

**Hierarchical phylogenetics as a quantitative analytical framework for Evolutionary
Developmental Biology**

Jeanne M. Serb¹ and Todd H. Oakley*

Ecology, Evolution and Marine Biology, University of California – Santa Barbara, Santa
Barbara, California 93106 USA

¹ Present address: Ecology, Evolution and Organismal Biology, Iowa State University,
Ames, Iowa 50011 USA

*Correspondence to: T. H. Oakley, Ecology, Evolution and Marine Biology, University
of California – Santa Barbara, Santa Barbara, California 93106 USA.

Email: oakley@lifesci.ucsb.edu

Summary

Many topics in evolutionary developmental biology (EDB) involve comparison of evolutionary histories of modules from different levels of biological organization.

Although these modules are hierarchically arranged in an organism, their evolutionary histories may differ between levels. Phylogenetics has inherent utility in EDB as it is an established methodology for estimating evolutionary relationships of modules at one level of organization and for making comparisons between levels. However, phylogenetic methods generally have been limited to species and gene levels in EDB.

We demonstrate that these methods can be applied to other levels of biological organization, such as morphological structures or cell types, to identify their evolutionary patterns. We present examples of these types of phylogenetic methods at and between different hierarchical levels of organization to address questions central to EDB. We argue that this application of “hierarchical phylogenetics” can be a unifying analytical approach to the field of EDB.

Introduction

The field of evolutionary developmental biology (EDB), or evo-devo, aims to understand the evolution of the developmental processes underlying phenotype. Achieving this aim requires examination of evolutionary changes at multiple levels of biological organization, from nucleotides, genes, and cells to genetic networks, organs, and species. In fact, many of the central questions in EDB hinge on understanding how historical changes at one level of organization relate to changes at other levels of organization⁽¹⁾. For example, are particular developmental or genetic changes correlated

with phenotypic changes during evolution? ⁽²⁻⁴⁾ What is the tempo of gene expression evolution? ⁽⁵⁻⁸⁾ To what extent are gene duplications associated with diversification of genetic networks, organs, or species? ^(9,10) Such historical comparisons are implicitly phylogenetic in nature and although phylogenetics already has a presence in EDB ⁽¹¹⁻¹⁶⁾, that presence is mostly focused on only two levels of biological organization - gene phylogenies and species phylogenies. Here we point out that explicit phylogenetic methods can also be used at other levels of biological organization, besides genes and species. This realization indicates that entire suites of established phylogenetic/comparative tools - such as character mapping and reconciled trees - can be used in new ways in EDB to test and generate general evolutionary hypotheses that are central to the field.

Multiple levels of modular organization

Biologists identify and employ multiple levels of hierarchical organization to describe natural systems. This is apparent in our understanding of how genes form genetic networks, genetic networks enable differentiation of cells, and cells constitute organs. One may also describe the components of an organism using the module concept, a central theme to EDB ⁽¹⁷⁻¹⁹⁾. Modules are the distinct processes or units that act cooperatively in an organism, but have a degree of dissociability due to their internal integration ⁽²⁰⁻²⁶⁾. These modules are spatiotemporally bounded with discrete origins, histories, and deaths through evolutionary time and can be treated as individuals ⁽²⁷⁻²⁹⁾. Modules are genetically coherent units that are heritable. They may contain elements from different levels of biological organization and are hierarchical ^(18,21,22,24,30). For

example, modules can be part of a single level of biological organization (such as gene or cell type) or can include components from multiple levels (for example, morphogenetic field or species). Regardless of their degree of biological inclusiveness, modules are dynamic. They may vary during ontogeny or transform over evolutionary time as their internal components accumulate change^(18,25). As a result of this evolvability⁽³¹⁾, modules at all levels have a distinct evolutionary history of "descent with modification" that can be represented as a tree diagram with a series of lineages (branches) diverging from a common ancestor (node).

We can reconstruct the pattern of descent by analyzing module characters using methods of phylogenetic analysis. These characters can include any heritable trait with measurable differences that can be coded as a series of variables (states). Traits are assessed at levels below the organizational level of phylogenetic interest. For example, a species tree must be generated from module characteristics below the species-level, such as the DNA sequence of a gene or the morphological elements of a particular structure. To construct a tree of morphological structures, such as arthropod wings, phylogenetically relevant traits may include temporally or spatially defined gene expression patterns or spatial arrangement of cell types. Thus, all aspects of the module can be used to generate a data set for phylogenetic analysis, resulting in a "module tree". We will use the term "module tree" generally to describe any phylogenetic hypothesis at a single organizational level, such as species, structure, gene, or cell type. We will distinguish between different types of module trees by specifying the level of biological organization (i.e. "structure tree" or "gene tree").

Module trees

Traditionally, researchers have used phylogenetic methods solely for the construction of gene trees or species trees. More recently, new levels of biological organization have been considered by researchers as distinct modules that can be placed in a phylogenetic context⁽³²⁻³⁷⁾. These studies either implicitly or explicitly used phylogenetic methods to generate module trees of morphological structures, cell types, developmental fields⁽³⁵⁾, or metabolic pathways⁽³⁶⁾.

