

Minireview

Comparative embryology without a microscope: using genomic approaches to understand the evolution of development

David A Garfield and Gregory A Wray

Address: Biology Department and Institute for Genome Science & Policy, Duke University, Durham, NC 27708, USA.

Correspondence: David Garfield. Email: dag23@duke.edu. Gregory Wray. Email: gwray@duke.edu

Abstract

Until recently, understanding developmental conservation and change has relied on embryological comparisons and analyses of single genes. Several studies, including one recently published in *BMC Biology*, have now taken a genomic approach to this classical problem, providing insights into how selection operates differentially across the life cycle.

It is an idea as old as the study of development itself: embryos of different species are more similar during earlier stages than later ones [1]. The first detailed descriptions came from Karl Ernst von Baer, whose meticulous observations of vertebrate embryos invented the field of developmental biology. Although von Baer's eponymous 'law' stating that early development is more conserved than later development was formulated without an explicit evolutionary context, Charles Darwin considered embryological similarity to be one of the most powerful forms of evidence for common descent when he wrote the *Origin of Species*. During the 150 years that have passed since that momentous publication, embryologists have uncovered numerous exceptions to von Baer's generalization [2]. Yet it remains true that early development is often remarkably conserved among even distantly related species. Exactly why this should be so remains unclear.

Alternative views of developmental constraint

One view is that developmental similarity is the result of functional constraint, and thus maintained by negative selection. The basic idea is that the processes of early development influence many later processes, so functional changes in genes underlying early development will generally be deleterious because of extended pleiotropy (Figure 1b, grey line). A variant promoted by Rudy Raff and others [2] argues that constraint is highest somewhat later in embryogenesis, when the adult body plan is laid down (Figure 1b, black line). But the two views share at their core the idea that there are points in development at which the effects of mutations will have disproportionate, and usually deleterious, consequences.

An alternative view is that conservation in early development results from developmental 'buffering'. Following the pioneering work of Conrad Waddington, the argument is that development is an inherently self-correcting process that buffers the effects of variation early in the life cycle more than later. As a result, mutations affecting genes acting at some developmental stages are less likely to have phenotypic consequences. This leads to the appearance of conservation in early development despite a seemingly paradoxical relaxation of constraint for developmental genes expressed during conserved stages (Figure 1c).

A third view is that divergence in developmental programs happens only when there is an adaptive reason. Walter Garstang was among the first to argue that embryonic development can readily evolve when ecological circumstances demand it. Conservation is common, he argued, simply because embryonic development tends to happen away from the influence of the environment (for instance, within a uterus or an egg case), and is thus largely screened from positive selection. In this view, it is patterns of developmental divergence rather than conservation that are most interesting (Figure 1d).

These three views offer strikingly different pictures of the role that developmental changes have in origin of adaptations and organismal diversity. Discovering which view, or more likely which combination, best explains patterns of conservation and divergence in development is central to understanding the origins of animal diversity.

Taking it to the genome

The advent of genome-scale datasets provides an exciting new approach for evaluating these views. This approach, which has now been applied in several studies [3-7], begins with measurements of transcript abundance throughout the genome (based on microarray or expressed sequence tag (EST) data) for several stages of development taken from one or (ideally) more species. Next, for each gene for which orthology can be confidently determined, the ratio of non-synonymous (dN) to synonymous (dS) substitutions

