014)

suggested that the high evolution rate of testis transcriptome could be caused by both sex-related positive selection and reduced constraint on transcription.

As far as we know, few studies have tried to distinguish the roles of adaptation vs. relaxation of constraints in late development (Artieri et al. 2009), and no evidence has shown stronger adaptive evolution in late development. Yet there is an intuitive case for adaptation to act on phenotypes established in late development, because they will be present in the juvenile and adult, and interact with a changing environment.

In the case of detecting individual gene adaptation, one of the best established methods is using the ratio ω of non-synonymous (dN) to synonymous (dS) substitutions (Yang and Nielsen 1998; Hurst 2002). Because synonymous changes are assumed to be functionally neutral, $\omega>1$ indicates evidence of positive selection. As adaptive changes probably affect only a few codon sites and at a few phylogenetic lineages, branch-site models allow the ω ratio to vary both among codon sites and among lineages (Yang and Nielsen 2002; Zhang et al. 2005). Polymorphism based methods such as frequency spectrum, linkage disequilibrium and population differentiation can also be used to identify changes due to recent positive selection (Vitti et al. 2013).

Since several genes with slight effect mutations can act together to have a strong effect, adaptive evolution can act on the pathway level as well (Daub et al. 2013; Berg et al. 2014). In the case of polygenic adaptation, a gene set enrichment test has successfully been applied to detect gene sets with polygenic adaptive signals (Daub et al. 2013; Daub et al. 2017). This gene set enrichment analysis allows to detect weak but consistent adaptive signals from whole genome scale, unlike traditional enrichment tests which only consider top scoring genes with an arbitrary significance threshold.

In order to estimate the contribution of positive selection to the evolution of highly divergent late development, we have adopted three approaches. First, we used modularity analysis to obtain distinct sets of genes (modules) which are specifically expressed in different meta developmental stages (Piasecka et al. 2013; Levin et al. 2016), and compared the signal of positive selection across modules. Second, we applied a modified "transcriptome index" (Domazet-Loso and Tautz 2010) to measure evolutionary adaptation on the whole transcriptome level. Finally, we used a gene set enrichment approach to detect polygenic selection on pathways, and studied the

expression of these gene sets over development. Each approach was applied to developmental transcriptomes from *D. rerio*, *M. musculus*, and *D. melanogaster* and to results of the branch-site test for positive selection in lineages leading to these species. All the analyses found a higher rate of adaptation in late and in some stages of post-embryonic development, including adult.

Results

In order to characterize the signal of positive selection, we used the log-likelihood ratio test statistic (ΔlnL) of H_1 to H_0 models with or without positive selection, from the branch-site model (Zhang et al. 2005) as precomputed in Selectome on filtered alignments (Moretti et al. 2014), and as used in Roux et al. (2014) and Daub et al. (2017). Briefly, ΔlnL represents the evidence for positive selection, thus a branch in a gene tree with a higher value indicates higher evidence for positive selection for this gene over this branch.

Modularity analysis

For the modularity analysis, we focused on different sets of specifically expressed genes (modules) in each developmental period. Our expectation is that genes in each module have specific involvement during embryonic development (Piasecka et al. 2013), so different adaptation rates of these genes should reflect a stage specific impact of natural selection. In addition, since the modules decompose the genes into different meta development stages, they allow to avoid the potential bias caused by imbalanced time points in each meta development stage from our transcriptome datasets; e.g., many more "late development" samples in fly than in the other two species studied. For *D. rerio*, we obtained seven modules from our previous study (Piasecka et al. 2013) (Figure S1). For *M. musculus* and *D. melanogaster*, we identified three and six modules respectively (see Methods; Figure S1).

Because not all genes have any evidence for positive selection, we first compared the proportion of genes either with strong evidence (q-value < 0.2) or with weak evidence (no threshold for q-value; ΔlnL > 0) of positive selection across modules. For strong evidence, the proportion is not significantly different across modules in *M. musculus* and *D. melanogaster* (Figure S2). In *D. rerio*, however, there is a higher proportion in the juvenile and adult modules. For the weak evidence, *D. melanogaster* has a higher

proportion in pupae and adult modules, but there is no significant difference in *D. rerio* and *M. musculus* (Figure S3).

We then compared the values of ΔlnL for genes with weak evidence of positive selection (Figure 1). In order to improve the normality of non-zero ΔlnL, we transformed ΔlnL with fourth root (Hawkins and Wixley 1986; Roux et al. 2014; Daub et al. 2017).

In *D. rerio*, we detected an hourglass pattern of ΔlnL, at its highest in late modules. Specifically, in the juvenile module, the mean ΔlnL is significantly higher than the mean ΔlnL for all genes (*p*-values reported in Table 1). We note that the adult module also has higher mean ΔlnL, even though it's not significant. In the pharyngula module, the mean ΔlnL is significantly lower than the mean ΔlnL for all genes, as expected under the hourglass model. In the other modules, the mean ΔlnL is not significantly different from the mean for all genes.

In *M. musculus*, similarly, we found an hourglass pattern of ΔlnL. The late embryo module has a higher mean ΔlnL than all genes, while the middle embryo module has a lower mean ΔlnL than all genes.

In *D. melanogaster*, however, we observed an early conservation pattern of ΔlnL. Specifically, in the early embryo module, the mean ΔlnL is lower than the mean ΔlnL for all genes. In the adult module, the mean ΔlnL is higher than the mean ΔlnL for all genes. There is no significant difference for the other modules.

It should be noted that the patterns reported in this modularity analysis are relatively weak, especially in *D. melanogaster*. After multiple test correction, some of the reported differences are not significant anymore (Table S1).

Overall, these findings suggest that positive selection is stronger on genes expressed in late development or in adult than in early and middle development. It also indicates that ΔlnL on gene modules in different phyla supports different Evo-Devo models (hourglass vs. early conservation).

Transcriptome index analysis

Although modularity analysis guarantees independence between the sets of genes which are compared, it only considers a subset of genes. This leaves open whether the higher adaptive evolution in late development and adult holds true for the whole transcriptome as well, or just for these modular genes. Additionally, while trends were

163 detected, significance is weak. To consider the composition of the whole
164 transcriptome and to increase our power to detect a signal of positive selection in
165 development, we used a modified "Transcriptome Age Index" (Domazet-Loso and
166 Tautz 2010) to calculate the weighted mean of ΔlnL for the transcriptome. Notably,
167 all expression levels were log-transformed before use, unlike in Domazet-Loso and
168 Tautz (2010). See discussion in Piasecka et al. (2013) and Liu and Robinson-Rechavi
169 (2018), but briefly log-transformation provides insight on the overall transcriptome
170 rather than a small number of highly expressed genes. We named this modified index
171 "Transcriptome Likelihood Index" (TLI). A higher index indicates that the
172 transcriptome has higher expression of transcripts from genes with high ΔlnL between
173 models with and without positive selection.
174 In *D. rerio*, generally, the pattern resembles an hourglass like pattern (Figure 2). The
175 TLI first decreases and reaches a minimum in the late stage of gastrula (8h), and then
176 progressively increases until adult (ninth month), with finally a slight decline. In
177 addition, in the adult stage, female has higher TLI than male, although the difference
178 is weak. To test whether TLIs are different between developmental periods, we
179 compared the mean TLI of all stages within a period, between each pair of periods
180 (see Methods). We found that middle development has low TLI, early development
181 has medium TLI, late development and maternal stage have very similar high TLI,
182 and adult has the highest TLI. Except late development and maternal stage (p=0.24),
183 all pairwise comparisons are significant: p<5.7e-07.
184 In *M. musculus*, we observed a clear hourglass-like pattern of TLI. For the mean TLI
185 comparison, we found low TLI in middle development, medium TLI in early
186 development, high TLI in late development, and the highest TLI in maternal stage (all
187 pairwise comparisons are significant: p<2e-16). Of note, unlike in *D. rerio*, the "late
188 development" here only contains late embryo stages, but no post embryo stages. This
189 may explain why late development has lower TLI than the maternal stage in this
190 dataset.
191 In *D. melanogaster*, we found the TLI progressively increasing over development,
192 suggesting an early conservation model. Unlike in *D. rerio*, we found that male has
193 higher TLI than female in the adult stage. For the mean TLI comparison, early
194 development has low TLI, middle development has medium TLI, late development