One module that has received considerable attention is the eye. Oakley⁽³⁴⁾ discusses how this complex structure can be replicated and subsequently modified or lost. He argues this pattern of evolution can be reconstructed phylogenetically as an “eye tree”, where differences in eye structure or gene expression can identify “eye families” of diverse morphological types. Liu and Friedrich⁽³⁵⁾ extend the “eye tree” concept. They place eye developmental fields in a phylogenetic context and present evidence for subdivision of the eye field as a mechanism for duplication. This “developmental field tree” of eyes is then used to understand changes in developmental mechanisms that correspond to differences in insect larval and adult compound eyes.

Trees can also be constructed from modules at lower levels of biological organization, such as cells. Arendt⁽³²⁾ suggests that comparing cell types using molecular characters can provide insight into their diversification and evolution, and should be the focal point of EDB analysis. This cellular focus in EDB also has been championed by others^(38,39). Arendt⁽³²⁾ refers to this as “comparative molecular cell biology” and uses the approach to examine evolutionary relationships among vertebrate retinal cell types. In that study, similarities in the gene expression data across cell types

were treated as homologous traits (characters), suggesting a single evolutionary history with a common ancestral cell type that diversified into the extant cell types. Although Arendt ⁽³²⁾ did not treat gene expression as explicit characters and character states for phylogenetic analysis, his goal, to generate evolutionary relationships among retinal eye types, was phylogenetic in nature: He envisioned what we are calling here a module tree of retinal cell types.

Geeta ⁽³³⁾ goes a step farther explicitly analyzing modules phylogenetically. She demonstrates that a structure tree can be estimated from the variation of morphological characters during development of the module. She then compares the evolutionary patterns between two organizational levels: the leaf primordium module tree to the angiosperm species tree. Although the primary goal of Geeta's paper was to determine "individuality" (= modularity) of leaf primordium as a "developmentally integrated structure", she also demonstrates two other important points. First, she shows that it is possible to reconstruct the evolutionary relationships of morphologically variable structures and represent them as a branching pattern. Second, she illustrates the usefulness of multiple level comparisons that incorporate phylogenetic analysis to address EDB questions.

Like Geeta ⁽³³⁾, Coates et al. ⁽³⁷⁾ also explicitly generate a structure phylogeny from morphological data. The resulting tree estimates the historical relationship of morphological characters distributed between a serially repeated structure found within a single organism: pectoral and pelvic appendages of the appendicular skeleton in early tetrapods. Only by reconstructing the structure phylogeny can the authors address an important concept in EDB: the influence of concerted evolution on vertebrate limbs. This

is an important point, as concerted evolution (= developmental integration) can only be understood in phylogenetic terms by simultaneously comparing two levels of biological organization.

Module trees can also be constructed from a module network. For example, Cunchillos and Lecointre⁽³⁶⁾ used explicit phylogenetic methods to infer the evolutionary history of metabolic pathways. This phylogenetic framework provides the means to address questions involving metabolism evolution, such as the relative timing of glycolysis and amino acid biosynthesis origins.

These papers demonstrate the potentially broad applications of phylogenetic methods in EDB that go beyond gene tree and species tree levels. Researchers are applying phylogenetic concepts implicitly and explicitly to module structure data, illustrating the need for quantitatively estimated module trees to address questions on the evolution and the history of modules. Therefore, we advocate that modules at all organizational levels can be treated under an explicit phylogenetic framework. Phylogenetic methods provide insight into the evolution and history of modules through module trees, explicit tests of character homology, and an impartial analytical method.

EDB applications for module trees

In the previous section, we demonstrated that modules at various levels of biological organization have an evolutionary pattern that can be reconstructed using phylogenetic methods. In addition to their phylogeny, each level of biological organization has specific historical associations with other levels as a result of being components of the same organism. Like genes within a genome, these levels can be

organized into degrees of hierarchical inclusiveness. However, inclusiveness does not necessarily equate with identical patterns of descent. Comparing phylogenetic patterns between different organizational levels, such as species tree vs. gene tree or species tree vs. structure tree, identifies discrepancies in the reconstructed evolutionary patterns. Indeed, it is the incongruence between tree patterns across different biological levels that can be used to develop hypotheses of the specific evolutionary processes that created the recovered patterns. Here we describe two existing sets of phylogenetic tools, character mapping and reconciled trees, that compare evolutionary histories between different levels of organization. We argue that these methods can be generalized beyond current common usage to test many central hypotheses in EDB.

Character Mapping

Character mapping is one set of phylogenetic tools that makes comparisons between levels of biological organization. Most commonly, character mapping methods are used to evaluate how traits below the level of species evolve in the context of a phylogenetic tree of species. Consider a phylogenetic tree of beetles. The distribution of characters with respect to a phylogenetic tree of the beetle species can be used to understand the evolution of particular traits. For example, if two closely related species possess large horns, we would probably hypothesize that they evolved from a common ancestor with large horns. In contrast, distantly related species with large horns may have evolved those horns separately, perhaps through convergent events of sexual selection. The interplay between species' phylogenetic relationships and the character traits of those species can be informative about the order, rate, and timing of trait evolution. We

maintain that character mapping can be used at many other levels of biological organization besides the common use of mapping characters on species trees.

