

RESEARCH

Open Access



# Stability in gene expression and body-plan development leads to evolutionary conservation

Yui Uchida<sup>1\*</sup>, Hiroyuki Takeda<sup>2</sup>, Chikara Furusawa<sup>1,3</sup> and Naoki Irie<sup>2,3\*</sup>

## Abstract

**Background** Phenotypic evolution is mainly explained by selection for phenotypic variation arising from factors including mutation and environmental noise. Recent theoretical and experimental studies have suggested that phenotypes with greater developmental stability tend to have a constant phenotype and gene expression level within a particular genetic and environmental condition, and this positively correlates with stronger evolutionary conservation, even after the accumulation of genetic changes. This could reflect a novel mechanism that contributes to evolutionary conservation; however, it remains unclear whether developmental stability is the cause, or whether at least it contributes to their evolutionary conservation. Here, using Japanese medaka lines, we tested experimentally whether developmental stages and gene expression levels with greater stability led to their evolutionary conservation.

**Results** We first measured the stability of each gene expression level and developmental stage (defined here as the whole embryonic transcriptome) in the inbred F0 medaka population. We then measured their evolutionary conservation in the F3 generation by crossing the F0 line with the distantly related Japanese medaka line (Teradomori), followed by two rounds of intra-generational crossings. The results indicated that the genes and developmental stages that had smaller variations in the F0 generation showed lower diversity in the hybrid F3 generation, which implies a causal relationship between stability and evolutionary conservation.

**Conclusions** These findings suggest that the stability in phenotypes, including the developmental stages and gene expression levels, leads to their evolutionary conservation; this most likely occurs due to their low potential to generate phenotypic variation. In addition, since the highly stable developmental stages match with the body-plan-establishment stage, it also implies that the developmental stability potentially contributed to the strict conservation of animal body plan.

**Keywords** Developmental stability, Evolutionary conservation, Canalization, Phenotypic evolution, Phylotypic period, Hourglass model, Transcriptome

\*Correspondence:

Yui Uchida

yui.uchida@riken.jp

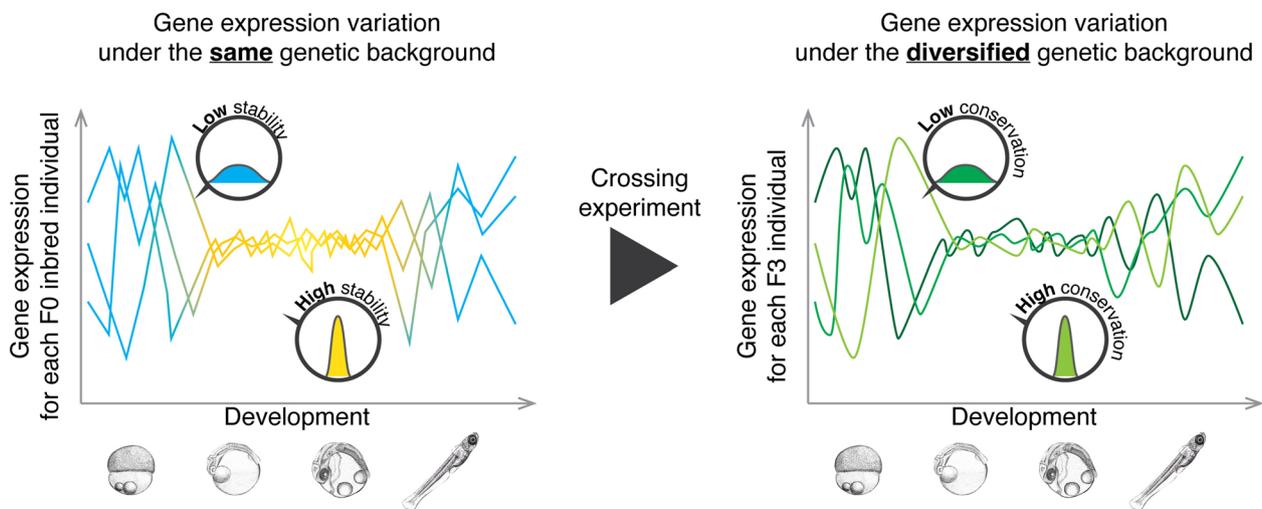
Naoki Irie

irie@bs.s.u-tokyo.ac.jp

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



**Fig. 1** The hypothetical relationship between developmental stability in the ancestors and conservation in the descendants. Schematic representation of gene expression variation under the same genetic background (variation between F0 inbred embryos, left) and predicted variation under the diversified genetic background (variation between F3 hybrid embryos, right). Previous study(22) using medaka showed lower transcriptomic variation (high developmental stability) during the body-plan developing phase than the earlier and later stages. In the present study, we hypothesized that developmental stages with higher stability would result in more conserved status after the crossing experiment

## Introduction

Phenotypic evolution depends fundamentally on phenotypic variation, which gives rise to the next generation's phenotype via selection and is affected by population genetics-based effects [1]. Phenotypic variation emerges from various factors, including genetic mutation, environmental perturbation, epigenetic effects, and even stochastic noises in developmental process itself. This in turn means that factors that limit phenotypic variation potentially leads to restricted diversity after evolution [1–12]. For instance, developmental processes may limit phenotypic variation by eliminating lethal phenotypes [11] or buffering against phenotypic change due to mutation and environmental perturbation [5, 12, 13]. Developmental processes, therefore, intrinsically bias the resulting phenotype and further evolutionary outcome [2, 9]. Understanding this bias will help in establishing a predictive theory for phenotypic evolution [14–16].