195 has high TLI, and adult has the highest TLI (all pairwise comparisons are significant:
196 p<2e-16).
197 As in the modularity analysis, but with much stronger signal, both *D. rerio* and *M.*
198 *musculus* support the hourglass model, while *D. melanogaster* follows an early
199 conservation model. Again, from whole transcriptome level, these results indicate that
200 genes with evidence for positive selection are more highly expressed in late
201 development and adult. Interestingly, the maternal stage has a comparable high TLI to
202 late development. This could be related to the maternal stage being dominated by
203 adult transcripts (Tadros and Lipshitz 2009). In this respect (transcriptome evolution),
204 the maternal stage should maybe be regarded as a special adult stage rather than as an
205 early embryonic stage.
206
207 **Polygenic selection analysis**
208 Positive selection can be detected at the biological pathway level, even when
209 individual genes within the pathway only fix small effect mutations (Daub et al. 2013;
210 Berg et al. 2014; Daub et al. 2017). Thus, we searched for such signals of positive
211 selection on pathways. Briefly, we calculated the sum of ΔlnL (SUMSTAT statistic)
212 for a pathway, and inferred the significance of this SUMSTAT with an empirical null
213 distribution (Tintle et al. 2009; Daub et al. 2013; Daub et al. 2017). In total, we
214 identified 10, 4 and 9 pathways with a significant signal of positive selection,
215 respectively in lineages leading to *D. rerio*, *M. musculus* and *D. melanogaster* (q-
216 value<0.2, Table2).
217 The function of these pathways, while not our primary focus, is consistent with
218 adaptive evolution of juvenile or adult phenotypes. First, we found metabolism related
219 pathways in all three species, suggesting pervasive adaptation, possibly related to diet;
220 this is consistent with previous results in primates (Daub et al. 2017). Second, in *D.*
221 *rerio* and *D. melanogaster*, several pathways are involved in morphogenesis and
222 remodelling of organs (e.g., laminin interactions, extra cellular matrix, ECM-receptor
223 interaction), suggesting potential adaptive evolution of morphological development.
224 Third, there are several pathways involved in aging in *D. melanogaster* and *M.*
225 *musculus* (e.g., reactive oxygen detoxification, longevity regulation, mitochondrial
226 translation), suggesting potential role of natural selection on modulating lifespan or on
227 metabolic activity. Forth, in *D. rerio,* we detected one pathway related to

228 environmental adaptation: visual phototransduction; adaptations in vision are
229 expected for aquatic species which under a wide variety of visual environments
230 (Sabbah et al. 2010).
231 If late development and adult are under stronger positive selection at the pathway
232 level as well, we expect genes involved in pathways with a signal of positive selection
233 to be more highly expressed at these periods. Thus we computed the ratio of median
234 expression between positively selected pathway genes and genes included in
235 pathways not positively selected. Since the median expression in the first time point
236 of *M. musculus* is 0, we removed it from our analysis.
237 In *D. rerio*, the ratio of median expression keeps increasing until the juvenile stage.
238 Then, it slightly decreases (Figure 3). In *M. musculus*, except the first time point, the
239 ratio of median expression also progressively increases. In *D. melanogaster*, there is a
240 small peak in the first time point, but it quickly decreases to minimum within the
241 same developmental period. Then, it keeps increasing until the middle of the larval
242 stage. Finally, for the last development stages, it resembles a wave pattern: decrease,
243 increase and decrease again. Again, we also tested the difference between male and
244 female in adult stages for *D. rerio* and *D. melanogaster*. Unlike the observation in the
245 transcriptome index analysis, here we found that male has higher ratio of median
246 expression than female in both species.
247 Overall, consistent with previous results, we found that late development and adult
248 tend to express genes involved in pathways enriched for signal of positive selection,
249 indicating that adaptive evolution at the pathway level mainly affects these stages.
250 While there is some signal of early development adaptive evolution on single genes,
251 the later developmental signal is more consistent at the pathway level. Because
252 pathways link genes to phenotypes (Müller 2007; Wray 2007; Tickle and Urrutia
253 2017), this suggests stronger phenotypic adaptation in late development and adult.
254

255 ## Discussion

256 ### Correcting confounding factors

257 Since some non-adaptive factors (such as gene length, tree size (number of branches),
258 and branch length) can be correlated with ΔlnL and affect our results (Daub et al.
259 2017), we investigated the correlation between ΔlnL and these potential confounding
260 factors. Generally, we found a small correlation between ΔlnL and tree size, but a

261 larger correlation between ΔlnL and gene length or branch length (Figure S4). One
262 explanation for this high correlation between ΔlnL and gene length is that long genes
263 could accumulate more mutations than short genes, so we have more power to detect
264 positive selection with higher number of mutations (Fletcher and Yang 2010; Gharib
265 and Robinson-Rechavi 2013). So, we checked the influence of gene length on our
266 results. Because branch length is inferred from the number of mutations, and higher
267 branch length can be driven by higher evolutionary rate due to positive selection, we
268 did not check further the correlation between ΔlnL and branch length.
269 In order to investigate whether gene length might have affected our results, for
270 modularity and TLI analysis, we tested whether patterns purely based on gene length
271 are similar to those based on ΔlnL or not. Surprisingly, we found an opposite pattern
272 of gene length, relative to ΔlnL. For modularity analysis, the modules with higher
273 ΔlnL have significantly lower mean gene length than all genes (Figure S5). For
274 transcriptome index analysis, the stages with higher TLI trend to have lower
275 transcriptome index for gene length (Figure S6), suggesting that these stages trend to
276 express shorter genes. These findings imply that the detection of higher positive
277 selection in late development is not driven by gene length.
278 Immune system genes can bias positive selection analyses, since they evolve under
279 pervasive positive selection (Flajnik and Kasahara 2010). To control for this, we also
280 confirmed our findings after removing immune genes from our analysis (Figure S7).
281
282 **Developmental constraint hypothesis and Darwin's selection opportunity**
283 **hypothesis**
284 Despite the repeated observation that late development is highly divergent for diverse
285 genomic properties (sequence evolution, duplication, gene age, expression divergence)
286 in diverse animal species (Roux and Robinson-Rechavi 2008; Domazet-Loso and
287 Tautz 2010; Kalinka et al. 2010; Irie and Kuratani 2011; Levin et al. 2012; Piasecka et
288 al. 2013; Drost et al. 2015; Liu and Robinson-Rechavi 2018), the underlying
289 evolutionary forces driving such a pattern remain obscure. The “developmental
290 constraint” hypothesis (Raff 2000; Brakefield 2006) suggests that this high divergence
291 is due to relaxed purifying selection, whereas Darwin's "selection opportunity"
292 hypothesis proposes stronger positive selection (as discussed in Artieri et al. 2009;
293 Kalinka and Tomancak 2012).

294 Several studies have found evidence, direct or indirect, to support the importance of
295 developmental constraints (Castillo-Davis and Hartl 2002; Roux and Robinson-
296 Rechavi 2008; Artieri et al. 2009; Kalinka et al. 2010). For example, we (Roux and
297 Robinson-Rechavi 2008) found that genes expressed earlier in development contain a
298 higher proportion of essential genes, and Uchida et al. (2018) found strong embryonic
299 lethality from random mutations in early development. Weaker purifying selection in
300 late development would imply that genes expressed in this period have less fitness
301 impact, which is consistent with the paucity of essential genes. Here and in Liu and
302 Robinson-Rechavi (2017) the branch-site codon model allows us to isolate the
303 contribution of purifying selection to coding sequence evolution. We found indeed
304 that genes under weaker purifying selection on the protein sequence trend to be
305 expressed in late development (Liu and Robinson-Rechavi 2018). This provides direct
306 evidence of relaxed purifying selection in late development.
307 To the best our knowledge, there has been no direct test of Darwin's "selection
308 opportunity" hypothesis. One such study, in *D. melanogaster*, was proposed by Artieri
309 et al. (2009). Unfortunately, they only had relatively poor expression data (ESTs) and
310 limited time points (embryonic, larval/pupal and adult), and they did not find any
311 direct evidence of higher positive selection in late development. Since they noticed
312 that the accelerated sequence evolution of genes expressed at adult stage was
313 confounded by male-biased genes, they argued that the rapid evolution observed in
314 late development could be due to specific selective pressures such as sexual selection.
315 A recent study, in *D. melanogaster*, provides indirect evidence: using in situ
316 expression data and population genomic data to map positive selection to different
317 embryonic anatomical structures, Salvador-Martínez et al. (2018) found larva stage
318 enriched with signal of positive selection. Our results clearly provide a quantitative
319 test which supports a role of positive selection in the high divergence of late
320 development. While our sampling is very far from covering the diversity of
321 developmental modes of animals, we show consistent patterns in a placental mammal,
322 a direct development ray-finned fish, and a holometabolous insect. While it is possible
323 that other patterns will be found in species with different development, this shows that
324 adaptation in late development is not limited to one model. We show that this is not
325 due to testis-expressed genes (Figure S8). In addition, in vertebrates, we also found
326 some evidence of adaptive evolution in early development on single genes. This