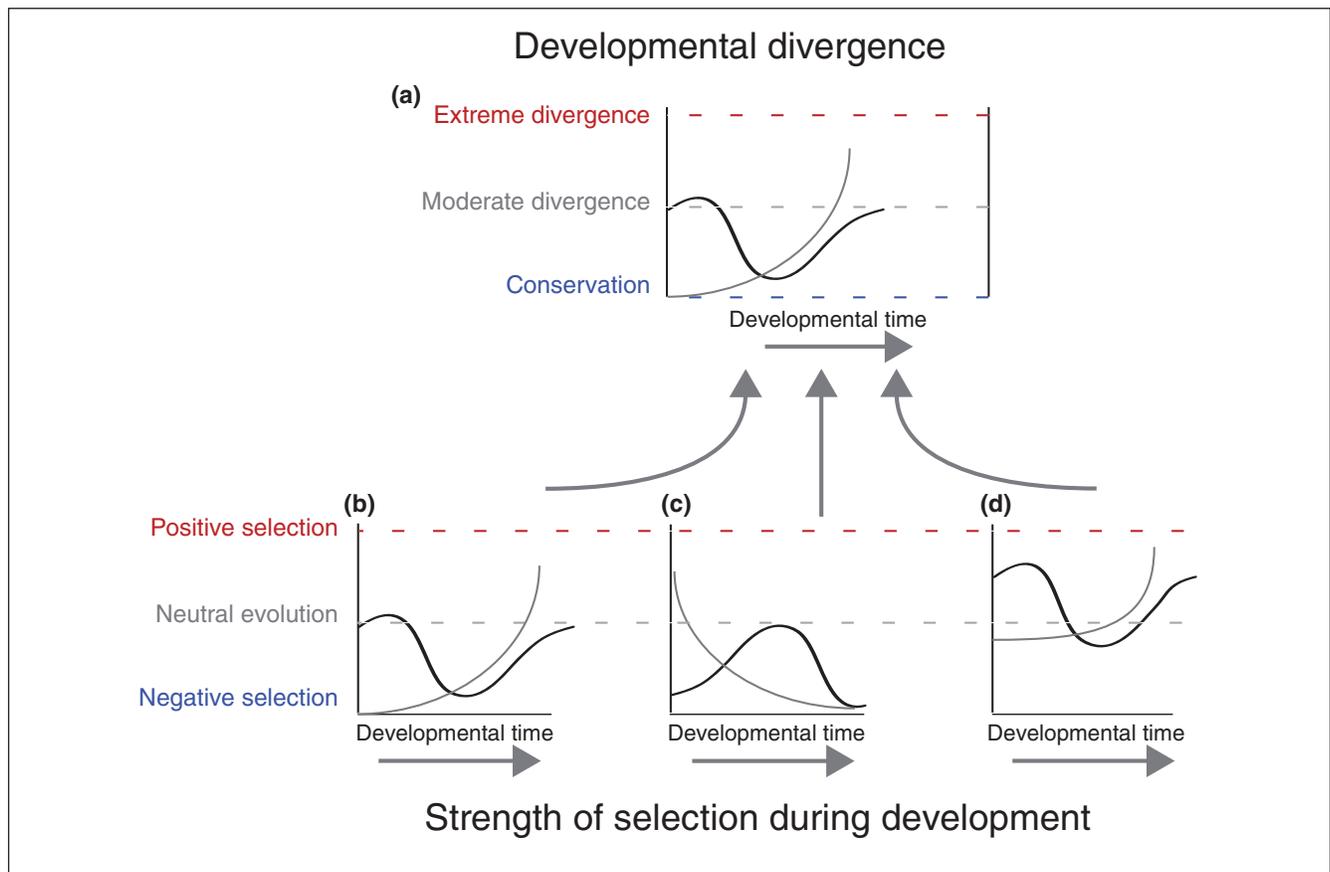


Figure 1

Expected relationships between developmental divergence and the strength of selection. **(a)** Two potential patterns of conservation at the level of embryonic development. The grey line represents the classical pattern of early developmental conservation described by von Baer with the earliest stages of development being more constrained than later stages. The black line shows the 'hour-glass' pattern of conservation described by Raff [1]. **(b-d)** Three predictions about the relationship between these patterns of divergence and natural selection (see text). Grey and black lines show, respectively, the action of natural selection acting on the genome at different times during development under von Baer's and Raff's models of embryological divergence and constraint. (b) Constraint results from negative selection on developmental variation. (c) Developmental buffering results in relaxed constraint on highly buffered developmental stages. (d) Divergence in development is the result of adaptation.

between each species is calculated as a proxy for functional divergence relative to the underlying mutation rate. One can then search for differences in the average value of dN/dS for genes expressed at different stages of development, and test for stage-specific patterns of positive selection ($dN/dS > 1$), negative selection ($dN/dS < 1$), or drift ($dN/dS = 1$) predicted by the different views (Figure 1).

