



RESEARCH ARTICLE

**REVISED** An explanatory evo-devo model for the developmental hourglass [version 2; referees: 3 approved]

Saamer Akhshabi<sup>1</sup>, Shrutii Sarda<sup>2</sup>, Constantine Dovrolis<sup>1</sup>, Soojin Yi<sup>3</sup>

<sup>1</sup>School of Computer Science, Georgia Institute of Technology, Atlanta, GA, 30332, USA  
<sup>2</sup>Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD, 20740, USA  
<sup>3</sup>School of Biology, Georgia Institute of Technology, Atlanta, GA, 30332, USA

**v2** First published: 08 Jul 2014, 3:156 (doi: 10.12688/f1000research.4583.1)  
 Latest published: 18 Dec 2014, 3:156 (doi: 10.12688/f1000research.4583.2)

**Abstract**

The "developmental hourglass" describes a pattern of increasing morphological divergence towards earlier and later embryonic development, separated by a period of significant conservation across distant species (the "phylotypic stage"). Recent studies have found evidence in support of the hourglass effect at the genomic level. For instance, the phylotypic stage expresses the oldest and most conserved transcriptomes. However, the regulatory mechanism that causes the hourglass pattern remains an open question. Here, we use an evolutionary model of regulatory gene interactions during development to identify the conditions under which the hourglass effect can emerge in a general setting. The model focuses on the hierarchical gene regulatory network that controls the developmental process, and on the evolution of a population under random perturbations in the structure of that network. The model predicts, under fairly general assumptions, the emergence of an hourglass pattern in the structure of a temporal representation of the underlying gene regulatory network. The evolutionary age of the corresponding genes also follows an hourglass pattern, with the oldest genes concentrated at the hourglass waist. The key behind the hourglass effect is that developmental regulators should have an increasingly specific function as development progresses. Analysis of developmental gene expression profiles from *Drosophila melanogaster* and *Arabidopsis thaliana* provide consistent results with our theoretical predictions.

**Open Peer Review**

Referee Status:

	Invited Referees		
	1	2	3
<b>REVISED</b>			
<b>version 2</b>	report		report
published 18 Dec 2014			
<b>version 1</b>			
published 08 Jul 2014	report	report	report

- 1 **Marcel Quint**, Leibniz Institute of Plant Biochemistry, Germany  
**Hajk-Georg Drost**, Martin Luther University Halle–Wittenberg, Germany
- 2 **Dmitri Krioukov**, Northeastern University, USA
- 3 **Gourab Ghoshal**, Harvard University, USA

**Discuss this article**

Comments (0)

**Corresponding authors:** Constantine Dovrolis ([constantine@gatech.edu](mailto:constantine@gatech.edu)), Soojin Yi ([soojinyi@gatech.edu](mailto:soojinyi@gatech.edu))

**Competing interests:** No competing interests were disclosed.

**How to cite this article:** Akhshabi S, Sarda S, Dovrolis C and Yi S. **An explanatory evo-devo model for the developmental hourglass [version 2; referees: 3 approved]** *F1000Research* 2014, **3**:156 (doi: [10.12688/f1000research.4583.2](https://doi.org/10.12688/f1000research.4583.2))

**Copyright:** © 2014 Akhshabi S *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

**Grant information:** This work was supported by the National Science Foundation (Grant No. 0831848 to C.D.). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

**First published:** 08 Jul 2014, **3**:156 (doi: [10.12688/f1000research.4583.1](https://doi.org/10.12688/f1000research.4583.1))

**REVISED** Amendments from Version 1

The main changes in Version-2 of this paper are:

- New text in the Introduction and Model Description sections, providing connections between modularity (see review by G.P.Wagner *et al.*) and our assumption of increasing specificity as development progresses.
- Included additional context and insight in the Introduction about the results shown later in the paper.
- Clarified several points in the Model Description, especially about how we model the specificity function and the probability of regulatory failure.
- A new paragraph in the Discussion section, focusing on Raff's hypothesis in the context of this work.
- Provided a connection between this work (the computational results about lethality probability) and the work of Galis and Metz about perturbations in the phylotypic stage of rodents.
- Added several new references.
- Re-formatted all equations that were previously embedded in the main text ("inline") as separate lines.
- The quantitative results of the paper (simulations and data analysis) have not been modified in any way.

See referee reports

## Introduction

The evolutionary mechanism of conservation during embryogenesis, and its connection to the gene regulatory networks that control development, are fundamental questions in systems biology<sup>1-3</sup>. Several models have been presented in the context of morphological, molecular, and genetic developmental patterns. The most widely discussed model is the "developmental hourglass", which places the strongest conservation across species in the "phylotypic stage". The first observations supporting the hourglass model go back to von Baer when he noticed that there exists a mid-developmental stage in which embryos of different animals look similar<sup>4</sup>. On the other hand, the "developmental funnel" model of conservation predicts increasing diversification as development progresses<sup>5,6</sup>.

Recently, the hourglass model has come under new light. Multiple studies have observed the hourglass pattern across diverse biological processes, including transcriptome divergence<sup>7-12</sup>, transcriptome age<sup>7,13,14</sup>, molecular interaction<sup>15</sup>, and evolutionary selective constraints<sup>10,15,16</sup>. Despite these observations the genomic basis and even the existence of the developmental hourglass effect have been the subject of an intense debate<sup>1,6,14,17-22</sup>. More importantly, the underlying mechanism that can shape the developmental process in the hourglass or funnel forms is still unknown.

We aim to understand the conditions under which the hourglass effect can emerge in a general setting, based on an abstract model for the evolution of embryonic development. The model focuses on a hierarchical network that represents the temporal "execution" of the underlying Gene Regulatory Network (GRN) during development. Each layer of the network corresponds to a developmental stage. The nodes at each layer represent regulatory genes (i.e., genes encoding transcription factors or signaling molecules) that undergo significant activity change at that corresponding stage. The edges from genes at one layer to genes at the next layer represent regulatory

interactions that cause those activity changes. We refer to this hierarchical network as Developmental Gene Execution Network (DGEN) to distinguish it from the corresponding GRN. A DGEN is subject to evolutionary perturbations (e.g., gene deletions, rewiring, duplication) that may be lethal, or that may impede development, for the corresponding organism.

The model predicts that the evolutionary process shapes the DGENs of a population in the form of an hourglass, under fairly general assumptions. Specifically, the number of genes at each developmental stage follows an hourglass pattern, with the smallest number at the "waist" of the hourglass. The main condition for the appearance of the hourglass pattern is that the DGEN should gradually get sparser as development progresses, with general-purpose regulatory genes at the earlier developmental stages and highly specialized regulatory genes at the later stages. This assumption is motivated from the well established patterns of increasing modularity as the embryo develops<sup>23</sup>. Under the previous condition, the model predicts that gene regulatory changes or rewiring in mid-development are more likely to cause cascades of removing non-essential genes from the DGEN, compared to early or late developmental stages. Another model prediction is that the evolutionary age of DGEN genes also follows an hourglass pattern, with the oldest genes concentrated at the waist.

We have examined the aforementioned predictions using transcriptome data from the development of *Drosophila melanogaster* and *Arabidopsis thaliana*. This data is insufficient to reconstruct the complete DGEN of these species but it allows to estimate the number of genes at each developmental stage, given an activity variation threshold. Under a wide range of this threshold, the inferred DGEN shape follows an hourglass pattern, the waist of that hourglass roughly coincides with the previously reported phylotypic stage for these species, and the age of the corresponding genes follows the predicted hourglass pattern.

## Developmental gene execution networks

As a first-order approximation, a regulatory gene can be modeled in one of several discrete *functional states*<sup>24</sup>. In the simplest case, a regulatory gene can act as a binary switch ("on" and "off") but in general a gene may have more than two functional states. The transition of a regulatory gene X from one functional state to another is often (but not always) caused by one or more upstream regulators of X that go through a functional state transition before X. We use the term *transitioning gene* to refer to a regulatory gene that goes through a functional state transition at a given developmental time anywhere at the developing embryo.

A DGEN is a directed and acyclic network; see [Figure 1a](#) for an abstract example. The vertical direction refers to developmental time, from the zygote at the top to the developed organism at the bottom. In the horizontal direction we can represent different spatial domains, even though this is not necessary and it is not done in our model. For instance, the zygote at the top of the DGEN would be a single domain, while the organism at the bottom of the DGEN would have the largest number of spatial domains.

Development is often approximated (conceptually and experimentally) as a succession of discrete developmental stages. The duration

of a developmental stage can be thought of as the typical time that is required for a gene's functional state transition, and it does not need to be the same for all stages. Each layer of a DGEN refers to a developmental stage, and it includes only the transitioning genes during that stage anywhere in the embryo. The same gene can appear in more than one stage if it goes through several functional state transitions during development. Additionally, a DGEN edge from a gene  $X$  at stage  $l$  to a gene  $Y$  at stage  $l + 1$  implies that the functional transition of  $X$  caused the functional transition of  $Y$  at the next stage. If gene  $Y$  has more than one incoming edge, its functional state transition was caused by the coupled effect of more than one transitioning genes at the earlier stage. Any upstream regulators of  $Y$  that remained at the same functional state at stage  $l$  are not included in that stage of the DGEN.

The sequence of developmental stages is denoted by  $l = 1 \dots L$ . The set of transitioning genes at stage  $l$  is  $G(l)$ . A gene  $g$  at stage  $l < L$  regulates a set of *downstream* genes at stage  $l+1$  denoted by  $D(g)$  (outgoing edges from  $g$ ). Similarly, a gene  $g$  at stage  $l > 1$  is regulated by a set of *upstream* genes  $U(g)$  at stage  $l-1$  (incoming edges to  $g$ ). The functional transitions at the first stage are assumed to be caused by regulatory maternal genes that initiate the developmental process.

### Model description

The model captures certain aspects of both the developmental process, in the form of a given DGEN for each embryo, and of the evolutionary process, as random perturbations in the structure of individual DGENs in a population. The model does not need to capture the actual functional state transitions or the regulatory input function of each gene. It does capture however the dynamic and stochastic effect of structural network perturbations (gene deletion, rewiring and duplication) on the success of the developmental process, as explained in the following.

Similar to the Wright-Fisher model<sup>25</sup>, we consider a population of  $N$  individuals, each represented by a DGEN. In each generation, individuals reproduce asexually, inheriting the DGEN of their parent. Various evolutionary events can cause structural changes in the DGEN of an individual that may result in "developmental failure". Such individuals (and their DGENs) are replaced with developmentally successful individuals so that the population size remains constant.

The model is meant to be as general as possible and so the regulatory interactions between genes of successive stages are determined probabilistically, as follows. Each stage  $l$  is assigned a *regulatory specificity*, or simply *specificity*  $s(l)$  with  $0 \leq s(l) \leq 1$ . A gene  $g$  at stage  $l$  acts as upstream regulator for a gene  $g'$  at stage  $l + 1$  with probability  $s'(l) = 1 - s(l)$ . So, the specificity of a developmental stage determines how likely it is for regulatory genes of that stage to cause a state transition of the genes at the next stage; a higher specificity decreases that likelihood.

Our major assumption is that the regulatory specificity increases substantially as development progresses. In other words, the DGEN becomes gradually sparser along the developmental time axis, with  $s(1) \ll s(L)$ . This assumption is plausible for the following reasons.

First, as development progresses the embryo grows in size forming distinct spatial domains. So, extracellular gene regulation becomes more difficult, especially across different domains. Additionally, as development progresses the transitioning genes become more organ- or tissue-specific, implying that their downstream interactions become sparser. Unfortunately, a direct and empirical investigation of the increasing specificity assumption requires knowledge of the complete DGEN for a given species; this is currently not feasible for even the most well-studied model organisms. However, this assumption is plausible if we consider the well established patterns of increasing modularity as the embryo develops at the morphological, signaling, and genomic levels<sup>23</sup>.

The DGEN structural changes we consider are gene deletions, gene duplications, and gene rewiring:

**Deletions (DL):** This event removes a gene from the DGEN, including its incoming and outgoing edges. There are many genetic mechanisms that may cause such events. A DL event deletes each gene of an individual and at each generation with probability  $P_{DL}$ .

**Duplication (DP):** This event creates an identical copy of a gene  $g$  with the same downstream and upstream regulators and at the same developmental stage as  $g$ . The two genes may have different fates if one of them is subject later to deletion or rewiring. Otherwise, the two genes are considered identical. A DP event duplicates each gene of an individual and at each generation with probability  $P_{DP}$ .

**Rewiring (RW):** This event changes the upstream and/or downstream regulators of a gene. Changes in the upstream versus downstream regulators may have different biological basis. The former occur, for instance, as a result of mutations in the transcription factor binding sites in a gene's promoter or mutations in distal regulatory elements such as enhancers, while the latter may be mostly caused by coding sequence mutations. The details of the rewiring process do not affect the results qualitatively as long as the average density of edges in each stage remains consistent with the specificity of that stage. The specific rewiring mechanisms we use are presented next.

Suppose that a RW event affects gene  $g$  at stage  $l$ . The upstream regulators of  $g$  are recomputed based on the specificity of the previous stage, i.e., by choosing each distinct gene of stage  $l - 1$  with probability  $s'(l - 1)$ . For the downstream regulators of  $g$ , we randomly remove  $N_-$  existing outgoing edges of  $g$ , and then add  $N_+$  outgoing edges to randomly chosen genes of stage  $l + 1$  that  $g$  is not already connected to. If  $B(N, p)$  denotes a Binomial random variable with  $N$  trials and success probability  $p$ , the random variables  $N_-$  and  $N_+$  are independent and they both follow the  $B(|D(g)|, s'(l))$  distribution ( $|X|$  denotes the cardinality of set  $X$ ). Thus, the number of downstream edges of gene  $g$  after a Rewiring event becomes:

$$|D_{new}(g)| = |D(g)| - N_- + N_+,$$

which is at least 0 and at most  $2 \times |D(g)|$ . This captures that the downstream regulators of  $g$  are derived by incremental changes in  $D(g)$ , instead of giving  $g$  a completely new network configuration (thereby, new regulatory function). The higher the regulatory

specificity of a stage, the less likely these incremental changes are. An RW event rewires each gene of an individual and at each generation with probability  $P_{RW}$ .

A gene deletion or rewiring event at stage  $l$  can remove an upstream regulator from genes at stage  $l + 1$ . A loss of incoming edges may trigger the *regulatory failure* of a gene, as described next.

**Regulatory failures (RF):** A gene  $g$  may not be able to change functional state if some of its upstream regulators  $U(g)$  are lost due to DL or RW events (see Figure 1b). Even though regulatory networks are often robust to structural perturbations, even a partial gene loss in  $U(g)$  may disable  $g$  causing a *regulatory failure*. It is plausible that the probability of a regulatory failure increases with the fraction of lost upstream regulators. So, if  $U'(g)$  is the new set of upstream regulators and  $|U(g)| > |U'(g)| > 0$ , gene  $g$  is removed with probability:

$$P_{RF}(r) = 1 - e^{-zr}, \quad 0 < r = 1 - \frac{|U'(g)|}{|U(g)|} < 1$$

while if  $|U'(g)| = 0$  we set  $P_{RF}(1) = 1$ .  $z$  is the *RF parameter* and it depends on the robustness of regulatory interactions to gene loss (see Figure 2).

When a DL or RW event causes one or more RF events, the latter can trigger additional RF events in subsequent developmental stages, leading to *cascades of regulatory failures*. Such RF cascades may cause *developmental failure*, meaning that the developed embryo is unable to survive or reproduce.

**Developmental failure (DF):** The last stage of a DGEN represents the fully developed embryo. If that stage includes  $\Gamma$  transitioning genes at a successfully developed embryo, the simplest assumption is that an individual with less than  $\Gamma$  genes at stage- $L$  has failed to develop properly. Such DGENs are removed from the population and they are replaced with randomly chosen but successfully developed DGENs. We have also experimented with two variations of the DF event: first, the individual is removed if its last stage has less than  $\Gamma - \gamma$  genes, where  $\gamma$  is small relative to  $\Gamma$ , and second, the probability of a DF event increases as the number of genes at

stage- $L$  decreases below  $\Gamma$ . The qualitative results, as described next, do not change with these two model variations.

**Methods**

**Hourglass score H.** Suppose that  $w(l)$  denotes the number of transitioning genes in stage  $l$ , also referred to as the width of stage  $l$ . Let  $b$  be the stage with the minimum number of such genes. We construct the sequences  $X = \{w(l), l = 1, \dots, b\}$  and  $Y = \{w(l), l = b, \dots, L\}$ .  $\tau_X$  and  $\tau_Y$  denote the normalized univariate Mann-Kendall statistic for monotonic trend in each sequence, respectively ( $\tau$  is -1 for decreasing, +1 for increasing and almost 0 for random sequences)<sup>26</sup>. The H score is defined as

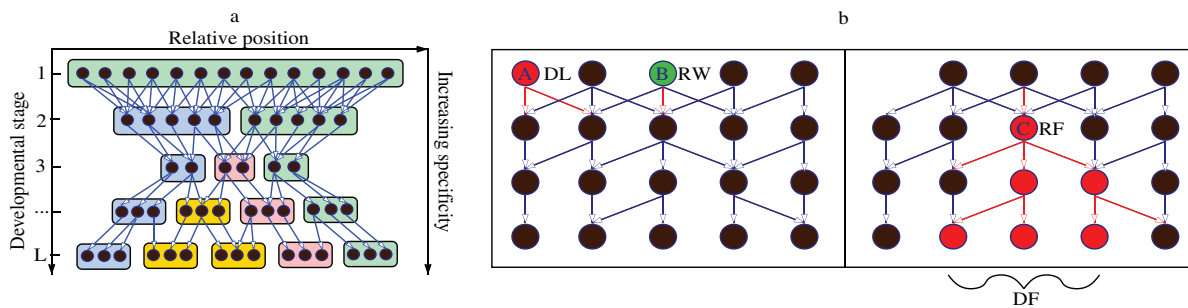
$$H = \frac{\tau_Y - \tau_X}{2}$$

See Figure 3 for an illustration, and for the definition of a more robust version of H.

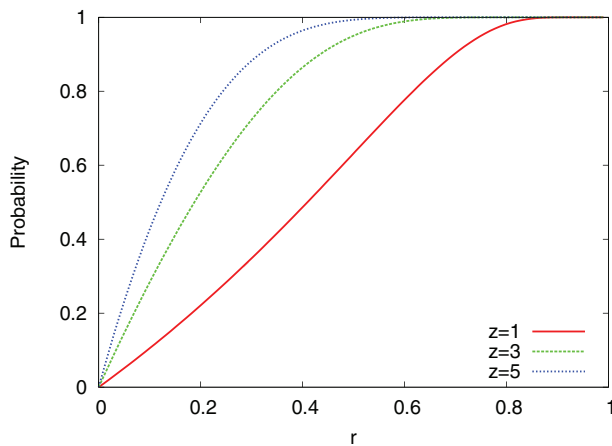
**Statistical analyses.** All biological data processing and analyses were performed using custom scripts written in Java (JDK v1.6) [Dataset 14].