327 indicates that some changes in early development might be adaptive consequences to
328 diverse ecological niches, as proposed by Kalinka and Tomancak (2012). It should be
329 noted that our results also provide counter evidence to the adaptive penetrance
330 hypothesis, which argues that adaptive evolution mainly occurs in the middle
331 development (Richardson 1999).
332
333 **Re-unification of structuralist and functionalist comparative biology**
334 There have been two major approaches to comparative biology since the late 18th
335 century: the structuralist approach (which gave rise to Evo-Devo) emphasizes the role
336 of constraints, and often focuses on investigating spatial and timing variations of
337 conserved structures in distantly related species. In a modern context, the focus is
338 often on comparing developmental genes' expression between species. The
339 functionalist or adaptationist approach (which gave rise to the Modern Synthesis and
340 most of evolutionary biology) emphasizes the role of natural selection. In a modern
341 context, the focus is often on investigating adaptive mutations. It has been suggested
342 that these two approaches could not be reconciled (Amundson 2007), since the former
343 underscores how mutations generate morphological diversity, while the later
344 underscores whether mutations are fixed by positive selection or not. A good example
345 of the differences between structuralist and adaptationist comes from the debate
346 between Hoekstra and Coyne (2007) and Carroll (2008). As a structuralist, Carroll
347 suggested that mutations affecting morphology largely occur in the *cis*-regulatory
348 regions. However, as adaptationists, Hoekstra and Coyne argued that this statement is
349 at best premature. Their main argument was that they didn't find that adaptive
350 evolution was more likely occur in *cis*-regulatory elements, but rather in protein
351 coding genes, from both genome-wide surveys and single-locus studies. It is
352 important to note that Carroll's theory is specific to morphological evolution, but not
353 directly related to evolutionary adaptation. Basically, both sides could be correct, and
354 were mostly discussing different things.
355 Since both adaptation and structure are part of biology, we should be able to explain
356 both in a consistent manner. Here, we try to bridge positive selection and
357 morphological evolution by combining developmental time-series transcriptomes,
358 positive selection inference on protein coding genes, modularity analysis,
359 transcriptome index analysis, and gene set analysis. From both modularity analysis

360 and transcriptome index analysis, we found that genes highly expressed in late
361 development and adult have higher evidence for positive selection. From polygenic
362 analysis, we found that the expression of positively selected pathways is higher in late
363 development and adult. Overall, these results suggest that higher morphological
364 variation in late development could be at least in part driven by adaptive evolution. In
365 addition, coding sequence evolution might also make a significant contribution to the
366 evolution of morphology, as suggested by Hoekstra and Coyne (2007) and Burga et al.
367 (2017). This is also supported by the observation of tissue-specific positive selection
368 in *D. melanogaster* development (Salvador-Martínez et al. 2018). It should be noted
369 that we do not test here whether regulatory sequence evolution plays a similar or
370 greater role, since we do not have equivalent methods to test for positive selection in
371 regulatory regions.
372

373 ## Materials and Methods

374 Data files and analysis scripts are available on our GitHub repository:
375 https://github.com/ljljolinq1010/Adaptive-evolution-in-late-development-and-adult

376 **Expression data sets**

377 For *D. rerio*, the log-transformed and normalized microarray data was downloaded
378 from our previous study (Piasecka et al. 2013). This data includes 60 stages from egg
379 to adult, which originally comes from Domazet-Loso and Tautz (2010).
380 For *M. musculus*, the processed RNA-seq (normalized but non-transformed) data was
381 retrieved from Hu et al. (2017). This data includes 17 stages from 2cells to E18.5. We
382 further transformed it with $\log_2$.
383 For *D. melanogaster*, we obtained processed (normalized but non-transformed) RNA-
384 seq data from http://www.stat.ucla.edu/~jingyi.li/software-and-data.html Li et al.
385 (2014), which originally comes from Graveley et al. (2011). This data has 27 stages
386 from embryo to adult. For the last three stages, since data were available for male and
387 female, we took the mean. We further transformed it with $\log_2$.

388 **Branch-site likelihood test data**

389 The log-likelihood ratio (ΔlnL) values of a test for positive selection were retrieved
390 from Selectome (Moretti et al. 2014), a database of positive selection based on the
391 branch-site likelihood test (Zhang et al. 2005). One major advantage of this test is
392 allowing positive selection to vary both among codon sites and among phylogenetic

393 branches. The branch-site test contrasts two hypotheses: the null hypothesis is that no
394 positive selection occurred (H0) in the phylogenetic branch of interest, and the
395 alternative hypothesis is that at least some codons experienced positive selection (H1).
396 The log likelihood ratio statistic (ΔlnL) is computed as 2*(lnLH1-lnLH0).
397 Importantly, in order to mitigate false positives due to poor sequence alignments,
398 Selectome integrates filtering and realignment steps to exclude ambiguously aligned
399 regions.
400 We used ΔlnL from the Clupeocephala branch, the Murinae branch and the
401 Melanogaster group branch for *D. rerio*, *M. musculus* and *D. melanogaster*
402 respectively. One gene could have two ΔlnL values in the focal branch because of
403 duplication events. In this case, we keep the value of the branch following the
404 duplication and exclude the value of the branch preceding the duplication.
405 **Pathways**
406 We downloaded lists of 1,683 *D. rerio* gene sets, 2,269 *M. musculus* gene sets and
407 1365 *D. melanogaster* gene sets of type "pathway" from the NCBI Biosystems
408 Database (Geer et al. 2009). This is a repository of gene sets collected from manually
409 curated pathway databases, such as BioCyc (Caspi et al. 2014), KEGG (Kanehisa et al.
410 2014), Reactome (Croft et al. 2014), The National Cancer Institute Pathway
411 Interaction Database (Schaefer et al. 2009) and Wikipathways (Kelder et al. 2012).
412 **Coding sequence length**
413 We extracted coding sequence (CDS) length from Ensembl version 84 (Yates et al.
414 2016) using BioMart (Kinsella et al. 2011). For genes with several transcripts, we
415 used the transcript with the maximal CDS length.
416 **Testis specific genes**
417 Testis specific genes for *M. musculus* and *D. melanogaster* were obtained from a
418 parallel study (Liu and Robinson-Rechavi 2018). The testis specific genes were
419 defined as genes with highest expression in testis and with tissue specificity value ≥
420 0.8.
421 **Immune genes**
422 To control for the impact of immune system genes, we downloaded all genes involved
423 in the "immune response" term (GO:0006955) from AmiGO (Carbon et al. 2009)
424 (accessed on 25.04.2018), and repeated analyses with these genes excluded.
425 **Phylotypic period**

426 The definition of phylotypic period is based on previous morphological and genomic
427 studies. For *D. melanogaster*, the phylotypic period defined as extended germband
428 stage (Sander 1983; Kalinka et al. 2010); for *D. rerio*, the phylotypic period defined
429 as segmentation and pharyngula satges (Ballard 1981; Wolpert 1991; Slack et al. 1993;
430 Domazet-Loso and Tautz 2010); for *M. musculus*, the phylotypic period defined as
431 Theiler Stage 13 to 20 (Ballard 1981; Wolpert 1991; Slack et al. 1993; Irie and
432 Kuratani 2011).