Just as species have characters (such as horns) that can be mapped on to species trees, modules at other levels of organization possess characters that can be mapped on to module trees. As a familiar example, consider genes. Gene trees are ubiquitous in the literature and they are estimated by phylogenetic comparison of their characters, namely nucleotides, amino acids, or intron positions. These characters of genes can be mapped on to the gene tree using phylogenetic tools. Genes also have other characteristics that can be mapped on to gene trees: gene expression is among the most important of these for EDB. Since gene expression patterns can be considered traits of genes, phylogenetic methods can be used to map expression patterns on to gene trees. For example, “hindlimb expression” is a trait of some *Tbx* genes and it can be mapped on to a gene tree derived from amino acid sequences [Fig. 1C].

Importantly, character mapping can be generalized beyond the more familiar examples of species trees and gene trees. For example, phylogenetic trees may be constructed for morphological structures, such as paired fins/limbs [Fig. 1E]. These trees are generated from characters specific to the structures, which may include gene expression patterns, skeletal elements, cell types, or developmental fields. These or other characters subsequently can be mapped on to the resulting structure trees. Expression of *Tbx4* is one example of a characteristic of some fins/limbs. This character of gene expression can be mapped on to a phylogeny of the structures themselves [Fig. 1F]. Indeed, character mapping can be generally applied to module trees from many different levels of biological organization to examine many kinds of evolutionary hypotheses.

These considerations raise the possibility that the entire toolkit of phylogenetic methods can be extended to modules at multiple levels of biological organization, with many applications for questions in EDB. Detailing all the methods of the character mapping toolkit is beyond the scope of this article and have been reviewed elsewhere^(40,41). Nevertheless, a brief overview should be useful as an introduction to the types of hypotheses that can be addressed. Here, we are using the term “character mapping” to refer broadly to phylogenetic comparative methods that include 1) correlative methods, 2) ancestral state reconstruction, and 3) rate parameter estimation. All three methods can be used to address questions in EDB.

Correlative methods seek to test hypotheses of correlation between different characters within a particular module/level. Common ancestry causes characters to be non-independent, violating an assumption of most statistical tests of correlation (like regression). This phylogenetic non-independence can be taken into consideration using methods like independent contrasts⁽⁴²⁾ or analogous discrete-state tests^(43,44).

Demonstrating evolutionary correlation is important for EDB, and phylogenetic methods should be used to account for the non-independence introduced by the common ancestry of species, genes, structures, cell types, and other modules. Correlative hypotheses of interest to EDB include changes in gene expression with species/structural/cellular phenotype, changes in specific amino acids with phenotype, and changes in promotor sequences with changes in gene expression.

Ancestral state reconstruction methods seek to estimate the state(s) of a character(s) for a common ancestor of a particular group of related modules. These methods have been used to estimate and recreate vocalizations of ancestral frog species

⁽⁴⁵⁾ and amino acid sequences of ancestral proteins ⁽⁴⁶⁾. An important use of these methods in EDB includes inferring how gene expression patterns evolved. Figure 1C shows that the ancestor of *Frog-Tbx5* and *Newt-Tbx5* genes was probably not expressed in hindlimbs, and hindlimb expression was most likely gained recently in the lineage leading to *Newt-Tbx5*. In this example, gene expression is considered a trait of different *Tbx* genes, but ancestral state reconstruction can also be used at other levels of biological organization. In Figure 1F, we instead treated gene expression as a trait of a structure module (the fin/limb). This allows the inference (given the presently available data) that *Tbx4* expression is a trait of the structure ancestral to zebrafish pectoral fins and mouse/chicken forelimbs. For illustrative purposes we presented here simple examples that can be solved by inspection. However, many EDB data sets will require algorithmic estimation of ancestral states using parsimony ⁽⁴⁷⁾, likelihood ⁽⁴⁸⁾, or Bayesian ⁽⁴⁹⁾ inference techniques.

Rate parameters usually comprise the simple models used by character mapping. Testing hypotheses based on estimates of these rate parameters using maximum likelihood ⁽⁴¹⁾ or Bayesian ⁽⁵⁰⁾ statistical techniques is an important use of character mapping methods. Previous studies have shown, for example, that complex morphology has been more often gained than lost during evolution of fungal species ⁽⁵¹⁾ and generalists species more commonly evolved from specialists than the reverse ⁽⁵²⁾. By examining modules at multiple levels of biological organization, rate estimates will be important for EDB: Wilkins ⁽¹¹⁾ cited rates of developmental evolution as one of three central topics for future work in EDB. For example, we may be interested to learn whether gene expression domains are more commonly lost than gained during evolution,

a hypothesis could be examined by constructing gene trees and mapping gene expression to compare the rate of gain versus loss of expression domains. Similar hypotheses could also be examined at other levels of biological organization.