**Drosophila data and treatment.** Developmental gene expression profiles for *D. melanogaster* are obtained from two sources. First, we obtained microarray data from Kalinka *et al.*<sup>9</sup> for 3,610 genes. The expression level of each gene is calculated as the median of probes mapping to that gene. Each stage represents a 2-hr interval during the first 20 hours of embryogenesis (10 stages). The second source is RNA-Seq data from Graveley *et al.*<sup>27</sup>. Raw data are processed to RPKM values. Each stage represents a 2-hr interval during the first 24 hours (12 stages). Genes with zero RPKM value in all stages are discarded, resulting in 14,110 genes.

**Arabidopsis data and treatment.** Genome-wide expression profiles of a complete developmental series from the zygote to the mature embryo in *A. thaliana* were obtained from Xiang *et al.*<sup>28</sup>. This comprised of microarray expression levels for 25,207 genes across seven developmental stages. Signal background correction and normalization of raw expression values was carried out by Xiang *et al.*<sup>28</sup>.



**Figure 1. An abstract DGEN. (A)** The circles denote state-transitioning genes, edges represent directed regulatory interactions, and colored boxes refer to spatial domains that form during development. If regulatory genes become increasingly function-specific as development progresses, the network gradually becomes sparser in that direction. **(B)** Evolutionary perturbations on a DGEN's structure: Gene A is deleted (DL), while gene B is rewired (RW) losing an outgoing edge. This RW event causes the regulatory failure (RF) of gene C, which then causes a cascade of five more RF events. This cascade causes developmental failure (DF). Note that the removal of some upstream regulators does not always cause an RF event (e.g., genes regulated by A).



**Figure 2. Probability of Regulatory Failure (RF) for three values of the parameter  $z$ .**  $r$  is the fraction of upstream regulating edges that are lost due to a DL or RW event.

**Transitioning gene identification.** Suppose that the reported expression value of gene  $i$  at stage  $l$  is  $e_{i,l}$ . We analyze both these “absolute” expression values as well as the normalized expression values, given by

$$e'_{i,l} = \frac{e_{i,l}}{\sum_j e_{j,l}}.$$

The identification of transitioning genes follows the same method for both absolute and normalized expression levels. In the case of absolute expressions we define  $\delta_{i,l} = e_{i,l} - e_{i,l-1}$  for each gene  $i$  and at each stage  $l = 2 \dots L$ , while in the case of normalized expressions, we similarly define  $\delta_{i,l} = e'_{i,l} - e'_{i,l-1}$ . Gene  $i$  is considered “transitioning” at the stage-pair  $(l-1, l)$  if

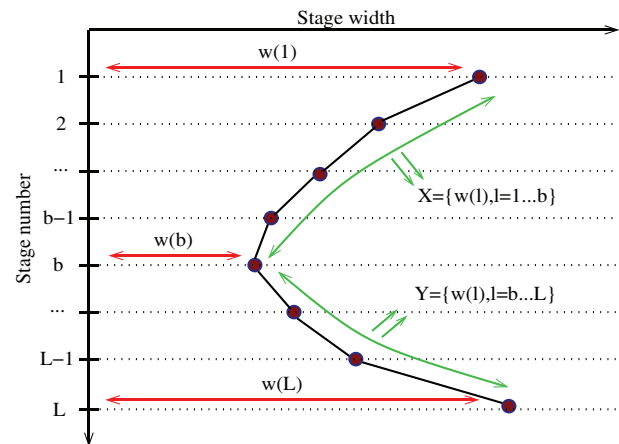
$$|\delta_{i,l}| > c,$$

where  $c$  is a given transition threshold. This condition is more robust to noise than a ratio-based rule ( $e'_{i,l}/e'_{i,l-1}$ ) for the identification of transitioning genes. Note that a gene may be transitioning at more than one stage-pair, but it may also not be transitioning at any stage-pair.

**Transcriptome age index (TAI).** We collected the groups of orthologs for each gene in *Drosophila* using two databases, OrthoDB<sup>29</sup> and OrthoMCL<sup>30</sup>. The Eumetazoa data were taken from OrthoDB, while Fungi and Plants species were retrieved from OrthoMCL, and the two datasets were merged. Using these orthologs we then assigned an age index to each gene based on its absence and presence in a phylogenetic tree of 24 well-diverged species (see Figure 4)<sup>11,13,31</sup> [Dataset 7].

The transcriptome age index (TAI) values for *Arabidopsis* genes were obtained from<sup>11</sup>.

**Age index for each stage-pair.** Suppose that we identify transitioning genes based on the normalized expression levels, and that  $n(l)$



**Figure 3. Illustration of the  $H$  score calculation.** Let  $w(l)$  be the width of stage  $l$ . Let  $w_b$  be the minimum width across all stages, and suppose that this minimum occurs at stage  $l = b$ ; this is the *waist* of the network (ties are broken so that the waist is closer to  $\lfloor L/2 \rfloor$ ). Consider the sequence  $X = \{w(l)\}$ ,  $l = 1, \dots, b$  and the sequence  $Y = \{w(l)\}$ ,  $l = b, \dots, L$ . We denote the normalized univariate Mann-Kendall statistic for monotonic trend on the sequences  $X$  and  $Y$  as  $\tau_x$  and  $\tau_y$ , respectively. The Mann-Kendall statistic varies between -1 (decreasing) and 1 (increasing), and it is approximately zero for a random sequence. We define  $H = (\tau_y - \tau_x)/2$ ;  $H$  is referred to as the hourglass score.  $H = 1$  if the DGEN is structured as an hourglass, with a decreasing sequence of  $b$  stages followed by an increasing sequence of  $L - b$  stages. In the computational modeling results, we do not consider the width of the first stage because it can never decrease in Models-1,2,3. We also define a variation of the hourglass score in which we do not take into account adjacent stages in calculating the two Mann-Kendall statistics. That statistic is denoted by  $\hat{H}$  and is referred to as the robust hourglass score.

genes are assigned to stage-pair  $(l-1, l)$ . Denote by  $p_i$  the phylogenetic rank (TAI value) of gene  $i$ . The age index assigned to that stage-pair is given by

$$\text{TAI}(l) = \frac{\sum_{i=1}^{n(l)} p_i e'_{i,l}}{\sum_{i=1}^{n(l)} e'_{i,l}}.$$

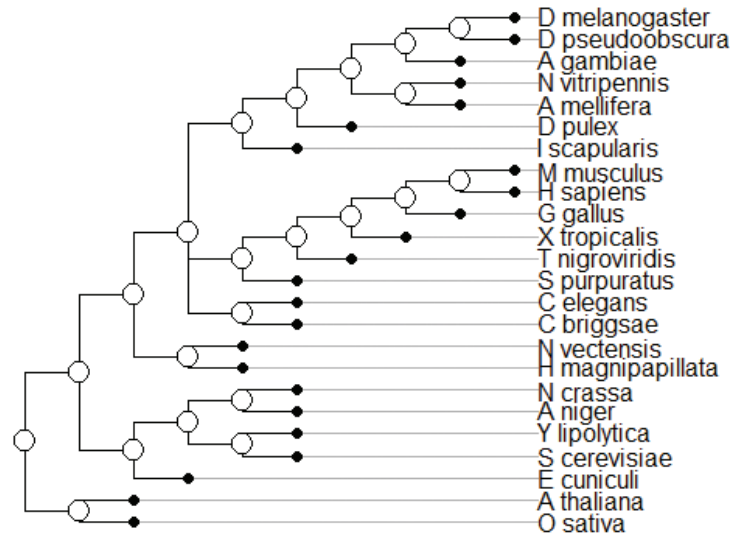
The same method is used when transitioning genes are identified based on absolute expression levels.

## Results

### Simulation

We simulate the presented model to examine the properties of the surviving DGEnS as evolutionary time progresses. The initial population consists of  $N$  identical DGEnS with  $\Gamma$  genes at each stage. The edges between genes are constructed probabilistically based on the specificity of each stage, as described previously. Simulating the complete model would not show the significance of individual mechanisms such as the increasing specificity assumption. For this reason we construct a sequence of four models with increasing complexity, presenting results separately for each of them:

**Model-1: Constant specificity.** Each stage has the same specificity,  $s(l) = 0.5$  for  $l = 1 \dots L - 1$ . Further, this model does not



**Figure 4.** The phylotypic tree that we use to calculate the age index of *Drosophila*'s genes. Each gene is assigned to one of the following six ages: Level-1: Common ancestor to Fungi, Plants and Eumetazoa. Level-2: Common ancestor to Fungi and Eumetazoa. Level-3: Common ancestor to all Eumetazoa. Level-4: Common ancestor to all Bilateria. Level-5: Common ancestor to all Arthropoda. Level-6: Common ancestor to all Diptera.

include gene deletion and duplication. Gene rewiring can cause RF and DF events even if there are no DL or DP events. We have also experimented with other values of the specificity probability, and the results are qualitative the same (namely, when the specificity is constant there is no hourglass effect).

**Model-2: Increasing specificity.** The difference from Model-1 is that the specificity is gradually increasing across developmental stages. Unless noted otherwise, the specificity is linear,  $s(l) = l/L$  for  $l = 1 \dots (L - 1)$ ; a nonlinear specificity function is also considered, which we describe later. We do not claim that these particular specificity functions are realistic or that there are experimental results that suggest them. They are just the simplest and most parsimonious models that lead to the emergence of the hourglass pattern.

**Model-3: With gene duplications.** Model-3 adds DP events in Model-2. The duplication probability  $P_{DP}$  is set so that the average size of a DGEN, across the entire population, stays within a given range (70%–80% of  $L \times \Gamma$  genes).

**Model-4: With gene deletions.** Model-4 adds DL events in Model-3 (complete model). The deletion probability  $P_{DL}$  is set so that the average size of a DGEN, across the entire population, stays within the same range as in Model-3.

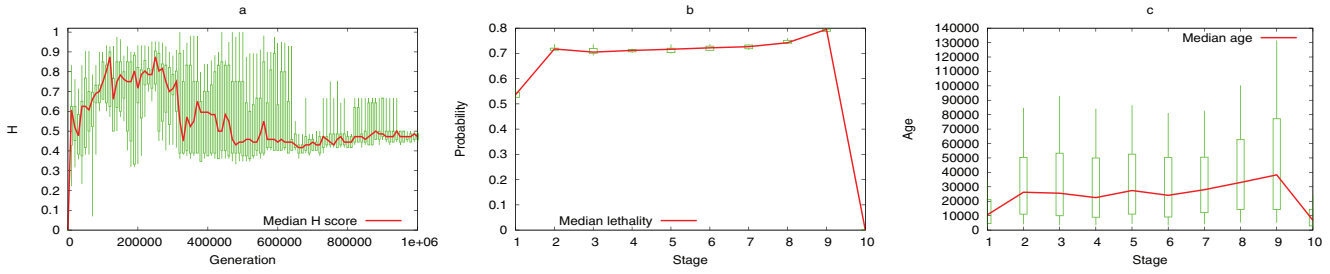
In Model-1 and Model-2, genes can be only removed (due to RW events, potentially followed by RF cascades) and so the average DGEN size decreases as evolutionary time progresses, which is unrealistic. Model-3 and Model-4 are more realistic because they can maintain a roughly constant DGEN size in the long-term. However, as will be shown next, all aspects of the developmental hourglass effect can already be seen with Model-2 (but not with Model-1). This highlights that the increasing specificity assumption is sufficient to generate the hourglass effect. Further, the inclusion of additional biological mechanisms in the model, namely gene duplication and

gene deletion, even though they make the model more realistic, they are not necessary for the emergence of the hourglass effect.

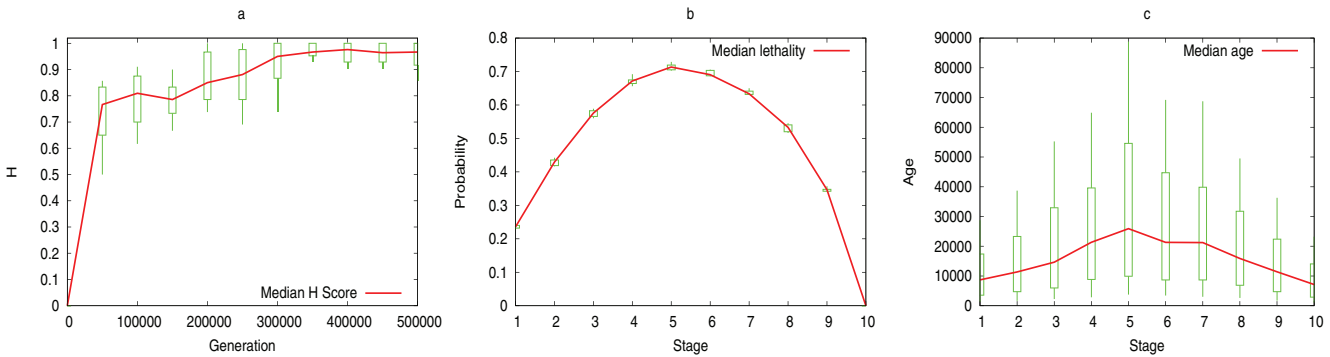
**Hourglass shape.** A first observation is that as evolutionary time progresses, DGENs acquire an “hourglass-like” shape in Models-2,3,4. This means that the width of each stage first decreases until a certain stage (referred to as the waist of the hourglass) and then gradually increases. The hourglass may not be symmetric with respect to the waist. To quantify this observation, we define an “hourglass score”  $H$  (see Methods and Figure 3) that is equal to 1 if the sequence of  $L$  stage widths consists of two segments: a decreasing sub-sequence of  $k \geq 2$  stages followed by an increasing sub-sequence of  $L - k \geq 2$  stages. Figure 6a shows the hourglass score for the population of DGENs in Model-2. Similar graphs for the three other models are shown in Figure 5a, Figure 7a, and Figure 8a. The  $H$  score quickly increases in the three models that have increasing specificity, and it fluctuates close to 1 afterwards.

What is the reason behind the hourglass shape of DGENs? When a gene  $g$  is rewired at stage  $l$ , it may trigger RF events in stage  $l + 1$  depending on the number of its lost outgoing edges. In the first few stages, where specificity is low, it is unlikely that a gene would lose a large fraction of its (typically many) incoming edges. In the last few stages, where specificity is high, edges are unlikely to get rewired in the first place. In the mid-stages however, where the specificity is close to 50%, there is higher variability in the number of outgoing edges lost or gained due to RW events. The loss of several outgoing edges due to an RW event at stage  $l$  can trigger several RF events and gene removals in the subsequent stage. Thus, the probability of RF events in mid-stages is higher than in early/late stages, making the removal of genes more likely in the former.

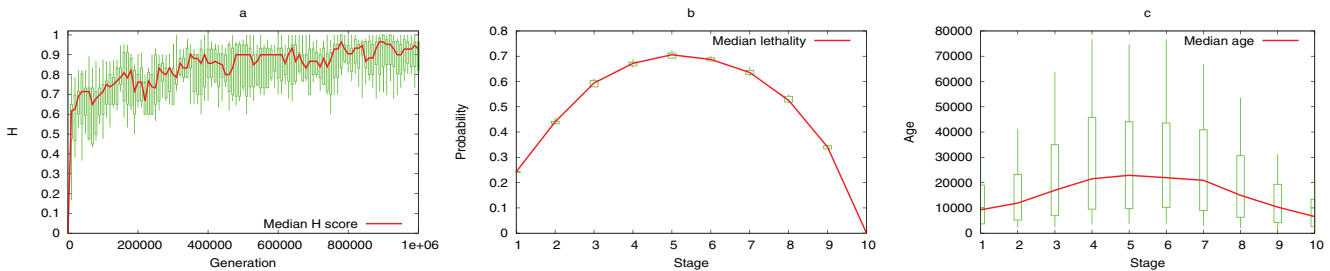
The constant specificity of Model-1 does not result in an hourglass pattern [Dataset 1] (see Figure 5a) for the following reason. RW events at stage  $l$  can cause RF events at the next stage with the same



**Figure 5. Computational results with Model-1.** Parameters: 10 runs with different initial populations,  $N = 10$  individuals,  $L = 10$  stages, specificity function  $s(l) = 0.5$  for all stages,  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 1,000,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The red line is the median and the green boxes are the 10th, 25th, 75th, and 90th percentiles, across all individuals and all runs. **(a)** The hourglass score  $H$  across evolutionary time. **(b)** Lethality probability at each stage. **(c)** Age of existing genes at the last generation.

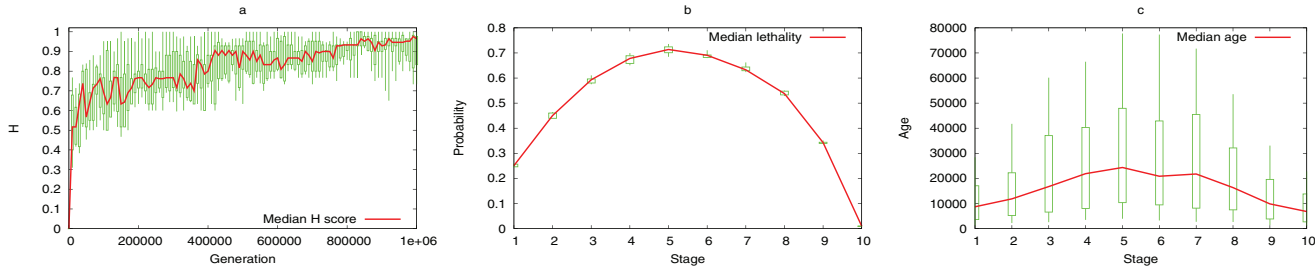


**Figure 6. Computational results with Model-2.** Parameters: 10 runs with different initial populations,  $N = 1000$  individuals,  $L = 10$  stages, specificity function  $s(l) = l/L$  ( $l = 1 \dots L - 1$ ),  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 500,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The red line is the median and the green boxes are the 10th, 25th, 75th, and 90th percentiles, across all individuals and all runs. **(A)** The hourglass score  $H$  across evolutionary time. **(B)** Lethality probability at each stage. **(C)** Age of existing genes at the last generation.