A recent study by Artieri *et al.* [3] is among the largest of this kind so far. It makes use of EST data from three developmental stages of *Drosophila melanogaster* (embryonic, pooled larval/pupal, and adult), sequence data from five species in the melanogaster subgroup, and thousands of randomly selected genes to examine how selective constraint changes as a function of developmental time. The results are intriguing: whereas genes expressed primarily in adults show some evidence of positive selection,

genes classified as embryonic seem to be under strong negative selection, suggesting both that early development is highly constrained (the first viewpoint mentioned above; Figure 1b) and that adult adaptations are primarily the result of genes acting in adults (the third viewpoint; Figure 1d).

Similar studies in nematodes [4,5] and vertebrates (zebrafish and mice [6]) failed to find differences in average dN/dS values at different points in embryonic development, whereas a previous study in *Drosophila* [7] found evidence for significantly stronger constraint (low dN/dS) among genes expressed in late embryogenesis than those expressed earlier in development or later in pupae or adults. These discrepancies suggest that differences in the modes of development used by these three taxa have evolutionary consequences. Two of the studies [4,6] also found evidence that selection on copy number variation

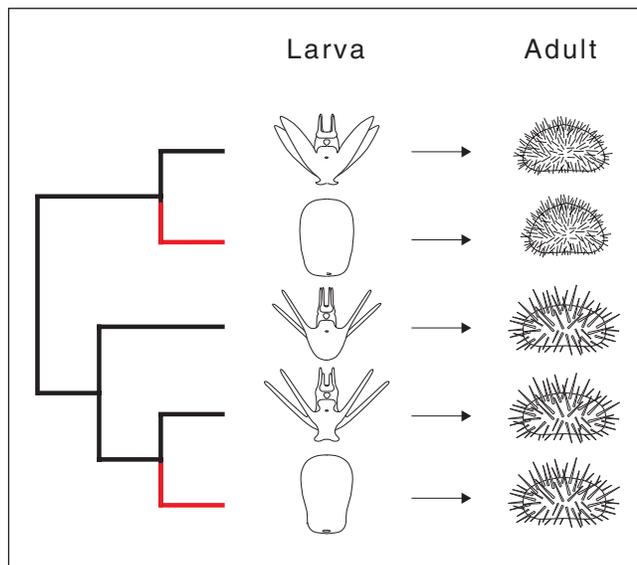


Figure 2

Exploring the effects of changes in life history on development. The lineages in red show two independent shifts to lecithotrophic development (in which the larva does not feed and thus has a much simpler morphology) in euechinoid sea urchins as a result of increases in maternal contributions [1,7]. By comparing convergent changes along the red lineages with those along the black lineages, we can get a sense of the ways in which changes in maternal contribution influence the evolution of development at a genetic level.

was stronger for genes expressed early in development, suggesting that changes in gene expression levels may disproportionately affect early development (a topic to which we return later). In addition, with one exception [4], the studies all found that genes expressed primarily in adults showed less evidence of constraint and more evidence of positive selection than genes expressed in earlier stages. In each case this particular trend was driven by positive selection acting on genes expressed specifically in testes.

Challenges and future directions

Although these genome-scale analyses [3-7] are beginning to provide insights into the action of selection across development, the approach faces several challenges. One is that expression levels of many regulatory genes rise during development, whereas 'housekeeping' genes are more constant. As a result, early development can appear to be enriched for housekeeping genes, such as those required for mitosis. Housekeeping genes are, not surprisingly, highly conserved; this can result in a misleading picture of conservation in very early development, as one study has pointed out [6]. Similarly, failing to correct for testis-specific genes can lead to a false impression, because positive selection on these genes is probably driven by sperm competition rather than stage-specific differences in selection *per se*.

