nature ecology & evolution

ARTICLES

DOI: 10.1038/s41559-017-0318-0

Constrained vertebrate evolution by pleiotropic genes

Haiyang Hu[1,2], Masahiro Uesaka[3], Song Guo[1], Kotaro Shimai[4], Tsai-Ming Lu[5], Fang Li[6], Satoko Fujimoto[7], Masato Ishikawa[8], Shiping Liu[6], Yohei Sasagawa[9], Guojie Zhang[6,10], Shigeru Kuratani[7], Jr-Kai Yu[5], Takehiro G. Kusakabe[4], Philipp Khaitovich[1], Naoki Irie[3,11]* and the EXPANDE Consortium

Despite morphological diversification of chordates over 550 million years of evolution, their shared basic anatomical pattern (or 'bodyplan') remains conserved by unknown mechanisms. The developmental hourglass model attributes this to phylum-wide conserved, constrained organogenesis stages that pattern the bodyplan (the phylotype hypothesis); however, there has been no quantitative testing of this idea with a phylum-wide comparison of species. Here, based on data from early-to-late embryonic transcriptomes collected from eight chordates, we suggest that the phylotype hypothesis would be better applied to vertebrates than chordates. Furthermore, we found that vertebrates' conserved mid-embryonic developmental programmes are intensively recruited to other developmental processes, and the degree of the recruitment positively correlates with their evolutionary conservation and essentiality for normal development. Thus, we propose that the intensively recruited genetic system during vertebrates' organogenesis period imposed constraints on its diversification through pleiotropic constraints, which ultimately led to the common anatomical pattern observed in vertebrates.

Over the past 550 million years, the basic anatomical features (or 'bodyplan') of animals at the phylum level have been conserved[1,2]. However, potential mechanisms underlying this conservation on the macro-evolutionary time scale remain poorly understood[2–6]. In this regard, the phylotype hypothesis[7–9] of the developmental hourglass model[7,9] provides an attractive framework for analysing this problem because it predicts that the bodyplan for each animal phylum can be defined by anatomical features observed during a mid-embryonic phase (the phylotypic period) that are conserved among species across a phylum (a process also known as phylotypic progression[7–10]). The model further predicts that the conservation of this phylotype arises from gene regulatory networks that are vulnerable to changes in molecular developmental signalling[7,9]. These ideas inspired scientists to test the hourglass model from a molecular perspective[5,11–20] and these efforts have provided support for hourglass-like, mid-embryonic conservation during development not only in animals (for example, vertebrates[11–13,15,18], *Drosophila* species[14], molluscs[20] and nematodes[17]), but also in plants (*Arabidopsis thaliana*[16]) and fungi (*Coprinopsis cinerea*[19]). A recent study showed that hourglass-like conservation was consistently not observed among ten species from different animal phyla and concluded that the phylotype hypothesis should be applied to species within the same phylum[21]. However, whether or not the phylotype hypothesis can be applied to animals at the phylum level still remains to be verified[4,22] as no studies have made phylum-wide comparisons of species to confirm that the most transcriptomically conserved mid-embryonic period actually accounts for phylotype or bodyplan-defining stages. For the phylum Chordata, it is known that ascidians do not have developmental stages that match the phylotypic period (as there is no developmental stage that shows chordate bodyplan elements simultaneously)[23,24] and the phylotype hypothesis may be applied only to the subphylum Vertebrata[7]. Nevertheless, the phylotypic period is still possibly a bodyplan-defining phase that once existed in a common ancestor of the phylum Chordata, and the period was modified secondarily during ascidian evolution[25], but remained the most conserved mid-embryonic phase in other species. In this study, we asked whether this phylotype hypothesis stands for phylum Chordata by analysing the gene expression profiles of chordates. In addition, since no consensus has been reached regarding the potential mechanisms that caused animals to follow hourglass-like conservation[3,4,6], we further sought to elucidate these by analysing the expression dataset.

Results

The phylotype hypothesis is supported not only for chordates but also vertebrates. Given that the phylotype hypothesis can be applied to the phylum Chordata[7,9], it is expected that the majority of chordate embryos exhibit hourglass-like conservation and that the conserved stages comprise a phase that shows a set of chordate bodyplan[26] features; namely, segmental muscles, a dorsal nerve cord, pharyngeal

[1]Key Laboratory of Computational Biology, Chinese Academy of Sciences–Max Planck Society Partner Institute for Computational Biology, Chinese Academy of Sciences, Shanghai 200031, China. [2]School of Life Science and Technology, China Pharmaceutical University, Nanjing 210009, China. [3]Department of Biological Sciences, University of Tokyo, Tokyo 113-0033, Japan. [4]Institute for Integrative Neurobiology and Department of Biology, Faculty of Science and Engineering, Konan University, Kobe 658-8501, Japan. [5]Institute of Cellular and Organismic Biology, Academia Sinica, Taipei 115, Taiwan. [6]China National GeneBank, BGI-Shenzhen, Shenzhen 518083, China. [7]Evolutionary Morphology Laboratory, RIKEN, Kobe 650-0047, Japan. [8]Department of Bioinformatics and Systems Biology, School of Science, University of Tokyo, Tokyo 113-0032, Japan. [9]Bioinformatics Research Unit, Advanced Center for Computing and Communication, RIKEN, Saitama 351-0198, Japan. [10]Centre for Social Evolution, Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark. [11]Universal Biology Institute, University of Tokyo, Tokyo 113-0033, Japan. A list of the members of the EXPANDE Consortium and their and affiliations appears in the Supplementary Information. Haiyang Hu, Masahiro Uesaka and Song Guo contributed equally to this work. *e-mail: irie@bs.s.u-tokyo.ac.jp

gill slits and a notochord[25]. To test this in chordates, we obtained staged, early-to-late whole embryonic gene expression profiles with biological replicates from eight chordate species[1,2]. Furthermore, we analysed their cross-species similarities using orthologous gene expression profiles with a view to represent the homologous cell composition of the embryos[15] (261 samples; Supplementary Figs. 1–2, Supplementary Tables 1–8 and Supplementary Note 1). To evaluate the conserved embryonic stages, we calculated the expression distance that reflected the phylogenetic scale of interest (expDist; Fig. 1 and Methods). After calculating the expDists of stage combinations of the target species (Supplementary Note 2) using 1:1 gene orthologs (1:1_expDists), we extracted the top 1% of stage combinations that were similar (that is, those with the lowest expDists) and visualized the contributions of each developmental stage to these top 1% stage combinations (that is, the percentage chance of being included in the top 1% of similar combinations (P_{top}); see Fig. 1 and Supplementary Information.

The chordate P_{top} signals calculated using the 1:1_expDists suggested that all chordate embryos have conserved stages around the mid-embryonic phase with no peak signals in the earliest or latest stages (chordate P_{top} graphs in Fig. 2). This was supported by similar analyses using different P_{top} thresholds (Supplementary Fig. 4).