**Figure 7. Computational results with Model-3.** Parameters: 10 runs with different initial populations,  $N = 10$  individuals,  $L = 10$  stages, specificity function  $s(l) = l/L$  ( $l = 1 \dots L - 1$ ),  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 1,000,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The probability of gene duplication  $P_{Dp}$  is adjusted dynamically so that the average DGEN size stays between 700 and 800 genes. The red line is the median and the green boxes are the 10th, 25th, 75th, and 90th percentiles, across all individuals and all runs. **(a)** The hourglass score  $H$  across evolutionary time. **(b)** Lethality probability at each stage. **(c)** Age of existing genes at the last generation.





**Figure 8. Computational results with Model-4.** Parameters: 10 runs with different initial populations,  $N = 10$  individuals,  $L = 10$  stages, specificity function  $s(l) = l/L$  ( $l = 1 \dots L - 1$ ),  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 1,000,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The probability of gene duplication  $P_{Dup}$  is adjusted dynamically so that the average DGEN size stays between 700 and 800 genes. The probability of gene deletion (DL) is  $P_{DL} = 10^{-6}$ . The red line is the median and the green boxes are the 10th, 25th, 75th, and 90th percentiles, across all individuals and all runs. **(a)** The hourglass score  $H$  across evolutionary time. **(b)** Lethality probability at each stage. **(c)** Age of existing genes at the last generation.

probability, independent of  $l$ . However, after the occurrence of an RF event, the size of the potential cascade increases as  $l$  decreases simply because there are more subsequent stages to affect. This gives DGENs a “funnel-like” shape with a gradually increasing number of genes after stage-1;  $H$  fluctuates around 0.5, as expected for an increasing sequence.

**Stage lethality.** Another aspect of the developmental hourglass is in terms of the significance of each stage for the survival of the embryo. We define *lethality of stage  $l$*  as the probability that a RW or DL event at stage  $l$  starts a RF cascade that eventually leads to a DF event. We estimate this probability at generation  $i$  as the fraction of RW and DL events, during the past  $i$  generations, that occurred at stage  $l$  and led to a DF.

In Model-1, there is no clear trend for the stage lethality probability (see Figure 5b); with the exception of the last stage (in which RW events cannot result in gene loss), the lethality probability is roughly the same at all stages. In the three models with increasing specificity, however, we observe a clear pattern: the lethality gradually increases until the waist of the hourglass, and then it decreases [Dataset 2, Dataset 3 and Dataset 4] (see Figure 6b, Figure 7b, and Figure 8b). The reason, as explained earlier, is that the probability of RF events in mid-stages is higher than in early/late stages. Additionally, after the formation of the hourglass shape the mid-stages have relatively few genes and so an RF event in those genes is more likely to initiate a lethal RF cascade.

These computational results for the lethality probability across development are consistent with the empirical observations of Galis and Metz<sup>15</sup> about the increased mortality of rodents due to perturbations in the phylotypic stage (see also Discussion section).

**Age of genes.** A third aspect of the developmental hourglass effect is related to the evolutionary age of genes. The age of a gene  $g$  at generation  $i$  is defined as

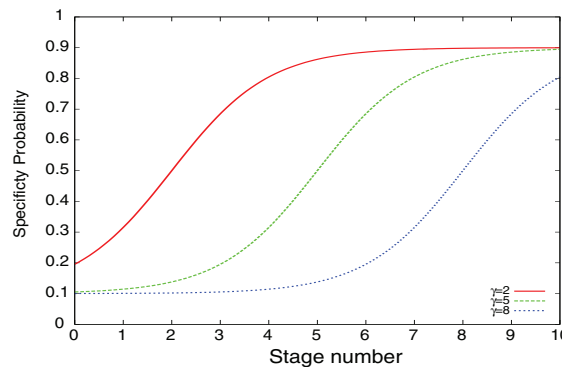
$$A(g) = i - t_0(g),$$

where  $t_0(g)$  is the generation at which  $g$  was most recently rewired (and 0, if it was not rewired earlier). The rationale behind this definition is that a rewiring event may give that gene a new function, at least in terms of its upstream and downstream regulators<sup>32,33</sup>.

In the case of Model-2, Figure 6c shows the median age of the genes at each stage, considering the population of all genes across all individuals at a given generation. See Figure 5c, Figure 7c, and Figure 8c for the same results with the three other models. The evolutionary age at stage  $l$  follows the same pattern as the lethality probability: it gradually increases until we reach the waist of the hourglass, and then it gradually decreases. Genes at intermediate stages tend to be older because, as explained earlier, they are fewer and their rewiring is more likely to be lethal. When one of those genes  $g$  is rewired or deleted from a DGEN, the corresponding individual is often replaced (DF event) by another individual that has the same gene  $g$ . So, the genes at the waist of a DGEN tend to be more conserved than genes at earlier or later stages.

**Location of waist.** What controls the location of the hourglass waist in the developmental process?

The location of the waist is mostly affected by two parameters of the model: the shape of the specificity function and the RF parameter  $z$ . To examine the effect of the former we use the nonlinear function shown in Figure 9.  $\gamma$  is the stage at which the specificity is 50%, and so that stage has the maximum variance in the number of outgoing regulatory edges. RW events at this stage can cause the largest extent of rewiring and so, the highest likelihood of RFs in genes of the next stage. Figure 10a shows that the location of the



**Figure 9. A nonlinear specificity function,  $s(l) = 0.9 - 0.8/(1 + e^{-\gamma(l-\gamma)})$ , for three values of the parameter  $\gamma$ .** This function allows us to control the stage  $\gamma$  at which the specificity is 50%.

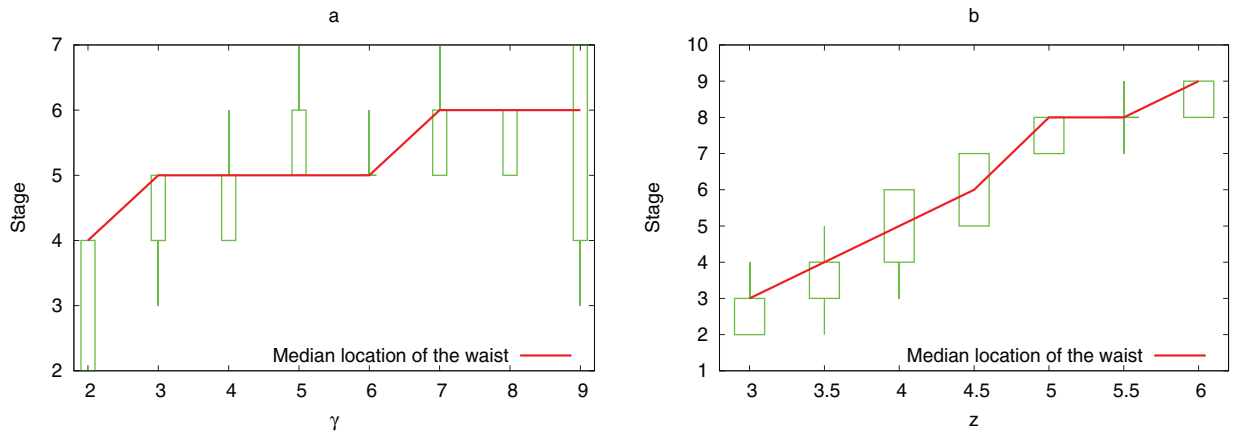
hourglass waist is "pushed" towards stage  $\gamma$ , even though it does not coincide always with that stage. Parameter  $z$  controls the shape of the RF probability: increasing  $z$  makes RF events more likely, also increasing the likelihood of lethal RF cascades. Figure 10b shows that as we increase  $z$  the hourglass waist moves towards later developmental stages [Dataset 5].

**Gene prevalence.** We introduce a "gene prevalence" metric for gene  $g$  at time  $t$  as the fraction of individuals that include  $g$  at evolutionary time  $t$  [Dataset 6]. Figure 11a shows the gene prevalence metric across developmental stages after 500,000 generations, while

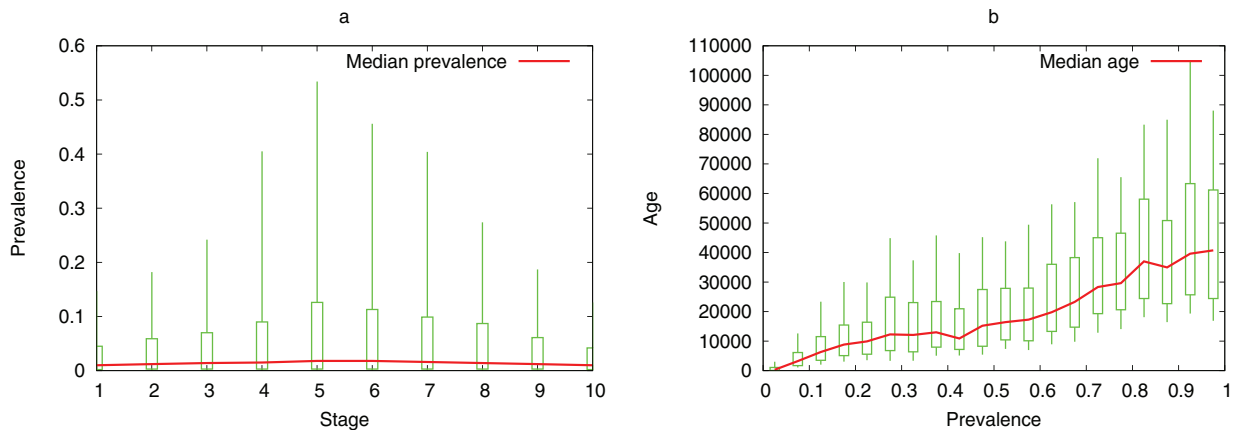
Figure 11b shows the relation between gene prevalence and gene age. The genes at the waist of the developmental hourglass are not only the oldest but also the most prevalent across the population. The implication of this simulation result is that we can expect that those genes that are transitioning near the waist of the developmental hourglass will be the most conserved genes in a population.

**Data analysis**

We have examined the predictions of the previous model using transcriptome data for *Drosophila melanogaster* and *Arabidopsis thaliana* [Dataset 14]. We summarize results from both species here;



**Figure 10. We examine the effect of the two model parameters that affect the location of the DGEN hourglass waist.** The first is the specificity function. To examine its effect, we use a sigmoid-like mathematical function that controls the stage  $\gamma$  at which the specificity is 50% (see Figure 9). This is the stage with the maximum variance in the number of outgoing regulatory edges. RW events at this stage can cause the largest extent of rewiring and so, the highest likelihood of RFs in genes of the next stage. Graph (a) shows that the location of the hourglass waist is "pushed" towards stage  $\gamma$ , even though it is not always exactly at that stage. The second way to affect the location of the hourglass waist is the parameter  $z$  that controls the shape of the RF probability. Increasing  $z$  makes RF events more likely, also increasing the likelihood of lethal RF cascades. Graph (b) shows that as we increase  $z$  the hourglass waist moves towards later developmental stages. These results are obtained using Model-2. Parameters: 10 runs with different initial populations,  $N = 1000$  individuals,  $L = 10$  stages, specificity function  $s(l) = l/L$  ( $l = 1 \dots L - 1$ ),  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 500,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The graphs show the median (red lines) and the 10th, 25th, 75th, and 90th percentiles (green boxes) of the location of the waist.



**Figure 11. The prevalence of a gene  $g$  in a population of  $N$  individuals is the fraction of individuals in which gene  $g$  appears.** These results are obtained using Model-2. Parameters: 10 runs with different initial populations,  $N = 1000$  individuals,  $L = 10$  stages, specificity function  $s(l) = l/L$  ( $l = 1 \dots L - 1$ ),  $\Gamma = 100$  genes at each stage initially, RF parameter  $z = 4$ , 500,000 generations, probability of RW event  $P_{RW} = 10^{-4}$ . The graphs show the median (red lines) and the 10th, 25th, 75th, and 90th percentiles (green boxes) for: (a) prevalence of genes in each stage after 500,000 generations, and (b) gene age as a function of gene prevalence. As expected, older genes tend to be more prevalent in the population.

the corresponding figures for which are [Figure 12](#) and [Figure 13](#) [Dataset 8 and Dataset 9]. For *Drosophila*, we analyze Microarray<sup>9</sup> and RNA-Seq<sup>27</sup> temporal expression profiles during the first 20 hours of development, taken at 10 stages of 2-hr intervals. We examine whether a) the number of transitioning genes follows an hourglass pattern, b) the waist of that hourglass coincides with the *Drosophila* phylotypic stage, and c) the evolutionary age of the transitioning genes follows a similar hourglass pattern. The two datasets are described in more detail in the Methods section. With such limited data, we cannot infer the regulatory edges between transitioning genes and we cannot reconstruct the underlying DGEN. However, we can identify the transitioning genes at each developmental stage given a “transition threshold”  $c$  (see Methods). Even though the correct value of this threshold is not known, the following results are robust in a wide interval of  $c$ , which includes most of the expression variation range across successive developmental stages (see [Figure S3](#) for the CDFs of expression level variations across successive stages [Dataset 12]).

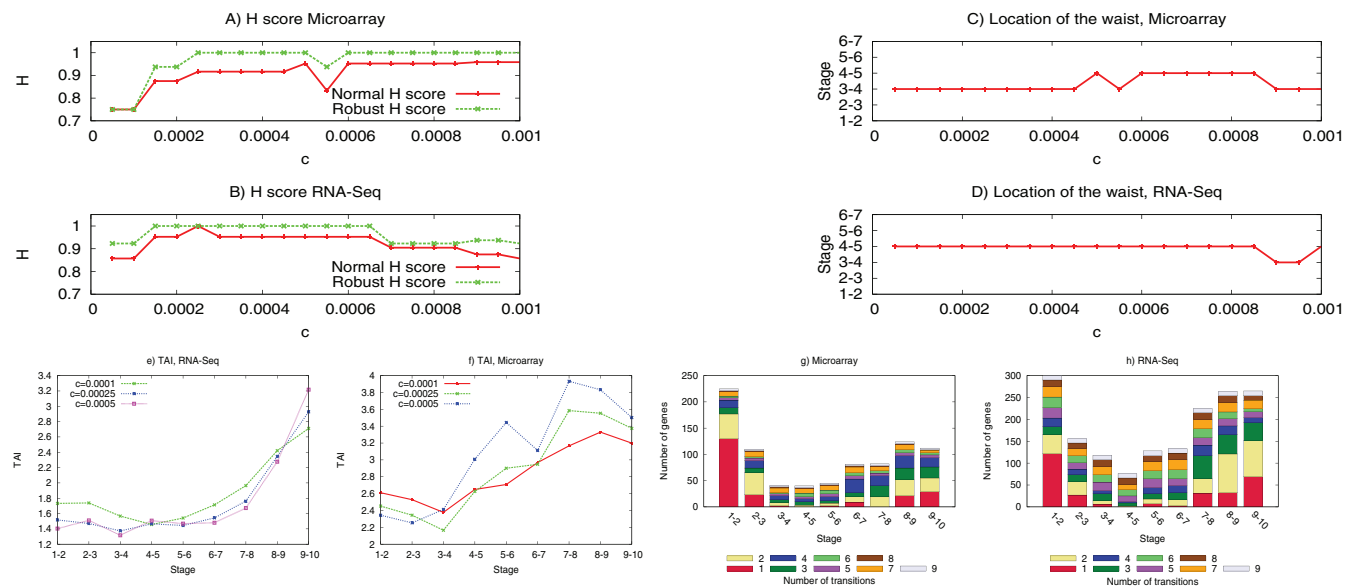
[Figure 12a](#) and [Figure 12b](#) show the hourglass resemblance score  $H$  (and its more robust variant) as function of  $c$ . Note that the  $H$  score is close to 1 for a wide range of  $c$ , confirming the presence of an hourglass-like structure in terms of the number of transitioning genes. [Figure 12g](#) and [Figure 12h](#) exhibit this pattern more clearly in the number of transitioning genes for a specific value of  $c$ . The two

datasets also show reasonable agreement in terms of the assignment of transitioning genes to stage-pairs [Dataset 10] (see [Figure S1](#)).

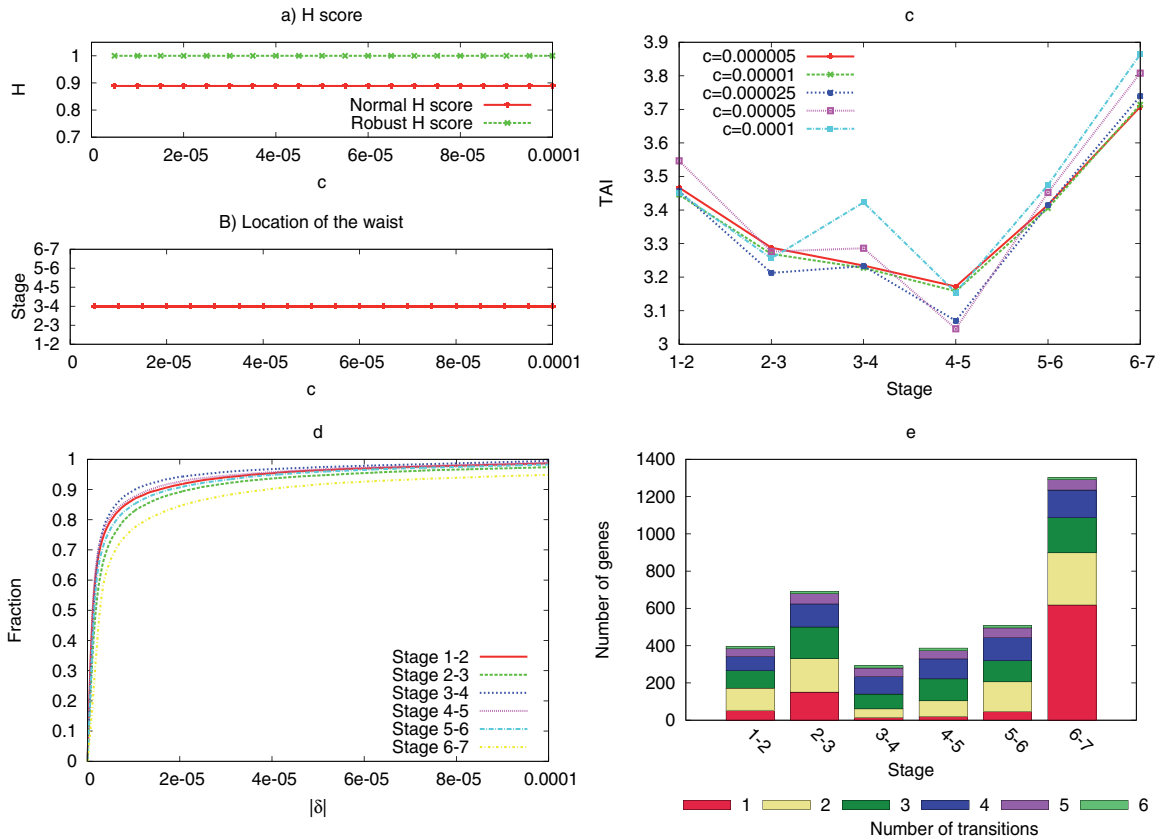
Second, the location of the waist in this hourglass pattern, shown in [Figure 12c](#) and [Figure 12d](#), occurs at the stage-pair (3,4) or (4,5), depending on  $c$ . This is roughly 8 hours after the formation of the zygote, and it includes the phylotypic stage for *Drosophila melanogaster*<sup>9</sup>.