Reconciled trees

Reconciled tree analysis (RTA) is another phylogenetic tool that makes comparisons between levels of biological organization. The analysis compares and contrasts the patterns of bifurcation between two trees that are historically and hierarchically related (such as species trees and gene trees), allowing inference of evolutionary processes. RTA was originally described⁽⁵³⁻⁵⁵⁾ as a generalized analytical approach that can be applied to the diverse disciplines of molecular evolution, parasitology, and biogeography. Researchers within these disciplines have used RTA to examine the historical relationship of genes within species, parasites within hosts, or species within geographic areas⁽⁵⁵⁾. We propose to expand the use of RTA to include lineage comparisons relevant to EDB such as gene trees to cell type trees, cell type trees to structure trees, or structure trees to species trees. We can justify the use of RTA for these finer scale comparisons because the same evolutionary processes occur at multiple levels of biological organization, like paralogous duplication and concerted evolution/developmental integration (Fig. 2). The patterns elucidated by phylogenetics and compared by RTA can help us identify these processes. For example, the patterns of a species tree and a gene tree will not be congruent if a gene lineage of the gene tree was lost (Fig. 2). However, similar incongruence between species tree and gene tree may be observed if a gene copy is horizontally transferred. RTA is as a rigorous method to

differentiate between evolutionary processes when examining complex data sets from multiple hierarchical levels simultaneously.

One good example of a complex data set where RTA could be applied is the study of Nordström et al. ⁽⁵⁶⁾ on modules of the vertebrate retina. The vertebrate retina is a highly organized neural structure composed of seven major cell types, including two photoreceptor cell types: rods and cones. Many gene families are involved with phototransduction and are expressed restrictively in specific retinal cell types. Nordström et al. ⁽⁵⁶⁾ phylogenetically examined a suite of these vertebrate gene families and identified multiple duplication events at the cell, chromosome, and gene levels that occurred in a vertebrate ancestor. This correlation may suggest causation; evolutionary changes (such as duplication events) at one level may induce or facilitate such changes at other levels ^(29,57,58). For example, gene duplications resulting from chromosome or genome duplications may have facilitated the differentiation of two retinal cell types, rods and cones, that evolved specific functions, scotopic and photopic vision, respectively (see Nordström et al. ⁽⁵⁶⁾ for a more detailed discussion).

Although Nordström et al. ⁽⁵⁶⁾ did not use explicit phylogenetic tools to identify correlations when comparing hierarchical levels of biological organization, we advocate the application of methods, such as RTA, to these types of data. For example, we can explicitly examine the historical associations among three hierarchical levels (species, cell type, and gene) in the vertebrate retina by generating phylogenetic trees for each, then comparing their historical patterns with RTA. Our species tree followed the well-supported phylogeny of the major vertebrate lineages (see review in Meyer and Zardoya ⁽⁵⁹⁾) (Fig. 3A). The cell type tree was constructed from gene expression data (compiled by

Arendt⁽³²⁾) and amino acid sequence changes of opsin homologs expressed in specific cell types (Fig. 3B). The gene tree of PDE6 homologs was generated from amino acid sequences across five species representing major vertebrate lineages: pufferfish, frog, chicken, mouse, and human (from Nordström et al.⁽⁵⁶⁾). We then compared two levels of historic associations at a time (vertebrate species vs. retinal cell type, species vs. PDE6 gene, retinal cell type vs. PDE6 gene) using a reconciled tree illustration (sensu Page⁽⁶⁰⁾) (Fig. 3D, E). Finally, we depict one possible reconciled tree including all three levels of biological organization (Fig. 3F). Our results using RTA reflect the conclusions of Nordström et al.⁽⁵⁶⁾, where there is an association of duplication events at different levels of biological organization in a vertebrate ancestor. Unlike Nordström et al.⁽⁵⁶⁾, our approach uses explicit phylogenetic methods to generate module trees at each biological level. Furthermore, we are able to make rigorous comparison between hierarchical levels with phylogenetic methods.

Although, our simple example of a reconciled tree (Fig. 3F) can be solved by inspection quite easily, many EDB data sets are more complex. For example, the number of gene copies in a gene tree may not match the number of taxa in a species tree or the pattern of the gene tree may have little or no similarity to the species tree pattern. We argue that RTA is an object tool that can be applied to complex data sets to develop evolutionary hypotheses in EDB.