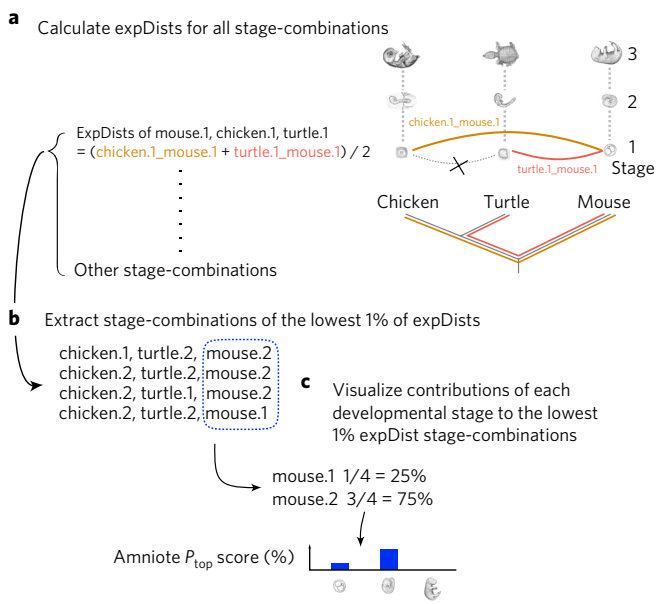

Fig. 1 | Basic strategy for evaluating evolutionarily conserved developmental stages. a, To find conserved stages among amniote species, for example, similarities in whole embryonic gene expression profiles of developmental stages from the chicken–mouse and turtle–mouse pairs were calculated. Previous approaches contained distances that did not reflect evolutionary conservation from the common ancestor of interest (the chicken–turtle pair in this case); however, these were specifically omitted in our study. ExpDists were calculated for all the possible stage combinations among the pairs of species (the chicken–mouse and turtle–mouse pairs in this case) and we took an average of the gene expression distances of the chicken–mouse and turtle–mouse stage pairs to evaluate mean distances between the chicken–turtle group and the mouse. To avoid computational barriers, one million random stage combinations were selected for identifying chordate and vertebrate conserved stages. **b**, Stage combinations with the lowest 1% expDists were then extracted to represent evolutionarily conserved stage combinations. **c**, By calculating the frequency of occurrence within the lowest 1% of stage combinations, the contributions of each developmental stage were visualized as P_{top} (percentage of stage included in the top 1% of similar stage combinations).

Although the chordate P_{top} signals were widened to include earlier stages in some species (such as chickens, turtles and both frogs), these signals included stages that showed the four bodyplan elements in all the chordates, except the ascidian (around HH16 in the chicken, TK11 in the turtle, E9.0 in the mouse, st.28 to st.31 in the western clawed frog and African clawed frog, and Prim-5–6 in the zebrafish; see also Supplementary Note 3). In ascidians, the stages that developed notochord and segmental muscle cells (st.19 to st.22) showed relatively high P_{top} signals and the stage in which the pharyngeal gills developed (the juvenile stage) also showed these signals, albeit weakly. The widened and shifted tendencies were also observed when highly derived organisms such as ascidians or the sauropsids (chickens and turtles) were removed from the analysis (Supplementary Figs. 5a and 6, respectively). Meanwhile, these could be due to biases arising from using 1:1 orthologs for the comparison, as this method omits large numbers of ortholog counterparts, particularly when distantly related species are being compared (Supplementary Note 4). In fact, only 1,704 orthologs (which is about one-tenth of all genes in each chordate species) were covered for the analysis with 1:1_expDists and the gene repertoire showed significant bias in gene ontology terms compared with the entire gene set (Supplementary Note 5). We therefore developed an alternative method using ortholog group expDists to overcome the problem associated with the 1:1 method. In brief, the method assigns a '0' expression value to genes of species that have no orthologous counterpart and sums the expression levels of genes within the same ortholog group (in-paralogs). This ortholog group method covers all the genes in each species (there is no bias in gene repertoire) and the resolution of the ortholog groups was high enough to successfully distinguish between genes in the same family belonging to different ortholog groups (Supplementary Note 6). Importantly, we confirmed that a sample relationship tree reconstructed from expression similarity with ortholog group expDists better recapitulated known phylogenetic relationships than that of 1:1_expDists (Supplementary Fig. 7), suggesting a possible advantage of the ortholog group method. In contrast with the results obtained with 1:1_expDists, those using ortholog group expDists did not support mid-embryonic conservation in five out of eight chordate species; major P_{top} signals in chickens (around prim), western clawed frogs (st.2), African clawed frogs (st.2 to st.5), zebrafish (32cell) and ascidians (st.1) occurred in much earlier stages than when the bodyplan elements first appeared (Supplementary Fig. 8a). However, the early signals could have been a bias from the inclusion of genes specific to protochordates (amphioxus and ascidian) because vertebrates have undergone massive post-2R gene loss events[27] and these lost genes are evaluated to have zero expression levels using the ortholog group method. However, we confirmed that early signals can also be observed even after excluding protochordate-specific orthologs from the chordate ortholog groups (Supplementary Fig. 5b). The early chordate P_{top} signals were also obtained for the results obtained with different P_{top} thresholds (Supplementary Fig. 4b), when ascidians or the sauropsids were excluded from the analysis (Supplementary Figs. 5c and 6b,d) and when ortholog group expDists were calculated with mean expression levels of genes in the same ortholog group rather than their summation (mean ortholog group expDists in Supplementary Fig. 8b). The results from evaluating transcriptomic conservation by the P_{tops} of ortholog group expDists do not support mid-embryonic or hourglass-like conservation for chordates. For vertebrate-conserved stages, this kind of dramatic difference in P_{top} signals was not observed between the 1:1_expDists (vertebrate P_{top} graphs in Fig. 2) and ortholog group expDists (vertebrate P_{top} graphs in Supplementary Fig. 8b) and both results consistently covered mid-embryonic stages, with no peak signals at the earliest stages. The results also coincide with previous studies[13,15,18,28]; however, it should be noted that somewhat earlier stages in chickens, turtles and frogs were also found to have vertebrate P_{top}