We have also estimated the evolutionary age of most of the transitioning genes at each developmental stage-pair using the Transcriptome Age Index (TAI) metric<sup>11</sup> (see Methods). TAI is lower for older genes. [Figure 12e](#) and [Figure 12f](#) show the average TAI for transitioning genes, weighted by the expression level of each gene, at each stage-pair and for each dataset using three values of  $c$ . The TAI index follows the pattern that the model predicts, with older genes (lower TAI values) close to the waist of the hourglass. This result appears consistent with the main observation of Domazet-Lozo and Tautz<sup>13</sup>, even though that study did not analyze transitioning genes.

The same approach was extended using transcriptome data<sup>28</sup> and TAI profiles<sup>11</sup> from *Arabidopsis*. The results obtained from the analysis ([Figure 13](#)) were consistent with the predictions of the model, as well as the results obtained from *Drosophila* data.



**Figure 12. *Drosophila* results using normalized expression levels.** Graphs (A) and (B) show the hourglass score (normal and robust) as a function of the transition threshold  $c$  for the two datasets. Graphs (C) and (D) show the location of the hourglass waist (stage-pair) as a function of the transition threshold  $c$  for the two datasets. Graphs (E) and (F) show the Transcriptome Age Index of transitioning genes for three different values of  $c$  (chosen so that the number of genes with known age index assigned to each stage is at least 25) for the two datasets. Graph (G) shows the transitioning genes for the Microarray dataset with  $c = 0.0005$ . The transitioning genes constitute 11% of all genes in that dataset. 53% of those genes transition in a single stage-pair. Of the remaining, 64% transition only in consecutive stage-pairs. Note that if a gene transitions  $n$  times, it is counted in  $n$  stage-pairs. Similarly, graph (H) shows the transitioning genes for the RNA-Seq dataset with  $c = 0.00025$ . The transitioning genes constitute 5% of all genes in that dataset. 45% of those genes transition in a single stage-pair. Of the remaining, 52% transition only in consecutive stage-pairs.



**Figure 13. *Arabidopsis thaliana* results using normalized expression levels.** Graph (a) shows the hourglass score (normal and robust) as a function of the transition threshold  $c$ . Graph (b) shows the location of the hourglass waist (stage-pair) as a function of the transition threshold  $c$ . Graph (c) shows the Transcriptome Age Index of transitioning genes for five different values of  $c$  (chosen so that the number of genes with known age index assigned to each stage-pair is at least 290). Graph (d) shows the CDFs of the expression level absolute variations  $|\delta|$  across successive stage-pairs. Note that in this case the hourglass waist (in terms of number of transitioning genes) appears in stage-pair (3,4), while the oldest genes appear in the next stage-pair. Graph (e) shows the transitioning genes with  $c = 0.0001$ . The transitioning genes constitute 7% of all genes in that dataset. 49% of those genes transition in a single stage-pair. Of the remaining, 36% transition only in consecutive stage-pairs.

## Discussion

Early studies of the developmental hourglass effect mostly analyzed morphological and phenotypic similarities across species<sup>2,34</sup>. Recently, the focus has shifted towards genomic and molecular comparative studies<sup>8,9,11,13,20</sup> that investigate conservation of gene expression variation, sequence conservation, selective constraint on coding sequences, and evolutionary gene “age”. These studies often report contradicting observations: some support strong conservation in earlier developmental stages<sup>5,6,20</sup>, while others support that strongest conservation occurs at a mid-developmental stage<sup>7-11,13-16</sup>. Nevertheless, the fact that the hourglass effect is observed in highly divergent species across deep phylogenetic scales (including fish, flies and plants), suggests that this observed pattern of conservation may stem from fundamental organization principles.

What these principles are has remained elusive. Earlier stages may be conserved because any changes therein could have large cascading effects in later stages<sup>5,35,36</sup>. Later stages may experience less constraint because as development progresses gene interactions become more modular, and so it is plausible that perturbations there have only local effects<sup>1</sup>. We refer to them as the “temporal”

constraint model and the “spatial” constraint model, respectively, following Tian *et al.*<sup>37</sup>.

In this paper, we developed an evolutionary model of development that combines some aspects of the previous two models. Regulatory perturbations at a certain stage can cause cascades of regulatory failures at subsequent stages (temporal model), while the likelihood that a gene regulates genes at a subsequent stage decreases as development progresses (spatial model).

Our computational results lead to the following testable predictions: a) the number of transitioning genes during development follows an hourglass pattern, b) the evolutionary age of the transitioning genes also follows an hourglass pattern, with the oldest genes being at the waist of the hourglass, and c) the genes at the waist of that hourglass are the most essential, in the sense that their deletion maximizes the probability of developmental failure.

We have relied on developmental gene expression profiles of *Drosophila melanogaster* and *Arabidopsis thaliana* to examine the predictions of the model. The analysis of that data agrees with the

first two theoretical model predictions. The third prediction of the model is in direct agreement with the empirical observations of Galis and Metz regarding the increased mortality caused by perturbations during the phylotypic stage<sup>15</sup>. That study has shown, based on the teratological literature for rodent development, that disturbances in the phylotypic stage lead to much higher mortality than in other stages. Further, such disturbances lead to the co-occurrence of several distinct anomalies in the developing embryo and so the increased mortality cannot be due to a single particularly vulnerable process that takes place at that stage. Further, our simulations confirm that the details of these regulatory perturbations, such as the probabilities of gene duplication and deletion or the parameter  $z$  in the regulatory failure probability, do not affect the results of the model, at least at the qualitative level.

It is interesting to examine our results in the context of Raff's hypothesis<sup>1</sup>. Raff proposed an interesting explanation for the developmental hourglass effect based on the notions of "interconnectivity between body elements" and "developmental flexibility". Specifically, his hypothesis states that the level of interaction (interconnectivity) between the elements of the developing embryo is maximized in mid-development. Early development is flexible because it governs robust and general global patterning processes, late development is also flexible because "signaling events within the primordia are little influenced by events in other primordia", while mid-development (phylotypic stage) is least flexible because of the "high interconnectivity between elements that will later come to represent separate modules". This hypothesis may be viewed initially as a contradiction with our increasing specificity assumption, which states that the density of regulatory interactions is maximized at the earliest stages of development. Note however that Raff's hypothesis was not stated in terms of gene regulatory interactions – he was referring more generally to "developmental flexibility". If we interpret the notion of "developmental flexibility" as the ability of an embryo to survive gene mutations and rewiring at different stages of the developmental process, then Raff's hypothesis is actually consistent with our results regarding the "lethality probability" at each developmental stage (see [Figure 6-B](#), [Figure 7-B](#) and [Figure 8-B](#)). The lethality probability is maximized at the phylotypic stage, and it is significantly lower at early and late developmental stages, following exactly the same pattern with Raff's developmental flexibility.

The use of DGEnS in this work was only as an abstract tool to study the effect of gene regulatory perturbations in the developmental process. In future work, it is important to infer the actual DGEN of model organisms. This will require information about gene regulatory interactions across time and space, but it should be possible for at least some developmentally well studied species<sup>24</sup>. Such DGEnS would help to identify the specific genes that form the hourglass waist and their function. Additionally, an inferred DGEN would allow to directly test the increasing specificity assumption.

Finally, we note that the hourglass effect (sometimes referred to as the "bow-tie" effect) has also been observed in other complex biological and technological systems that exhibit hierarchical modularity and that are subject to evolutionary pressure or optimization

tradeoffs<sup>38–41</sup>. Recently, Friedlander *et al.* have proposed a different mechanism for the emergence of hourglass-like patterns in evolving biological or technological networks, based on a linear system formulation<sup>42</sup>. They showed that if a system can be represented as a hierarchical linear transformation of an input vector to an output vector, and the desired transformation matrix is rank-deficient, then an evolutionary process that selects that particular transformation can, under certain conditions, converge to an hourglass-like structure. It is not clear yet how to adapt this linear model in the context of inherently non-linear systems, such as gene regulatory networks.

Another example of a hierarchical system that is structured as an hourglass is the Internet "protocol stack"<sup>43</sup>; this pattern was not designed but it emerged through the competition between protocols that serve roughly the same function at each communication layer, during the last 30–40 years. In earlier work, we proposed an abstract model (EvoArch) that captures the evolution of protocol architectures and that predicts the emergence of an hourglass structure. Interestingly, both EvoArch and the model of this paper share the same principle: the underlying hierarchical networks that control both systems should be increasingly sparser as complexity increases, i.e., the specificity of each complexity stage (or layer) should be increasing. In the future, we will further investigate this common organization principle between biological and technological systems.

#### Data availability

"ZENODO: An evo-devo model for the developmental hourglass: Code and data", DOI: [10.5281/zenodo.1057944](https://doi.org/10.5281/zenodo.1057944), License: GNU GPLv3.

#### Author contributions

S.A., S.S., C.D., and S.Y. designed research; S.A. and C.D. developed the computational model; S.A. and S.S. analyzed data; and S.A., S.S., C.D., and S.Y. wrote the paper.

#### Competing interests

No competing interests were disclosed.

#### Grant information

This work was supported by the National Science Foundation (Grant No. 0831848 to C.D.). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

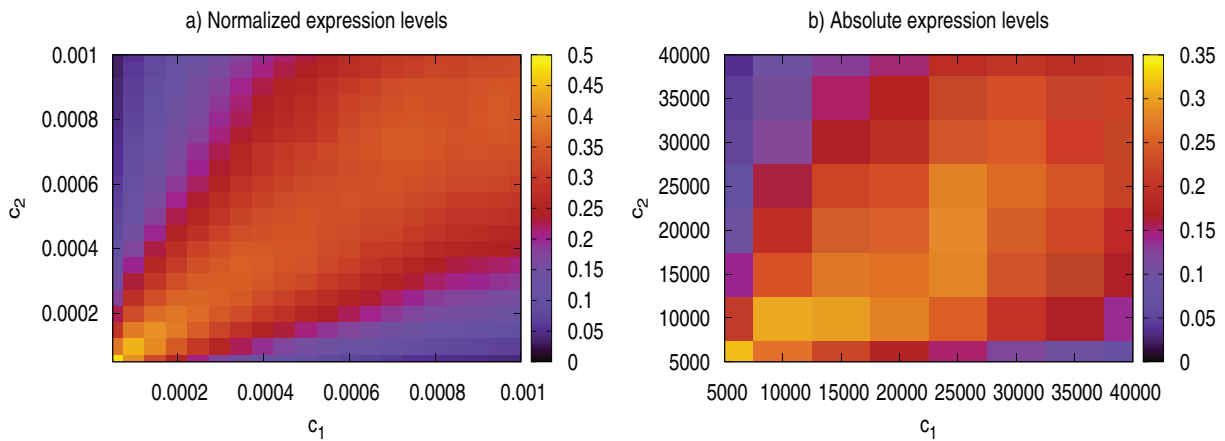
#### Acknowledgements

We thank Alex Kalinka for providing time-series gene expression data for *Drosophila melanogaster*. We also thank Joshua Weitz, Sridhar Hannenhalli and Todd Strelman for helpful comments on the manuscript.

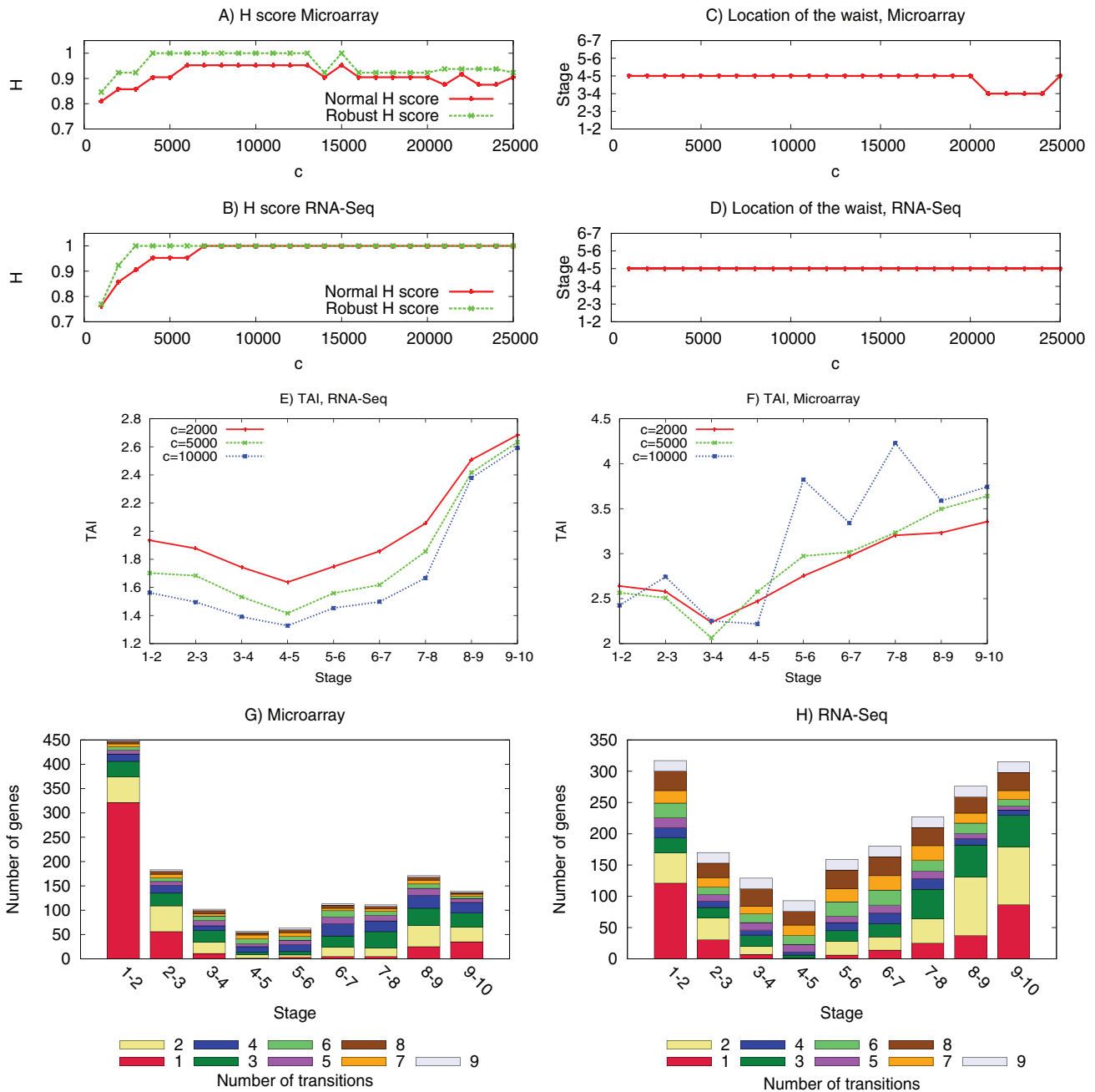
#### Note added in proof

Dedicated to the memory of the first author, Saamer Akhshabi, who passed away on March 6, 2014.

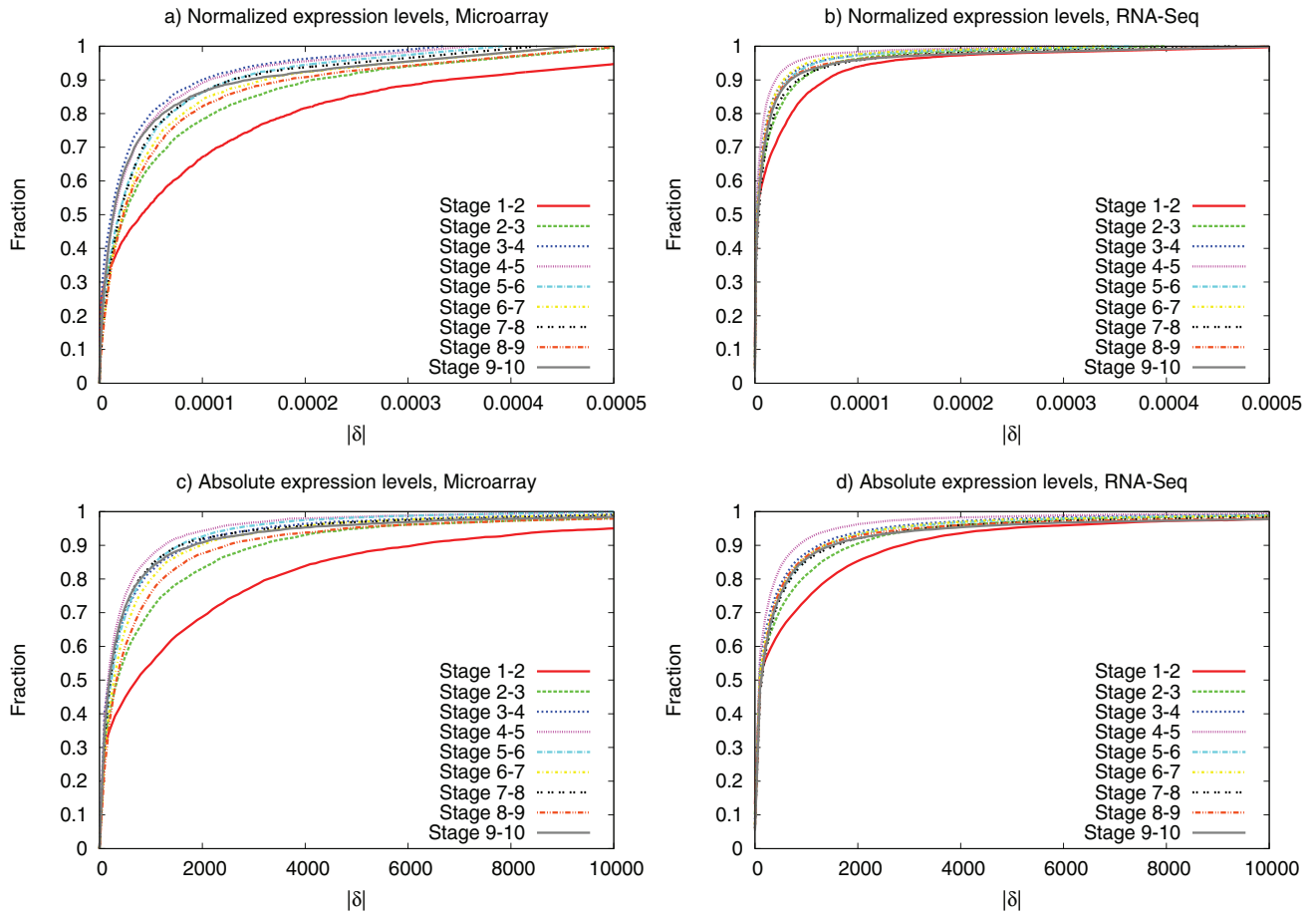
## Supporting information



**Figure S1.** We evaluated the agreement between the two *Drosophila* datasets in terms of the transitioning genes assigned to each stage-pair, considering only those genes that appear in both datasets. Because the appropriate transition threshold may be different at each dataset, we use a different threshold for each dataset, say  $c_1$  and  $c_2$ . For each pair  $(c_1, c_2)$ , we determine the transitioning genes at each stage-pair with the corresponding dataset,  $L - 1$  pairs of gene sets), and then calculate the average Jaccard similarity across these  $L - 1$  pairs. The Jaccard similarity maps show this average across all stage-pairs for various threshold pairs  $(c_1, c_2)$ . In graph (a) with normalized expression levels, when the two thresholds are roughly equal, the average Jaccard similarity is as high as 50%; this means that about 2/3 of the genes assigned to a certain stage-pair using one dataset are also assigned to the same stage-pair using the other dataset. Graph (b) shows a similar Jaccard similarity map for the case of absolute expression levels. [Dataset 10].