433 **Module detection**

434 For D. rerio, we obtained seven modules from our previous study (Piasecka et al.
435 2013). This is based on the Iterative Signature Algorithm, which identifies modules
436 by an iterative procedure (Bergmann et al. 2003). Specifically, it was initialized with
437 seven artificial expression profiles, similar to presented in Figure S11. Each profile
438 corresponds to one of the zebrafish meta developmental stages. Next, the algorithm
439 will try to find genes with similar expression profiles to these artificial ones through
440 iterations until the processes converges. This method has proven to be very specific,
441 but lacks power with medium or small datasets (<30 time points). For mouse and fly,
442 the sample size is not enough, so we used the method introduced by Levin et al.
443 (2016). Firstly, we generated standardized gene expression for each gene
444 by subtracting its mean (across all stages) and dividing by its standard deviation. Next,
445 we calculated the first two principal components of each gene based on the
446 standardized expression across development. Since the expression was standardized,
447 the genes form a circle with scatter plot (Figure S9). Then, we computed the four-
448 quadrant inverse tangent for each gene based on its principal components, and sort
449 these values to get gene expression order from early to late (Figure S10). Next, we
450 performed Pearson correlation of the standardized expression and idealized
451 expression profile of each module (Figure S11). Finally, for each module, we defined
452 genes with correlation coefficient rank in top 10% as modular genes. Clearly, the
453 genes in earlier modules have higher gene orders (Figure S9).

454

455 **Randomization test of modularity analysis**

456 For each module, we randomly choose the same number of ΔlnL from all modular
457 genes (genes attributed to any module in that species) without replacement and
458 calculated the mean value. We repeated this 10000 times and approximated a normal
459 distribution for the mean value of ΔlnL. The *p*-value that the mean value of interested

460 module is higher (or lower) than the mean value from all modular genes is the
461 probability that the randomly sampled mean value of ΔlnL is higher (or lower) than
462 the original mean value of ΔlnL. In the same way, we also estimated the *p*-value of
463 the median ΔlnL value.

464 **Transcriptome index of log-likelihood ratio (TLI)**

465 The TLI is calculated as:

$$\mathrm{TLI}_s = \frac{\sum_{i=1}^{n} \sqrt[4]{\triangle \mathrm{lnL}_i}\, e_{is}}{\sum_{i=1}^{n} e_{is}},$$

466

467 s is the developmental stage, $\Delta\mathrm{lnL}_i$ is the value of log-likelihood ratio for gene i, n the
468 total number of genes and e_{is} is the log-transformed expression level of gene i in
469 developmental stage s. Here, we used all ΔlnL values without applying any cut-off on
470 ΔlnL or the associated *p*-value. For genes with ΔlnL<0, we replaced it with 0. For *M.*
471 *musculus*, we calculated the TLI from a merged data set, instead of computing it on
472 two data sets separately.

473 **Polynomial regression**

474 For polynomial regression analysis, we keep increasing the degree of polynomial
475 model until no further significant improvement (tested with ANOVA, $p<0.05$ as a
476 significant improvement). For *M. musculus*, since the development time points in
477 transcriptome data set are close to uniformly sampled, we used the natural scale of
478 development time for regression. For *C. elegans*, *D. melanogaster* and *D. rerio*,
479 however, we used the logarithmic scale, to limit the effect of post-embryonic time
480 points.

481 **Bootstrap approach for transcriptome index of ΔlnL (TLI) comparison between**
482 **developmental periods**

483 Firstly, we randomly sampled the same size of genes from original gene set (with
484 replacement) for 10,000 times. In each time, we calculated the TLI of each
485 development stage. Then, we calculated the mean TLI (mean TLI of all stages within
486 a period) for each developmental period (maternal stage, early development, middle
487 development, late development, and adult). Thus, each developmental period contains
488 10,000 mean TLI. Finally, we performed pairwise Wilcoxon test to test the
489 differences of mean TLI between developmental periods.

490 **Detection of polygenic selection**

491 We performed a gene set enrichment approach to detect polygenic signals of positive
492 selection on pathways (Ackermann and Strimmer 2009; Daub et al. 2013; Daub et al.
493 2017). For each pathway, we calculated its SUMSTAT score, which is the sum of
494 ΔlnL of all genes within this pathway. The ΔlnL values were fourth-root transformed.
495 This approach makes the distribution of non-zero ΔlnL approximate normal
496 distribution (Canal 2005; Roux et al. 2014; Daub et al. 2017). So, with fourth-root
497 transformation, we limit the risk that the significant pathways we found be due to a
498 few outlier genes with extremely high ΔlnL. The SUMSTAT score of a pathway is
499 calculated as:

$$\mathrm{SUMSTAT}_p = \sum_{i \in p} \sqrt[4]{\triangle \mathrm{lnL}_i},$$

500

501 where p represents a pathway, and $\Delta\mathrm{lnL}_i$ represents the value of log-likelihood ratio
502 for gene i within pathway p. *P*athways less than 10 ΔlnL values were excluded from
503 our analysis. Like in TLI analysis, we used all ΔlnL values and replaced <0 values
504 with 0.

505 **Empirical null distribution of SUMSTAT**

506 We used a randomization test to infer the significance of the SUSMTAT score of a
507 pathway. To correct for the potential bias caused by gene length, we firstly created
508 bins with genes that have similar length (Figure S12). Secondly, we randomly
509 sampled (without replacement) the same number of genes from each bin, to make the
510 total number of genes equal to the pathway being tested. Thirdly, we computed the
511 SUMSTAT score of the randomly sampled ΔlnL values. We repeated the second and
512 third processes one million times. Fourthly, we approximated a normal distribution
513 for SUMSTAT score of the interested pathway. Finally, the *p*-value was calculated as
514 the probability that the expected SUMSTAT score is higher than the observed
515 SUMSTAT score.

516 **Removing redundancy in overlapping pathways ("pruning")**

517 Because some pathways share high ΔlnL value genes, the identified significant
518 pathways might be partially redundant. In other words, shared genes among several
519 pathways can drive all these pathways to score significant. We therefore removed the
520 overlap between pathways with a "pruning" method (Daub et al. 2013; Daub et al.
521 2017). Firstly, we inferred the *p*-value of each pathway with the randomization test.

522 Secondly, we removed the genes of the most significant pathway from all the other
523 pathways. Thirdly, we ran the randomization test on these updated gene sets. Finally,
524 we repeated the second and third procedures until no pathways were left to be tested.
525 With this "pruning" method, the randomization tests are not independent and only the
526 high scoring pathways will remain, so we need to estimate the False Discovery Rate
527 (FDR) empirically. To achieve this, we applied the "pruning" method to pathways
528 with permuted ΔlnL scores and repeated it for 300 times. So, for each pathway, we
529 obtained one observed *p*-value (*p**) and 300 empirical *p*-values. The FDR was
530 calculated as follow:

531 $$\hat{FDR}(p^*) = \frac{\pi_0 \hat{V}(p^*)}{R(p^*)},$$

532 where π_0 represents the proportion of true null hypotheses, $\hat{V}(p^*)$ represents the
533 estimated number of rejected true null hypotheses and $R(p^*)$ represents the total
534 number of rejected hypotheses. For π_0, we conservatively set it equal to 1 as in Daub
535 et al. (2017). For $\hat{V}(p^*)$, in each permutation analysis, we firstly calculated the
536 proportion of *p*-value (from permutation analysis) ≤ *p**. Then, the value of $\hat{V}(p^*)$ was
537 estimated by the mean proportion of *p*-value (from permutation analysis) ≤ *p** for the
538 300 permutation tests. For $R(p^*)$, we defined it as the number of *p*-value *(from*
539 *original analysis)* ≤ *p**. For *q*-value, we determined it from the lowest estimated FDR
540 among all *p*-values (from original analysis) ≥ *p**.
541
542
543

Acknowledgements

We thank Sébastien Moretti for help with Selectome data retrieval, the Bgee team for help with expression data retrieval, Barbara Piasecka for help with programming, Josephine T. Daub for help with polygenic selection analysis, and Julien Roux, Elsa Guillot, Iakov Davydov and all members of the Robinson-Rechavi lab for helpful discussions. Part of the computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. JL is supported by Swiss National Science Foundation grant 31003A_153341 / 1.

AUTHOR CONTRIBUTIONS

JL and MRR designed the work. JL performed the data gathering and analysis. JL and MRR interpreted the results. JL wrote the first draft of the paper. JL and MRR finalized the paper.

Tables

Table 1

P-values of randomization test for modular analysis.