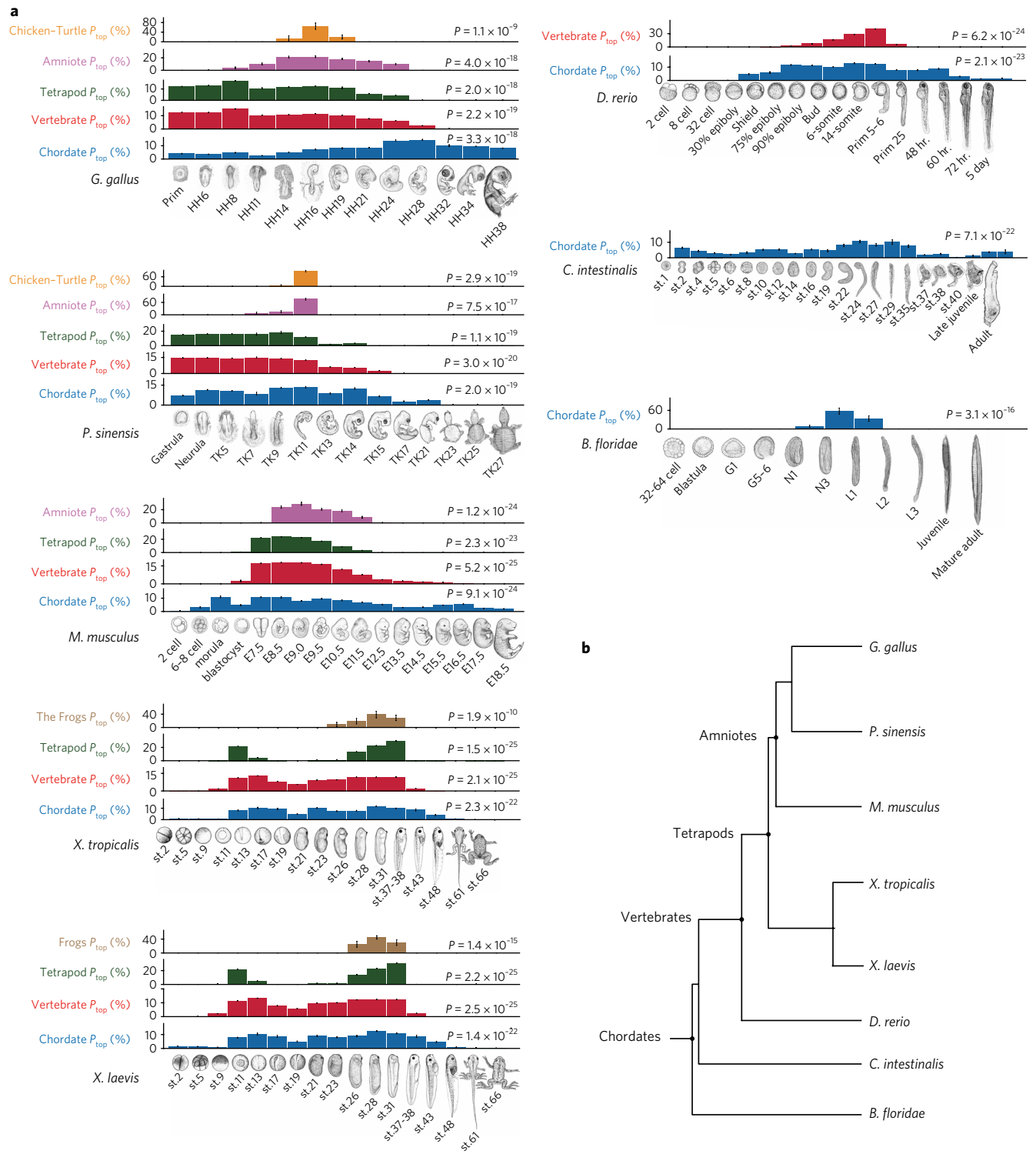

Fig. 2 | Conserved embryonic stages identified by gene expression similarity. a, The percentage of developmental stages included in the most similar (lowest 1% of 1:1_expDists) stage combinations are shown as P_{top}. Various taxonomic levels (that is, chordates, vertebrates, tetrapods, amniotes, the frogs and the chicken–turtle group) of expDists were calculated to evaluate evolutionary conservation at different evolutionary timescales. Whole embryonic expression levels of 1,704, 4,814, 6,347, 10,718, 11,957 and 15,207 1:1 orthologous genes were used to calculate the expDists of the chordate, vertebrate, tetrapod, amniote, frogs and chicken–turtle group, respectively. Changes in the P_{top} scores of the developmental stages were all significant (Friedman test; *P* values indicated in each plot, tested with 100 randomly picked-up BRI-exp expression tables for each species). The bestHit FPKM data were used for the analysis. Error bars represent s.d. of P_{top} values in 100 randomly picked-up biological replicate (BRI-exp) data. Note that the vertebrate, tetrapod and amniote conserved stages largely overlapped with each other despite their different evolutionary time scales. The overall tendency did not change for results obtained by the lowest 3 or 5% expDist thresholds (see Supplementary Fig. 4 for the results by the 5% threshold). **b**, Phylogenetic relationships of chordate species—a mammal (mouse *M. musculus*), bird (chicken *G. gallus*), diapsid (turtle *P. sinensis*), amphibians (African clawed frog *X. laevis* and western clawed frog *X. tropicalis*), bony fish (zebrafish *D. rerio*), tunicate (ascidian *C. intestinalis*) and cephalochordate (amphioxus *B. floridae*)— referenced[1] in calculating expDists. In this study, representative chordate species with sequenced genomes were selected.

signals (Fig. 2 and Supplementary Fig. 8), suggesting that conservation of the mid-embryonic stages may not be as specific as considered previously. Meanwhile, both the 1:1_expDist and ortholog group expDist vertebrate P_{top} signals included stages with morphological features, including the neural tube, neural crest cells, somite, rhombomere, placodes and aortic arches, that potentially explain the vertebrates' shared anatomical pattern[29]. These stages were around HH16 in the chicken, TK11 in the turtle, E9.0 in the mouse, st.28 to st.31 in the western clawed frog and African clawed frog, and 14-somite to Prim-5–6 in the zebrafish. This suggests that the phylotype hypothesis may also stand for the subphylum Vertebrata. Further studies are needed to reach a consensus on how embryonic conservation should be evaluated transcriptomically; however, our results, together with the exceptional case in ascidians imply that the phylotype hypothesis of the hourglass model would be better suited when applied to the subphylum Vertebrata than to the phylum Chordata. In addition, our results indicate that conservation of the mid-embryonic stages in each species is not strictly specific to the organogenesis phase (for both chordate- and vertebrate-wide conservations) and further imply that conservation of this phase becomes more obviously recognized when commonly conserved developmental stages are extracted from different species; however, a consensual method for this remains to be developed[30,31].

Vertebrates' pharyngula stages show persistent conservation throughout evolution. In addition to the above findings, we noticed that vertebrate-conserved stages largely overlapped with those of less inclusive groups within vertebrates such as tetrapods, amniotes, chickens and turtles, and frogs (Fig. 2, green, purple, orange and brown graphs, respectively), indicating that these embryonic stages have been a major target of conservation (or persistently conserved[4]) over all evolutionary time scales during evolution from the vertebrate common ancestor to the current species. This was also observed in the analyses done with different P_{top} thresholds (Supplementary Fig. 4a) and ortholog group expDists (Supplementary Figs. 4b and 8). Based on this persistence of vertebrates' mid-embryonic conservation we next focused on the common features of the conserved developmental system to try to explain the 'frozen' vertebrate basic anatomical pattern through the macro-evolutionary timescale[2]. Although active morphogenesis occurs during these vertebrate-conserved stages, we found significantly lower numbers and ratios of stage-specific up- and downregulated genes at mid-embryonic stages, especially in vertebrates (Fig. 3a and Supplementary Fig. 9a), suggesting regulatory quiescence or relative lulls of expression changes[17] in these phases. This tendency was also corroborated by a tau-based analysis (Supplementary Fig. 9b).

Since the number of expressed genes per stage increased almost constantly during development (Fig. 3b and Supplementary Fig. 2c), we speculated that the fewer stage-specific genes around the mid-embryonic phase was due to repeated expression or temporally pleiotropic expression in other developmental stages. As expected, genes expressed in the mid-embryonic stage of vertebrates (defined as those genes expressed in more than 50% of developmental stages) were highly temporally pleiotropic and the tendency was consistently observed in vertebrate embryos (Fig. 3c, Supplementary Fig. 10a and Supplementary Note 7). However, in ascidians and amphioxi, this tendency was blurred as the earlier stages had the highest proportion of pleiotropic genes. A caveat of this interpretation, however, is that sampled stages would not have equal distances in terms of expression profiles, and temporally densely/sparsely sampled stages would have biased the results. As this potential temporal bias seems to be insurmountable either by sampling all the morphologically defined stages or by sampling temporally equal intervals in real time along embryogenesis (because there is no guarantee that transcriptomic changes occur