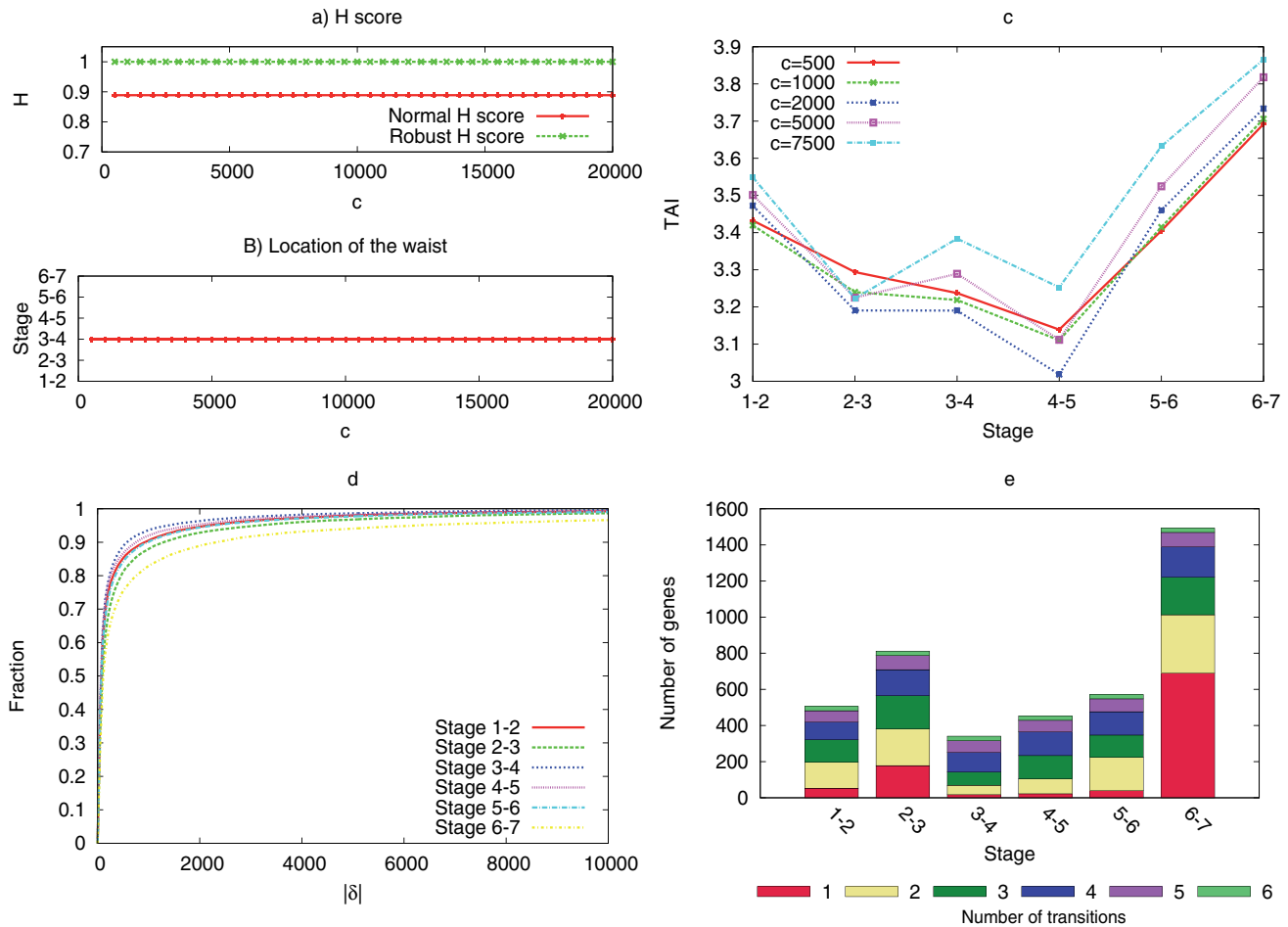


**Figure S2. *Drosophila* results using absolute expression levels.** Graphs (a) and (b) show the hourglass score (normal and robust) as a function of the transition threshold  $c$  for the two datasets. Graphs (c) and (d) show the location of the hourglass waist (stage-pair) as a function of the transition threshold  $c$  for the two datasets. Graphs (e) and (f) show the Transcriptome Age Index of transitioning genes for three different values of  $c$  (chosen so that the number of genes with known age index assigned to each stage is at least 10) for the two datasets. Graph (g) shows the transitioning genes for the Microarray dataset with  $c = 5000$ . The transitioning genes constitute 21% of all genes in that dataset. 62% of those genes transition in a single stage-pair. Of the remaining, 58% transition only in consecutive stage-pairs. Similarly, graph (h) shows the transitioning genes for the RNA-Seq dataset with  $c = 10000$ . The transitioning genes constitute 5% of all genes in that dataset. 45% of those genes transition in a single stage-pair. Of the remaining, 53% transition only in consecutive stage-pairs. [Dataset 11].



**Figure S3. *Drosophila* data: CDFs of the expression level absolute variations  $|\delta|$  across successive stage-pairs.** (a) normalized expressions, Microarray, (b) normalized expressions, RNA-Seq, (c) absolute expressions, Microarray, absolute expressions, RNA-Seq. [Dataset 12].





**Figure S4. *Arabidopsis thaliana* results using absolute expression levels.** Graph (a) shows the hourglass score (normal and robust) as a function of the transition threshold  $c$ . Graph (b) shows the location of the hourglass waist (stage-pair) as a function of the transition threshold  $c$ . Graph (c) shows the Transcriptome Age Index of transitioning genes for five different values of  $c$  (chosen so that the number of genes with known age index assigned to each stage-pair is at least 170). Graph (d) shows the CDFs of the expression level absolute variations  $|\delta|$  across successive stage-pairs. Graph (e) shows the transitioning genes with  $c = 5000$ . The transitioning genes constitute 8% of all genes in that dataset. 48% of those genes transition in a single stage-pair. Of the remaining, 38% transition only in consecutive stage-pairs. [Dataset 13].

## References

- Raff RA: **The shape of life: genes, development, and the evolution of animal form.** University of Chicago Press. *Trends Ecol Evol.* 1996; **11**(10): 441–442. [Publisher Full Text](#)
- Richardson MK, Keuck G: **Haeckel's ABC of evolution and development.** *Biol Rev Camb Philos Soc.* 2002; **77**(4): 495–528. [PubMed Abstract](#) | [Publisher Full Text](#)
- Davidson EH: **The regulatory genome: gene regulatory networks in development and evolution.** Academic Press, 2nd edition. 2010. [Reference Source](#)
- Duboule D: **Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony.** *Dev Suppl.* 1994; **1994**(Supplement): 135–142. [PubMed Abstract](#)
- Rasmussen N: **A new model of developmental constraints as applied to the *Drosophila* system.** *J Theor Biol.* 1987; **127**(3): 271–299. [PubMed Abstract](#) | [Publisher Full Text](#)
- Roux J, Robinson-Rechavi M: **Developmental constraints on vertebrate genome evolution.** *PLoS Genet.* 2008; **4**(12): e1000311. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Irie N, Kuratani S: **Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis.** *Nat Commun.* 2011; **2**: 248. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Irie N, Sehara-Fujisawa A: **The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information.** *BMC Biol.* 2007; **5**(1): 1. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kalinka AT, Varga KM, Gerrard DT, et al.: **Gene expression divergence recapitulates the developmental hourglass model.** *Nature.* 2010; **468**(7325): 811–814. [PubMed Abstract](#) | [Publisher Full Text](#)
- Levin M, Hashimshony T, Wagner F, et al.: **Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo.** *Dev Cell.* 2012; **22**(5): 1101–1108. [PubMed Abstract](#) | [Publisher Full Text](#)
- Quint M, Drost HG, Gabel A, et al.: **A transcriptomic hourglass in plant embryogenesis.** *Nature.* 2012; **490**(7418): 98–101. [PubMed Abstract](#) | [Publisher Full Text](#)
- Schep AN, Adryan B, et al.: **A comparative analysis of transcription factor expression during metazoan embryonic development.** *PLoS One.* 2013; **8**(6): e66826. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Domazet-Lošo T, Tautz D: **A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns.** *Nature.* 2010; **468**(7325): 815–818. [PubMed Abstract](#) | [Publisher Full Text](#)
- Hazkani-Covo E, Wool D, Graur D: **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.* 2005; **304**(2): 150–158.  
[PubMed Abstract](#) | [Publisher Full Text](#)
15. Galis F, Metz JA: **Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation.** *J Exp Zool.* 2001; **291**(2): 195–204.  
[PubMed Abstract](#)
  16. Cruickshank T, Wade MJ: **Microevolutionary support for a developmental hourglass: gene expression patterns shape sequence variation and divergence in *Drosophila*.** *Evol Dev.* 2008; **10**(5): 583–590.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  17. Comte A, Roux J, Robinson-Rechavi M: **Molecular signaling in zebrafish development and the vertebrate phylotypic period.** *Evol Dev.* 2010; **12**(2): 144–156.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  18. Hall BK: **Phylotypic stage or phantom: is there a highly conserved embryonic stage in vertebrates?** *Trends Ecol Evol.* 1997; **12**(12): 461–463.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  19. Kalinka AT, Tomancak P: **The evolution of early animal embryos: conservation or divergence?** *Trends Ecol Evol.* 2012; **27**(7): 385–393.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  20. Piasecka B, Lichocki P, Moretti S, et al.: **The hourglass and the early conservation models—co-existing patterns of developmental constraints in vertebrates.** *PLoS Genet.* 2013; **9**(4): e1003476.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  21. Richardson MK, Minelli A, Coates M, et al.: **Phylotypic stage theory.** *Trends Ecol Evol.* 1998; **13**(4): 158.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  22. Bininda-Emonds OR, Jeffery JE, Richardson MK: **Inverting the hourglass: quantitative evidence against the phylotypic stage in vertebrate development.** *Proc Biol Sci.* 2003; **270**(1513): 341–346.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  23. Wagner GP, Pavlicev M, Cheverud JM: **The road to modularity.** *Nat Rev Genet.* 2007; **8**(12): 921–931.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  24. Peter IS, Faure E, Davidson EH: **Predictive computation of genomic logic processing functions in embryonic development.** *Proc Natl Acad Sci U S A.* 2012; **109**(41): 16434–16442.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  25. Ewens WJ: **Mathematical Population Genetics 1: I. Theoretical Introduction,** Springer, 2004; **27**.  
[Reference Source](#)
  26. Gibbons JD, Chakraborti S: **Nonparametric statistical inference.** Springer, 2011.  
[Publisher Full Text](#)
  27. Graveley BR, Brooks AN, Carlson JW, et al.: **The developmental transcriptome of *Drosophila melanogaster*.** *Nature.* 2011; **471**(7339): 473–479.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  28. Xiang D, Venglat P, Tibiche C, et al.: **Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*.** *Plant Physiol.* 2011; **156**(1): 346–356.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  29. Waterhouse RM, Tegenfeldt F, Li J, et al.: **OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs.** *Nucleic Acids Res.* 2013; **41**(Database issue): D358–D365.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  30. Li L, Stoeckert CJ Jr, Roos DS: **OrthoMCL: identification of ortholog groups for eukaryotic genomes.** *Genome Res.* 2003; **13**(9): 2178–2189.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  31. Domazet-Loso T, Brajković J, Tautz D: **A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages.** *Trends Genet.* 2007; **23**(11): 533–539.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  32. Guet CC, Elowitz MB, Hsing W: **Combinatorial synthesis of genetic networks.** *Science.* 2002; **296**(5572): 1466–1470.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  33. Kim J, Kim I, Han SK: **Network rewiring is an important mechanism of gene essentiality change.** *Sci Rep.* 2012; **2**: 900.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  34. Richardson MK, Hanken J, Gooneratne ML, et al.: **There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development.** *Anat Embryol (Berl).* 1997; **196**(2): 91–106.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  35. Riedl R, Jefferies RP: **Order in living organisms: a systems analysis of evolution.** Wiley New York. 1978; 283–293.  
[Reference Source](#)
  36. Arthur W: **Mechanisms of morphological evolution: a combined genetic, developmental, and ecological approach.** Wiley New York. 1984.  
[Reference Source](#)
  37. Tian X, Strassmann JE, Queller DC: ***Dictyostelium* development shows a novel pattern of evolutionary conservation.** *Mol Biol Evol.* 2013; **30**(4): 977–984.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  38. Beutler B: **Inferences, questions and possibilities in Toll-like receptor signalling.** *Nature.* 2004; **430**(6996): 257–263.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  39. Csete M, Doyle J: **Bow ties, metabolism and disease.** *Trends Biotechnol.* 2004; **22**(9): 446–450.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  40. Doyle JC, Csete M: **Architecture, constraints, and behavior.** *Proc Natl Acad Sci U S A.* 2011; **108**(Suppl 3): 15624–15630.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  41. Tieri P, Grignolio A, Zaikin A, et al.: **Network, degeneracy and bow tie. Integrating paradigms and architectures to grasp the complexity of the immune system.** *Theor Biol Med Model.* 2010; **7**: 32.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  42. Friedlander T, Mayo AE, Tlustý T, et al.: **Evolution of bow-tie architectures in biology.** *arXiv preprint arXiv:1404.7715.* 2014.  
[Reference Source](#)
  43. Akhshabi S, Dovrolis C: **The evolution of layered protocol stacks leads to an hourglass-shaped architecture.** *ACM SIGCOMM.* 2011; **41**(4): 206–217.  
[Publisher Full Text](#)
  44. Dovrolis C, Yi S, Akhshabi S, et al.: **An evo-devo model for the developmental hourglass: code and data.** *Zenodo.* 2014.  
[Data Source](#)


# Open Peer Review

Current Referee Status:   

Version 2

Referee Report 14 January 2015

doi:10.5256/f1000research.6375.r7316

 **Marcel Quint** <sup>1</sup>, **Hajk-Georg Drost** <sup>2</sup>

<sup>1</sup> Department of Molecular Signal Processing, Leibniz Institute of Plant Biochemistry, Halle, Germany

<sup>2</sup> Institute of Computer Science, Martin Luther University Halle–Wittenberg, Halle, Germany

We highly appreciate that the authors carefully addressed all our concerns and we now have no further reservations.

We also add a detailed response to the rebuttal letter:

1. **Point#1 of our summarized comments to the first version of the manuscript:** *“The authors do cite Raff in the first sentence of their introduction, but should state more clearly in which manner their model was inspired by Raff’s hypothesis.”*

**Authors’ response:** *The reviewers are right to highlight the connection between our “increasing regulatory specificity” assumption and Raff’s hypothesis. We have included a paragraph in the Discussion section to put Raff’s hypothesis in the context of this work.*

*A first remark is that our assumption is consistent with Raff’s basic premise that modularity increases as the embryo develops. Second, Raff’s hypothesis (see chapter-6 of [1], and in particular Figures 6.6 and 6.7) also states that the “level of interaction” or the “interconnectivity between (body) elements” is maximized at the phylotypic stage. This may be viewed initially as a contradiction between Raff’s hypothesis and our increasing specificity assumption, which states that the density of regulatory interactions is maximized at the earliest stages of development. Note however that Raff’s hypothesis was not stated in terms of gene regulatory interactions -- he was referring more generally to “developmental flexibility”, arguing that early development is flexible because it governs robust and general global patterning processes, late development is also flexible because “signaling events within the primordia are little influenced by events in other primordia”, while mid-development (phylotypic stage) is least flexible because of the “high interconnectivity between elements that will later come to represent separate modules.” If we think of “developmental flexibility” as the ability of an embryo to survive gene mutations and rewiring at different stages of the developmental process, Raff’s hypothesis is actually consistent with our results regarding the “lethality probability” at each developmental stage (see Figs 6-B, 7-B and 8-B). The lethality probability is maximized at the phylotypic stage, and it is significantly lower at early and late developmental stages, following the same pattern with Raff’s “developmental flexibility”.*

**Our comment to authors' response:** We appreciate that the authors now differentiate between Raff's hypothesis concerning the modularization (interconnectivity) between body elements (modules) and their "increasing regulatory specificity" assumption. We agree that Raff's hypothesis is consistent with the "lethality probability" proposed by the authors in terms of the observed pattern of maximal interconnectivity and maximal probability of lethal effects during the phylotypic stage (mid embryogenesis). However, we would like to point out that a causal relationship between the observed patterns remains to be tested.