D. melanogaster	Early embryo	Middle embryo	Late embryo	Larva	Pupae	Adult
p-value	0.014	0.484	0.263	0.435	0.213	0.018

D. rerio	Cleavage/ Blastula	Gastrula	Segmentation	Pharyngula	Larva	Juvenile	Adult
p-value	0.166	0.238	0.448	0.005	0.273	0.003	0.066

M. musculus	Early embryo	Middle embryo	Late embryo
p-value	0.094	0.043	0.001

564 **Table 2**

565 Candidate pathways enriched with signal of positive selection. We reported all
566 pathways with *q-value* <0.2 after removing overlapping genes ("pruning") for *D.*
567 *rerio*, *D. melanogaster* and *M. musculus*.

Species	Rank	Pathway	Pathway size before /after pruning	*p* value before pruning	*q* value before pruning	*p* value after pruning	*q* value after pruning
D. rerio	**1**	Laminin interactions	12/12	2.00E-06	1.32E-03	2.00E-06	0.00E+00
	2	Phenylalanine metabolism	10/10	7.70E-05	1.51E-02	7.80E-05	8.78E-03
	3	Visual phototransduction	33/33	9.10E-05	1.51E-02	8.30E-05	8.78E-03
	4	Metabolism of carbohydrates	119/118	2.02E-04	2.23E-02	2.37E-04	2.97E-02
	5	Gamma carboxylation, hypusine formation and arylsulfatase activation	18/18	1.46E-03	8.05E-02	1.24E-03	1.16E-01
	6	Extracellular matrix organization	75/61	2.00E-05	6.62E-03	1.50E-03	1.16E-01
	7	Acyl chain remodelling of PE	10/10	5.12E-03	1.79E-01	3.79E-03	1.66E-01
	8	Base excision repair	24/24	4.94E-03	1.79E-01	3.82E-03	1.66E-01
	9	Aminoacyl-tRNA biosynthesis	30/30	6.23E-03	1.97E-01	3.93E-03	1.66E-01
	10	Phase II conjugation	37/30	1.08E-03	6.92E-02	4.00E-03	1.66E-01
D. melanogaster	**1**	Triglyceride Biosynthesis	59/59	7.60E-05	1.57E-02	7.60E-05	3.32E-02
	2	Glycosaminoglycan degradation	16/16	6.99E-04	4.70E-02	6.37E-04	9.44E-02
	3	Metabolism of porphyrins	12/12	1.57E-03	7.31E-02	1.47E-03	9.62E-02
	4	Detoxification of Reactive Oxygen Species	17/17	1.45E-03	7.31E-02	1.48E-03	9.62E-02
	5	Longevity regulating pathway	43/28	2.83E-02	3.12E-01	3.37E-03	1.54E-01
	6	ECM-receptor interaction	10/10	4.54E-03	1.66E-01	4.10E-03	1.54E-01
	7	Lysine degradation	25/15	1.84E-02	2.86E-01	5.15E-03	1.54E-01
	8	Metabolic pathways	813/767	3.04E-04	3.11E-02	5.23E-03	1.54E-01
	9	Glutathione metabolism	55/32	2.59E-02	3.12E-01	5.66E-03	1.54E-01
M. musculus	**1**	Pantothenate and CoA biosynthesis	16/16	9.10E-05	7.20E-02	9.10E-05	5.01E-02
	2	Mineralocorticoid biosynthesis	10/10	1.52E-04	7.20E-02	1.40E-04	5.01E-02
	3	Mitochondrial translation	72/72	2.91E-04	7.86E-02	2.73E-04	5.25E-02
	4	Cytokine-cytokine receptor interaction	100/100	3.41E-04	7.86E-02	2.78E-04	5.25E-02

568
569
570
571
572
573
574
575
576
577
578

579 **Table S1**

580 Multiple test corrected p-values (Benjamini–Hochberg method) of randomization test
581 for modular analysis.

582

D. melanogaster	Early embryo	Middle embryo	Late embryo	Larva	Pupae	Adult
p-value	0.056	0.484	0.336	0.478	0.336	0.058

D. rerio	Cleavage/ Blastula	Gastrula	Segmentation	Pharyngula	Larva	Juvenile	Adult
p-value	0.295	0.336	0.478	0.027	0.336	0.024	0.151

M. musculus	Early embryo	Middle embryo	Late embryo
p-value	0.188	0.115	0.016

583
584
585

586 ## Figures legends

587 **Figure 1: Variation of ΔlnL in different modules.**

588 For each module, dots are values of ΔlnL for individual genes and the black line is the
589 mean of ΔlnL. Red (respectively blue) dots indicate modules for which the mean of
590 ΔlnL is significantly (p<0.05) higher (respectively lower) than the mean of ΔlnL from
591 all modules. The green dashed line denotes the mean value of ΔlnL from all modular
592 genes.

593

594 **Figure 2: Transcriptome index of ΔlnL (TLI) across development.**

595 For sub-figure A, B and C:

596 Orange, blue, red, green and purple time points represent stages within the
597 developmental periods of maternal stage, early development, middle development,
598 late development, and adult, respectively. For the adult stage, the black solid circle
599 represents TLI from average expression between male and female; the purple solid
600 triangle and square represent TLI from only males or females, respectively.

601 For sub-figure D, E and F:

602 Comparison of the TLI (mean TLI of all stages within a period) between any two
603 different periods. Each period has 10,000 pseudo-TLIs which come from random
604 resampling with replacement.

605

606 **Figure 3: Expression in development for genes involved in pathways enriched**
607 **with signal of positive selection.**

608 Each solid circle represents the ratio of the median expression for genes involved in
609 pathways enriched with signal of positive selection to the median expression for genes
610 involved in pathways without signal of positive selection. Orange, blue, red, green
611 and purple time points represent stages within the developmental periods maternal
612 stage, early development, middle development, late development, and adult,
613 respectively. In adult samples, black solid circles represent ratios generated from
614 average expression of males and females; purple solid triangles and squares represent
615 ratios generated from only males or only females, respectively.

616

617

Figure S1: Expression profiles of different modules across development.

The bold black line represents median expression of modular genes, the two gray lines represent 25th and 75th quantiles of expression of modular genes respectively. The blue, red and green named stages represent early, middle and late modules respectively.

Figure S2: Proportion of genes with strong evidence of positive selection in each module.

The number of genes in each module is indicated below each box. The *p*-value from chi-square goodness of fit test is reported in the top-left corner of each graph.

Figure S3: Proportion of genes with weak evidence of positive selection in each module.

Legend as in Figure S2.

Figure S4: Spearman's correlation between gene properties and ΔlnL.

Spearman's correlation coefficient (rho) and adjusted *p*-value are indicated in the top-right corner of each graph. Loess regression lines are plotted in red.

Figure S5: Variation of gene length in different modules.

Legend as in Figure 1.

Figure S6: Transcriptome index of gene length across development.

Legend as in Figure 2.

Figure S7: Transcriptome index of ΔlnL (TLI) for non-immune genes.

Legend as in Figure 2.

Figure S8: Transcriptome index of ΔlnL (TLI) for non-testis genes in *M. musculus* and *D. melanogaster*.

Legend as in Figure 2.

Figure S9: Scatter plot of genes based on principal component analysis

Each dot represents one gene, grey dots represent genes not assigned to any modules, blue dots represent genes in early embryo module, red dots represent genes in middle embryo module, blue dots represent genes in late embryo module, pink dots represent genes in larva module, and purple dots represent genes in pupae/adult module. Arrow indicates the gene expression order (from early to late).

Figure S10: Heat map of gene expression across development

Genes arranged by the order of expression (from early to late). Generally, the earlier genes have higher expression in earlier stages. Color bar represents the standardized expression value.

652 **Figure S11: idealized expression profile for each module.**

653 **Figure S12: Correction of gene length for polygenic selection.**

654 The red lines indicate the boundaries of bins which contains genes with similar length.

References

Ackermann M, Strimmer K. 2009. A general modular framework for gene set enrichment analysis. BMC Bioinformatics 10:47.

Amundson R. 2007. The changing role of the embryo in evolutionary thought : roots of evo-devo. 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. 7:42.

Ballard WW. 1981. Morphogenetic Movements and Fate Maps of Vertebrates. Am. Zool. 21:391–399.

Berg JJ, Coop G. 2014. A Population Genetic Signal of Polygenic Adaptation. PLoS Genet. 10:e1004412.

Bergmann S, Ihmels J, Barkai N. 2003. Iterative signature algorithm for the analysis of large-scale gene expression data. Phys Rev E Stat Nonlin Soft Matter Phys 67:31902.