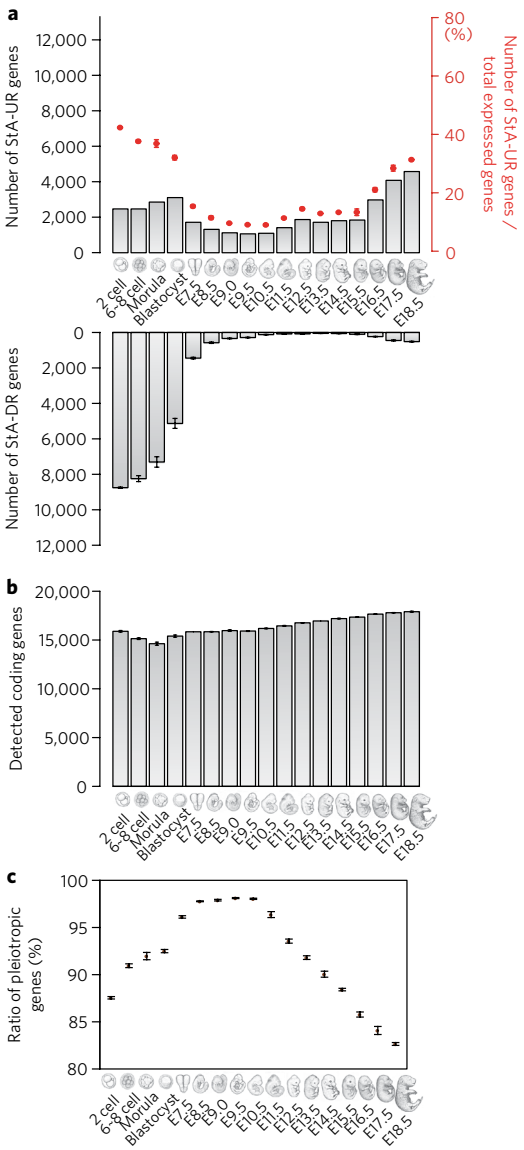

Fig. 3 | Regulatory quiescence and highly pleiotropic genes enriched in the mid-embryonic period. a, Numbers of stage-associated upregulated (StA-UR) and stage-associated downregulated genes (StA-DR) defined by 1> Z-score for the mouse embryonic data (see Supplementary Information for other species) and ratio of StA-UR to the number of expressed genes at each stage (red dots, right axis). Z-scores were calculated using the 100 BRI-exp data created with biological replicates (see Supplementary Table 1 for more details). Error bars indicate the s.d. of biological replicates at each developmental stage (100 BRI-exp). The changes in these values along developmental stages were statistically significant (analysis of variance without assuming equal variance, $P < 2.2 \times 10^{-16}$). Essentially, the same tendency was observed for the other chordate species (Supplementary Fig. 9a). In addition, the tendencies were also corroborated using a tau-based analysis (Supplementary Fig. 9b). Error bars indicate the s.d. of two or three biological replicates at each developmental stage (see Supplementary Table 1 for more details). **b**, Number of detected (>1 FPKM) genes at each developmental stage. **c**, Ratio of detected pleiotropic genes (genes detected in more than 50% of all developmental stages analysed; more than eight stages for the mouse) to total numbers of detected genes in each stage. Error bars indicate the s.d. of the ratios in 100 biological replicates (100 BRI-exp) data. The bestHit TMM-normalized FPKM values were used for these analyses. Note that the analyses used in **a**–**c** were based on a gene-by-gene analysis and not by ortholog groups. The results were also confirmed in the sampling bias corrected, simulated pilot data set (Supplementary Fig. 10b).

in concert with changes in embryonic morphology or actual time), we made a simulated expression dataset with equal expression distances between the samples (Supplementary Fig. 10b). These results also supported mid-embryonic stages of vertebrates having a higher ratio of genes with temporally pleiotropic expressions; however, it should be noted that further studies are required to find a consensus on how this bias should be controlled. With this caveat in mind, we next analysed additional aspects of temporally pleiotropic genes and found that the temporally pleiotropic genes tended to show widely distributed expression patterns (Supplementary Fig. 11a,b) and had higher numbers of predicted transcription factor binding sites (Supplementary Fig. 11c,d). In addition, temporally pleiotropic genes detected at the mid-embryonic stages were functionally enriched with anatomical structure developmental processes (Supplementary Fig. 11g), implying their functional contributions to anatomical pattern formation.

Genes expressed in vertebrates' mid-embryonic phases are pleiotropic, essential and conserved. Next, we asked if the temporally pleiotropic gene expression could explain the hourglass-like conservation of vertebrate mid-embryonic stages. Based on our quantitative analyses, we found that developmental stages expressing higher ratios of temporally pleiotropic genes significantly correlated with their evolutionary conservation estimated using 1:1_expDists (Fig. 4a), which was also supported by ortholog group expDists (Supplementary Fig. 12a). The correlation was also observed when different criteria were used to define temporally pleiotropic genes (Supplementary Fig. 13). While the correlation was significant in all the phylogenetic groups tested, it tended to be weaker in large phylogenetic groups (Fig. 4a and Supplementary Fig. 12a). To tease apart the factors driving these correlations, we performed simple and partial correlation analyses, together with general linear model analysis, on pleiotropic expression, total expression levels, the dN/dS ratio of genes, the ratio of ancestral genes and 1:1_expDists in the turtle and chicken stage combinations (Supplementary Fig. 12b–e and Supplementary Tables 14 and 15). While total gene expression levels and the dN/dS ratio were correlated with evolutionary conservation (Supplementary Fig. 12b–e), their confounding effects were not large enough to cancel the correlation between the temporally pleiotropic expression and evolutionary conservation, and the temporally pleiotropic gene ratio was the largest explanatory factor among the factors tested (Supplementary Tables 14–16). In addition to these gene set level results, the correlation between pleiotropic (including ubiquitous) expressions and evolutionary conservation was also observed at a gene-by-gene level. By analysing ortholog expression at the corresponding developmental stages of different species (Supplementary Fig. 12j), we found that genes expressed in more developmental stages tended to show conserved expression levels between the mouse and chicken (Fig. 4b and Supplementary Fig. 12g–i) and between the two *Xenopus* species (Supplementary Fig. 12f). Further analyses revealed that the temporally pleiotropic mouse genes tended to be essential for normal embryogenesis with a stronger tendency for developmental genes (Fig. 4c) and were highly connected protein–protein interaction nodes (Supplementary Fig. 11e,f), suggesting their functional importance.

Discussion

Our results indicate that hourglass-like conservation can be observed for chordates when 1:1 ortholog sets are used for comparison and that those conserved stages include developmental phases that exhibit bodyplan elements, except for in the ascidian (Fig. 2 and Supplementary Figs. 4–6). These results support the idea that a bodyplan-defining phase once existed in the chordate common ancestor and extant species still retain the phase as a highly conserved mid-embryonic period (as well as the idea that the phylotype hypothesis stands for phylum-wide species)[9]. However, the results were based on expression profiles of only 1,704 orthologs (for the 1:1_expDists method) and hourglass-like conservation for chordates was not supported by the results from ortholog group expDists analyses (Supplementary Figs. 4b, 5b and 6b,d). Considering the potential advantage of ortholog group expDists over 1:1_expDists (Supplementary Fig. 7), the fact that major chordate P_{top} signals were considerably shifted and widened in some species (Fig. 2 and Supplementary Figs. 4a, 5a and 6a,c) and the fact that ascidian embryogenesis lacks stages that show all the bodyplan elements simultaneously[23,24], the idea that the phylotype hypothesis stands for phylum-wide species may not be as robust as was previously considerd[9,21] (Supplementary Note 8). In contrast, none of the earliest or latest stages in vertebrates was found to have the highest vertebrate P_{top} scores and this tendency was consistently observed both in 1:1_expDists (Fig. 2) and ortholog group_expDists (Supplementary Fig. 8), suggesting that hourglass-like conservation may stand for vertebrates. However, similar to chordates, earlier stages were also identified to be vertebrate-conserved in the chicken, turtle and *Xenopus* species (Fig. 2 and Supplementary Fig. 8), indicating that mid-embryonic conservation in vertebrates was not as specific compared with smaller phylogenetic groups. Of note, despite widened conservation of stages in vertebrates, stages that exhibit organ primordials of the vertebrates' shared anatomical pattern[29] were fairly covered. Although further studies are required as we did not include cyclostomes, such as hagfish and lampreys, our results imply that the phylotype hypothesis of the hourglass model can only weakly be supported for chordates, but can be applied to vertebrates (Supplementary Note 9).