2. **Point#2 of our summarized comments to the first version of the manuscript:** *"Correlate and reference major assumptions of the DGEN model that seem plausible with findings from experimental studies that validate this plausibility."*

**Authors' response to Point #2:** *First, to the extent of our knowledge there is no prior work that directly validates the assumption of increasing regulatory specificity as formulated in our paper (i.e., with the specificity of a developmental stage defined as the topological density of regulatory edges in that stage). We are planning to examine the validity of this assumption in the future using various approaches:*

*First we plan to analyze the aforementioned developmental GRNs for the sea-urchin that have been directly examined by Dr. Davidson's lab at Caltech.*

*Second, we are in the process of examining other genomic metrics, which can be viewed as "proxies" to regulatory specificity. For example, we have defined new measures to quantify how specific a gene's expression profile in a developmental stage is compared to other developmental stages, which we have tentatively named as "developmental stage specificity index (DSI)".*

*Similarly, we defined "tissue specificity index (TSI)", measuring the bias of a gene's expression in a specific tissue compared to many other tissues. Analyses of these measures using genomic data from *Drosophila* so far have revealed generally increasing trends, which are consistent with the idea of increasing specificity. However we are still working on generalizing these analytical tools to data sets generated by different methods and from different species. Importantly, our model focuses DGENs underlying development, and genomic data can obscure the true signals from genes constituting DGENs. Thus we are also working on examining these specificity measures using inferred DGENs. Additionally, we are developing other methods to approximate functional specificity. We hope to complete these analyses as a follow-up paper. The current paper provides theoretical motivation.*

*Third, as the reviewers also point out, there is a large body of prior work in developmental biology that confirms the increasing modularity in the developing embryo (for instance, see [Wagner et al., 2007](#)) as well as the increasing specificity of the developmental process at the signaling or genomic level. The connections between this pattern of increasing modularity and the structure of the underlying gene regulatory networks are still not well understood however. We think that it would be misleading if we had argued that these modularity patterns provide a direct validation for our "increasing regulatory specificity" assumption. Please note that we have cited the paper by [Wagner et al.](#) at the Introduction and Model Description sections, providing some connections between this work and earlier work on the role of modularity in developmental biology.*

**Our comment to authors' response:** The authors point out that "the connections between this pattern of increasing modularity and the structure of the underlying gene regulatory network are still not well understood". This statement is valid for experimental studies that aim to address this

presumptive causal relationship in a systematic manner and we appreciate that the authors share their future plans to introduce sufficient measures (DSI and TSI) to better understand the nature of this relationship.

- Point#3 of our summarized comments to the first version of the manuscript:** *“Correlate the output of the DGEN simulation with existing findings of studies addressing similar questions of the causality of the developmental hourglass phenomenon and discuss potential simplifications of the model assumptions that do not fit experimental results or that are not derivable via experiments yet.”*

**Authors' response to Point #3:** *This is a very interesting idea and we thank the reviewers for this suggestion. We are planning to pursue this investigation in a follow-up paper that will also include the previously mentioned analysis of sea-urchin GRNs as well as the DSI/TSI results. We believe that it would be distracting to try to include all these results in this first paper, given that the main focus of this work has been on the computational model and the computational results. We are also open to collaborate in this investigation with the reviewers or other researchers working in this area, if they are interested.*

**Our comment to authors' response:** We are delighted to hear that the authors address this question in a follow up paper. We gladly offer our expertise to collaboratively investigate the evolutionary process on the regulatory level that might have led to a correlation between the maximized evolutionary age of expressed genes during mid-embryogenesis and the interconnectivity of (body) modules that are hypothesized to cause the morphological patterns during the phylotypic stage.

- Points#4 and 5 of our summarized comments to the first version of the manuscript:** *Include a clear statement how the exact values of the stage specificity (regulatory specificity)  $s(l)$  have been derived. State more clearly that  $s(l)$  is the crucial parameter determining the observed phenomena.”*

**Authors' response to Points #4 and 5:** *The regulatory specificity functions we experimented with are arguably the simplest one could think of. Specifically, we start (Model-1) with the simplest hypothesis that the regulatory specificity does not vary across developmental stages (in the paper we show results for  $s=0.5$  but we have generated results for all values 0.1, 0.2, ... 0.9 and the conclusion remains that when the specificity is constant the network does not evolve into an hourglass shape). Then, we consider (Model-2) a linearly increasing specificity function. We do not argue that this is realistic or that there are experimental results that suggest this linearity or even the increasing trend. It is just the simplest and most parsimonious model that leads to the emergence of an hourglass pattern. Finally, in Figures-9 and 10 we consider a more general specificity function that increases in a non-linear manner, providing a simple way to control the stage at which it gives the mid-range value 0.5. This allows us to examine how the shape of the specificity function  $s(l)$  affects the location of the hourglass waist. We hope that follow-up work will reveal experimentally the actual shape of the regulatory specificity function  $s(l)$  and there is no doubt that it will not be identical to any of these simple mathematical functions.*

*Please note that we have revised the paragraphs about “Model-1” and “Model-2” in the Simulation section to clarify these points.*

**Our comment to authors' response:** We thank the authors for clarifying the paragraphs about

"Model-1" and "Model-2".

5. **Specific reviewer comment on the first version of the manuscript** "For the probabilities  $P_{\{DL\}}$  and  $P_{\{DP\}}$  the explicit formula to compute the corresponding probabilities could also be included within the paragraph analogous to  $P_{\{RF\}}$ . This would enable a clearer reproducibility."

**Authors' response:** Actually there is no formula for the gene deletion and duplication probabilities. They are given probability values that do not depend on any other parameters. At each round of the simulation, a gene is duplicated with probability  $P_{\{DP\}}$ . If not rewired or duplicated, the gene is deleted with probability  $P_{\{DL\}}$ .

**Our comment to authors' response:** We thank the authors for clarifying this point for us.

6. **Point#7 of our summarized comments to the first version of the manuscript:** "Explain why merging OrthoDB results with OrthoMCL results might be plausible and why comparing it with results obtained from phylostratigraphy is reliable."

**Authors' response to Point#7:** The phylostratigraphy technique as employed by Domazet-Loso et al. and Quint et al. involves the assignment of an evolutionary age to each gene in a given species, by tracing the most distant ancestral node containing at least one immediate daughter species with a detectable homologue. A phylogenetic tree of divergent species ranging from cellular organisms transitioning into simple and more complex eukaryota served as a reference to divide phyla radiating from consecutive ancestral points into distinct phylostratum layers. To determine the homologues, BLAST sequence similarity searches were done against complete genomes reliably annotated across the different phylostrata. Similarly, we obtained known homologues of each *Drosophila* gene from OrthoDB5 and OrthoMCL5 (that also use BLAST based searches) in the species depicted in Figure 4, and mapped them to six different "phylostrata" (marked in Figure 4). Akin to the method described in the above references, we assigned each gene to a specific "phylostratum" based on the farthest detectable homologue of that gene, and this is referred to as "age index" in the paper. This made it possible to compare the age index based results we obtain in *Drosophila* to the phylostratigraphy based assignments in *Arabidopsis*.

Homologues for *Drosophila* genes were sourced from these two databases, to increase coverage across the species depicted in the tree in Figure 4. Since these two databases use slightly different parameters to identify orthologs, we examined whether the usage of specific databases introduce bias by analyzing data from each database separately and observed that age index of each gene remained the same.

**Our comment to authors' response:** We are glad that the authors carefully addressed our initial concerns about the reliability of results obtained from merging methods of gene homology detection. Since the observed age indices are not influenced by the database used to assign gene homology relationships, we agree that merging OrthoDB5 and OrthoMCL5 results do not bias this specific analysis.

### Reviewer summary

All remaining concerns were carefully revised and we thank the authors for investing the additional time to respond to our review in such detail.

**Competing Interests:** No competing interests were disclosed.

**We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Referee Report 29 December 2014

doi:10.5256/f1000research.6375.r7095



**Gourab Ghoshal**

Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

The authors have addressed all my concerns appropriately. I approve this manuscript in its current form with no reservations.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

### Version 1

Referee Report 06 August 2014

doi:10.5256/f1000research.4902.r5391



**Gourab Ghoshal**

Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

The manuscript by Akhshabi *et al.* proposes and describes in detail an evolutionary model that qualitatively reproduces the "Hourglass Effect"---a phenomenon (in the biological context) whereby there is increased morphological divergence at early and late stages of embryonic development separated by increased conservation in the phylotypic stage.

The proposed model is based on a hierarchical network representation of the Gene Regulatory Network - the so-called Developmental Gene Execution Network (DGEN) - and evolution proceeds through a series of stochastic perturbations involving gene duplication, re-wiring and deletion. The manuscript suggests that the key factor that reproduces the hourglass effect is the monotonicity in "time" (increasing) of a function they call specificity  $s(i)$  (a measure of likelihood of a gene regulating descendant genes at later times) that competes against perturbative effects such that there is a "waist" of the number of transitions genes  $w(l)$  at intermediate times. They then test predictions of the model in two datasets of developmental data *Drosophila melanogaster* and *Arabidopsis Thaliana*, finding good qualitative agreement.

The manuscript is properly motivated and well written (with some caveats, see comments below). The model proposed is quite intuitive and compelling and represents an excellent first attempt in beginning to uncover the mechanism behind this intriguing phenomenon. I anticipate that this will have high impact (considering the hourglass effect transcends biological phenomena and is also found in "designed" systems such as the Internet *protocol stack*) in multiple fields. Consequently I strongly endorse this manuscript.

I do however have some concerns about some of the details about the model and therefore propose the following that may improve the readability of the manuscript and the plausibility and robustness of the model:

1. The section describing the model is a bit dense, and can benefit from some rewriting. For example it is a bit hard to follow what the significance of a spatial domain is and its relevance to the model. In addition a bewildering plethora of parameters are introduced, that can be a bit overwhelming for the reader. Consequently I suggest putting all the important and relevant parameters in a table for easy access to the reader. The section on Rewiring (RW) is quite hard to read and will benefit from the introduction of explicit equations similar to that for  $P_{\{RF\}}$ .
2. Which brings me to my second point about the assumptions of the model. It seems to me that two key factors lead to the hourglass effect: a) the monotonicity of  $s(l)$  with stage  $l$  and likewise b) of  $P_{\{RF\}}$  with  $r$ . While the choice of this dependence for the specificity is a simple linear dependence (which makes sense as a first pass), the choice of function for  $P_{\{RF\}}$  is highly non-trivial. Why not for example choose a simple linear dependence such as  $P_{\{RF\}} \sim r$ ? The reason why I say this is that the authors would like to propose the most general and minimal model to explain this phenomenon. And it is not clear to me whether the specific choice of the functional dependences matter. For example with a non-linear choice of  $P_{\{RF\}}$  we see that a non-linear  $s(l)$  has the effect of shifting the waist. But what about a linear choice of  $P_{\{RF\}}$  and  $s(l)$ ? Does that still preserve the hourglass effect? In my opinion a truly robust model will state that the main things that determine the hourglass are the monotonic dependencies in combination with non-linear or linear choices of the functions independent of their specific forms. This will allow the model to be applied to multiple settings.

At the moment I'm a bit concerned that the non-trivial choice of  $P_{\{RF\}}$ , as it stands in the manuscript, might be key to the observed effect. If it is indeed so, then one must motivate why one must make this particular choice. Same goes for the slightly peculiar choice for the non-linearity of  $s(l)$  introduced later in the model. The authors will do well to explain in detail what motivated them to make such a choice.

Additionally it would be helpful if the authors make it clearer (than they already have) that effects such as duplication and deletion are there only to make the model more realistic in terms of maintaining the number density of genes across different stages and play no part in the hourglass effect (or at least that's how I understand it).

3. Some minor nitpicks:
  1. In the section Model Description, the authors refer to the Wright-Fisher model. I was not aware of what this was and had to look it up, so a reference for the reader should be provided.
  2. The Hourglass score  $H$  is determined through Mann-Kendall statistics. A brief description of the method should be provided for the benefit of the reader (maybe in the supplementary material). Additionally, why this particular choice? Presumably any statistical test for monotonic trends should be robust to the parameters of the model. Would there be much of a difference is one used the Spearman rho for example?



3. Finally I think the authors will do well to highlight the importance of Specificity right at the outset of the manuscript and move some simplified variant of the segment "What is the reason behind the hourglass shape of DGENS?" (page 7, line 11, second column) to somewhere in the introduction.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 11 Dec 2014

**Constantine Dovolis**, Georgia Institute of Technology, USA

First, we would like to thank the author of this review for his very thorough reading of our paper. His comments and suggestions have been extremely helpful in revising the paper (please look at Version-2). We would also like to apologize for not revising the paper earlier.

- **Review:** *"The section describing the model is a bit dense, and can benefit from some rewriting. For example it is a bit hard to follow what the significance of a spatial domain is and its relevance to the model."*

**Response:** We hope our extensive revision has mitigated this problem of the previous version of the paper.

We explain what we mean by "spatial domains" early in the paper (and in Fig-1) because it is an important and necessary concept when we justify later in the paper the assumption of increasing regulatory specificity (see paragraph that starts with "The major assumption is that the regulatory specificity..."). It is the formation of those distinct spatial domains (or "modules") that explains why the regulatory specificity is probably decreasing as development progresses.

- **Review:** *"The section on Rewiring (RW) is quite hard to read and will benefit from the introduction of explicit equations similar to that for  $P_{RF}$ ."*

**Response:** We have revised that paragraph to make it more clear.

- **Review:** *It seems to me that two key factors lead to the hourglass effect: a) the monotonicity of  $s(l)$  with stage  $l$  and likewise b) of  $P_{RF}$  with  $r$ . While the choice of this dependence for the specificity is a simple linear dependence (which makes sense as a first pass), the choice of function for  $P_{RF}$  is highly non-trivial. Why not for example choose a simple linear dependence such as  $P_{RF} \sim r$ ? For example with a non-linear choice of  $P_{RF}$  we see that a non-linear  $s(l)$  has the effect of shifting the waist. But what about a linear choice of  $P_{RF}$  and  $s(l)$ ? Does that still preserve the hourglass effect? In my opinion a truly robust model will state that the main things that determine the hourglass are the monotonic dependencies in combination with non-linear or linear choices of the functions independent of their specific forms."*

**Response:** We chose the non-linear  $P_{RF}$  function shown in Fig-2 because it is quite flexible and it allows us to examine how the shape of this curve affects the behavior of the

model by simply varying the parameter  $z$ . Please note that when  $z$  is equal to one this function is almost linear. As discussed in Fig-10 however, as  $z$  decreases towards one the location of the hourglass waist also decreases. Consequently, we expect that a linearly increasing  $P_{\{RF\}}$  function would result in a shape that resembles a “funnel” rather than an hourglass.

We hope that new experimental work will provide some direct evidence for the shape of this function in the near future.

**Review:** *“In the section Model Description, the authors refer to the Wright-Fisher model. I was not aware of what this was and had to look it up, so a reference for the reader should be provided.”*

- **Response:** We have added a reference for this model:

“Ewens WJ. Mathematical Population Genetics. Springer-Verlag Berlin New York. 1979. [Reference Source](#)”

- **Review:** *“The Hourglass score  $H$  is determined through Mann-Kendall statistics. A brief description of the method should be provided for the benefit of the reader (maybe in the supplementary material). Additionally, why this particular choice? Presumably any statistical test for monotonic trends should be robust to the parameters of the model. Would there be much of a difference is one used the Spearman rho for example?”*

**Response:** We have added a reference for this statistic:

“Gibbons JD, Chakraborti S: Nonparametric Statistical Inference. Marcel Dekker New York. 2003. [Reference Source](#)”

The Mann-Kendall statistical test is one of the most widely used statistics to assess the significance of trends in data (Gibbons and Chakraborti 2003). Mann-Kendall and Spearman’s tests perform highly similar in data analyses (e.g., [Yue et al., 2002](#); [Shadmani et al., 2012](#)).

- **Review:** *“Finally I think the authors will do well to highlight the importance of Specificity right at the outset of the manuscript and move some simplified variant of the segment “What is the reason behind the hourglass shape of DGENS?” (page 7, line 11, second column) to somewhere in the introduction.”*

**Response:** The reviewer is right. We revised the fourth paragraph of the introduction to include a “hint” about the main reason behind the hourglass effect. Obviously, we cannot write much more at that early point of the paper because we still haven’t defined what we mean by DGEN, RFs, cascades, etc.

**Competing Interests:** No competing interests were disclosed.

doi:10.5256/f1000research.4902.r5650



**Dmitri Krioukov**

Department of Physics, Northeastern University, Boston, MA, USA

The article proposes and analyzes a network model that is capable of reproducing the hourglass effect in the number of genes that undergo a functional state transition as a function of the developmental stage. The hourglass effect refers to higher morphological divergence at earlier or later stages of embryonic developmental, compared to medium stages. The main idea of the model is to explain the waist of this hourglass (the number of "transitioning genes"  $w(l)$  having a minimum at medium developmental stages  $l$ ) by an interplay between the impact of random rewiring events in the developmental gene execution network (DGEN) and stage specificity. The DGEN is a network representing regulatory interactions between genes at different developmental stages, while stage specificity  $s(l)$  defines the probability  $1-s(l)$  with which a gene at stage  $l$  regulates a gene at stage  $l+1$ . One of the key assumptions in the model is that  $s(l)$  is an increasing function of  $l$ . As a result of the interplay between a decreasing impact of DGEN perturbations as a function of  $l$ , and increasing  $s(l)$ , the lethality of perturbations is maximized at mid  $l$ 's, leading to evolutionary incentives to minimize  $w(l)$  at those mid stages. The article consists of three parts. The first part describes the model in detail. The second part discusses extensive simulation results of the model. The third part presents an extensive analysis of existing developmental data on *Drosophila melanogaster* and *Arabidopsis thaliana* in the model context.

The most interesting aspect of the article is that provides a possible intriguing explanation of the hourglass effect in developmental biology, which may foster future creative thinking and research in this important direction. The most obvious reservation one can express about the article is that the model cannot be currently refuted since it is formulated at the DGEN level, and there is currently no data from which a reliable DGEN reconstruction would be possible. The fact that some outcomes of model simulations qualitatively agree with available data does not directly validate either the model, or its main assumptions. The article does not contain such claims, however, that would not be supported by the data.

My minor comments that may help to improve the article are:

1. The main idea of the model is buried somewhere on page 6 in the middle of a paragraph after sentence "What is the reason behind the hourglass shape of DGENs?" Why not explain this as early as possible, certainly in the Introduction, if not in the abstract?
2. The third paragraph in the Introduction reads somewhat unclear and imprecise, as well as the "Developmental gene execution networks" section.
3. When  $D(g)$  and  $\Gamma$  notations first appear, it's not spelled out what they are.
4. DGEN removals upon DF events have a clear meaning. But what do replacements correspond to in reality?
5. The Mann-Kendall statistics could be briefly described, or at least a reference could be provided.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 11 Dec 2014

**Constantine Dvrolis**, Georgia Institute of Technology, USA

First, we would like to thank the author of this review for his thorough reading of our work. His comments have been very helpful in revising this paper (please see Version 2). We would also like to apologize for not revising the paper earlier.