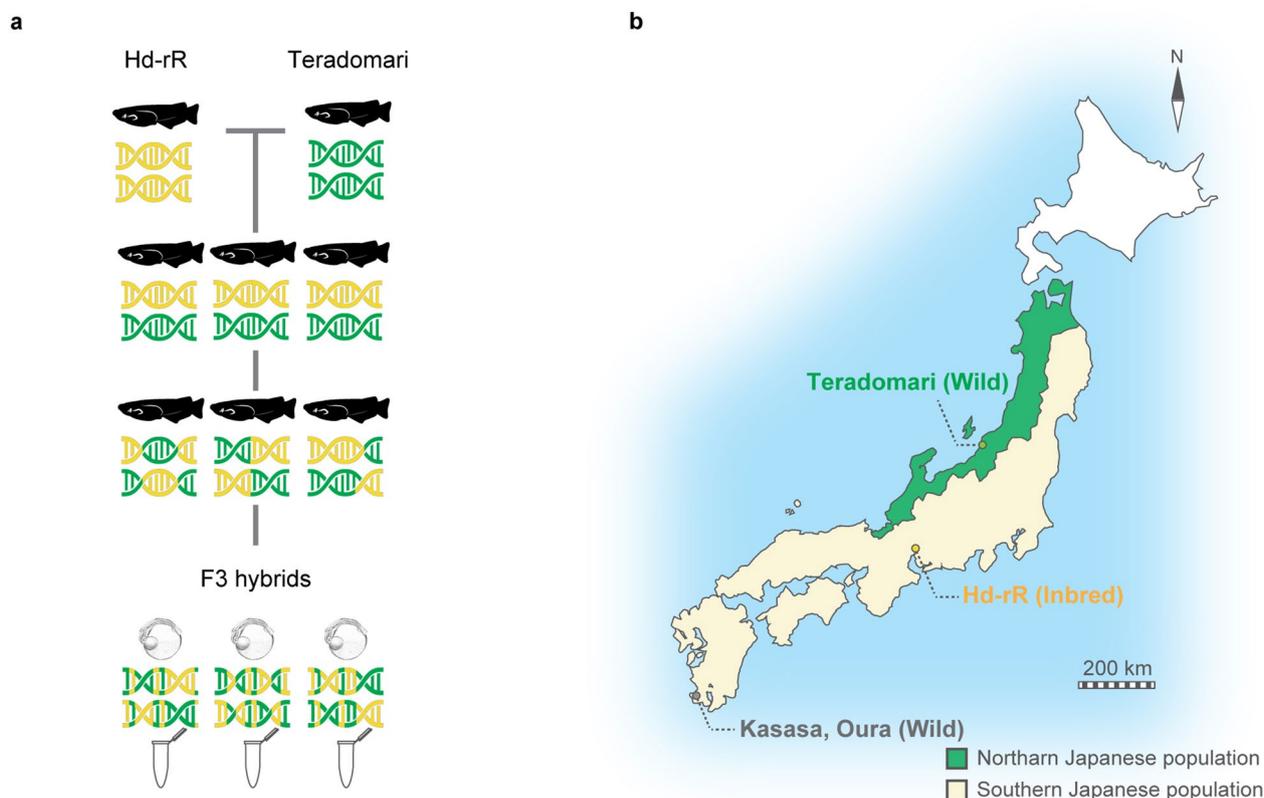
Developmental stability has recently been identified as another factor that may potentially limit evolutionary diversity [17–22]. As proposed by Hallgrímsson et al. [17], developmental stability reflects the low variability of phenotypes (including gene expressions) under a particular set of developmental conditions (i.e., the same genotype and environment). While the exact mechanism requires clarification, both theoretical [18, 19] and experimental studies [20–23] have shown the existence of a correlative relationship between stability and evolutionary conservation. Similarly, we have previously found that gene expression profiles are more stable for

the mid-embryonic phase, the phase of body-plan development (the establishment of the basic anatomical pattern), than for other developmental phases; this suggests that phenotypic stability could be included to explain the strict and continuous conservation of animal body plans [24–30]. Although previous studies highlighted phenotypic stability is associated with evolutionary conservation, it remains unclear whether it functions as an effect or cause.

Here, we used Japanese medaka *Oryzias latipes* to determine whether the developmental stability in gene expression levels and developmental stages measured by the whole embryonic transcriptome in the F0 population led to their conservation in descendants (Fig. 1). In brief, after measuring the developmental stability of genes and developmental stages in the highly inbred Hd-rR line (F0 generation) raised under the same conditions, we crossed the F0 generation with a distantly related lineage to generate genetically diverse descendants and then evaluated the conservation of their developmental stages and genes. Our results indicated that genes and developmental stages of greater stability in the F0 generation led to less diversity in the F3 generation. This, therefore, coincides well with the idea that developmental stability of phenotype limits their evolutionary diversification.

## Results

To test whether the developmental stability of the whole embryonic transcriptome and gene expression contributes to evolutionary conservation, we first evaluated