Published studies have considered species with relatively similar ecologies and life-histories, a limitation imposed by the fact that the current genome projects on model organisms cover a restricted phylogenetic distribution. However, it has long been clear that early development can differ enormously among even closely related species [8]. Understanding these exceptions to the conservation of early development poses an important challenge. Cases of extreme divergence in early development are generally interpreted as adaptations driven by changes in life histories; such as modification in embryonic nutrition, altered larval dispersal and defense mechanisms, or by changes in the embryonic environment. Investigating how environmental factors drive the evolution of early development is now possible as new technologies bring genome-scale sequence and expression data from virtually any organism within reach. One possible approach involves comparing species with different life histories or inhabiting different environments (Figure 2). Parallel changes in developmental divergence and conservation along branches leading to evolutionarily derived life histories can provide a deeper understanding of the role that adaptation has in shaping development.

Another important challenge arises from the fact that natural selection can operate on noncoding as well as coding sequences. Indeed, it is in the noncoding regulatory sequences around each gene that we might expect to find an important part of the genetic basis for divergence in expression among species. All the studies published so far have contrasted selection on coding sequences with gene expression across the life cycle. Methods now exist to test for selection in noncoding sequences [9], opening the door to analyses that incorporate selection on regulatory elements. This could provide insights that might be missed by analyses that consider only coding sequences.

These are exciting times for evolutionary biologists, as genome-scale datasets are applied to an ever-expanding range of problems. Understanding how and why natural selection operates differentially across development is among the first instances in which sequence and functional comparisons across the genome have been brought together to address a classical problem in evolutionary biology. The studies carried out so far [3-7] highlight some intriguing trends, especially concerning the potential impacts of changes in gene expression during early development. But this is just the beginning. Although there are some thorny technical problems that need to be addressed, the real promise lies in applying genome-scale data to a much wider range of species contrasts. How does the genome-wide distribution of selection across development change when closely related species occupy very different habitats or differ markedly in their life history? Sampling a wider range of species comparisons may solve one of the oldest conundrums in evolutionary developmental

biology: why development is so often conserved across vast phylogenetic gulfs and yet sometimes spectacularly diverged among closely related species.

References

1. Richards RJ: *The Meaning of Evolution: the Morphological Construction and Ideological Reconstruction of Darwin's Theory*. Chicago: University of Chicago Press; 1992.
2. Raff RA: *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. Chicago: University of Chicago Press; 1996.
3. Artieri CG, Haerty W, Singh RS: **Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of *Drosophila***. *BMC Biology* 2009, **7**:42.
4. Castillo-Davis CI, Hartl DL: **Genome evolution and developmental constraint in *Caenorhabditis elegans***. *Mol Biol Evol* 2002, **19**:728-735.
5. Cutter AD, Ward S: **Sexual and temporal dynamics of molecular evolution in *C. elegans* development**. *Mol Biol Evol* 2005, **22**:178-188.
6. Roux J, Robinson-Rechavi M: **Developmental constraints on vertebrate genome evolution**. *PLoS Genet* 2008, **4**:e1000311.
7. Davis JC, Brandman O, Petrov DA: **Protein evolution in the context of *Drosophila* development**. *J Mol Evol* 2005, **60**:774-785.
8. Wray GA: **The evolution of embryonic patterning mechanisms in animals**. *Semin Cell Dev Biol* 2000, **11**:385-393.
9. Haygood R, Fedrigo O, Hanson B, Yokoyama KD, Wray GA: **Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution**. *Nat Genet* 2007, **39**:1140-1144.

Published: 21 July 2009
doi:10.1186/jbiol161
© 2009 BioMed Central Ltd