Conclusions

Many central topics in evolutionary developmental biology involve the comparison of evolutionary histories of modules from different levels of biological

organization^(1,12,61). This hierarchical and phylogenetic view of evolutionary developmental biology is not completely new: phylogenetics has been routinely applied to hierarchical questions in EDB⁽¹²⁾. However, previous phylogenies have been constructed only for species and genes, representing only two of many possible levels of biological organization. In fact, usually only one level - species trees - have been used for mapping developmental characters. Here we point out that researchers in EDB have begun to construct phylogenetic trees from units (modules) of other levels of organization, indicating that phylogenetics can be generalized beyond genes and species to bifurcating modules from any level. Several phylogenies already have been constructed at the level of structure or organ: Coates et al.⁽³⁷⁾ estimated a tree of tetrapod appendages, Geeta⁽³³⁾ estimated a tree of leaf primordia and Oakley⁽³⁴⁾ and Liu and Friedrich⁽³⁵⁾ illustrated the concept of structure trees for photoreceptors. In addition to structure-level trees, Arendt⁽³²⁾ conceptualized phylogenetic trees of retinal cell types and Cunchillos and Lecointre⁽³⁶⁾ explicitly estimated phylogenetic trees of metabolic pathways. These separate articles provide examples for module trees of structures, cell-types, and metabolic pathways, in addition to the common trees of species and genes. These examples already represent module trees from five different levels of biological organization and point to the generalization of phylogenetic methods whereby trees can be constructed for modules at any level of biological organization using traits of those modules. The generalized hierarchical phylogenetics advocated here indicates the exciting possibility of co-opting the entire phylogenetic toolkit to address fundamental questions in EDB.

The interest in phylogenies comes not from building them, rather from what they tell us about evolution ⁽⁶²⁾. Just as species trees inform us about the evolution of species' characteristics, module trees can inform us about the evolution of modules' characteristics, which is perhaps the fundamental goal of EDB. We have pointed out that module evolution can be investigated using two suites of phylogenetic tools, character mapping and reconciled trees. We use "character mapping" to refer to phylogenetic comparative methods, including correlational methods, ancestral state reconstruction, and rate parameter estimation. Applying these methods to new levels of biological organization will allow testing of central hypotheses in EDB, such as correlations between particular genotypic and phenotypic changes during evolution, tempo and mode of gene expression ⁽⁷⁾ and developmental pathway evolution, co-option, and developmental integration (concerted evolution). Similarly, reconciled trees – when generalized to new levels of biological organization - will provide new insights into EDB by revealing processes of evolution that span levels of organization. These processes include orthologous and paralogous duplication (of for example, cell types and structures), horizontal transfer, and loss (Fig. 2). Often, the timing of such evolutionary events may suggest the primary origin of evolutionary novelty, such as when gene duplications precede cell type or organ differentiation.

Evolutionary developmental biology aims to understand the history of hierarchical and modular systems. Phylogenetics is an established system of methods for dealing with the history of hierarchical systems. We maintain that phylogenetics is currently under utilized because it primarily focuses only on species and gene levels. Generalizing phylogenetics to bifurcating modules at all levels of biological organization

can provide a valuable analytical framework for testing fundamental EDB hypotheses using well established, versatile, and sophisticated statistical methodologies.

Acknowledgments

We thank E. Abouheif, J. Endler, R. Geeta, J. Jeffery, R. Mayden, K. Nordström, S. Parker, K. Roe, and members of the Oakley lab group for valuable comments that greatly improved this article. J.M.S. was supported by the University of California President's Postdoctoral Fellowship Program. T.H.O. was supported by a National Science Foundation Grant (DEB-0316330).

References

1. Arthur W. 2002. The emerging conceptual framework of evolutionary developmental biology. *Nature* 415:757-64.
2. Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305:1462-5.
3. Averof M, Patel NH. 1997. Crustacean appendage evolution associated with changes in Hox gene expression. *Nature* 388:682-6.
4. Gompel N, Carroll SB. 2003. Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. *Nature* 424:931-5.
5. Wagner A. 2000. Decoupled evolution of coding region and mRNA expression patterns after gene duplication: implications for the neutralist-selectionist debate. *Proc Natl Acad Sci U S A* 97:6579-84.
6. Gu Z, Nicolae D, Lu HH, Li WH. 2002. Rapid divergence in expression between duplicate genes inferred from microarray data. *Trends Genet* 18:609-13.
7. Oakley TH, Gu Z, Abouheif E, Patel NH, Li WH. 2005. Comparative Methods for the Analysis of Gene-Expression Evolution: An Example Using Yeast Functional Genomic Data. *Mol Biol Evol* 22:40-50.

8. Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* 20:1377-419.
9. Ohno S. 1970. *Evolution by gene duplication*. New York: Springer-Verlag.
10. Williams NA, Holland PW. 2000. An amphioxus Emx homeobox gene reveals duplication during vertebrate evolution. *Mol Biol Evol* 17:1520-8.
11. Wilkins A. 2002. *The Evolution of Developmental Pathways*. Sunderland, MA: Sinauer.
12. Abouheif E. 1999. Establishing homology criteria for regulatory gene networks: prospects and challenges. *Novartis Found Symp* 222:207-21.
13. Richardson MK, Jeffery JE, Coates MI, Bininda-Emonds ORP. 2001. Comparative methods in developmental biology. *Zoology-Analysis of Complex Systems* 104:278-283.
14. Telford MJ, Budd GE. 2003. The place of phylogeny and cladistics in Evo-Devo research. *Int J Dev Biol* 47:479-90.
15. Mabee PM, Crotwell PL, Bird NC, Burke AC. 2002. Evolution of median fin modules in the axial skeleton of fishes. *J Exp Zool* 294:77-90.
16. Wray GA, Bely AE. 1994. The evolution of echinoderm development is driven by several distinct factors. *Dev Suppl*:97-106.
17. Raff RA. 2000. Evo-devo: the evolution of a new discipline. *Nat Rev Genet* 1:74-9.
18. Winther RG. 2001. Varieties of modules: Kinds, levels, origins, and behaviors. *Journal of Experimental Zoology* 291:116-129.
19. Schlosser G, Wagner GP, editors. 2004. *Modularity in Development and Evolution*. Chicago: University of Chicago Press. 600 p.
20. Needham J. 1933. On the dissociability of the fundamental process in ontogenesis. *Biol. Rev.* 8:180-223.
21. Raff RA. 1996. *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. Chicago: University of Chicago Press.
22. Wagner GP. 1996. Homologues, natural kinds and the evolution of modularity. *American Zoologist* 36:36-43.

23. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402:C47-52.
24. Raff EC, Raff RA. 2000. Dissociability, modularity, evolvability. *Evolution & Development* 2:235-237.
25. Raff RA, Sly BJ. 2000. Modularity and dissociation in the evolution of gene expression territories in development. *Evol Dev* 2:102-13.
26. Schlosser G, Wagner GP. 2004. Introduction: The modularity concept in developmental and evolutionary biology. In: Schlosser G, Wagner GP, editors. *Modularity in Development and Evolution*. Chicago: University of Chicago Press. p 1-11.
27. Ghiselin MT. 1974. A radical solution to the species problem. *Syst Biol* 25:536-544.
28. Hull DL. 1976. Are species really individuals? *Syst Biol* 25:174-191.
29. Vrba ES, Eldredge N. 1984. Individuals, hierarchies and processes: towards a more complete evolutionary theory. *Paleobiology* 10:146-171.
30. Schlosser G. 2004. The role of modules in development and evolution. In: Schlosser G, Wagner GP, editors. *Modularity in Development and Evolution*. Chicago: University of Chicago Press. p 519-582.
31. Kirschner M, Gerhart J. 1998. Evolvability. *Proc Natl Acad Sci U S A* 95:8420-7.
32. Arendt D. 2003. Evolution of eyes and photoreceptor cell types. *Int J Dev Biol* 47:563-71.
33. Geeta R. 2003. Structure trees and species trees: what they say about morphological development and evolution. *Evol Dev* 5:609-21.
34. Oakley TH. 2003. The eye as a replicating and diverging, modular developmental unit. *Trends in Ecology and Evolution* 18:623-627.
35. Liu Z, Friedrich M. 2004. The *Tribolium* homologue of glass and the evolution of insect larval eyes. *Dev Biol* 269:36-54.
36. Cunchillos C, Lecointre G. 2005. Integrating the Universal Metabolism into a Phylogenetic Analysis. *Mol Biol Evol* 22:1-11.
37. Coates MI, Jeffery JE, Ruta M. 2002. Fins to limbs: what the fossils say. *Evolution & Development* 4:390-401.

38. Hall BK. 2001. The gene is not dead, merely orphaned and seeking a home. *Evol Dev* 3:225-8.
39. Atchley WR, Hall BK. 1991. A model for development and evolution of complex morphological structures. *Biol Rev Camb Philos Soc* 66:101-57.
40. Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401:877-84.
41. Oakley TH. 2003. Maximum likelihood models of character trait evolution. *Comments on Theoretical Biology* 8:1-17.
42. Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1-15.
43. Maddison WP. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* 44:539-557.
44. Pagel MD. 1994. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B* 255:37-45.
45. Ryan MJ, Rand AS. 1995. Female responses to ancestral advertisement calls in Tungara frogs. *Science* 269:390-392.
46. Chang BS, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP. 2002. Recreating a functional ancestral archosaur visual pigment. *Mol Biol Evol* 19:1483-9.
47. Swofford DL, Maddison WP. 1992. Parsimony, character-state reconstructions, and evolutionary inferences. In: Mayden RL, editor. *Systematics, Historical Ecology, and North American Freshwater Fishes*. Stanford, California: Stanford University Press. p 186-223.
48. Schluter D, Price T, Mooers AØ, Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699-1711.
49. Pagel M, Meade A, Barker D. 2004. Bayesian estimation of ancestral character States on phylogenies. *Syst Biol* 53:673-84.
50. Huelsenbeck JP, Rannala B, Masly JP. 2000. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* 288:2349-2350.