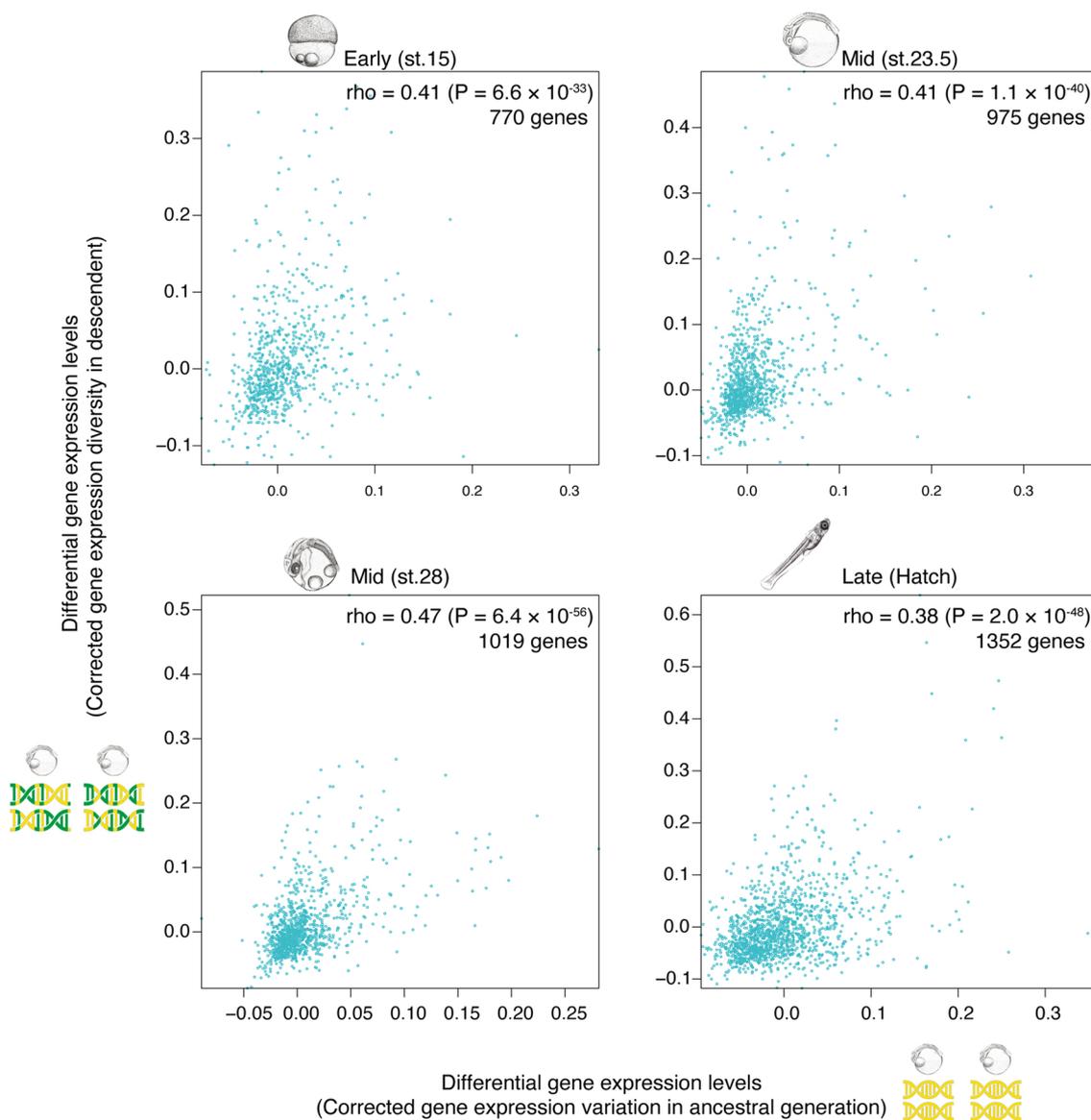


**Fig. 2** Crossing strategy applied in this study **(a)** Experimental evolution was achieved by crossing distantly related medaka lines, the highly inbred (hereafter, Hd-rR) line (from the southern Japanese population), and the wild Teradomari line (from the northern Japanese population). We used two individuals as the F0 generation: an Hd-rR male (yellow DNA) and a Teradomari female (green DNA); the single nucleotide polymorphism (SNP) rate between their genomes is as much as 3.2 % [33]. The F1 generation was then further crossed to obtain the F2 generation, and the evolutionary outcome was evaluated in the F3 generation (illustrated as embryos). The genomic sequence of the F3 individual is expected to be a mix of those of the two original populations (illustrated beneath the embryos). **(b)** Geographical distribution of the four medaka populations used in this study. The Hd-rR line was used for evaluating developmental stability. The map is modified from a previous study [39]

developmental stability using our previously published data for medaka embryos [22].

An accurate measurement of developmental stability should be made under the same environmental conditions and without genetic mutations [17]. As an indicator of the developmental stability of each gene, we measured the differential gene expression levels in gender-matched sibling pairs (hereafter referred to as the “gene expression variation”) at four developmental stages (stages 15, 23.5, 28, and hatching [31], with 13, 25, 24, and 23 embryo pairs analyzed per stage, respectively). To minimize genetic and environmental bias and achieve this condition, we used pairs of highly inbred (Hd-rR) sibling embryos raised under the same environment. Bias from technical error were also minimized when measuring the phenotypic and gene expression stability; we used only those genes with deviation in expression that was significantly greater than the technical error [22].

We then crossed the F0 generation with the distantly related northern Japanese Teradomari population (Fig. 2), which diverged approximately four million years ago [32, 33]. The produced F1 generation was further crossed twice among siblings to obtain the F3 generation, which had lineage-mixed DNA sequences at the chromosome level. Since the Hd-rR belongs to the Southern population of Japanese medaka, whereas the Teradomari population belongs to the Northern medaka population, this crossing strategy was expected to produce highly diversified and heterogeneous individuals in the F3 generation (Fig. 2a). Intra-species evolutionary conservation in gene expression and in the whole embryonic transcriptome were then evaluated among the F3 individuals. As an indicator of evolutionary conservation, we measured differential gene expression in all possible F3 embryo pairs (hereafter represented as “diversity”). Of note, since the absolute gene expression levels are a potential confounding factor between gene expression variation



**Fig. 3** Gene expression with less variation in ancestral generation tended to be conserved within descendant generation In the scatter plots, the x-axis shows variation in gene expression levels in the F0 generation, reflecting developmental stability: higher scores reflect lower stability. The y-axis shows variations, or diversity of gene expression levels in F3 generation: higher scores reflect greater diversity. Negative values for the gene expression variations in F0 and F3 are due to corrections for technical errors and mean expression levels. The results are shown for the four developmental stages. For the F0 generation, the numbers of sibling pairs analyzed were 23 for stage 15, 24 for stage 23.5, 25 for stage 28, and 13 for hatching (22); for the F3 generation, the numbers of embryonic pairs used were 10 for stage 15 (the number of embryos = 5), 15 for stage 23.5 ( $n = 6$ ), 15 for stage 28 ( $n = 6$ ), and 15 for hatching ( $n = 6$ ). Spearman's correlation coefficients and the number of genes used are shown in each plot (stage 15,  $\rho = 0.41$ ,  $P = 6.6 \times 10^{-33}$ ; stage 23.5,  $\rho = 0.41$ ,  $P = 1.1 \times 10^{-40}$ ; stage 28,  $\rho = 0.47$ ,  $P = 6.4 \times 10^{-56}$ ; hatching,  $\rho = 0.38$ ,  $P = 2.0 \times 10^{-48}$ ). The  $P$  values are derived from the test of no correlation

and diversity [34–38], we used corrected values to calculate their correlation. For each developmental stage, genes were sorted by their absolute expression levels and then corrected for their gene expression diversity using the running median of 500 genes with similar expression levels [22]. The corrected gene expression variation/

diversity was then obtained by subtracting the median from the original values, and some of the genes thus have negative values (Fig. 3).

We first examined whether stability in the ancestors (F0 generation) correlated with conservation in the descendants (F3 generation) on a gene-by-gene level. The results

indicated that genes with less expression level variations in the F0 generation tended to show less diversity in their expression levels in the F3 generation, with a moderate but significant Spearman's correlation coefficient (of approximately  $\rho=0.4$ ) in all the developmental stages tested (Fig. 3). This result suggests that gene expressions with greater stability in the ancestors showed higher conservation in the descendants.