Another finding was that the vertebrate-conserved mid-embryonic stages largely overlapped more with those in the less inclusive groups than the vertebrates, such as tetrapods and amniotes. This implies that these stages have been the major evolutionary conservation target throughout vertebrate evolution (Fig. 2 and Supplementary Figs. 4 and 8) even though vertebrates experienced major environmental changes during their evolution, such as exposure to sea, freshwater, land and air. To investigate the 'frozen' vertebrate basic anatomical pattern through the macroevolutionary timescale[2], we next focused on the common features of the conserved developmental system and found that the mid-embryonic genetic system of vertebrates consists of genes repeatedly used (pleiotropically expressed) at various developmental stages (Fig. 3 and Supplementary Figs. 9 and 10). As expected, these temporally pleiotropic genes strongly correlated with spatial distribution (Supplementary Fig. 11a,b), and those detected at mid-embryonic stages were enriched with genes involved in anatomical patterning (Supplementary Fig. 11g). Given that the temporally pleiotropic expressions were predicted to have complex regulation (Supplementary Fig. 11c,d) and highly connected protein nodes (Supplementary Fig. 11e,f) and to contain more essential genes for normal development (Fig. 4c), they may also be multifunctional genes (Supplementary Note 10). Importantly, we found that the developmental stages enriched with temporally pleiotropic genes tended to be evolutionarily conserved, and the correlation tended to be weaker when the analysis was performed for higher taxonomic groups (Fig. 4a,b and Supplementary Figs. 12 and 13). This correlation was observed even after controlling for effects from total gene expression levels, dN/dS and ancestral gene ratio in turtle and chicken (Supplementary Tables 14–16). Given these findings, we propose that one of the factors that contributed to the conservation of mid-embryonic vertebrate developmental systems (including those involved in establishing the vertebrate Bauplan) is pleiotropic constraints[32,33] (Fig. 4d). As has been predicted by theoretical studies, traits that share an abundance of the same genetic machinery (such as arms and legs or serially homologous traits) would be pleiotropically constrained and thus undergo individuation with difficulty[34] (Supplementary Note 11). While our results

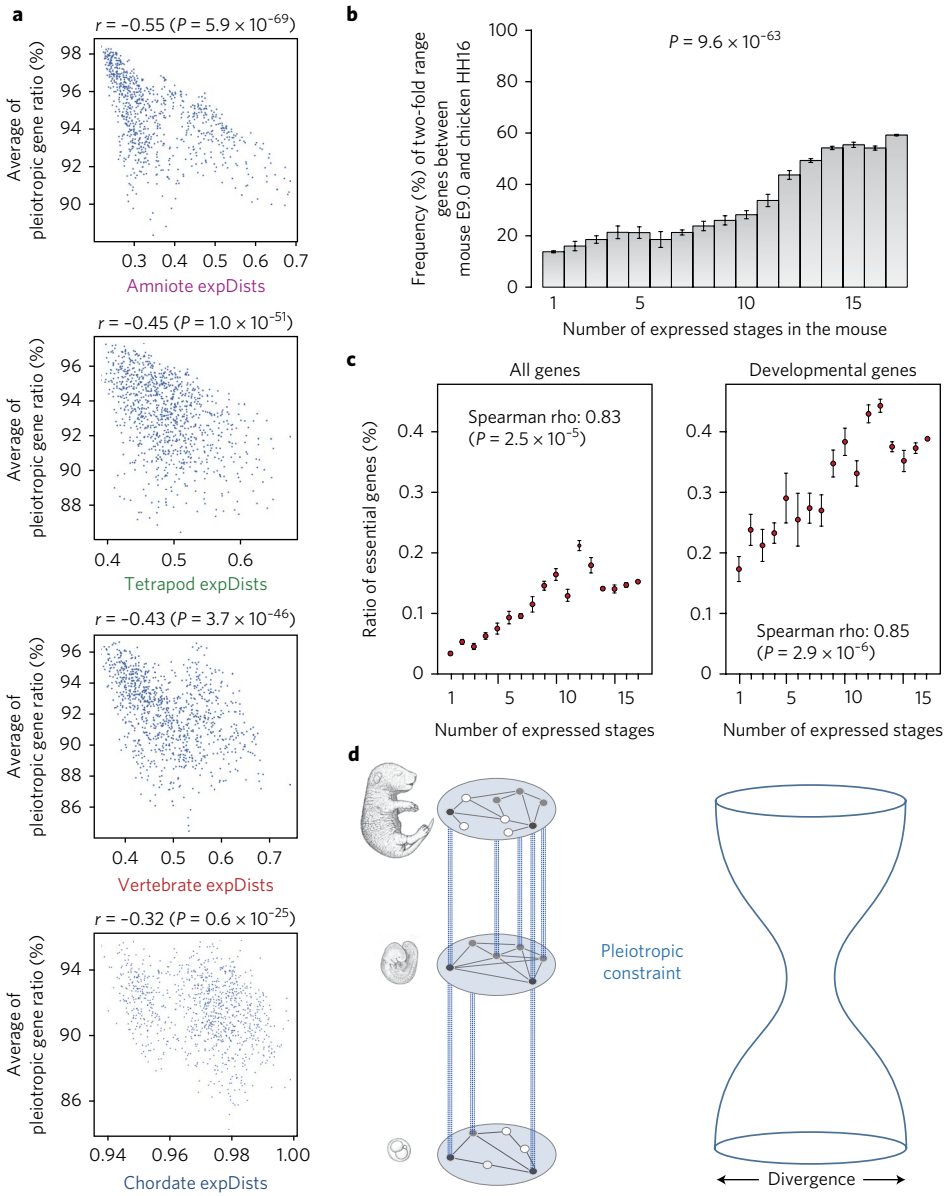

Fig. 4 | Temporally pleiotropic genes tend to be evolutionarily conserved and essential for normal development. a, Developmental stages with more pleiotropic genes (defined as genes expressed in >50% of analysed stages) tend to be conserved and the tendencies are stronger in those analysed with smaller phylogenetic groups. The *x* axis indicates 1:1 expDists of 1,000 randomly picked-up stage combinations. Pearson's correlation coefficients and *P* values (test of no correlation) are shown for these relationships. *y* axis shows the average of the pleiotropic gene ratio of each stage combination. Scores of 1,000 randomly picked-up stage combinations are shown for each result. The pleiotropic gene ratio (detected pleiotropic genes/total detected genes) in each stage was calculated using mean expression levels of biological replicates. The average of these ratios was used to calculate the average pleiotropic ratio in each stage combination. **b**, Similarity between expression levels of 1:1 orthologs detected in mouse E9.0 and chicken HH16 were analysed for the relationship between temporally pleiotropic expression (see also Supplementary Fig. 12 for stage correspondences defined by expression similarity). Mouse orthologs were initially categorized by the number of detected developmental stages, and each of these categorized genes was further analysed for the frequency of genes that showed conserved expression levels (within a two-fold range) between these two mouse–chicken stages. Error bars indicate the s.d. of the frequencies calculated from 100 BRI-exp data. Changes in frequencies significantly differed among the number of expressed stages in the mouse (analysis of variance; *P* value calculated using 100 BRI-exp data). **c**, Genes expressed in many stages tend to be essential for normal embryogenesis in the mouse (Spearman rho > 0.8; *P* values represent tests of no-correlation). The ratio of essential genes (2,247 genes in total) to the number of genes and developmental genes expressed (>1 FPKM) at each stage was calculated. Error bars indicate the s.d. of the ratio of essential genes calculated from the 100 BRI-exp data. **d**, Potential contribution of hourglass-like conservation of vertebrate mid-embryos by pleiotropic constraints. Vertebrate mid-embryonic molecular components are highly pleiotropic (grey and black circles) and are shared (blue vertical line) by other developmental process. This 'chained' status makes it difficult for pleiotropic genes to change even at the gene regulation level, leading to conservation of a mid-embryonic developmental system that is largely composed of pleiotropic genes.