- **Review:** *"The most obvious reservation one can express about the article is that the model cannot be currently refuted since it is formulated at the DGEN level, and there is currently no data from which a reliable DGEN reconstruction would be possible. The fact that some outcomes of model simulations qualitatively agree with available data does not directly validate either the model, or its main assumptions. The article does not contain such claims, however, that would not be supported by the data."*

**Response:** It is true that a direct validation of the model is not possible with any currently available data (at least to the extent of our knowledge). Such a direct validation would require the reconstruction of the complete DGEN for specific model organisms, i.e., knowing not only the expression profile of each gene as a function of developmental time and at different tissues of the embryo but also knowing the regulatory inputs for each gene during development. However, there are research groups that are working in that direction (e.g., Prof. Davidson's group at Caltech is gradually "reverse-engineering" the DGEN of the sea-urchin). Our hope is that a direct validation of our model will be possible within the next few years.

Having said that however, even though our model cannot be directly validated, we are providing indirect evidence that certain predictions of the model are true in *Drosophila* and *Arabidopsis* (see Results - Data Analysis section). Consequently, the model should not be viewed as completely hypothetical either.

- **Review:** *"The main idea of the model is buried somewhere on page 6 in the middle of a paragraph after sentence "What is the reason behind the hourglass shape of DGGENs?" Why not explain this as early as possible, certainly in the Introduction, if not in the abstract?"*

**Response:** The reviewer is right. We revised the fourth paragraph of the introduction to include a "hint" about the main reason behind the hourglass effect. Obviously, we cannot write much more at that early point of the paper because we still haven't defined what we mean by DGEN, RFs, cascades, etc. The revised paragraph is:

"The model predicts that the evolutionary process shapes the DGGENs of a population in the form of an hourglass, under fairly general assumptions. Specifically, the number of genes at each developmental stage follows an hourglass pattern, with the smallest number at the "waist" of the hourglass. The main condition for the appearance of the hourglass pattern is that the DGEN should gradually get sparser as development progresses, with general-purpose regulatory genes at the earlier developmental stages and highly specialized regulatory genes at the later stages. Under this assumption, the model predicts that gene regulatory changes or rewiring in mid-development are more likely to cause cascades of removing non-essential genes from the DGEN, compared to early or late developmental stages. Another model prediction is that the evolutionary age of DGEN genes also follows an hourglass pattern, with the oldest genes concentrated at the waist."

- **Review:** “When  $D(g)$  and Gamma notations first appear, it's not spelled out what they are.”

**Response:** We checked this. The notation  $D(g)$  is introduced at the end of the “Developmental gene execution networks” section, and that is the first time it appears in the paper. The notation Gamma first appears in the “Developmental Failure (DF)” paragraph, and that is also where it is defined.

- **Review:** “DGEN removals upon DF events have a clear meaning. But what do replacements correspond to in reality?”

**Response:** DGEN replacements model the effect of selection: an embryo that did not develop properly (DF) dies and its DGEN is removed from the population, while healthy embryos develop and eventually give birth to new a embryo with the same genotype as its parent (asexual reproduction).

- **Review:** “The Mann-Kendall statistics could be briefly described, or at least a reference could be provided.”

**Response:** We have added a reference for this statistic:

“Gibbons JD, Chakraborti S: Nonparametric Statistical Inference. Marcel Dekker New York. 2003. [Reference Source](#)”

**Competing Interests:** None

Referee Report 24 July 2014

doi:10.5256/f1000research.4902.r5563



**Marcel Quint**<sup>1</sup>, **Hajk-Georg Drost**<sup>2</sup>

<sup>1</sup> Department of Molecular Signal Processing, Leibniz Institute of Plant Biochemistry, Halle, Germany

<sup>2</sup> Institute of Computer Science, Martin Luther University Halle–Wittenberg, Halle, Germany

This study by Akhshabi *et al.* aims to provide the community with an explanatory theoretical model for the developmental hourglass phenomenon. Motivated by the question concerning the regulatory mechanism causing the developmental hourglass pattern during embryogenesis, the authors developed an evolutionary model of regulatory gene interactions during development to identify the conditions under which the molecular hourglass effect might emerge in general.

For this purpose their evolutionary model focuses on hierarchical gene regulatory networks that control the corresponding developmental processes. Based on the output of the model the authors predict the emergence of an hourglass pattern in the structure of a temporal representation of the underlying gene regulatory network and correlate this effect with the evolution of protocol architecture found for the Internet’s “protocol stack”.

Based on this universal finding the authors speculate that the developmental hourglass pattern might be a causal effect of a common organization principle during the establishment of complexity and is based on

the underlying hierarchical network structure.

**The paper is very interesting, well written, and will potentially help the community to better understand the developmental hourglass phenomenon. However, we see some issues with the model assumptions and strongly suggest including the following aspects in a revised version.**

Summary of the points that have to be addressed:

1. Include a clear reference and paragraph to Raff's hypothesis and how it influenced the modeling of DGEN.
2. Correlate and reference major assumptions of the DGEN model that seem plausible with findings from experimental studies that validate this plausibility.
3. Correlate the output of the DGEN simulation with existing findings of studies addressing similar questions of the causality of the developmental hourglass phenomenon and discuss potential simplifications of the model assumptions that do not fit experimental results or that are not derivable via experiments yet.
4. Include a clear statement how the exact values of the stage specificity (regulatory specificity)  $s(l)$  have been derived.
5. State more clearly that  $s(l)$  is the crucial parameter determining the observed phenomena.
6. Correct the TAI formula.
7. Explain why merging OrthoDB results with OrthoMCL results might be plausible and why comparing it with results obtained from phylostratigraphy is reliable.
8. Discuss more clearly how the DGEN model can help to find causal processes of the developmental hourglass phenomenon based on recent knowledge: the three predictions of the DGEN model a) – c) mentioned in the Discussion section seem not integrative enough to allow future studies to correlate existing results provided by the community with the predictions obtained by the DGEN model.

Specific comments to the separate sections:

In general, most of our criticisms concern the model description and the methods section. Both are integral to the validity of the obtained results, which are nicely reported and discussed.

## Introduction

Modeling the evolution of embryonic development as directed acyclic graph (DAG) in which the nodes correspond to state-transitioning genes and the edges model directed regulatory interactions causing significant activity change at the corresponding stage seems plausible. The authors refer to this model as DGEN. A testable hypothesis proposed by the authors is that if regulatory genes become increasingly function-specific as development progresses, the network gradually should become sparser in that direction and evolutionary older genes should be concentrated at the waist of the hourglass (Figure 1).

This hypothesis is essential to detect possible regulatory mechanisms that cause the developmental hourglass pattern. We therefore suggest including references to existing biological studies investigating the phenomenon of morphological complexity being correlated with regulatory specialization to support the validity and plausibility of their DGEN model assumptions.

A biological motivation of this hypothesis comes from Rudolf A. Raff who proposed that early embryogenesis is governed by global pattern formation processes and that during later development, the embryo is being organized into developmental modules, allowing each module to differentiate autonomously. Raff furthermore proposed that the phylotypic stage marks the transition from early global

development to the modular mode later in development<sup>1</sup>.

The authors do cite Raff in the first sentence of their introduction, but should state more clearly in which manner their model was inspired by Raff's hypothesis.

Numerous studies investigated the phenomenon that increasing morphological complexity is correlated with increasing modularity and specialization of the underlying regulatory apparatus<sup>2</sup>. Here the authors predict that the main condition leading to the appearance of the hourglass pattern is that the DGEN should gradually get sparser as development progresses, with general-purpose regulatory genes at the earlier developmental stages and highly specialized regulatory genes at the later stages, expressing the oldest genes during mid development.

We suggest to also include findings investigating the functionality and age of the genes that have been shown to be active during mid embryogenesis to correlate potential outcomes of the DGEN prediction that evolutionary older genes should be concentrated at the waist of the hourglass with results obtained by recent studies<sup>3 4 5 6 7 8 9 10 11 12</sup>. These aforementioned studies were able to show that during mid embryogenesis evolutionary old genes or evolutionary old processes that are shared within and between phyla are most active.

In our opinion it would be very interesting to test the degree of intersection between genes predicted to be causal for the waist of the DGEN hourglass and evolutionary conserved genes that have been found to be expressed during mid embryogenesis. We are aware that the model uses the number of genes at each developmental stage combined with the age of the corresponding genes to observe constraints during development. Hence, only the authors can estimate how much effort would be needed to identify such intersecting genes.

If the search for intersecting genes returned by the DGEN model during mid development and genes found by previous studies should be possible, the authors could integrate a motivation section to find these intersecting genes within the Introduction section.

### Model description

The model description is quite intuitive. However, as the validity of the whole study depends on the model parameters chosen, several points need clarification including the rationale and motivation to choose exactly these parameters.

One of the most crucial parameters of the model is the regulatory specificity  $s(l)$ , with  $0 \leq s(l) \leq 1$ . We think, that adding a more detailed paragraph discussing how  $s(l)$  is being estimated, would be highly beneficial due to the importance for the parameter  $s(l)$  for the entire model and predictions. In Model-1-4 the authors choose three different values for  $s(l)$ : In Model-1:  $s(l) = 0.5$  and in Model-2-4:  $s(l) = l/L$ , and in Figure 9 a non-linear function  $s(l) = 0.9 - (0.8 / (1 + \exp(\gamma - l)))$ . A gene  $g$  at stage  $l$  is modeled to act as upstream regulator for a gene  $g'$  at stage  $l + 1$  with probability  $s'(l) = 1 - s(l)$ . From this follows that each potential upstream regulator in stage  $l$  has the same probability  $s'(l)$  to regulate genes  $g'$  in stage  $l + 1$ . In other words, Model-1 assigns each potential upstream regulator with the same probability  $s'(l) = 1 - 0.5 = 0.5$  to regulate genes  $g'$  in stage  $l + 1$ , in Model-2:  $s'(l) = 1 - l/L = \text{const.}$ , and for the non-linear function  $s(l)$ :  $s'(l) = 1 - (0.9 - (0.8 / (1 + \exp(\gamma - l)))) = \text{const.}$

It would be advantageous, if the authors could include motivations for the three cut-offs, especially how they derived the sigmoid-like mathematical function (Figure 10, legend caption) they later denote as

non-linear function  $s(l)$ . We do see that modeling the stage specificity with a constant value, an increasing value, and a non-linear function is plausible, but a clear motivation or derivation of the exact values would be beneficial for a better understanding of the modeling process.

For the probabilities  $P_{DL}$  and  $P_{DP}$  the explicit formula to compute the corresponding probabilities could also be included within the paragraph analogous to  $P_{RF}$ . This would enable a clearer reproducibility.

The authors discuss that their major assumption is that the regulatory specificity increases substantially as development progresses, hence the DGEN becomes gradually sparser along the developmental time axis starting with  $s(1) \sim 0, \dots, s(L) \sim 1$ . Here the authors should include references to experimental observations that support their assumption. Since Models 2-4 rely on this assumption, published experimental studies in line with this assumption need to be referenced. Otherwise, the models applied would seem largely academic.

## Methods

### Hourglass score H

Here  $w(l)$  is defined as the number of transitioning genes in stage  $l$ . In Figure 3 and in section 'Hourglass shape'  $w(l)$  is referred to as *stage width*. In case the definition of  $w(l)$  is equivalent to the *stage width*, it would be beneficial for the reader to see the term *stage width* in the same sentence as the initial definition of  $w(l)$ . A correct understanding of the *stage width* is important to understand the measurement of  $H$ .

Furthermore, the univariate Mann-Kendall statistic should be referenced to allow a non-statistical community to better understand this way of statistical modeling.

### Transcriptome age index (TAI)

In this section the authors describe and perform two different methods to assign each gene a corresponding evolutionary age. For *D. melanogaster* they collected groups of orthologs from OrthoDB and OrthoMCL, whereas for *A. thaliana* they obtained phylostratum assignments from Quint et al.<sup>10</sup> that was derived by phylostratigraphy[ref-13.] We do not understand the motivation of merging orthologous gene groups obtained from OrthoDB and OrthoMCL, as both approaches apply different parameters and heuristics to determine orthologous genes. Was this done to maximize the number of potential orthologs to be included in the study? In addition, why did the authors use a different approach for the plant gene set? This seems somewhat inconsistent and statement concerning the plausibility to compare results returned by the DGEN model based on age assignments obtained from phylostratigraphy (*A. thaliana*) with OrthoDB and OrthoMCL merged orthologs (*D. melanogaster*) would be appreciated.

Furthermore, the database specific parameters they used in OrthoDB and OrthoMCL should be mentioned as well as the database version they used, e.g. OrthoDB2, or OrthoDB3, ... , or OrthoDB6.

### Age index for each stage-pair

There seems to be a typo within the TAI formula. As defined by Domazet-Lošo and Tautz<sup>6</sup> and adapted to the author's notation the formula of the TAI should begin  $TAI(l)$  not  $TAI(1)$ .

## Results



The authors show based on their findings returned by Model-2-4 that the increasing specificity assumption is the key property behind the developmental hourglass effect. We think that this statement is not clear enough. The authors could phrase more clearly what they mean by “*key property behind the hourglass effect*” inferred from Model-2-4.

### Age of genes

Here the authors motivate that defining the age of a gene as  $A(g) = i - t_0(g)$  is plausible, because a rewiring event may give a gene a new function, at least in terms of its upstream and downstream regulators. This statement needs validation from the literature and from experimental studies that already tested this hypothesis for specific rewiring events.

Furthermore, the authors show (Figure 8c) that the evolutionary age at stage I follows the same pattern as the lethality probability: gradually increasing until the waist of the hourglass, and then gradually decreasing. This finding should be correlated with the findings by Galis and Metz<sup>3</sup> and should be discussed more clearly in the Discussion.

In the first sentence of the last paragraph before the Discussion, “... and TAI profiles from *Arabidopsis*” should be referenced to reference 11 and not 12.

### **Discussion**

The authors discuss that the observed pattern of conservation (developmental hourglass) may stem from fundamental organization principles and that the exact origin of these principles remains elusive. Here the authors should include existing studies that also aim at predicting a universal organization principle during development (modularization, etc.) and what exact problems are preventing studies to elucidate this universal phenomenon. We do note, that the authors cite studies that aim to explain the effects causing the early phase of the developmental hourglass<sup>141516</sup> and studies aiming to explain effects causing the late phase of the developmental hourglass<sup>1</sup>, but we think that it should be explicitly stated that the model is designed to test aspects of Raff’s initial hypothesis and that also more recent studies have been investigating the underlying reasons for the pattern in the early phase of the hourglass<sup>17</sup>.

### **References**

1. Raff RA: The Shape of Life: Genes, Development and the Evolution of Animal Form. (Univ. Chicago Press). 1996.
2. Wagner GP, Pavlicev M, Cheverud JM: The road to modularity. *Nature Reviews Genetics*. 2007; **8** (12): 921-931 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Galis F, Metz JA: Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation. *J Exp Zool*. 2001; **291** (2): 195-204 [PubMed Abstract](#)
4. Hazkani-Covo E, Wool D, Graur D: 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*. 2005; **304** (2): 150-158 [PubMed Abstract](#) | [Publisher Full Text](#)
5. Roux J, Robinson-Rechavi M: Developmental constraints on vertebrate genome evolution. *PLoS Genet*. 2008; **4** (12): e1000311 [PubMed Abstract](#) | [Free Full Text](#) | [Publisher Full Text](#)
6. Domazet-Lošo T, Tautz D: A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature*. 2010; **468** (7325): 815-818 [PubMed Abstract](#) | [Publisher Full Text](#)
7. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, Ohler U, Bergman CM,

- Tomancak P: Gene expression divergence recapitulates the developmental hourglass model. *Nature*. 2010; **468** (7325): 811-814 [PubMed Abstract](#) | [Publisher Full Text](#)
8. Irie N, Kuratani S: Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat Commun*. 2011; **2** (248). [PubMed Abstract](#) | [Free Full Text](#) | [Publisher Full Text](#)
9. Levin M, Hashimshony T, Wagner F, Yanai I: Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Dev Cell*. 2012; **22** (5): 1101-1108 [PubMed Abstract](#) | [Publisher Full Text](#)
10. Quint M, Drost HG, Gabel A, Ullrich KK, Bönn M, Grosse I: A transcriptomic hourglass in plant embryogenesis. *Nature*. 2012; **490** (7418): 98-101 [PubMed Abstract](#) | [Publisher Full Text](#)
11. Schep AN, Adryan B: A comparative analysis of transcription factor expression during metazoan embryonic development. *PLoS One*. 2013; **8** (6): e66826 [PubMed Abstract](#) | [Free Full Text](#) | [Publisher Full Text](#)
12. Piasecka B, Lichocki P, Moretti S, Bergmann S, Robinson-Rechavi M: The hourglass and the early conservation models--co-existing patterns of developmental constraints in vertebrates. *PLoS Genet*. 2013; **9** (4): e1003476 [PubMed Abstract](#) | [Free Full Text](#) | [Publisher Full Text](#)
13. Domazet-Lošo T, Brajković J, Tautz D: A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet*. 2007; **23** (11): 533-539 [PubMed Abstract](#) | [Publisher Full Text](#)
14. Riedl R, Jefferies RP: Order in living organisms: a systems analysis of evolution. *Wiley New York*. 1978.
15. Arthur W: Mechanisms of morphological evolution: a combined genetic, developmental, and ecological approach. *Wiley New York*. 1984.
16. Rasmussen N: A new model of developmental constraints as applied to the *Drosophila* system. *J Theor Biol*. 1987; **127** (3): 271-299 [PubMed Abstract](#)
17. Kalinka AT, Tomancak P: The evolution of early animal embryos: conservation or divergence?. *Trends Ecol Evol*. 2012; **27** (7): 385-393 [PubMed Abstract](#)

**Competing Interests:** No competing interests were disclosed.

**We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

Author Response 11 Dec 2014

**Constantine Dovrolis**, Georgia Institute of Technology, USA

First, we would like to thank the two authors of this review for their very thorough reading of our paper. Their comments and suggestions have been extremely helpful in revising the paper (please look at Version-2). We would also like to apologize for not revising the paper earlier.