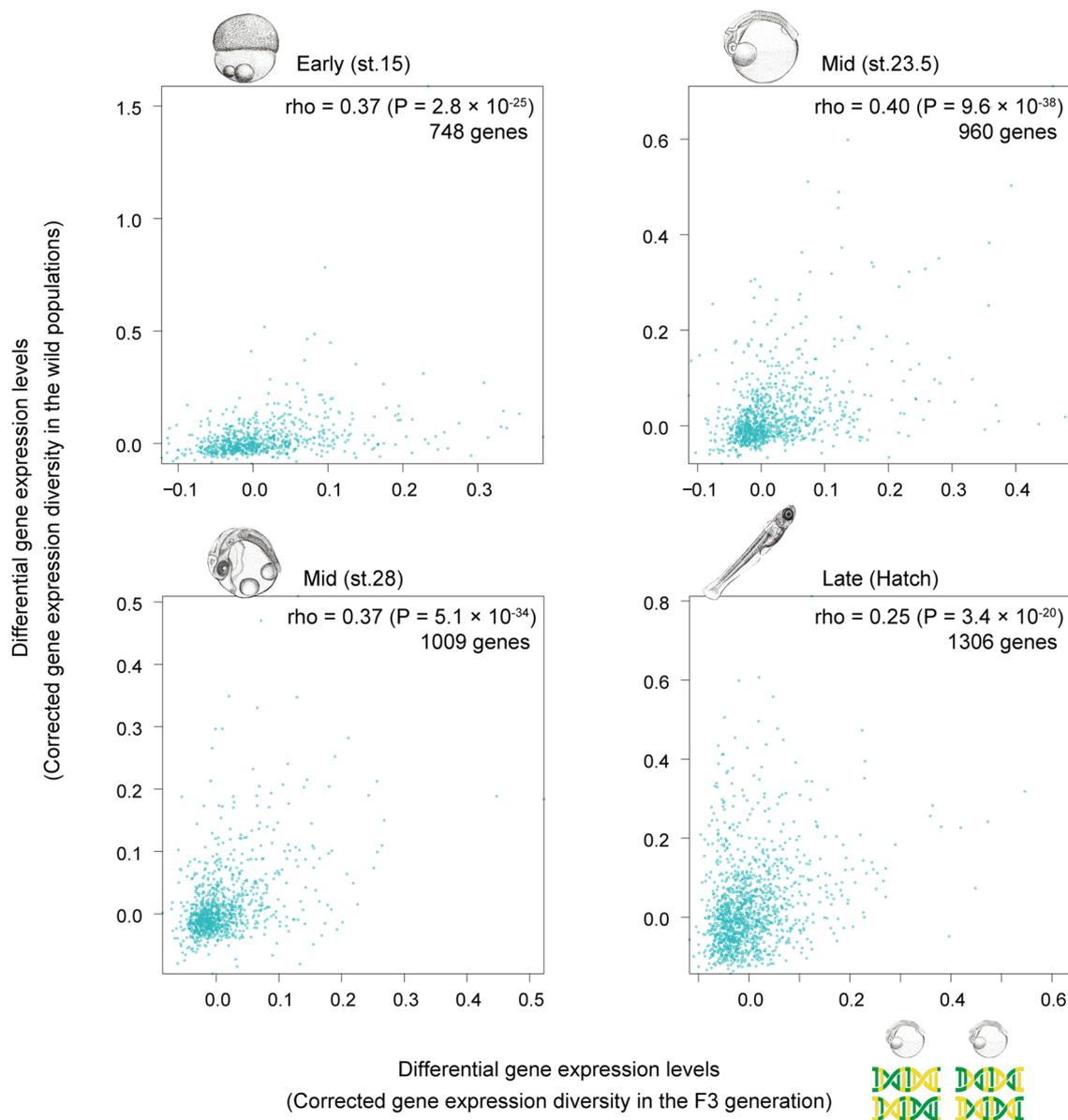
Although the results suggested that genes with more stable expression led to less evolutionary diversity, the experiment was done under controlled artificial conditions, with a small population size. To confirm if this artificial evolution mimicked evolutionary changes under natural conditions, we compared the F3 diversity with naturally occurring intra-species diversity. Kasasa and Oura wild populations were used for this purpose; they live in the same water system and were found to have genetic differences of only around 0.10% (Fig. 2b) [22]. Their diversity in gene expression was calculated using the same method as for the F3 generation. Correlation of the gene expression variations between the two groups was found to be moderate, but significant (Fig. 4; Spearman's  $\rho$ , 0.25–0.4). This supports that our experimental conditions were not substantially different from natural conditions. However, it is also of note that the result itself does not necessarily mean that genes with greater stability will also lead to conservation in a natural environment, as we were unable to measure the stability in the common ancestors of the Kasasa Oura lineages. In addition, all of the lineages and populations used here were from the Southern medaka population, and it is possible that this population has a particular genetic background that contributed to the observed correlation.

While developmental stability led to less diversity at gene-by-gene level, it still remains to be tested whether this also holds true for phenotypes, such as developmental stage measured by whole embryonic transcriptome. Here, whole embryonic transcriptome from individual embryos were defined as phenotypes of that developmental stage, as in the previous studies [26, 28, 29, 40]. We next evaluated whether the developmental stages of ancestors (F0 generation) with greater stability led to higher evolutionary conservation in the hybrid F3 descendants. Stability of the developmental stages was evaluated by measuring the differences in the whole embryonic transcriptome between sex-matched inbred Hd-rR siblings (stage 15,  $n=23$  pairs; stage 23.5,  $n=24$  pairs; stage 28,  $n=25$  pairs; and hatching,  $n=13$  pairs). Among the developmental stages analyzed, stage 28 exhibited significantly lower stability than the other stages (Fig. 5a, Steel–Dwass test). This result was consistent with a previous study [22], and it suggested that the potential phylotypic period is highly stable in medaka.

Given that developmental stability reduces evolutionary diversity, stage 28 should exhibit the lowest diversity in the hybrid F3 descendants. Consistent with this, the phenotypes of stage 28 showed less diversity than the earlier and later developmental stages in F3 generation (Fig. 5b). Notably, this stage is also the most conserved stage in intraspecies evolution of medaka under natural conditions [22], and at much larger evolutionary scales, such as in vertebrate evolution [24–30]. This is consistent with the hypothesis that developmental stability in this phase contributes to persistent conservation of the developmental period, which establishes the basic anatomical pattern of the phylum (the body plan) [24, 25, 40, 41].

