Evolution of Nervous Systems: a Comprehensive Reference - CONTRIBUTORS' INSTRUCTIONS

PROOFREADING
The text content for your contribution is in final form when you receive proofs. Read proofs for accuracy and clarity, as well as for typographical errors, but please DO NOT REWRITE.

Titles and headings should be checked carefully for spelling and capitalization. Please be sure that the correct typeface and size have been used to indicate the proper level of heading. Review numbered items for proper order – e.g., tables, figures, footnotes, and lists. Proofread the captions and credit lines of illustrations and tables. Ensure that any material requiring permissions has the required credit line and that we have the relevant permission letters.

Your name and affiliation will appear at the beginning of the article and also in a List of Contributors. Your full postal address appears on the non-print items page and will be used to keep our records up-to-date (it will not appear in the published work). Please check that they are both correct.

Keywords are shown for indexing purposes only and will not appear in the published work.

Any copy-editor questions are presented in an accompanying Author Query list at the beginning of the proofs. Please address these questions as necessary. While it is appreciated that some articles will require updating/revising, please try to keep any alterations to a minimum. Excessive alterations may be charged to the contributors.

Note that these proofs may not resemble the image quality of the final printed version of the work, and are for content checking only. Artwork will have been redrawn/relabelled as necessary, and is represented at the final size.

Cross-references
Please ensure that all cross-references to other articles are in place, adding details where necessary. 'See' references should appear within the main article text and will link directly with relevant articles. Please use volume number, article code and full article title for each reference. You will receive a table of contents attached to the proof email.

DESPATCH OF CORRECTIONS

PLEASE KEEP A COPY OF ANY CORRECTIONS YOU MAKE.

Proof corrections should be returned in one communication to your subject editor, by **24-07-06** using one of the following methods:

1. **PREFERRED**: Corrections should be listed in an e-mail to your subject editor **Georg F. Striedter** at **gstriedt@uci.edu** (copying in Cathy Succamore at Elsevier at NRVSproofs@elsevier.com).

The e-mail should state the article code number in the subject line. Corrections should be consecutively numbered and should state the paragraph number, line number within that paragraph, and the correction to be made.

2. If corrections are substantial, send the amended hardcopy by courier to **Cathy Succamore, Elsevier MRW Production Department, The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK.** If it is not possible to courier your corrections, please fax the relevant marked pages to the Elsevier MRW Production Department **(fax number: +44 (0)1865 843974)** with a covering note clearly stating the article code number and title.

Note that a delay in the return of proofs could mean a delay in publication. Should we not receive corrected proofs within 7 days, Elsevier may proceed without your corrections.

CHECKLIST

Author queries addressed/answered?	❑
Affiliations, names and addresses checked and verified?	❑
References section checked and completed?	❑
Permissions details checked and completed?	❑
Cross references inserted?	❑
Outstanding permissions letters attached/enclosed?	❑
Figures and tables checked?	❑

If you have any questions regarding these proofs please contact the Elsevier MRW Production Department at: **NRVSproofs@elsevier.com**

Non-print items

Author Query Form

Galley Proof

ELSEVIER

Title: Evolution of Nervous Systems

Article: 00108

Dear Author,

During the preparation of your manuscript for typesetting some questions have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list. Your responses should then be returned with your marked proof/list of corrections to Cathy Succamore at Elsevier **NRVSproofs@elsevier.com**

Queries and/or remarks

AU:1	Please check the long affiliation for accuracy. This is for Elsevier's records and will not appear in the printed work.
AU:2	Please provide cross-references to other articles from this Comprehensive where pertinent.
AU:3	The references supplied for 'Further Reading' section are already listed in 'References' section. Please supply additional references for 'Further Reading' section.
AU:4	Please supply publisher name.
AU:5	Reference 'Felsenstein 1973' is not cited in text. Please check.
AU:6	Reference 'Gittleman and Luh 1992' is not cited in text. Please check.
AU:7	Please supply initial(s) of editors page range.
AU:8	Reference 'Slowinski 1993' is not cited in text. Please check.
AU:9	Please supply the article title.
AU:10	Please supply page range.
AU:11	Valves are not clear. Please supply better quality figure for figure 3 & 4.

ELSEVIER FIRST PROOF

Non-print items

Author's Contact Information

AU1 **James S Albert**
Department of Biology
University of Louisiana
Lafayette
LA 70504-42451
USA

Keywords: Adaptation, Phylogenetic character, Comparative method, Convergence, Homology, Homoplasy, Parallelism, Parsimony optimization, Phenotypic evolution, Phylogenetic trees, Character polarity, Character state reconstruction, Reversal, Trait evolution

Abstract: This article provides an overview of recent methods of ancestral character state reconstruction and phenotypic evolution using examples drawn largely from vertebrate nervous systems. The evolutionary meaning of key concepts is reviewed, including homology, convergence, parallelism, reversal, adaptation, polarity, and phylogenetic tree. Parsimony-based methods for reconstructing ancestral states and analyzing the evolution of character or trait data include maximum parsimony optimization for discrete traits, and squared-change and linear parsimony for continous traits. Descriptions of likelihood, Bayesian and correlative comparative methods are also reviewed. The strengths and weaknesses of these alternative methods are evaluated in light of different classes of data being examined, and the different assumptions they make.