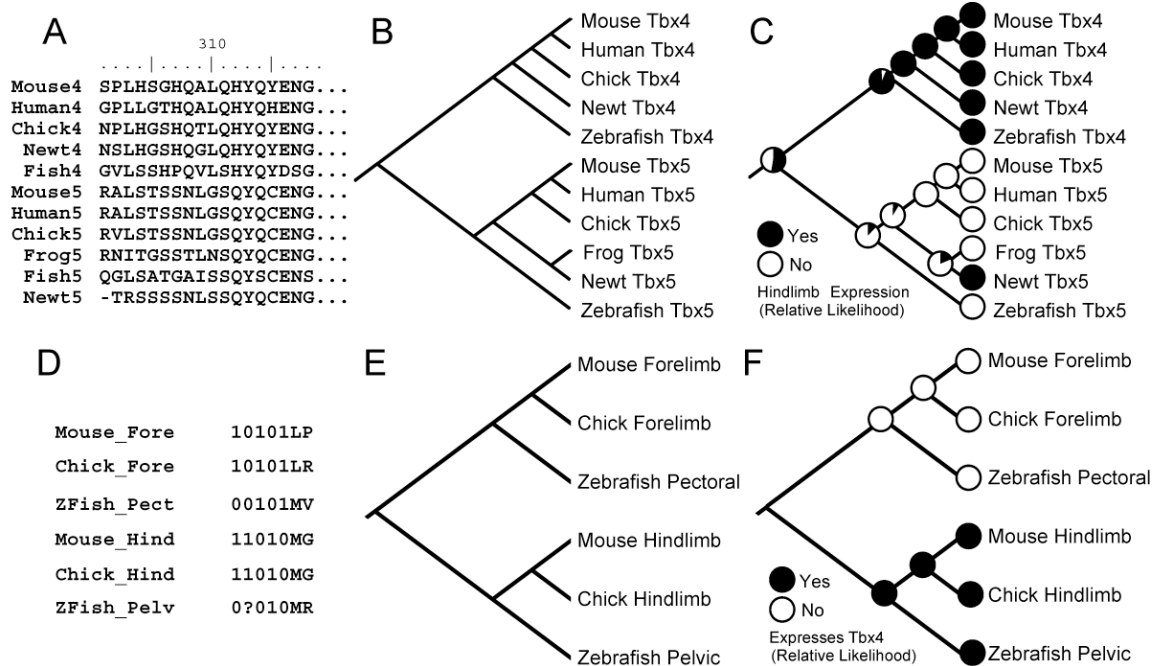
51. Hibbett DS, Binder M. 2002. Evolution of complex fruiting-body morphologies in homobasidiomycetes. *Proceedings of the Royal Society of London Series B-Biological Sciences* 269:1963-1969.
52. Nosil P. 2002. Transition rates between specialization and generalization in phytophagous insects. *Evolution* 56:1701-1706.
53. Page RDM. 1993. Genes, organisms, and areas: the problem of multiple lineages. *Syst Biol* 42:77-84.
54. Page RDM, Charleston MA. 1997. From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. *Mol Phylogenet Evol* 7:231-40.
55. Page RDM, Charleston MA. 1998. Trees within trees: phylogeny and historical associations. *Trends in Ecology and Evolution* 13:356-359.
56. Nordström K, Larsson TA, Larhammar D. 2004. Extensive duplications of phototransduction genes in early vertebrate evolution correlate with block (chromosome) duplications. *Genomics* 83:852-72.
57. Budd GE. 1999. Does evolution in body patterning genes drive morphological change - or vice versa? *Bioessays* 21:326-332.
58. West-Eberhard MJ. 2003. *Developmental Plasticity and Evolution*. Oxford: Oxford University Press. 794 p.
59. Meyer A, Zardoya R. 2003. Recent advances in the (molecular) phylogeny of vertebrates. *Ann. Rev. Ecol. Systematics* 34:311-338.
60. Page RDM. 1993. Parasites, Phylogeny and Cospeciation. *International Journal for Parasitology* 23:499-506.
61. de Beer GR. 1971. *Homology: an unsolved problem*. Oxford: Oxford University Press.
62. Cunningham CW, Omland KO, Oakley TH. 1998. Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology and Evolution* 13:361-366.