provide empirical support for this idea, a variety of mechanisms have been proposed as pleiotropic constraints[34,35] and further studies are required to elucidate the actual mechanism, particularly with regard to the causal relationship between pleiotropic gene expressions and conservation (or evolvability[35]) of these genes at a regulatory level. One potential mechanism of the pleiotropic constraint at the gene

regulatory level could be that changes in regulations of pleiotropic genes tended to be less adaptive. For example, genes utilized at the vertebrate mid-embryonic (especially pharyngula) phase are repeatedly used in various developmental processes and are considered to be highly responsible for (or burdened[36] by) a variety of biological processes. Because of this, changes in the regulation of these genes (for example, mutation to promoter) would cause disadvantageous effects in a set of tissues or cells using the same genetic system and this may have led to the evolutionary conservation of the mid-embryonic system and its phenotypes (Supplementary Note 12). It is also possible that the expression of pleiotropic genes tends to be stable against mutational and/or environmental changes, perhaps by having higher chances of functionally overlapping regulatory elements[37,38], and this may have led to less variation in gene regulation within a population, contributing to its evolutionary conservation (Supplementary Note 13). A report by Papakostas et al.[39] may be consistent with this idea; by studying grayling fish, they reported that expression levels of pleiotropic genes (defined by the number of protein interactions or gene biological processes) were less responsive to both environmental and evolutionary changes. In summary, although further studies are needed, our results highlight the potential contribution of pleiotropic constraints to hourglass-like conservation, particularly trait conservation on a macro-evolutionary scale (Supplementary Note 14). It would be interesting to determine whether the possible contribution of pleiotropic constraints can be observed in other phylogenetic groups, particularly nematodes, as potential contributions of developmental constraints (albeit by unknown mechanisms) towards the conserved mid-embryonic stages of *Caenorhabditis elegans*[40] have recently been reported. Considering that an increasing number of studies have reported major roles of rewiring gene regulatory networks in macroscale evolutionary transitions[41,42] (Supplementary Note 12), studying the relationships among pleiotropic constraint at regulatory level, co-options[43] and variation in pleiotorpy[44] towards phenotypic evolvability would be of special interest; as has also been predicted in the 'cost of complexity'[45,46] hypothesis, while recruitment of a limited genetic repertoire (for example, developmental toolkit genes[43]) allowed animals to adopt evolutionary novelties, the concept of pleiotropic constraint together with our findings implies that this may have in turn limited evolutionary diversification. Our findings also provide a basis for objectively defining the long-argued[47] circulatory concepts of phylum and bodyplan (phylum being defined as a mono-phyletic group of species that share the bodyplan, and bodyplan as a set of anatomical features shared among species in the same phylum) by grouping species that show persistent conservation. Taken together, we propose the possibility that the phylotype hypothesis of the hourglass model may better be applied to vertebrates, and that pleiotropic constraints of the developmental system in the pharyngula stages may have contributed to hourglass-like conservation on the macro-evolutionary scale, which ultimately led to the recognition of the conserved organogenesis stages in vertebrates (the vertebrates' Bauplan phase) or chordates (the phylotypic period). However, it remains unclear how genes with pleiotropic expressions have concentrated around organogenesis stages in the vertebrate species.

Methods

Animal care and use. Experimental procedures and animal care were conducted in strict accordance with guidelines approved by the RIKEN Animal Experiments Committee (approval ID: H14-23, H16-10 and AH21-08). All efforts were made to minimize suffering. Individual animals and embryos were selected blindly from wild types.

Embryo collection, RNA extraction and RNA sequencing. In addition to using previously published RNA sequencing (RNA-seq) datasets (DNA Data Bank of Japan accession DRA000567) for chicken (*Gallus gallus*) and turtle (*Pseudotsuga sinensis*) embryos, we generated a new RNA-seq dataset by collecting early-to-late stage whole embryos of *Mus musculus*, *G. gallus*, *P. sinensis*, *Xenopus tropicalis*, *Xenopus laevis*, *Danio rerio*, *Ciona intestinalis* and *Branchiostoma floridae*, which included the early-to-late developmental stages of each species. For each species, stage-determined embryos were pooled from 2–500 randomized individuals, and at least two of these biologically independent samples were prepared for each developmental stage to statistically represent the general population of each stage, giving a total of 260 samples. An RNeasy Lipid Tissue kit (QIAGEN), RNeasy Micro Kit (QIAGEN) or Quartz-Seq[48] was used for the RNA extraction depending on the sample (Supplementary Tables 1–8). After adjusting the amounts of extracted RNA between samples, RNA-seq libraries were constructed using the Illumina TruSeq protocol according to the manufacturer's instructions, except for mouse early embryos (two-celled blastocysts), which were prepared using the Quartz-Seq[48] protocol. An Agilent 2100 BioAnalyzer was used to check RNA quality before and during RNA-seq library construction. Illumina's Hiseq 2000 system was used for sequencing (100 base single read, non-strand specific). Quality checks of the RNA-seq raw reads were inspected using the FastQC programme (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Calculation of expression levels for each gene. RNA-seq reads were mapped to species-specific reference genomes using bowtie2 (version 2.2.3)[49] and Tophat2 (version 2.0.11)[50]. To obtain the best genome-mapped reads ('bestHit data'), the '-g 1' option of Tophat2 was used[50]. To fully encompass the genes expressed in each animal, we upgraded the coding gene annotations of *X. laevis* and *B. floridae*, which had a relatively poorly predicted gene set (Supplementary Fig. 3 and Supplementary Tables 9–12). We also confirmed that saturating numbers of gene repertoires were detected in each sample, even at later stages (Supplementary Fig. 2b). To make sure that our results and conclusions were not affected by mapping conditions, we obtained mapped data using multi-mapped criteria (referred to as 'multiHits data' in this study, which allowed, at most, 20 loci for each mapped read). Only the results or tendencies supported by both methods (bestHit data and multi-mapped) were considered. To obtain expression scores based on these mapped datasets, we used Cufflinks (version 2.2.1)[51] and calculated the fragments per kilobase of exon per million mapped fragments (FPKM) values (the expression data are available as Supplementary Data). In some analyses (where appropriate, as specified in each figure legend), we further normalized expression values among different developmental stages by trimmed mean of *M* values (TMM)[52] normalization to confirm that results or tendencies did not vary according to the normalization method used (that is, non-TMM normalized versus TMM normalized data).

Biological replicates of expression data. For each gene, single early-to-late expression patterns during development can be obtained by calculating the mean values of biological replicates for each developmental stage. However, this may cause bias (or false positive and false negative results) specific to the sample set obtained in this study. To avoid this potential bias, we incorporated deviations of gene expression levels among different biological replicates (of each developmental stage). For example, one early-to-late expression pattern could be obtained by randomly choosing one biological replicate sample for each developmental stage. Given that many combinations could be created using biological replicate samples, we repeated the process and created biological-replicate-included expression (BRI-exp) tables. As an example, species with five developmental stages and three biological replicates for each stage could have up to $3^5 = 243$ expression tables.

Gene annotation update of *B. floridae* and *X. laevis* genomes. We updated the gene annotations of *B. floridae* and *X. laevis* using the consensus gene prediction based on multiple methods (homology-based, de novo and RNA-seq predictions) by fully utilizing the very large RNA-seq dataset obtained in this study. A detailed description of gene prediction as well as quality evaluation is in the Supplementary Information.

Identification of 1:1 orthologs. To identify 1:1 orthologs that had a single ortholog counterpart in each species, peptide sequences of each species (Supplementary Table 13) were used. After removing entries shorter than 30 amino acids, the longest peptide sequence for each gene was selected and used to identify orthologs between different species. Before the identification of 1:1 orthologs, we initially evaluated the precision and prediction sensitivity of the previously used[15,18,28] reciprocal best blast hit (RBBH) method by comparing it with orthoMCL[53] and proteinortho[54], since these tools are considered to be among the top ortholog identification tools[54,55]. For the RBBH-based method, 1:1 orthologs were selected based on the combination of the lowest e-value and bit-score reported from BLASTP (blast+2.28). For orthoMCL and proteinortho, default parameters were used to identify 1:1 orthologs. On the basis of the *F*-scores that were calculated as the harmonic mean of the sensitivity and precisions of each method, we decided to use the RBBH-based method for defining the 1:1 orthologs. The quality of 1:1 orthologs obtained using the RBBH-based method was also validated by the high consistency with the 1:1 orthologs reported by Ensembl Compara, based on the gene sets of five species (the chicken, turtle, mouse, western clawed frog and zebrafish). The other three species were not available in the Ensembl database).