## Discussion

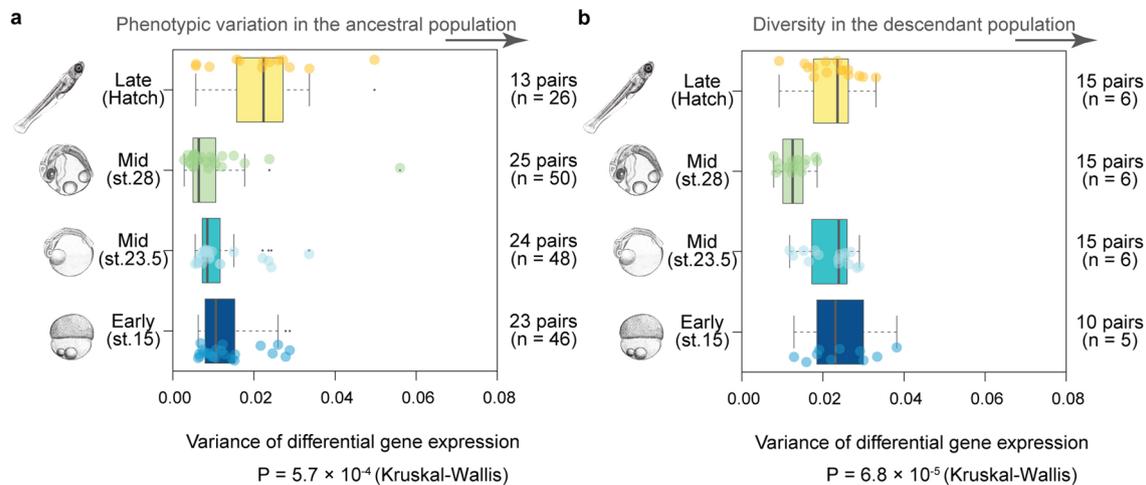
Previous studies highlighted possible contribution of developmental stability toward evolutionary conservation [17–22]. Nonetheless, causal relationships between stability and conservation have remained untested, especially for multicellular organisms. To address this, we crossed medaka lines to test whether gene expression levels and developmental stages with greater developmental stability led to lower diversity in descendants. The results indicated that genes with lower variation in their expression levels in the F0 population exhibited less expression diversity in the F3 population (Fig. 3). Similarly, developmental stages with the lowest variation in the F0 resulted in the lowest diversity in the F3 hybrid population (Fig. 5). While the diversity of gene expression levels was observed for the F3 population raised under artificial conditions, similar diversity was also observed for naturally separated and evolving Oura and Kasasa populations (Fig. 4). These findings coincide well with the hypothesis that developmental stability of certain phenotypes contributes to their evolutionary conservation in descendants. In other words, diversification of phenotypes (including individuals as a total phenotypic entity) would be intrinsically biased by their developmental stability. However, the detailed mechanisms remain unclear. One possible scenario is that stability reflects the low potential of creating phenotypic variations (including gene expression levels). As implied in theoretical studies [18, 19], given that stability also correlates with, or is a reflection of, robustness against mutational and environmental perturbations, it would be reasonable to assume that stable phenotypes eventually result in a less diversified status after accumulating mutations. In this context, the mid-developmental period (body-plan development phase) of vertebrates has been demonstrated to show low susceptibility to mutational and environmental perturbations [13]. Together, these imply that the body-plan development phase has less potential for generating phenotypic variation, even with or without mutations



**Fig. 4** Gene expression diversity under a laboratory crossing strategy significantly correlated with that under natural evolution. In the scatter plots, the x-axes show F3 gene expression diversity, and the y-axes show gene expression diversity between the two wild Kasasa and Oura populations. The same analysis was performed for all four developmental stages. The numbers of embryonic pairs from the two wild populations were 16 (stage 15,  $n = 4$  per population), 16 (stage 23.5,  $n = 4$ ), 16 (stages 28,  $n = 4$ ), and 4 (hatching,  $n = 2$ ). Spearman's correlation coefficients and the numbers of genes used in the analysis are shown in each plot (stage 15,  $\rho = 0.37$ ,  $P = 2.8 \times 10^{-25}$ ; stage 23.5,  $\rho = 0.40$ ,  $P = 9.6 \times 10^{-38}$ ; stage 28,  $\rho = 0.37$ ,  $P = 5.1 \times 10^{-34}$ ; hatching,  $\rho = 0.25$ ,  $P = 3.4 \times 10^{-20}$ ). The  $P$  values are derived from the test of no correlation

and environmental perturbations. However, as has been pointed out in previous studies [42–47], further studies are awaited to clarify the relationship between variations of each gene and whole embryonic transcriptome. Similarly, how conserved whole embryonic transcriptome is related with anatomical features in the phylotypic period require further studies.

One caveat of this study would be that the “experimental evolution” conducted in this study was only observed for three generations under controlled artificial conditions. Although the gene expression diversity among the F3 generation was similar (Fig. 4) to that observed for naturally evolved lineages (Oura-Kasasa), further studies are required to determine whether stability does indeed bias the evolutionary outcome under a natural



**Fig. 5** Stable developmental stages tended to be more conserved in the hybrid descendants Analyzing whole embryonic transcriptomes, we evaluated the stability for the four developmental stages in F0 generation, in terms of the variance in gene expression between pairs of siblings (**a**). Evolutionary diversity in the hybrid F3 embryos was evaluated using transcriptomic data, comparing all possible embryonic pairs within the same stages (**b**). The numbers of pairs used are shown on the right of each panel. A Kruskal–Wallis test ( $P$  values were shown below the panels) followed by multiple comparisons using the Steel–Dwass test suggested that, for the F0 generation, stage 28 exhibited significantly less phenotypic variation than the other stages (stage 15 vs. 23.5,  $P = 4.1 \times 10^{-1}$ ; stage 15 vs. 28,  $P = 2.2 \times 10^{-2}$ ; stage 15 vs. hatching,  $P = 1.8 \times 10^{-1}$ ; stage 23.5 vs. 28,  $P = 2.6 \times 10^{-1}$ ; stage 23.5 vs. hatching,  $P = 4.9 \times 10^{-2}$ ; stage 28 vs. hatching,  $P = 4.2 \times 10^{-3}$ ). Similarly, in the F3 descendants, stage 28 showed significantly less phenotypic diversity than the earlier and later stages (stage 15 vs. 23.5,  $P = 8.3 \times 10^{-1}$ ; stage 15 vs. 28,  $P = 1.8 \times 10^{-3}$ ; stage 15 vs. hatching,  $P = 9.5 \times 10^{-1}$ ; stage 23.5 vs. 28,  $P = 1.6 \times 10^{-1}$ ; stage 23.5 vs. hatching,  $P = 1.0$ , stage 28 vs. hatching,  $P = 7.2 \times 10^{-4}$ ). Box plots: center line, median; limits, upper and lower quartiles; whiskers,  $1.5 \times$  interquartile range; points, outliers

environment. Nonetheless, considering that the F3 generation with genetic heterogeneity showed a conserved expression status in the body plan establishing stages, it is possible that diversity or the variance of gene expression is largely stage specific, and has less to do with genomic configurations. In addition, this study focused on gene expression levels from whole embryos to analyze stability and conservation. Therefore, we were unable to address the evolutionary diversity that arises from temporal and spatial changes in gene expression, which may not change the entire gene expression levels in embryos. The phylotypic period was reported to show less variation in the timing of gene expression in *C. elegans* [45], and it will be of interest to see whether genes that contribute to the construction of body plans in vertebrates also exhibit spatial and temporal stability in vertebrates.