- **Review:** "Modeling the evolution of embryonic development as directed acyclic graph (DAG) in which the nodes correspond to state-transitioning genes and the edges model directed regulatory interactions causing significant activity change at the corresponding stage seems plausible. The authors refer to this model as DGEN. A testable hypothesis proposed by the authors is that if regulatory genes become increasingly function-specific as development progresses, the network gradually should become sparser in that direction and evolutionary older genes should be concentrated at the waist of the hourglass (Figure 1). This hypothesis is essential to detect possible regulatory mechanisms that cause the developmental hourglass pattern. We therefore suggest including references to existing biological studies investigating the phenomenon of morphological complexity being correlated with regulatory specialization to support the validity and plausibility of their DGEN

*model assumptions. A biological motivation of this hypothesis comes from Rudolf A. Raff who proposed that early embryogenesis is governed by global pattern formation processes and that during later development, the embryo is being organized into developmental modules, allowing each module to differentiate autonomously. Raff furthermore proposed that the phylotypic stage marks the transition from early global development to the modular mode later in development.”*

**Summary - Point#1:** *“The authors do cite Raff in the first sentence of their introduction, but should state more clearly in which manner their model was inspired by Raff’s hypothesis.”*

**Summary - Point#2:** *“Correlate and reference major assumptions of the DGEN model that seem plausible with findings from experimental studies that validate this plausibility.”*

**Response to Point #1:** The reviewers are right to highlight the connection between our “increasing regulatory specificity” assumption and Raff’s hypothesis. We have included a paragraph in the Discussion section to put Raff’s hypothesis in the context of this work. A first remark is that our assumption is consistent with Raff’s basic premise that modularity increases as the embryo develops. Second, Raff’s hypothesis (see chapter-6 of [1], and in particular Figures 6.6 and 6.7) also states that the “level of interaction” or the “interconnectivity between (body) elements” is maximized at the phylotypic stage. This may be viewed initially as a contradiction between Raff’s hypothesis and our increasing specificity assumption, which states that the density of regulatory interactions is maximized at the earliest stages of development. Note however that Raff’s hypothesis was not stated in terms of gene regulatory interactions -- he was referring more generally to “developmental flexibility”, arguing that early development is flexible because it governs robust and general global patterning processes, late development is also flexible because “signaling events within the primordia are little influenced by events in other primordia”, while mid-development (phylotypic stage) is least flexible because of the “high interconnectivity between elements that will later come to represent separate modules.” If we think of “developmental flexibility” as the ability of an embryo to survive gene mutations and rewiring at different stages of the developmental process, Raff’s hypothesis is actually consistent with our results regarding the “lethality probability” at each developmental stage (see Figs 6-B, 7-B and 8-B). The lethality probability is maximized at the phylotypic stage, and it is significantly lower at early and late developmental stages, following the same pattern with Raff’s “developmental flexibility”.

**Response to Point #2:** First, to the extent of our knowledge there is no prior work that **directly** validates the assumption of increasing regulatory specificity as formulated in our paper (i.e., with the specificity of a developmental stage defined as the topological density of regulatory edges in that stage). We are planning to examine the validity of this assumption in the future using various approaches:

First we plan to analyze the aforementioned developmental GRNs for the sea-urchin that have been directly examined by Dr. Davidson’s lab at Caltech.

Second, we are in the process of examining other genomic metrics, which can be viewed as “proxies” to regulatory specificity. For example, we have defined new measures to quantify how specific a gene’s expression profile in a developmental stage is compared to other developmental stages, which we have tentatively named as “developmental stage

specificity index (DSI)". Similarly, we defined "tissue specificity index (TSI)", measuring the bias of a gene's expression in a specific tissue compared to many other tissues. Analyses of these measures using genomic data from *Drosophila* so far have revealed generally increasing trends, which are consistent with the idea of increasing specificity. However we are still working on generalizing these analytical tools to data sets generated by different methods and from different species. Importantly, our model focuses DGEnS underlying development, and genomic data can obscure the true signals from genes constituting DGEnS. Thus we are also working on examining these specificity measures using inferred DGEnS. Additionally, we are developing other methods to approximate functional specificity. We hope to complete these analyses as a follow-up paper. The current paper provides theoretical motivation.

Third, as the reviewers also point out, there is a large body of prior work in developmental biology that confirms the increasing modularity in the developing embryo (for instance, see [Wagner et al., 2007](#)) as well as the increasing specificity of the developmental process at the signaling or genomic level. The connections between this pattern of increasing modularity and the structure of the underlying gene regulatory networks are still not well understood however. We think that it would be misleading if we had argued that these modularity patterns provide a direct validation for our "increasing regulatory specificity" assumption. Please note that we have cited the paper by Wagner et al. at the Introduction and Model Description sections, providing some connections between this work and earlier work on the role of modularity in developmental biology.

- **Review:** *"Numerous studies investigated the phenomenon that increasing morphological complexity is correlated with increasing modularity and specialization of the underlying regulatory apparatus<sup>2</sup>. Here the authors predict that the main condition leading to the appearance of the hourglass pattern is that the DGEN should gradually get sparser as development progresses, with general-purpose regulatory genes at the earlier developmental stages and highly specialized regulatory genes at the later stages, expressing the oldest genes during mid development. We suggest to also include findings investigating the functionality and age of the genes that have been shown to be active during mid embryogenesis to correlate potential outcomes of the DGEN prediction that evolutionary older genes should be concentrated at the waist of the hourglass with results obtained by recent studies<sup>3 4 5 6 7 8 9 10 11 12</sup>. These aforementioned studies were able to show that during mid embryogenesis evolutionary old genes or evolutionary old processes that are shared within and between phyla are most active. In our opinion it would be very interesting to test the degree of intersection between genes predicted to be causal for the waist of the DGEN hourglass and evolutionary conserved genes that have been found to be expressed during mid embryogenesis. We are aware that the model uses the number of genes at each developmental stage combined with the age of the corresponding genes to observe constraints during development. Hence, only the authors can estimate how much effort would be needed to identify such intersecting genes. If the search for intersecting genes returned by the DGEN model during mid development and genes found by previous studies should be possible, the authors could integrate a motivation section to find these intersecting genes within the Introduction section."*

**Summary:** *"Correlate the output of the DGEN simulation with existing findings of studies addressing similar questions of the causality of the developmental hourglass phenomenon and discuss potential simplifications of the model assumptions that do not fit experimental*

*results or that are not derivable via experiments yet.”*

**Response:** This is a very interesting idea and we thank the reviewers for this suggestion. We are planning to pursue this investigation in a follow-up paper that will also include the previously mentioned analysis of sea-urchin GRNs as well as the DSI/TSI results. We believe that it would be distracting to try to include all these results in this first paper, given that the main focus of this work has been on the computational model and the computational results. We are also open to collaborate in this investigation with the reviewers or other researchers working in this area, if they are interested.

- **Review:** *“The model description is quite intuitive. However, as the validity of the whole study depends on the model parameters chosen, several points need clarification including the rationale and motivation to choose exactly these parameters. One of the most crucial parameters of the model is the regulatory specificity  $s(l)$ , with  $0 \leq s(l) \leq 1$ . We think, that adding a more detailed paragraph discussing how  $s(l)$  is being estimated, would be highly beneficial due to the importance for the parameter  $s(l)$  for the entire model and predictions. In Model-1-4 the authors choose three different values for  $s(l)$ : In Model-1:  $s(l) = 0.5$  and in Model-2-4:  $s(l) = l/L$ , and in Figure 9 a non-linear function  $s(l) = 0.9 - (0.8 / (1 + \exp(\gamma - l)))$ . A gene  $g$  at stage  $l$  is modeled to act as upstream regulator for a gene  $g'$  at stage  $l + 1$  with probability  $s'(l) = 1 - s(l)$ . From this follows that each potential upstream regulator in stage  $l$  has the same probability  $s'(l)$  to regulate genes  $g'$  in stage  $l + 1$ . In other words, Model-1 assigns each potential upstream regulator with the same probability  $s'(l) = 1 - 0.5 = 0.5$  to regulate genes  $g'$  in stage  $l + 1$ , in Model-2:  $s'(l) = 1 - l/L = \text{const.}$ , and for the non-linear function  $s(l)$ :  $s'(l) = 1 - (0.9 - (0.8 / (1 + \exp(\gamma - l)))) = \text{const.}$  It would be advantageous, if the authors could include motivations for the three cut-offs, especially how they derived the sigmoid-like mathematical function (Figure 10, legend caption) they later denote as non-linear function  $s(l)$ . We do see that modeling the stage specificity with a constant value, an increasing value, and a non-linear function is plausible, but a clear motivation or derivation of the exact values would be beneficial for a better understanding of the modeling process.*

*The authors discuss that their major assumption is that the regulatory specificity increases substantially as development progresses, hence the DGEN becomes gradually sparser along the developmental time axis starting with  $s(1) \sim 0, \dots, s(L) \sim 1$ . Here the authors should include references to experimental observations that support their assumption. Since Models 2-4 rely on this assumption, published experimental studies in line with this assumption need to be referenced. Otherwise, the models applied would seem largely academic.*

*Include a clear statement how the exact values of the stage specificity (regulatory specificity)  $s(l)$  have been derived. State more clearly that  $s(l)$  is the crucial parameter determining the observed phenomena.”*

**Response:** The regulatory specificity functions we experimented with are arguably the simplest one could think of. Specifically, we start (Model-1) with the simplest hypothesis that the regulatory specificity does not vary across developmental stages (in the paper we show results for  $s=0.5$  but we have generated results for all values 0.1, 0.2, ... 0.9 and the conclusion remains that when the specificity is constant the network does not evolve into an hourglass shape). Then, we consider (Model-2) a linearly increasing specificity function. We

do not argue that this is realistic or that there are experimental results that suggest this linearity or even the increasing trend. It is just the simplest and most parsimonious model that leads to the emergence of an hourglass pattern. Finally, in Figures-9 and 10 we consider a more general specificity function that increases in a non-linear manner, providing a simple way to control the stage at which it gives the mid-range value 0.5. This allows us to examine how the shape of the specificity function  $s(l)$  affects the location of the hourglass waist. We hope that follow-up work will reveal experimentally the actual shape of the regulatory specificity function  $s(l)$  and there is no doubt that it will not be identical to any of these simple mathematical functions.

Please note that we have revised the paragraphs about “Model-1” and “Model-2” in the Simulation section to clarify these points.

- **Review:** *“For the probabilities  $P_{\{DL\}}$  and  $P_{\{DP\}}$  the explicit formula to compute the corresponding probabilities could also be included within the paragraph analogous to  $P_{\{RF\}}$ . This would enable a clearer reproducibility.”*

**Response:** Actually there is no formula for the gene deletion and duplication probabilities. They are given probability values that do not depend on any other parameters. At each round of the simulation, a gene is duplicated with probability  $P_{\{DP\}}$ . If not rewired or duplicated, the gene is deleted with probability  $P_{\{DL\}}$ .

- **Review:** *“Hourglass score  $H$ : Here  $w(l)$  is defined as the number of transitioning genes in stage  $l$ . In Figure 3 and in section ‘Hourglass shape’  $w(l)$  is referred to as stage width. In case the definition of  $w(l)$  is equivalent to the stage width, it would be beneficial for the reader to see the term stage width in the same sentence as the initial definition of  $w(l)$ . A correct understanding of the stage width is important to understand the measurement of  $H$ .”*

**Response:** The reviewer is right. We will use the “stage width” term when we first introduce the notation  $w(l)$ .

- **Review:** *“Furthermore, the univariate Mann-Kendall statistic should be referenced to allow a non-statistical community to better understand this way of statistical modeling.”*

**Response:** Please see how we addressed a similar comment (comment #5) in Dmitri Krioukov's review.

- **Review:** *“Transcriptome age index (TAI): In this section the authors describe and perform two different methods to assign each gene a corresponding evolutionary age. For *D. melanogaster* they collected groups of orthologs from OrthoDB and OrthoMCL, whereas for *A. thaliana* they obtained phylostratum assignments from Quint et al. [10](#) that was derived by phylostratigraphy[ref-13.] We do not understand the motivation of merging orthologous gene groups obtained from OrthoDB and OrthoMCL, as both approaches apply different parameters and heuristics to determine orthologous genes. Was this done to maximize the number of potential orthologs to be included in the study? In addition, why did the authors use a different approach for the plant gene set? This seems somewhat inconsistent and statement concerning the plausibility to compare results returned by the DGEN model based on age assignments obtained from phylostratigraphy (*A. thaliana*) with OrthoDB and OrthoMCL merged orthologs (*D. melanogaster*) would be appreciated. Furthermore, the*

*database specific parameters they used in OrthoDB and OrthoMCL should be mentioned as well as the database version they used, e.g. OrthoDB2, or OrthoDB3, ... , or OrthoDB6."*

**Summary:** *"Explain why merging OrthoDB results with OrthoMCL results might be plausible and why comparing it with results obtained from phylostratigraphy is reliable."*

**Response:** The phylostratigraphy technique as employed by Domazet-Loso et al. and Quint et al. involves the assignment of an evolutionary age to each gene in a given species, by tracing the most distant ancestral node containing at least one immediate daughter species with a detectable homologue. A phylogenetic tree of divergent species ranging from cellular organisms transitioning into simple and more complex eukaryota served as a reference to divide phyla radiating from consecutive ancestral points into distinct phylostratum layers. To determine the homologues, BLAST sequence similarity searches were done against complete genomes reliably annotated across the different phylostrata. Similarly, we obtained known homologues of each *Drosophila* gene from OrthoDB5 and OrthoMCL5 (that also use BLAST based searches) in the species depicted in Figure 4, and mapped them to six different "phylostrata" (marked in Figure 4). Akin to the method described in the above references, we assigned each gene to a specific "phylostratum" based on the farthest detectable homologue of that gene, and this is referred to as "age index" in the paper. This made it possible to compare the age index based results we obtain in *Drosophila* to the phylostratigraphy based assignments in *Arabidopsis*.

Homologues for *Drosophila* genes were sourced from these two databases, to increase coverage across the species depicted in the tree in Figure 4. Since these two databases use slightly different parameters to identify orthologs, we examined whether the usage of specific databases introduce bias by analyzing data from each database separately and observed that age index of each gene remained the same.

- **Review:** *"Age index for each stage-pair: There seems to be a typo within the TAI formula. As defined by Domazet-Loso and Tautz<sup>6</sup> and adapted to the author's notation the formula of the TAI should begin TAI(l) not TAI(1)."*

**Response:** We have fixed this problem. Thank you.

- **Review:** *"The authors show based on their findings returned by Model-2-4 that the increasing specificity assumption is the key property behind the developmental hourglass effect. We think that this statement is not clear enough. The authors could phrase more clearly what they mean by "key property behind the hourglass effect" inferred from Model-2-4."*

**Response:** The reviewer is right. We revised that paragraph as follows:

"In Model-1 and Model-2, genes can be only removed (due to RW events, potentially followed by RF cascades) and so the average DGEN size decreases as evolutionary time progresses, which is unrealistic. Model-3 and Model-4 are more realistic because they can maintain a roughly constant DGEN size in the long-term. However, as will be shown next, all aspects of the developmental hourglass effect can already be seen with Model-2 (but not with Model-1). This highlights that the increasing specificity assumption is sufficient to generate the hourglass effect. Further, the inclusion of additional biological mechanisms in

the model, namely gene duplication and gene deletion, even though they make the model more realistic, they are not necessary for the emergence of the hourglass effect."

- **Review:** "Age of genes: Here the authors motivate that defining the age of a gene as  $A(g) = i - t_0(g)$  is plausible, because a rewiring event may give a gene a new function, at least in terms of its upstream and downstream regulators. This statement needs validation from the literature and from experimental studies that already tested this hypothesis for specific rewiring events."

**Response:** We have added two references that provide some evidence to the previous hypothesis: [Guet et al. \(2002\)](#) and [Kim et al. \(2012\)](#).

- **Review:** "Furthermore, the authors show (Figure 8c) that the evolutionary age at stage  $l$  follows the same pattern as the lethality probability: gradually increasing until the waist of the hourglass, and then gradually decreasing. This finding should be correlated with the findings by Galis and Metz<sup>3</sup> and should be discussed more clearly in the Discussion."

**Response:** The reviewer is right. The findings of Galis and Metz are highly relevant and consistent with our computational results regarding the lethality probability. We modified the paper as follows:

In the Results section:

"These computational results for the lethality probability across development are consistent with the empirical observations of Galis and Metz \cite{galis2001testing} about the increased mortality of rodents due to perturbations in the phylotypic stage (see Discussion section)."

In the Discussion section:

"The third prediction of the model is in direct agreement with the empirical observations of Galis and Metz regarding the increased mortality caused by perturbations during the phylotypic stage \cite{galis2001testing}. That study has shown, based on the teratological literature for rodent development, that disturbances in the phylotypic stage lead to much higher mortality than in other stages. Further, such disturbances lead to the co-occurrence of several distinct anomalies in the developing embryo and so the increased mortality cannot be due to a single particularly vulnerable process that takes place at that stage."

- **Review:** "In the first sentence of the last paragraph before the Discussion, "... and TAI profiles from Arabidopsis" should be referenced to reference 11 and not 12."

**Response:** Thank you, we have fixed this problem.

**Review:** "Discussion: The authors discuss that the observed pattern of conservation (developmental hourglass) may stem from fundamental organization principles and that the exact origin of these principles remains elusive. Here the authors should include existing studies that also aim at predicting a universal organization principle during development (modularization, etc.) and what exact problems are preventing studies to elucidate this universal phenomenon. We do note, that the authors cite studies that aim to explain the



*effects causing the early phase of the developmental hourglass<sup>14 15 16</sup> and studies aiming to explain effects causing the late phase of the developmental hourglass<sup>1</sup>, but we think that it should be explicitly stated that the model is designed to test aspects of Raff's initial hypothesis and that also more recent studies have been investigating the underlying reasons for the pattern in the early phase of the hourglass<sup>17</sup>."*

**Response:** The reviewer is right. We have revised the Discussion section to also include a paragraph about Raff's hypothesis (see our response to comment 1). We have also added another paragraph (see below) to discuss a very recent paper that proposes a different mechanism for the emergence of the hourglass effect.

"Recently, Friedlander *et al.* have proposed a different mechanism for the emergence of hourglass-like patterns in evolving biological or technological networks, based on a linear system formulation [\cite{friedlander2014evolution}](#). They showed that if a system can be represented as a hierarchical and layered linear transformation of an input vector to an output vector, and the desired transformation matrix is rank-deficient, then an evolutionary process that selects that particular transformation can, under certain conditions, converge to an hourglass-like structure. It is not clear yet how to adapt this linear model in the context of inherently non-linear systems, such as gene regulatory networks."

**Competing Interests:** No competing interests were disclosed.

---