ELSEVIER FIRST PROOF

a0005 1.03 Phylogenetic Character Reconstruction

J S Albert, University of Louisiana, Lafayette, LA, USA

© 2007 Elsevier Inc. All rights reserved.

Glossary

g0005	*Adaptation*	A feature or phenotype or trait that evolved to serve a particular function or purpose.
g0010	*Anagenesis*	The origin of evolutionary novelties within a species lineage by changes in gene allele frequencies by the processes of natural selection and/or neutral genetic drift.
g0015	*Character Polarity*	The temporal direction of change between alternative (primitive and derived) states of a character.
g0020	*Character State Reconstruction*	The process of estimating the ancestral or primitive condition of a character at a given node (branching point) in a phylogenetic tree.
g0025	*Clade*	A complete branch of the tree of life. A monophyletic group.
g0030	*Cladogenesis*	The origin of daughter species by the splitting of ancestral species; may or may not occur under the influence of natural selection.
g0035	*Cladogram*	A branching tree-shaped diagram used to summarize comparative (interspecific) data on phenotypes or gene sequences. In contrast to a Phylogeny, a cladogram has no time dimension.
g0040	*Comparative Method*	The study of differences between species.
	Continuous Trait	A quantitatively defined feature with no easily distinguished boundaries between phenotypes (e.g., size, cell counts, and gene expression). g0045
	Convergence	Similarity of structure or function due to independent evolution from different ancestral conditions. g0050
	Discrete Trait	A qualitatively defined feature with only a few distinct phenotypes (e.g., polymorphism; presence vs. absence). g0055
	Homology	Similarity of structure or function due to phylogeny(common ancestry). g0060
	Homoplasy	Similarity of structure or function due to convergence, parallelism or reversal. g0065
	Monophyletic	A systematic category that includes an ancestor and all of its descendants; a complete branch of the tree of life; a 'natural' taxon; a clade. g0070
	Node	An internal branching point in a phylogenetic tree. g0075
	Optimization	Methods for estimating ancestral trait values on a tree. Commonly used optimization criteria are: maximum parsimony (MP) which minimizes the amount of trait change, and maximum likelihood (ML) which maximizes the likelihood of a trait at a node given likelihood values for trait evolution. g0080

ELSEVIER FIRST PROOF

g0085 *Parallelism* — Similarity of structure or function due to independent evolution from a common ancestral condition.

g0090 *Paraphyletic* — A systematic category that includes an ancestor and some but not all of its descendents (e.g., 'invertebrates', 'agnathans', 'fish', and 'reptiles' (*sans* birds)).

g0095 *Parsimony* — A principle of scientific inquiry that one should not increase, beyond what is necessary, the number of entities required to explain anything.

g0100 *Phenotypic Evolution* — Change in the developmental program descendents inherit from their ancestors.

g0105 *Phylogenetic Character* — A homologous feature or phenotype or trait of an organism or group of organisms.

g0110 *Phylogenetic Systematics* — A method for reconstructing evolutionary trees in which taxa are grouped exclusively on the presence of shared derived features.

g0115 *Phylogenetic Tree* — Genealogical map of interrelationships among species, with a measure of relative or absolute time on one axis. Also called a tree of life or a phylogeny.

g0120 *Phylogeny* — The evolutionary history of a species or group of species that results from anagenesis and cladogenesis.

g0125 *Polyphyletic* — A systematic category that includes taxa from multiple phylogenetic origins (e.g., 'homeothermia' consisting of birds and mammals).

g0130 *Reversal* — Change from a derived character state back to a more primitive state; an atavism. Includes evolutionary losses (e.g., snakes have 'lost' their paired limbs').

g0135 *Synapomorphy* — A shared, derived character used as a hypothesis of homology.

g0140 *Taxon* — A species or monophyletic group of species (plural taxa).

g0145 *Trait evolution* — The sequence of changes of a feature or phenotype on a phylogeny.

s0005 1.03.1 Introduction to Character State Reconstruction and Evolution

AU2

p0005 Comparisons among the features of living organisms have played a prominent role in the biological sciences at least since the time of Aristotle. The comparative approach takes advantage of the enormous diversity of organismal form and function to study basic biological processes of physiology, embryology, neurology and behavior. This approach has given rise to the widespread use of certain species as model systems, based on what has become known as the August Krogh Principle: "For many problems there is an animal on which it can be most conveniently studied (Krebs, 1975)." b0195

p0010 From an evolutionary perspective, interspecific (between species) comparisons allow for the systematic study of organismal design. Rensch (1959) conceived of phylogeny as being composed of two distinct sets of processes; anagenesis, the origin of phenotypic novelties within an evolving species lineage (from the Greek *ana* = up + *genesis* = origin), and cladogenesis, the origin of new species from lineage splitting (speciation) (from the Greek *clado* = branch). Anagenetic changes arise within a population by the forces of natural selection and genetic drift. Cladogenesis may or may not arise from these population-level processes, and in fact many (or perhaps most?) species on Earth are thought to have their origins from geographical (allopatric) speciation under the influence of landscape and geological processes (Mayr, 1963; Coyne and Orr, 1989). b0305 b0235 b0065