Brakefield PM. 2006. Evo-devo and constraints on selection. Trends Ecol. Evol. 21:362–368.

Burga A, Wang W, Ben-David E, Wolf PC, Ramey AM, Verdugo C, Lyons K, Parker PG, Kruglyak L. 2017. A genetic signature of the evolution of loss of flight in the Galapagos cormorant. Science 356:eaal3345.

Cai JJ, Petrov DA. 2010. Relaxed purifying selection and possibly high rate of adaptation in primate lineage-specific genes. Genome Biol. Evol. 2:393–409.

Canal L. 2005. A normal approximation for the chi-square distribution. Comput. Stat. Data Anal. 48:803–808.

Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, AmiGO Hub the A, Web Presence Working Group the WPW. 2009. AmiGO: online access to ontology and annotation data. Bioinformatics 25:288–289.

Carroll SB. 2008. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell 134:25–36.

Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A, et al. 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. 42:459–471.

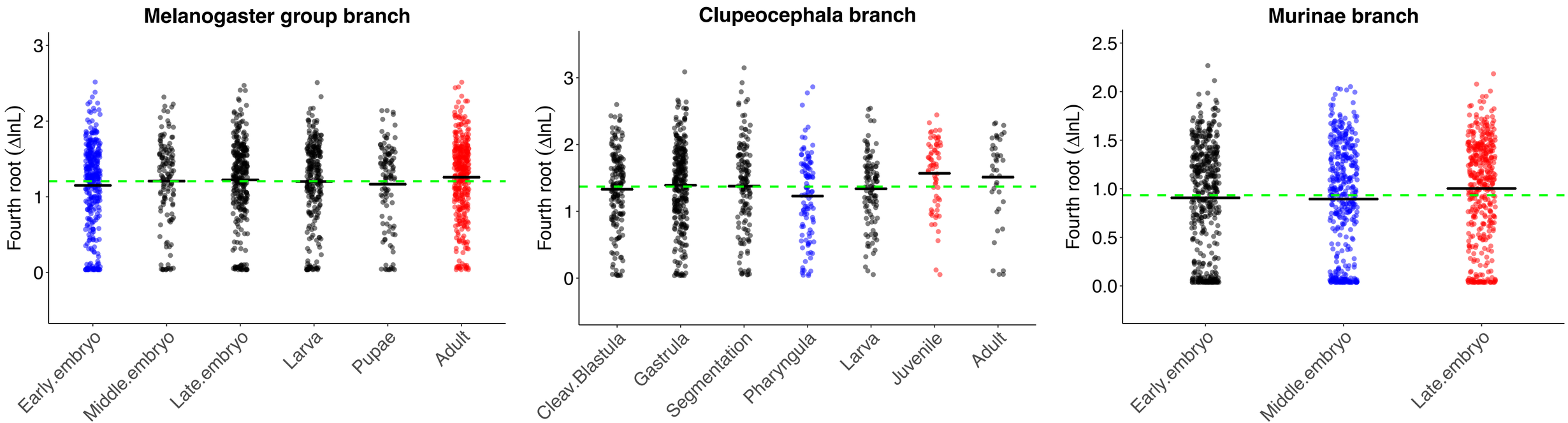
688 Castillo-Davis CI, Hartl DL. 2002. Genome evolution and developmental constraint
689 in Caenorhabditis elegans. Mol. Biol. Evol. 19:728–735.
690 Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P,
691 Gillespie M, Kamdar MR, et al. 2014. The Reactome pathway knowledgebase.
692 Nucleic Acids Res. 42:1–6.
693 Darwin C. 1871. The descent of man, and selection in relation to sex. Princeton
694 University Press
695 Daub JT, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L, Robinson-Rechavi M,
696 Excoffier L. 2013. Evidence for polygenic adaptation to pathogens in the human
697 genome. Mol. Biol. Evol. 30:1544–1558.
698 Daub JT, Moretti S, Davydov II, Excoffier L, Robinson-Rechavi M. 2017. Detection
699 of pathways affected by positive selection in primate lineages ancestral to
700 humans. Mol. Biol. Evol. 34:1391–1402.
701 Domazet-Loso T, Tautz D. 2010. A phylogenetically based transcriptome age index
702 mirrors ontogenetic divergence patterns. Nature 468:815–818.
703 Drost H-G, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of
704 phylotranscriptomic hourglass patterns in animal and plant embryogenesis. Mol.
705 Biol. Evol. 32:1221–1231.
706 Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the
707 stability of a vertebrate Bauplan and the evolution of morphologies through
708 heterochrony. Development 1994:135–142.
709 Flajnik MF, Kasahara M. 2010. Origin and evolution of the adaptive immune system:
710 genetic events and selective pressures. Nat. Rev. Genet. 11:47–59.
711 Fletcher W, Yang Z. 2010. The Effect of Insertions, Deletions, and Alignment Errors
712 on the Branch-Site Test of Positive Selection. Mol. Biol. Evol. 27:2257–2267.
713 Garstang W. 1922. The Theory of Recapitulation: A Critical Re-statement of the
714 Biogenetic Law. J. Linn. Soc. London, Zool. 35:81–101.
715 Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH.
716 2009. The NCBI BioSystems database. Nucleic Acids Res. 38:492–496.
717 Gharib WH, Robinson-Rechavi M. 2013. The Branch-Site Test of Positive Selection
718 Is Surprisingly Robust but Lacks Power under Synonymous Substitution
719 Saturation and Variation in GC. Mol. Biol. Evol. 30:1675–1686.
720 Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG,