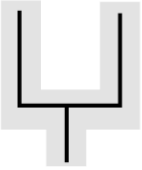
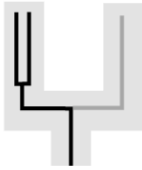
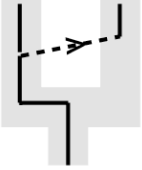
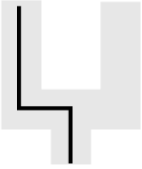
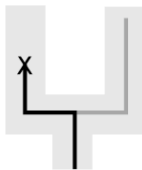
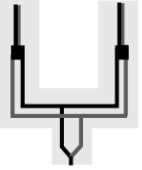
Figure 1. Phylogenetic relationships can be estimated at different levels of biological organization. **(A-C)** represent a character matrix and gene tree (i.e. the phylogenetic relationships of genes) in this case a phylogeny of *Tbox* genes. **(D-F)** represent a character matrix and "structure tree", (i.e. phylogenetic relationships of structures) in this case a phylogeny vertebrate fins and limbs. **A:** Matrices of amino acids are often used as characters for inferring relationships among genes. **B:** Phylogenetic relationships (a "gene tree") among *Tbox* genes inferred using amino acid characters. **C:** Other traits or characteristics of genes can be mapped on to gene trees using phylogenetic methods. Here we used maximum likelihood to map the presence and absence of hindlimb expression. Pies represent the relative likelihood that genes were (black) or were not (white) expressed in hindlimbs in the hypothetical common ancestors of the genes in question. According to this analysis, newt secondarily regained hindlimb expression of *Tbx5*. Just as different genes have different characters (amino acids), different structures also have different characters. As one example, **(D)** shows a non-exhaustive list of characters for different fins/limbs of vertebrates. Non-inclusive list of Characters: Digits (1=present, 0=absent); Pitx1 Expression (1=present, 0=absent); Anterior Position (1=present, 0=absent); Posterior Position (1=present, 0=absent); wnt2b Expression (1=present, 0=absent); Amino acid at site 7 of expressed Tbx gene; Amino acid at site 15 of expressed Tbx gene. **E:** A "structure tree" illustrates the phylogenetic relationships of structures like fins/limbs. Characteristics of structures, such as whether or not they express a particular gene, can be mapped on to structure trees using phylogenetic methods. **F:** We used maximum likelihood to reconstruct presence (black) and absence (white) of *Tbx4* expression in various fins/limbs.

Figure 2. Phylogenetic patterns and the inferred evolutionary process between different organizational levels, such as species and cell types, species and genes, or cell types and genes. Each pattern has a major, more inclusive lineage (light grey), such as species or cell type, and an ancillary lineage (black), such as gene or cell type. Listed processes are analogous to events that occur between host species and associated parasite lineages. “X’s” represent a lineage extinction or loss. Solid squares represent homogenizing mechanisms in concerted evolution or developmental integration.

Figure 3. Phylogenetic relationships are estimated from three levels of biological organization (A-C) and phylogenetic patterns are compared between the levels (D-F). **A:** A species tree of five major vertebrate lineages inferred from morphological and molecular data. **B:** A cell type tree of vertebrate photoreceptors inferred from gene expression data and amino acid sequences. Solid circles represent paralogous gene duplication events. **C:** A gene tree of PDE6 gene homologs (A, B, and C) inferred from amino acid sequences. Gene copies are expressed in either rods (red) or cones (blue) photoreceptors. Solid circles represent paralogous gene duplication events, while open circles and dotted lines indicate putative gene deletions or loss of gene expression in that lineage. **D:** Comparison of phylogenetic pattern between species tree (in grey) and cell type tree, where rod photoreceptors are in red and cone photoreceptors are in blue. **E:** Comparison of the species tree (in grey) to the PDE6 gene tree. The gene tree shows evolutionary relationships among PDE6 gene homologs (A, B, and C) expressed in either rod (red) or cone (blue) photoreceptors. Solid circles represent paralogous gene

duplication events, while “X’s” indicate putative gene deletions or loss of gene expression in that lineage. **F:** A possible reconciliation of all three organizational levels. The species tree is in grey; the cell type tree has wide, transparent colored lines; the gene tree is depicted as thin, opaque lines. “X’s” represent gene loss. Solid circles represent paralogous duplications between the gene tree and the species tree. Note that a paralogous duplication at one level of comparison (duplication of the PDE6 lineage in an early vertebrate ancestor) can be an orthologous duplication at a different level of comparison (duplication of the PDE6 lineage in the ancestral photoreceptor cell type).



Pattern	Process	Levels of organization (non-inclusive)		
		Species v. gene	Species v. cell	Organ v. cell
	Orthologous duplication	Gene tracks species at time of speciation; species and gene are congruent at that node	Cell type tracks species tree at time of speciation	Cell type tracks organ tree at time of organ diversification
	Paralogous duplication	Gene duplication occurs within a single species lineage	Cell type duplicates within a species	Cell type duplicates and is limited to a specific organ
	Horizontal transfer	Gene or genome is transferred ("jumping") or copied ("duplicative") between species. For example, transposable elements mediated by viruses or lateral gene transfer between unrelated bacterial strains	Cell type is transferred between species lineages. Cell lineage does not track species tree. For example, transfer of ascidian stem cell to another individual's germ line (Stoner et al. 1999)	Cell type migrates between two non-sister tissue types; cell type does not track organ tree
	Incomplete tracking	Gene or chromosomal set tracks only one species lineage after speciation event. For example, aneuploidy [monosomy] in plant species	Cell type does not track both lineages after a speciation event; cell type is specific to one lineage. For example, sperm loss in parthenogenetic species	Organs differentiate, but cell type is specific to only one organ lineage
	Loss	Gene is lost (i.e. deleted) in species lineage	Cell type is lost in species	Specific cell type is lost in organ lineage
	Concerted evolution	Members of a gene family in a species do not evolve independently, but in a concordant fashion. For example, ribosomal RNA genes in prokaryotic and eukaryotic organisms	Members of a cell family in a species do not evolve independently, but in a concordant fashion	Members of a cell family in an organ do not evolve independently, but in a concordant fashion