Predicting phenotypic evolution is an important but challenging problem in the field of evolutionary biology [48–50]. Further research on the mechanisms of phenotypic stability and robustness, and on how these factors promote evolutionary conservation will contribute to the development of an evolutionary theory of phenotypic predictability.

## Conclusions

The present study demonstrated that gene expressions and developmental stages with greater stability in the medaka F0 generation tended to be more conserved in the F3 generation.

The results suggest that the developmental stability leads to the phenotypic conservation in the course of evolution, possibly by reducing the phenotypic variations which becomes the target of selections and genetic drifts. Clarifying the mechanism behind this stability to conservation would open up a way for predicting phenotypic evolutions.

## Methods

### Animal care and sampling of embryos

A Northern Japanese wild strain of medaka, Teradomari (strain ID: WS1317), was supplied by the National BioResource Project of Utsunomiya University (Utsunomiya, Japan). Adult medaka were raised and maintained at 28 °C under a 14 h light:10 h dark cycle. Fertilized eggs were obtained [31] by natural mating, and were hatched in hatching buffer at 28 °C. Embryos were staged according to the standard developmental table. For the potential phylotypic period, the number of somites was used to accurately identify the stage (stage 23.5: 14 somites; stage 28: 30 somites).

### RNA extraction and RNA sequencing

Each staged embryo was homogenized in QIAzol reagent (Qiagen, Hilden, Germany), and its whole embryonic total RNA was extracted using an RNeasy Min Elute kit (Qiagen) in accordance with the manufacturer's protocol. RNA quality was checked using a Bioanalyzer (Agilent Technologies, Santa Clara, CA), applying the RNA Integrity Number (RIN)  $\geq 9$  threshold. RNA-Seq libraries were prepared using a TruSeq RNA sample preparation Kit v. 2 (Illumina, San Diego, CA), and sequenced on a HiSeq 1500 platform (Illumina, 100-bp single read, >20 million mapped reads).

### Estimation of gene expression

Trimmomatic v. 0.38 [51] was used to trim adapters. RNA-seq data quality was checked using FastQC v. 0.11.8 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). After removing the mitochondrial genomic sequences from the medaka reference genome (v. ASM223467v1), HISAT2 v. 2.1.0 [52] was used to map the sequenced reads onto the medaka reference genome. To avoid bias due to differences in read depth among samples, random subsampling from the total mapped reads was performed (up to 20 million mapped reads per sample). Relative gene expression (in transcripts per million, TPM) was calculated using StringTie v. 1.3.5, with default parameters [53, 54], then  $\log_{10}$  transformed, as  $\log_{10}(\text{TPM} + 1)$ , for all analyses. Here,  $x_j^i$  is the log-transformed gene expression level of gene  $j$  in individual  $i$ . We excluded genes with low relative expression ( $x_j^i < 0.1$ ) for all individuals. We obtained consistent results using  $x_j^i$  thresholds of 0–1.5.

### Gene set selection

For our inbred samples, we reduced bias from technical error in evaluating stability by selecting those genes for which the deviation in expression significantly exceeded that obtained for the technical replicates, as per our prior study [22]. In brief, for  $j$ th gene, the difference in expression  $|x_j^i - x_j^k|$  was first calculated (where  $i$ th and  $k$ th individuals were sex-matched siblings); the technical error in expression was calculated as the average of  $|x_j^{tech,i} - x_j^{tech,k}|$  over all possible combinations with  $i \neq k$  (six combinations in total) among the four replicates, where  $x_j^{tech,i}$  represents the expression level of  $j$ th gene in  $i$ th technical replicate. A one-sided Wilcoxon rank-sum test ( $\alpha=0.01$ ) was then conducted for each gene, to determine whether the differences in gene expression between inbred siblings were significantly greater than those between the technical replicates.

We further selected genes that exhibited statistically significantly different expression levels between the F0 and F3 generations. The mean expression level for each gene was calculated across all individuals in each population (regardless of sex or sibling relationship). The two-sided Wilcoxon rank-sum test ( $\alpha=0.01$ ) was then conducted for each gene and only those with significantly different mean expression levels between the two populations were selected.

### Evaluation of developmental stability for each gene expression level

Briefly, highly inbred (Hd-rR) medaka lines (i.e, those from the southern Japanese population) raised under the same water conditions were crossed, and sex-matched pairs of siblings were used. For each gene, the difference in the expression level (reflecting variation in expression) was calculated by taking the average of the difference in the expression levels  $|x_j^i - x_j^k|$  between sex-matched siblings. The absolute gene expression levels can be a confounding factor between gene expression variability and its evolutionary conservation [34–38]. To avoid this, we used corrected values to evaluate developmental stability, as previously described [22, 36]. In this respect, to correct these values, for each developmental stage, the genes were sorted by their absolute expression levels averaged over all inbred individuals. We then calculated the running median of the intra-pair differences in expression (window size: 501 genes; i.e.,  $\pm 250$  genes with similar expression levels). For window sizes  $< 250$  genes on either side (i.e, within the top or bottom 250 genes), window size was reduced to an equal number of genes on each side; the corrected difference in expression was then obtained by subtracting the running median from the average of the intra-pair difference in expression levels. This corrected value was used as an indicator of gene expression stability.

### Evaluation of diversity for each gene expression level

Diversity in gene expression levels in the F3 population, and between the wild Kasasa and Oura populations, was quantified. For the F3 population, the difference in expression  $|x_j^i - x_j^k|$  was calculated among for all possible pairs of descendant individual embryos, where  $x_j^i$  and  $x_j^k$  represent the expression of  $j$ th gene of  $i$ th and  $k$ th individuals. This value was then corrected to minimize expression-level dependency, as described in the previous section. Diversity in expression was then calculated as the average of the differences in expression for all pairs of F3 embryos. For the wild Kasasa and Oura populations, the difference in gene expression levels

$|x_j^{Kasasa,i} - x_j^{Oura,k}|$  was calculated within sex-matched pairs, where  $x_j^{Kasasa,i}$  and  $x_j^{Oura,k}$  are the expression levels of  $j$ th gene from  $i$ th Kasasa individual and  $k$ th Oura individual, respectively. This value was then corrected to minimize dependency on variation, as described in the previous section. Our previously reported wild population transcriptome data [22] are available in the DNA Data Bank of Japan (accession number DRA012427; experiment numbers DRX298419–DRX298634).