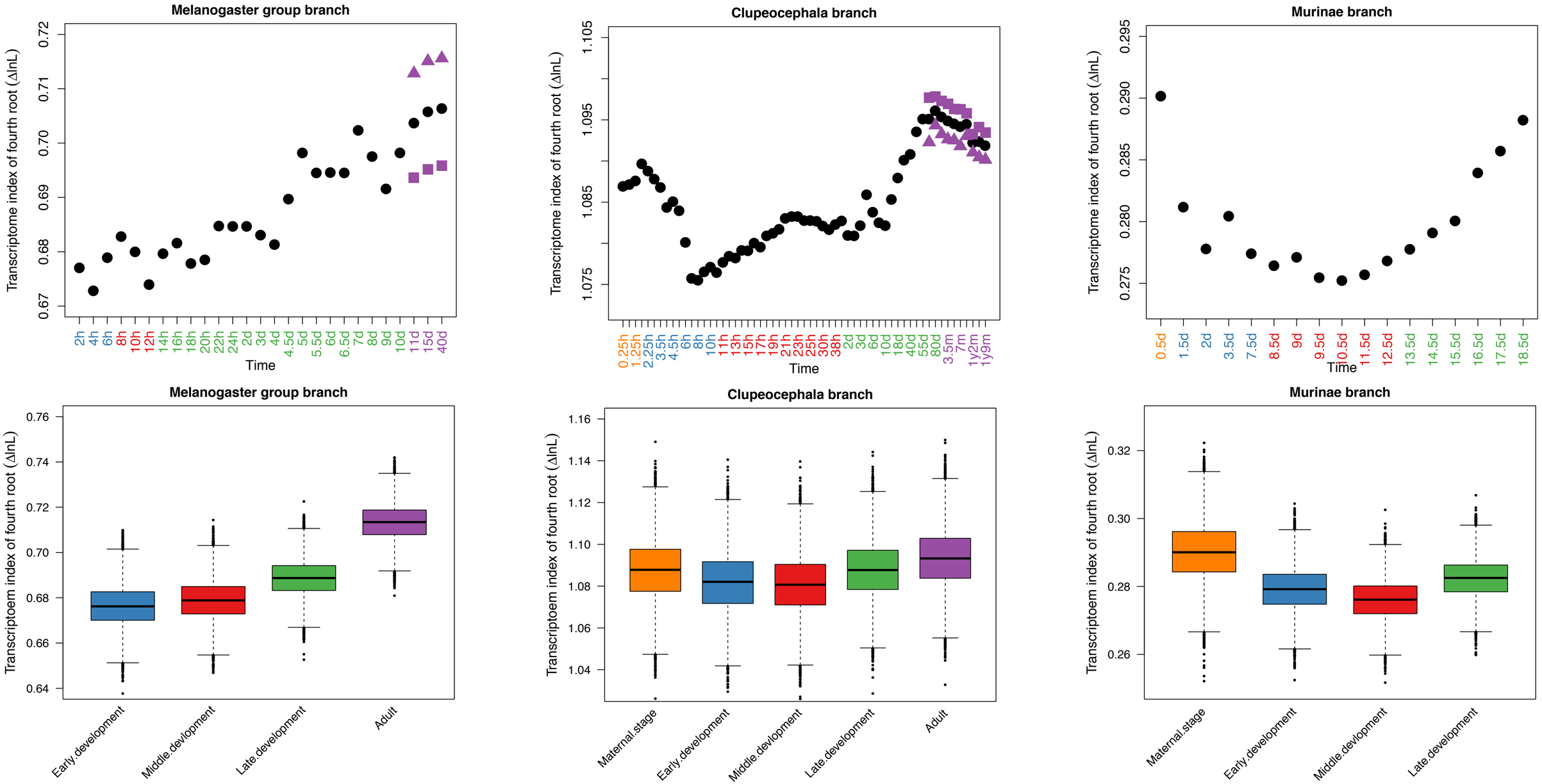
721 van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental
722 transcriptome of Drosophila melanogaster. Nature 471:473–479.
723 Hawkins DM, Wixley RAJ. 1986. A Note on the Transformation of Chi-Squared
724 Variables to Normality. Am. Stat. 40:296.
725 Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of
726 adaptation. Evolution 61:995–1016.
727 Hu H, Uesaka M, Guo S, Shimai K, Lu T-M, Li F, Fujimoto S, Ishikawa M, Liu S,
728 Sasagawa Y, et al. 2017. Constrained vertebrate evolution by pleiotropic genes.
729 Nat. Ecol. Evol. 1:1722–1730.
730 Hurst LD. 2002. The Ka/Ks ratio: Diagnosing the form of sequence evolution. Trends
731 Genet. 18:486–487.
732 Irie N, Kuratani S. 2011. Comparative transcriptome analysis reveals vertebrate
733 phylotypic period during organogenesis. Nat. Commun. 2:248.
734 Irie N, Kuratani S. 2014. The developmental hourglass model: a predictor of the basic
735 body plan? Development 141:4649–4655.
736 Kalinka AT, Tomancak P. 2012. The evolution of early animal embryos: conservation
737 or divergence? Trends Ecol. Evol. 27:385–393.
738 Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, Ohler U,
739 Bergman CM, Tomancak P. 2010. Gene expression divergence recapitulates the
740 developmental hourglass model. Nature 468:811–814.
741 Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data,
742 information, knowledge and principle: Back to metabolism in KEGG. Nucleic
743 Acids Res. 42:199–205.
744 Kelder T, Van Iersel MP, Hanspers K, Kutmon M, Conklin BR, Evelo CT, Pico AR.
745 2012. WikiPathways: Building research communities on biological pathways.
746 Nucleic Acids Res. 40:1301–1307.
747 Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudich G, Almeida-King J,
748 Staines D, Derwent P, Kerhornou A, et al. 2011. Ensembl BioMarts: a hub for
749 data retrieval across taxonomic space. Database 2011:bar030.
750 Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, Senderovich N, Kovalev
751 E, Silver DH, Feder M, et al. 2016. The mid-developmental transition and the
752 evolution of animal body plans. Nature 531:637–641.
753 Levin M, Hashimshony T, Wagner F, Yanai I. 2012. Developmental milestones

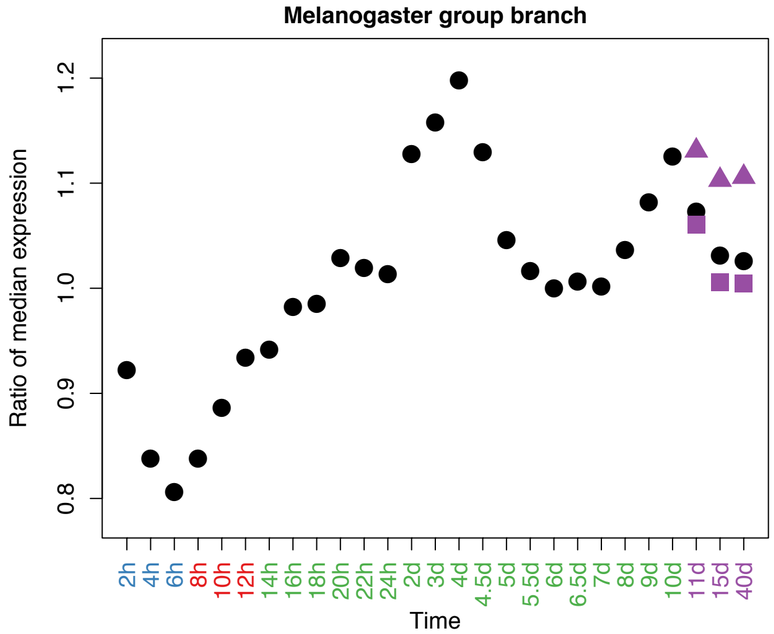
754 punctuate gene expression in the Caenorhabditis embryo. Dev. Cell 22:1101–
755 1108.
756 Li JJ, Huang H, Bickel PJ, Brenner SE. 2014. Comparison of D. melanogaster and C.
757 elegans developmental stages, tissues, and cells by modENCODE RNA-seq data.
758 Genome Res. 24:1086–1101.
759 Liu J, Robinson-Rechavi M. 2018. Developmental constraints on genome evolution in
760 four bilaterian model species. bioRxiv:doi.org/10.1101/161679.
761 Moretti S, Laurenczy B, Gharib WH, Castella B, Kuzniar A, Schabauer H, Studer RA,
762 Valle M, Salamin N, Stockinger H, et al. 2014. Selectome update: Quality
763 control and computational improvements to a database of positive selection.
764 Nucleic Acids Res. 42:917–921.
765 Müller GB. 2007. Evo–devo: extending the evolutionary synthesis. Nat. Rev. Genet.
766 8:943–949.
767 Necsulea A, Kaessmann H. 2014. Evolutionary dynamics of coding and non-coding
768 transcriptomes. Nat. Rev. Genet. 15:734–748.
769 Piasecka B, Lichocki P, Moretti S, Bergmann S, Robinson-Rechavi M. 2013. The
770 Hourglass and the Early Conservation Models—Co-Existing Patterns of
771 Developmental Constraints in Vertebrates. PLoS Genet. 9:e1003476.
772 Quint M, Drost H-G, Gabel A, Ullrich KK, Bönn M, Grosse I. 2012. A transcriptomic
773 hourglass in plant embryogenesis. Nature 490:98–101.
774 Raff RA. 1996. The shape of life : genes, development, and the evolution of animal
775 form. University of Chicago Press.
776 Raff RA. 2000. Evo-devo: the evolution of a new discipline. Nat. Rev. Genet. 1:74–
777 79.
778 Richardson MK. 1999. Vertebrate evolution: the developmental origins of adult
779 variation. Bioessays 21:604–613.
780 Riedl R. 1978. Order in living organisms. West Sussex, UK: Wiley-Interscience.
781 Roux J, Privman E, Moretti S, Daub JT, Robinson-Rechavi M, Keller L. 2014.
782 Patterns of positive selection in seven ant genomes. Mol. Biol. Evol. 31:1661–
783 1685.
784 Roux J, Robinson-Rechavi M. 2008. Developmental constraints on vertebrate genome
785 evolution. PLoS Genet. 4:e1000311.
786 Sabbah S, Laria R, Gray SM, Hawryshyn CW. 2010. Functional diversity in the color