Calculation of distances in gene expression profiles among embryos from different species. To identify evolutionarily conserved developmental stages, we initially calculated stage–stage pairwise expression distances between different species using expression profiles of 1:1 orthologous genes (defined by the reciprocal best BLAST hit; Supplementary Table 13) as in previous studies[15,18,28,56]. To avoid problems arising from ignoring the phylogenetic relationships of the compared species[57,58], we next extracted the pairwise distances that reflected the phylogenetic scale of interest, calculated their average values and defined these as 1:1 ortholog-based expression distances (expDists) or 1:1_expDists of the stages being compared (see 'Calculation of expDists'). After calculating and comparing the 1:1_expDists for all possible stage combinations of the target species (for stage combinations of vertebrates and chordate species, up to one million stage combinations were randomly selected to avoid computational barriers), we extracted the top 1% of stage combinations that were most similar (that is, those with the lowest expDists values) and visualized the contributions of each developmental stage to the top 1% of stage combinations (that is, the percentage being included in the top 1% of similar combinations, or P_{top}; see Fig. 1 and Supplementary Information for detailed methods) to evaluate conserved stages in each species.

Calculation of expDists. *Identification of the expression levels of each gene (step 1).* FPKM values (gene-length normalized values) and TMM-normalized FPKM values were used for the expression levels as the length of genes varied between different species. No cutoffs were applied to the expression levels.

Cross-species expression distances (step 2). On the basis of the whole embryonic gene expression levels identified in step 1, pairwise distances between pairs of embryos from different species were calculated as 1 – the Spearman's rank correlation coefficient. For the 1:1 ortholog-based method, only the expression levels of genes identified as 1:1 orthologs were compared between the different species to calculate the expression distance. In this case, genes that lacked 1:1 ortholog counterparts in any of the species (for example, by gene loss in any of the species being compared) were ignored. For the ortholog group-based method, gene expression levels of in-paralogs defined by orthoMCL were either summed or averaged (mean version of ortholog group method) and further compared between different species to calculate the expression distances.

Calculation of expDists (step 3). Using the pairwise distances between pairs of embryonic samples (calculated in steps 1 and 2) and the known phylogenetic relationship of species being compared, we calculated expDists for each phylogenetic level (chordates, vertebrates, tetrapods, amniotes, frogs, and chickens and turtles). For example, for clade 'C' consisting of species W, X, Y and Z with a phylogenetic relationship (W, (X, (Y, Z))), the average pairwise distances (clade C-expDists) of the three pairs W–X, W–Y and W–Z were calculated. Other combinations (X–Z, X–Y and Y–Z) were not calculated because they do not reflect the evolutionary timescale from the emergence of a common ancestor of W–X–Y–Z. Similarly, the expDists of clade D with the phylogenetic relationship ((W, X), (Y, Z)) would be calculated as the average pairwise distances of W–Y, W–Z, X–Y and X–Z. As in many other evolutionary studies based on comparisons of extant species, expDists largely depended on the outermost or earliest diverged species.

Estimation of the top 1% of conserved developmental stages and visualization by P_{top}. To evaluate the most conserved stages by expDists, we first identified the lowest 1% of expDist stage combinations from all the possible combinations, except for vertebrates and chordates (one million random stage combinations were selected for these two groups as all the stage combinations of these two groups far exceeded the computable capacity owing to combinatorial expansion). We then visualized contributions of each developmental stage to this lowest 1% of similar stage combinations by P_{top}. P_{top} denotes the percentage of stage included in the top similar 1% of stage combinations. For example, if 100 stage combinations from amniotes were identified as the lowest 1% of expDists and a stage of interest (for example, chicken_st.X) was included in 30 of these combinations, P_{top} would be calculated as $30/100 = 0.3$ (30%). We also took advantage of biological replicates for each stage in our analyses to draw statistically robust conclusions (for example, we created ten versions of gene expression tables for each species with biological replicates; see Supplementary Information for more detail).

Detection of stage-associated up- and downregulated genes. Stage-associated upregulated genes (genes that are highly expressed at certain stages) were detected using the method reported by Li et al.[59]. Briefly, expressed genes (FPKM > 1.0) with a Z-score > 1.0 (compared with expression levels among all the stages) were identified as stage-associated upregulated genes. While there are two methods to calculate Z-score (using standard deviation (s.d.) and standard error of the mean (s.e.m.)), the Z-scores in this study were calculated using s.d. to avoid unwanted bias from the different numbers determined for different developmental stages. A FPKM cutoff value of 1.0 was used for detecting stage-associated upregulated genes as in a previous study[59]. To identify stage-associated down-regulated genes (genes showing low expression levels at certain stages), a Z-score < –1.0 (exactly the opposite distance from the mean value compared with upregulated stage-associated genes) was used, and no FPKM cutoff was used to include genes that were downregulated even to a FPKM value of 0. Note that expression levels of the 'ortholog groups' were not used to identify stage-associated genes.

Detection of temporally pleiotropic genes. Genes expressed in more than 50% of developmental stages were defined as temporally pleiotropic genes. These included the genes expressed in all stages (ubiquitously expressed genes). Because the developmental system is not completely independent from ubiquitously expressed genes, we considered it appropriate to include them (including housekeeping genes) in analyses of the conserved developmental systems.

Gene ontology. A total of 6,203 developmental gene ontology terms[60,61] were extracted from go.obo (release 25 February 2015). Extraction of the developmental gene ontology terms (0032502 (developmental process) as an ancestor and self) was conducted using OBO edit (http://oboedit.org/), as described previously[12,18]. Developmental genes were defined as coding genes annotated with developmental gene ontologies. The gene ontology slim file was downloaded from http://geneontology.org on 14 March 2015. Gene ontology term enrichment analyses were conducted using gene ontology slim in the R package 'GSEAbase' and further analysed for enrichment. Essential genes in the mouse were adopted from Georgi et al.[62].

Statistical tests and sample size. To avoid inflating the type I error rate, an alpha level of 0.01 (further Bonferroni correction in case of multiple comparisons) was accepted as statistically significant throughout the analyses unless otherwise specified. Statistical methods were chosen to properly reflect the population of interest. A Welch two-sample *t*-test was used for two-sample comparisons when the data passed the Kolmogorov–Smirnov test for normal distribution; otherwise, a Wilcoxon signed-rank test was used. Statistically effective numbers of biological replicates for each stage were estimated from the chicken and turtle RNA-seq dataset obtained previously[18].

Software. In addition to the programmes specified elsewhere in this document, we also used Perl, Python, R (together with the GSEABase, latticeExtra, reshape, gplots, vegan, maptools, gtools, RColorBrewer, ggplot2, WGCNA, TCC and DEGseq packages) and shell scripts to create the analysis programmes.

Data availability. The RNA-seq data that support the findings of this study have been deposited in the DNA Data Bank of Japan and are accessible through accession number DRA003460 (experiment number DRX029460–DRX029731).

Received: 22 November 2016; Accepted: 16 August 2017;
Published online: 25 September 2017

References

1. Erwin, D. H. et al. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* **334**, 1091–1097 (2011).
2. Wallace, A. *The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology* (Cambridge Univ. Press, Cambridge, 2000).
3. Galis, F. & Metz, J. A. Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation. *J. Exp. Zool.* **291**, 195–204 (2001).
4. Irie, N. & Kuratani, S. The developmental hourglass model: a predictor of the basic body plan? *Development* **141**, 4649–4655 (2014).
5. 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.* **9**, e1003476 (2013).
6. Kalinka, A. T. & Tomancak, P. The evolution of early animal embryos: conservation or divergence? *Trends Ecol. Evol.* **27**, 385–393 (2012).
7. 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.* 135–142 (1994).
8. Sander, K. Specification of the basic body pattern in insect embryogenesis. *Adv. Insect Physiol.* **12**, 125–238 (1976).
9. Raff, A. *The Shape of Life: Genes, Development, and the Evolution of Animal Form* (Univ. of Chicago Press, Chicago, 1996).
10. Richardson, M. K., Minelli, A., Coates, M. & Hanken, J. Phylotypic stage theory. *Trends Ecol. Evol.* **13**, 158 (1998).
11. 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.* **304**, 150–158 (2005).
12. Irie, N. & Sehara-Fujisawa, A. The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. *BMC Biol.* **5**, 1 (2007).
13. Domazet-Loso, T. & Tautz, D. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* **468**, 815–818 (2010).