### Evaluation of developmental stability of developmental stages

As per our previous methods [22], we evaluated developmental stability by studying the differences between the whole embryonic gene expression profiles of sex-matched F0 siblings. This was quantified by calculating the variance of the differential gene expression between inbred siblings. Defining  $y_j^{ik}$  as the difference in the expression of  $j$ th gene between  $i$ th and  $k$ th individuals, where  $y_j^{ik} = (x_j^i - x_j^k)$ , the variance was  $V^{ik} = \frac{1}{N} \sum_j (y_j^{ik} - \bar{y}^{ik})^2$ , where  $\bar{y}^{ik}$  is the average of the difference in gene expression, and  $N$  is the number of genes analyzed.

### Evaluation of evolutionary diversity in the F3 phenotype

To quantify the diversity in developmental stages using whole-embryonic gene expression profiles, the variance in differential gene expression between pairs of  $i$ th and  $k$ th individual embryos,  $V^{ik}$ , was calculated for all possible F3 pairs.

### Statistics

The biological replicates comprised embryos from different parents and that were born on different days, to appropriately represent the population of interest. For statistical tests, the threshold for statistical significance was set at  $\alpha = 0.01$ . To avoid an inflated type-I error rate in multiple comparisons following the Kruskal–Wallis test, we performed the Steel–Dwass test in R, using the package NSM3 v. 1.12 [55].

### Abbreviation

TPM Transcripts per million

### Acknowledgements

We thank National BioResource Project (NBRP) Medaka (<https://shigen.nig.ac.jp/medaka/>) for providing the Teradomari strain (strain ID: WS1317).

### Author contributions

NI and YU conceived and designed the study. HT provided the facilities for raising the medaka. YU collected the samples and performed the RNAseq experiments. YU, CF, and NI analyzed the data. YU, CF, and NI edited the manuscript. NI supervised the study. All authors read and approved the final manuscript.

### Funding

This work was supported in part by a KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (17H06387) and by a RIKEN Research Fund for Special Postdoctoral Researcher (Project Code, 202001061033).

### Availability of data and materials

The data sets supporting the conclusions of this article, the developmental transcriptome data sets, are available in the DNA Data Bank of Japan (Accession Number DRA013804; experiment Numbers DRX343545–DRX343567).

### Declarations

#### Ethics approval and consent to participate

Animal care and experimental procedures were conducted in strict accordance with the guidelines approved by the Animal Experiments Committee of the Graduate School of Science, University of Tokyo (approval IDs AP19-8 and AP19-10).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Center for Biosystems Dynamics Research, RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan. <sup>2</sup>Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-0033, Japan. <sup>3</sup>Universal Biology Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-0033, Japan.

Received: 6 November 2022 Accepted: 1 March 2023

Published online: 14 March 2023

### References

- Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, et al. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc Royal Soc B Biol Sci.* 2015;282(1813):20151019. <https://doi.org/10.1098/rspb.2015.1019>.
- Smith JM, Burian R, Kauffman S, Alberch P, Campbell J, Goodwin B, et al. Developmental constraints and evolution: a perspective from the mountain lake development and evolution. *Q Rev Biol.* 1985;60(3):265–87.
- Brakefield PM. The power of evo-devo to explore evolutionary constraints: experiments with butterfly eyespots 1. *Zoology.* 2003;106(4):283–90.
- Fritz JA, Brancale J, Tokita M, Burns KJ, Hawkins MB, Abzhanov A, et al. Shared developmental programme strongly constrains beak shape diversity in songbirds. *Nat Commun.* 2014;5(1):3700. <https://doi.org/10.1038/ncomms4700>.
- Green RM, Fish JL, Young NM, Smith FJ, Roberts B, Dolan K, et al. Developmental nonlinearity drives phenotypic robustness. *Nat Commun.* 2017;8(1):1970. <https://doi.org/10.1038/s41467-017-02037-7>.
- Kavanagh KD, Evans AR, Jernvall J. Predicting evolutionary patterns of mammalian teeth from development. *Nature.* 2007;449(7161):427–32.
- Kavanagh KD, Shoval O, Winslow BB, Alon U, Leary BP, Kan A, et al. Developmental bias in the evolution of phalanges. *Proc Natl Acad Sci.* 2013;22:201315213.
- Payne JL, Wagner A. The causes of evolvability and their evolution. *Nat Rev Genet.* 2019;20(1):24–38. <https://doi.org/10.1038/s41576-018-0069-z>.
- Alberch P. Ontogenesis and morphological diversification. *Am Zool.* 1980;20(4):653–67.
- Furusawa C, Irie N. Toward understanding of evolutionary constraints: experimental and theoretical approaches. *Biophys Rev.* 2020;12(5):1155–61. <https://doi.org/10.1007/s12551-020-00708-2>.
- Galis F, Metz JAJ. Testing the vulnerability of the phylotypic stage: On modularity and evolutionary conservation. *J Exp Zool.* 2001;291(2):195–204.