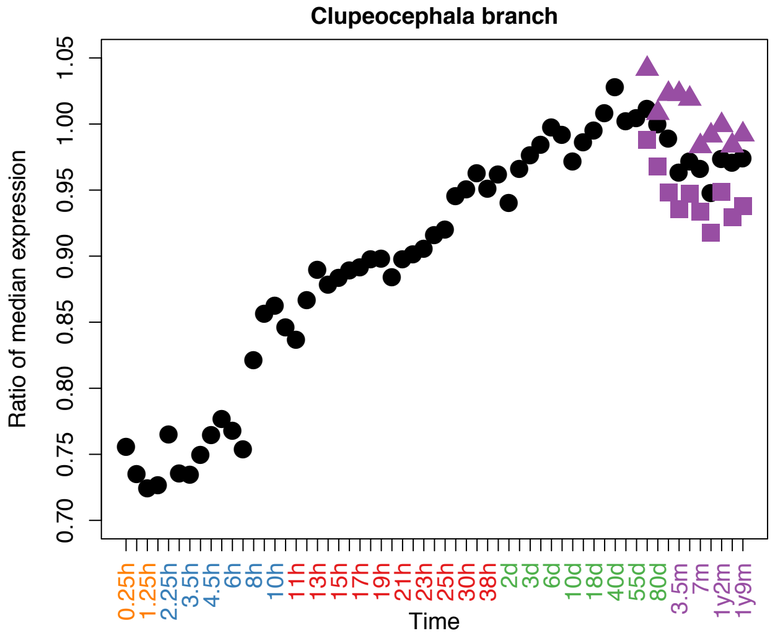
787 vision of cichlid fishes. BMC Biol. 8:133.
788 Salvador-Martínez I, Coronado-Zamora M, Castellano D, Barbadilla A, Salazar-
789 Ciudad I. 2018. Mapping Selection within Drosophila melanogaster Embryo's
790 Anatomy. Mol. Biol. Evol. 35:66–79.
791 Sander K. 1983. Development and evolution: the sixth Symposium of the British
792 Society for Developmental Biology. Cambridge University Press.
793 Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, Buetow KH. 2009.
794 PID: The pathway interaction database. Nucleic Acids Res. 37:674–679.
795 Slack JMW, Holland PWH, Graham CF. 1993. The zootype and the phylotypic stage.
796 Nature 361:490–492.
797 Tadros W, Lipshitz HD. 2009. The maternal-to-zygotic transition: a play in two acts.
798 Development 136:3033–3042.
799 Tickle C, Urrutia AO. 2017. Perspectives on the history of evo-devo and the
800 contemporary research landscape in the genomics era. Philos. Trans. R. Soc.
801 Lond. B. Biol. Sci. 372:20150473.
802 Tintle NL, Borchers B, Brown M, Bekmetjev A. 2009. Comparing gene set analysis
803 methods on single-nucleotide polymorphism data from Genetic Analysis
804 Workshop 16. BMC Proc. 3 (Suppl 7):S96.
805 Uchida Y, Uesaka M, Yamamoto T, Takeda H, Irie N. 2018. Embryonic lethality is
806 not sufficient to explain hourglass-like conservation of vertebrate embryos.
807 Evodevo 9:7.
808 Vitti JJ, Grossman SR, Sabeti PC. 2013. Detecting natural selection in genomic data.
809 Annu. Rev. Genet. 47:97–120.
810 Von-Baer KE. 1828. Über Entwickelungsgeschichte der Tiere: Beobachtung und
811 Reflexion. Königsberg: Gebrüder Bornträger.
812 Wolpert L. 1991. The Triumph of the Embryo. Oxford University Press.
813 Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev.
814 Genet. 8:206–216.
815 Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear
816 genes of mammals. J. Mol. Evol. 46:409–418.
817 Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular
818 adaptation at individual sites along specific lineages. Mol. Biol. Evol. 19:908–
819 917.

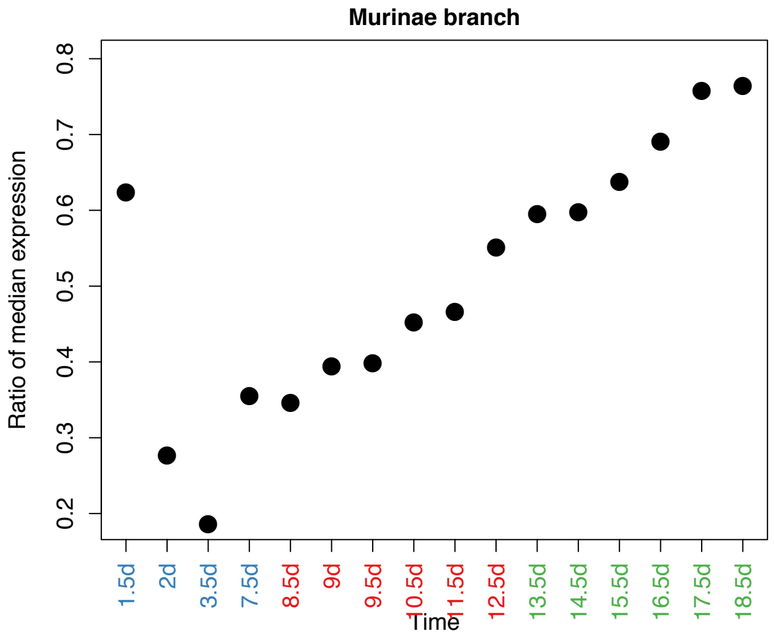
820 Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C,
821 Clapham P, Fitzgerald S, Gil L, et al. 2016. Ensembl 2016. Nucleic Acids Res.
822 44:D710–D716.
823 Zalts H, Yanai I. 2017. Developmental constraints shape the evolution of the
824 nematode mid-developmental transition. Nat. Ecol. Evol. 1:0113.
825 Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood
826 method for detecting positive selection at the molecular level. Mol. Biol. Evol.
827 22:2472–2479.
828


Melanogaster group branch
Fourth root (ΔlnL)
3
2
1
0
Early.embryo
Middle.embryo
Late.embryo
Larva
Pupae
Adult
Clupeocephala branch
Fourth root (ΔlnL)
3
2
1
0
Cleav.Blastula
Gastrula
Segmentation
Pharyngula
Larva
Juvenile
Adult
Murinae branch
Fourth root (ΔlnL)
2.5
2.0
1.5
1.0
0.5
0.0
Early.embryo
Middle.embryo
Late.embryo


Melanogaster group branch
Transcriptome index of fourth root (ΔlnL)
0.67
0.68
0.69
0.70
0.71
0.72
2h 4h 6h 8h 10h 12h 14h 16h 18h 20h 22h 24h 2d 3d 4d 4.5d 5d 5.5d 6d 6.5d 7d 8d 9d 10d 11d 15d 40d
Time
Clupeocephala branch
Transcriptome index of fourth root (ΔlnL)
1.075
1.085
1.095
1.105
0.25h 1.25h 2.25h 3.5h 4.5h 6h 8h 10h 11h 13h 15h 17h 19h 21h 23h 25h 30h 38h 2d 3d 6d 10d 18d 40d 55d 80d 3.5m 7m 1y2m 1y9m
Time
Murinae branch
Transcriptome index of fourth root (ΔlnL)
0.275
0.280
0.285
0.290
0.295
0.5d 1.5d 2d 3.5d 7.5d 8.5d 9d 9.5d 10.5d 11.5d 12.5d 13.5d 14.5d 15.5d 16.5d 17.5d 18.5d
Time
Melanogaster group branch
Transcriptoem index of fourth root (ΔlnL)
0.64
0.66
0.68
0.70
0.72
0.74
0.76
Early.development
Middle.devlopment
Late.development
Adult
Clupeocephala branch
Transcriptoem index of fourth root (ΔlnL)
1.04
1.06
1.08
1.10
1.12
1.14
1.16
Maternal.stage
Early.development
Middle.devlopment
Late.development
Adult
Murinae branch
Transcriptoem index of fourth root (ΔlnL)
0.26
0.28
0.30
0.32
Maternal.stage
Early.development
Middle.devlopment
Late.development


Melanogaster group branch
Ratio of median expression
0.8
0.9
1.0
1.1
1.2
2h
4h
6h
8h
10h
12h
14h
16h
18h
20h
22h
24h
2d
3d
4d
4.5d
5d
5.5d
6d
6.5d
7d
8d
9d
10d
11d
15d
40d
Time


Clupeocephala branch
Ratio of median expression
0.70
0.75
0.80
0.85
0.90
0.95
1.00
1.05
0.25h
1.25h
2.25h
3.5h
4.5h
6h
8h
10h
11h
13h
15h
17h
19h
21h
23h
25h
30h
38h
2d
3d
6d
10d
18d
40d
55d
80d
3.5m
7m
1y2m
1y9m
Time


Murinae branch
Ratio of median expression
0.2
0.3
0.4
0.5
0.6
0.7
0.8
1.5d
2d
3.5d
7.5d
8.5d
9d
9.5d
10.5d
11.5d
12.5d
13.5d
14.5d
15.5d
16.5d
17.5d
18.5d
Time