14. Kalinka, A. T. et al. Gene expression divergence recapitulates the developmental hourglass model. *Nature* **468**, 811–814 (2010).
15. Irie, N. & Kuratani, S. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat. Commun.* **2**, 248 (2011).
16. Quint, M. et al. A transcriptomic hourglass in plant embryogenesis. *Nature* **490**, 98–101 (2012).
17. Levin, M., Hashimshony, T., Wagner, F. & Yanai, I. Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Dev. Cell* **22**, 1101–1108 (2012).
18. Wang, 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.* **45**, 701–706 (2013).
19. Cheng, X., Hui, J. H., Lee, Y. Y., Wan Law, P. T. & Kwan, H. S. A "developmental hourglass" in fungi. *Mol. Biol. Evol.* **32**, 1556–1566 (2015).
20. Xu, F. et al. High expression of new genes in trochophore enlightening the ontogeny and evolution of trochozoans. *Sci. Rep.* **6**, 34664 (2016).
21. Levin, M. et al. The mid-developmental transition and the evolution of animal body plans. *Nature* **531**, 637–641 (2016).
22. Irie, N. Remaining questions related to the hourglass model in vertebrate evolution. *Curr. Opin. Genet. Dev.* **45**, 103–107 (2017).
23. Chiba, S., Sasaki, A., Nakayama, A., Takamura, K. & Satoh, N. Development of *Ciona intestinalis* juveniles (through 2nd ascidian stage). *Zool. Sci.* **21**, 285–298 (2004).
24. Hotta, K. et al. A web-based interactive developmental table for the ascidian *Ciona intestinalis*, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. *Dev. Dynam.* **236**, 1790–1805 (2007).
25. Holland, L. Z. Genomics, evolution and development of amphioxus and tunicates: the Goldilocks principle. *J. Exp. Zool. B* **324**, 342–352 (2015).
26. Janvier, P. Facts and fancies about early fossil chordates and vertebrates. *Nature* **520**, 483–489 (2015).
27. Albalat, R. & Canestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **17**, 379–391 (2016).
28. Yanai, I., Peshkin, L., Jorgensen, P. & Kirschner, M. W. Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Dev. Cell* **20**, 483–496 (2011).
29. Benton, J. M. *Vertebrate Palaeontology*. 4th edn (Wiley-Blackwell, Hoboken, NJ, 2014).
30. Roux, J., Rosikiewicz, M. & Robinson-Rechavi, M. What to compare and how: comparative transcriptomics for Evo-Devo. *J. Exp. Zool. B* **324**, 372–382 (2015).
31. Jeffery, J. E., Bininda-Emonds, O. R., Coates, M. I. & Richardson, M. K. Analyzing evolutionary patterns in amniote embryonic development. *Evol. Dev.* **4**, 292–302 (2002).
32. Galis, F. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. *J. Exp. Zool.* **285**, 19–26 (1999).
33. Duboule, D. & Wilkins, A. S. The evolution of 'bricolage'. *Trends Genet.* **14**, 54–59 (1998).
34. Pavlicev, M. & Cheverud, J. M. Constraints evolve: context dependency of gene effects allows evolution of pleiotropy. *Annu. Rev. Ecol. Evol. Syst.* **46**, 413–434 (2015).
35. Wagner, G. P. & Zhang, J. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* **12**, 204–213 (2011).
36. Riedl, R. *Order in Living Organisms: a Systems Analysis of Evolution* (Wiley, Hoboken, NJ, 1978).
37. Hong, J. W., Hendrix, D. A. & Levine, M. S. Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314 (2008).
38. Perry, M. W., Boettiger, A. N. & Levine, M. Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA* **108**, 13570–13575 (2011).
39. Papakostas, S. et al. Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. *Nat. Commun.* **5**, 4071 (2014).
40. Zalts, H. & Yanai, I. Developmental constraints shape the evolution of the nematode mid-developmental transition. *Nat. Ecol. Evol.* **1**, 0113 (2017).
41. Seki, R. et al. Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. *Nat. Commun.* **8**, 14229 (2017).
42. Lindblad-Toh, K. et al. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482 (2011).
43. Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* 9th edn (Blackwell, Malden, MA, 2005
44. Pavlicev, M. & Wagner, G. P. A model of developmental evolution: selection, pleiotropy and compensation. *Trends Ecol. Evol.* **27**, 316–322 (2012).
45. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
46. Orr, H. A. Adaptation and the cost of complexity. *Evolution* **54**, 13–20 (2000).
47. Satoh, N., Rokhsar, D. & Nishikawa, T. Chordate evolution and the three-phylum system. *Proc. R. Soc. B* **281**, 20141729 (2014).
48. Sasagawa, Y. et al. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol.* **14**, R31 (2013).
49. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
50. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
51. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
52. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
53. Li, L., Stoeckert, C. J. Jr. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).
54. Lechner, M. et al. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**, 124 (2011).
55. Chen, F., Mackey, A. J., Vermunt, J. K. & Roos, D. S. Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS ONE* **2**, e383 (2007).
56. Tena, J. J. et al. Comparative epigenomics in distantly related teleost species identifies conserved cis-regulatory nodes active during the vertebrate phylotypic period. *Genome Res.* **24**, 1075–1085 (2014).
57. Hejnol, A. & Dunn, C. W. Animal evolution: are phyla real? *Curr. Biol.* **26**, R424–R426 (2016).
58. Dunn, C. W., Zapata, F., Munro, C., Siebert, S. & Hejnol, A. Pairwise comparisons across species are problematic when analyzing funcional genomic data. Preprint at http://www.biorxiv.org/content/early/2017/02/09/107177 (2017).
59. Li, J. J., Huang, H., Bickel, P. J. & Brenner, S. E. Comparison of *D. melanogaster* and *C. elegans* developmental stages, tissues, and cells by modENCODE RNA-seq data. *Genome Res.* **24**, 1086–1101 (2014).
60. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucl. Acids Res.* **43**, D1049–D1056 (2015).
61. Ashburner, M. et al. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
62. Georgi, B., Voight, B. F. & Bucan, M. From mouse to human: evolutionary genomics analysis of human orthologs of essential genes. *PLoS Genet.* **9**, e1003484 (2013).

Acknowledgements

The EXPANDE (EXpression Profiling AloNg Development and Evolution) project and N.I. was supported in part by Grants in Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan (15H05603, 22128003, 24570243, 3902 and 17H06387) and the Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science and Technology, Japan. T.G.K. was supported in part by KAKENHI 16H04724. The research performed by J.-K.Y. was supported by grants from Academia Sinica and the Ministry of Science and Technology, Taiwan (AS-98-CDA-L06, 102-2311-B-001-011-MY3 and 104-2923-B-001-002-MY3). We thank K. Yamanaka for help collecting the embryos (mouse, chicken, turtle, frog and zebrafish), extracting RNAs and sample preparation using Quartz-Seq for early embryos from mice. We thank C. Tanegashima, K. Itomi, O. Nishimura and S. Kuraku for help constructing libraries, sequencing samples and quality checking some of the RNA-seq data. We thank F. Castellan for critically reading the manuscript. We thank Y. Uchida for providing high-resolution images of zebrafish embryos. We thank G. Renaud and J. Kelso for providing the robust demultiplexing method for analysing the ascidian sequencing data.

Author contributions

N.I., P.K. and S.K. conceived the study. S.F., K.S., T.-M.L., J.-K.Y., T.G.K., Y.S. and N.I. collected the samples. H.H., S.G., S.F., Y.S., T.G.K. and N.I. conducted the experiments needed for RNA-seq. F.L., S.L., G.Z. and H.H. made new gene sets in *X. laevis* and *B. floridae*. H.H., S.G., M.U., P.K., M.I. and N.I. analysed the data. N.I., J.-K.Y., M.U., T.G.K. and S.K. edited the paper. N.I. and P.K. supervised the project.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at doi:10.1038/s41559-017-0318-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.I.