12. von Dassow G, Meir E, Munro EM, Odell GM. The segment polarity network is a robust developmental module. *Nature*. 2000;406(6792):188–92.
13. Uchida Y, Uesaka M, Yamamoto T, Takeda H, Irie N. Embryonic lethality is not sufficient to explain hourglass-like conservation of vertebrate embryos. *EvoDevo*. 2018;9(7):1–11. <https://doi.org/10.1186/s13227-018-0095-0>.
14. Wagner GP, Altenberg L. Perspective: complex adaptations and the evolution of evolvability. *Evolution*. 1996;50(3):967–76. <https://doi.org/10.1111/j.1558-5646.1996.tb02339.x>.
15. Hendrikse JL, Parsons TE, Hallgrímsson B. Evolvability as the proper focus of evolutionary developmental biology. *Evol Dev*. 2007;9(4):393–401. <https://doi.org/10.1111/j.1525-142X.2007.00176.x>.
16. Pigliucci M. Is evolvability evolvable? *Nat Rev Genet*. 2008;9(1):75–82. <https://doi.org/10.1038/nrg2278>.
17. Hallgrímsson B, Willmore K, Hall BK. Canalization, developmental stability, and morphological integration in primate limbs. *Am J Phys Anthropol*. 2002;113:131–58. <https://doi.org/10.1002/ajpa.10182>.
18. Kaneko K, Furusawa C. An evolutionary relationship between genetic variation and phenotypic fluctuation. *J Theor Biol*. 2006;240(1):78–86.
19. Lehner B, Kaneko K. Fluctuation and response in biology. *Cell Mol Life Sci*. 2011;68(6):1005–10.
20. Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL. Genetic properties influencing the evolvability of gene expression. *Science*. 2007;317(5834):118–21.
21. Hayden L, Lochovska K, Sémon M, Renaud S, Delignette-Muller ML, Vilcot M, et al. Developmental variability channels mouse molar evolution. *Elife*. 2020;9:e50103. <https://doi.org/10.7554/eLife.50103>.
22. Uchida Y, Shigenobu S, Takeda H, Furusawa C, Irie N. Potential contribution of intrinsic developmental stability toward body plan conservation. *BMC Biol*. 2022;20(1):82. <https://doi.org/10.1186/s12915-022-01276-5>.
23. Sato K, Ito Y, Yomo T, Kaneko K. On the relation between fluctuation and response in biological systems. *Proc Natl Acad Sci USA*. 2003;100(SUPPL. 2):14086–90.
24. Denis D. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development*. 1994;120:135–42.
25. Raff RA. *The shape of life: genes, development, and the evolution of animal form*. Chicago: University of Chicago Press; 1996. p. 520.
26. Irie N, Kuratani S. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat Commun*. 2011;2(1):248.
27. Yanai I, Peshkin L, Jorgensen P, Kirschner MW. Mapping gene expression in two xenopus species: evolutionary constraints and developmental flexibility. *Dev Cell*. 2011;20(4):483–96. <https://doi.org/10.1016/j.devcel.2011.03.015>.
28. Wang Z, Pascual-Anaya J, Zadissa A, Li W, Niimura Y, Huang Z, et al. The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat Genet*. 2013;45(6):701–6.
29. Hu H, Uesaka M, Guo S, Shimai K, Lu TM, Li F, et al. Constrained vertebrate evolution by pleiotropic genes. *Nat Ecol Evol*. 2017;1(11):1722–30. <https://doi.org/10.1038/s41559-017-0318-0>.
30. Uesaka M, Kuratani S, Takeda H, Irie N. Recapitulation-like developmental transitions of chromatin accessibility in vertebrates. *Zoological Lett*. 2019;5:33.
31. Iwamatsu T. Stages of normal development in the medaka *Oryzias latipes*. *Mech Dev*. 2004;121(7–8):605–18.
32. Sakaizumi M, Moriwaki K, Egami N. 1983 Allozymic variation and regional differentiation in wild populations of the fish *Oryzias latipes*. *Copeia*. 1983;2:311–8.
33. Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, et al. The medaka draft genome and insights into vertebrate genome evolution. *Nature*. 2007;447(7145):714–9.
34. Csaba P, Papp B, Hurst LD. highly expressed genes in yeast evolve slowly. *Genetics*. 2001;158(2):927–31.
35. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science*. 2002;297(5584):1183–6.
36. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et al. Single-cell proteomic analysis of *S cerevisiae* reveals the architecture of biological noise. *Nature*. 2006;441(7095):840–6.
37. Gout JF, Kahn D, Duret L. The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. *PLoS Genet*. 2010;6(5):20.
38. Barroso GV, Puzovic N, Duthel JY. The evolution of gene-specific transcriptional noise is driven by selection at the pathway level. *Genetics*. 2018;208(January):173–89.
39. Spivakov M, Auer TO, Peravali R, Dunham I, Dolle D, Fujiyama A, et al. Genomic and phenotypic characterization of a wild medaka population: Towards the establishment of an isogenic population genetic resource in fish. *G3*. 2014;4(3):433–45.
40. Irie N, Kuratani S. The developmental hourglass model: a predictor of the basic body plan? *Development*. 2014;141(24):4649–55.
41. Onai T, Irie N, Kuratani S. the evolutionary origin of the vertebrate body plan: the problem of head segmentation. *Annu Rev Genomics Hum Genet*. 2014;15(1):443–59. <https://doi.org/10.1146/annurev-genom-091212-153404>.
42. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. Gene expression divergence recapitulates the developmental hourglass model. *Nature*. 2010;468(7325):811–6.
43. Levin M, Hashimshony T, Wagner F, Yanai I. Developmental milestones punctuate gene expression in the caenorhabditis embryo. *Dev Cell*. 2012;22(5):1101–8. <https://doi.org/10.1016/j.devcel.2012.04.004>.
44. Xu F, Domazet-Lošo T, Fan D, Dunwell TL, Li L, Fang X, et al. High expression of new genes in trochophore enlightening the ontogeny and evolution of trochozoans. *Sci Rep*. 2016;6(February):1–10.
45. Zalts H, Yanai I. Developmental constraints shape the evolution of the nematode mid-developmental transition. *Nat Ecol Evol*. 2017;1(5):1–7. <https://doi.org/10.1038/s41559-017-0113>.
46. Hogan JD, Keenan JL, Luo L, Ibn-Salem J, Lamba A, Schatzberg D, et al. The developmental transcriptome for *Lytechinus variegatus* exhibits temporally punctuated gene expression changes. *Dev Biol*. 2020;460(2):139–54.
47. Li Y, Omori A, Flores RL, Satterfield S, Nguyen C, Ota T, et al. Genomic insights of body plan transitions from bilateral to pentameral symmetry in Echinoderms. *Commun Biol*. 2020;3(1):371. <https://doi.org/10.1038/s42003-020-1091-1>.
48. Lässig M, Mustonen V, Walczak AM. Predicting evolution. *Nat Ecol Evol*. 2017;1(3):77. <https://doi.org/10.1038/s41559-017-0077>.
49. Mas A, Lagadeuc Y, Vandenkoornhuyse P. Reflections on the predictability of evolution: toward a conceptual framework. *iScience*. 2020;23(11):101736.
50. Blount ZD, Lenski RE. Contingency and determinism in evolution: replaying life's tape. *Science*. 2018. <https://doi.org/10.1126/science.aam5979>.
51. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
52. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
53. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT StringTie and Ballgown. *Nat Protoc*. 2016;11(9):1650–67. <https://doi.org/10.1038/nprot.2016-095>.
54. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33(3):290–5.
55. Schneider G, Chicken E, Becvarik R. NSM3. 2018. Functions and Datasets to Accompany Hollander, Wolfe, and Chicken—Nonparametric Statistical Methods. Third Edition.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.