p0015 Because species descend from common ancestors in a hierarchical fashion (i.e., from a branching, tree-like process of speciation) closely related species tend to resemble each other more than they do more distantly related species. Patterns in the diversification of phenotypes have therefore been described as mosaic evolution, in which different species inherit distinct combinations of traits depending on the position of that species in the tree of life (McKinney and McNamara, 1990). Under this view, character evolution is regarded as a process of historical transformation from a primitive to a derived state, and study of this process necessarily presumes knowledge of primitive or ancestral conditions. In other words, because character evolution is perceived of as trait change on a tree, it is necessary to estimate *ancestral trait values*. b0240

p0020 Direct observations of ancient phenotypes may be taken from fossils, which provide unique information on entirely extinct groups of organisms, and are usually associated with stratigraphic information pertaining to relative and absolute geological ages (Benton, 1993). Nonetheless, the fossil record has many well-known shortcomings, including the famously incomplete levels of preservation, and usually very limited information about the nature of soft tissues such as nerves and brains (but see Edinger (1941) and Stensiö (1963)). Paleontological information on ancient physiological and behavioral traits is even more scanty (but see Jerison (1976), MacLeod and Rose (1993), and Rogers (2005)). b0020 b0085 b0185 b0315

ELSEVIER FIRST PROOF

p0025 Recent years have seen great advances in the formulation of comparative methods to estimate or infer ancestral phenotypes from extant (living) species (b0120 Garland *et al.*, 1992,1999; b0130 b0230 Martins, 2000). These methods use patterns in the mosaic of traits present among species in the context of an explicit hypothesis of interrelationships. These methods also address new topics, such as whether rates of phenotypic evolution have differed among lineages (clades), the circumstances in which a phenotype first evolved, the selective and developmental mechanisms underlying the origin of new phenotypes, and the evolutionary lability of phenotypes (b0010 Albert *et al.*, 1998; b0025 Blomberg *et al.*, 2003; b0030 Blackledge and Gillespie, 2004).

p0030 In this article, I summarize the major recent developments in phylogenetically based methods of studying character evolution, with the goals of explaining both the strengths and weaknesses of alternative methods. Most of the empirical examples cited are among animals with the most complex central nervous systems (e.g., vertebrates) in which neurological and behavioral evolution has been (arguably) most extensively studied. A major goal of this article is to highlight some of the most exciting new developments in the study of character evolution now being explored in this fascinating area of comparative neurobiology.

s0010 1.03.2 Basic Concepts

s0015 1.03.2.1 Homology: Similarity due to Common Ancestry

p0035 All methods of ancestral character state reconstruction make explicit assumptions about the homology of the traits under study. In comparative biology the term homology refers to similarity in form or function arising from common ancestry. In other words, homologous features among organisms can be traced to a single evolutionary origin. In the language of b0135 Garstang (1922), a homologous trait is a unique historical change in the developmental program of an evolving lineage. Homologous similarities may be observed in any aspect of the heritable phenotype, from properties of genetic sequences (e.g., base composition and gene order), through aspects of development, including cellular, tissue, and organismal phenotypes, to aspects of behavior that emerge from the organization of the nervous system. Homology in behavioral traits has been examined in a number of taxa, and in a variety of contexts (b0075 de Queiroz and Wimberger, 1993; b0390 Wimberger and de Queiroz, 1996; b0035 Blomberg *et al.*, 2003). Taxa are individual branches of the tree of life, and may include species or groups of species that share a common ancestor (the latter are also referred to as clades or monophyletic groups).

p0040 It is important to note that developmental, structural, positional, compositional, and functional features of phenotypes are all useful in proposing hypotheses of homology. Yet by the evolutionary definition employed above, only features that can be traced to a common ancestor in an explicitly phylogenetic context are regarded as homologues. Because phylogenies are the product of comparative analyses using many traits, it is in fact congruence in the phylogenetic distribution of characters that serves as the ultimate criterion for homology. By this criterion homologous characters are said to have passed the test of congruence. In other words, congruence in the phylogenetic distribution of numerous character states is regarded to be the ultimate evidence for homology (b0295 Patterson, 1982).

s0020 1.03.2.2 Homoplasy: Convergence, Parallelism, and Reversal

p0045 All other forms of phenotypic similarity that arise during the course of evolution are referred to collectively as homoplasy (similarity due to causes other than homology). Homoplastic characters may arise from several sources: convergence due to similar functional pressures and natural selection, parallel (independent) evolution to a common structure or function from organisms with similar genetic and developmental backgrounds, or convergent reversal to a common ancestral (plesiomorphic) condition. Some well known examples of convergent evolution in the nervous system include: image forming eyes of cephalopod mollusks (e.g., squids and octopods) and vertebrates (b0280 Packard, 1972), and the evolution of G-protein-coupled receptors as odorant receptors in many animal phyla (b0090 Eisthen, 2002). Examples of parallel evolution in the nervous system of vertebrates have been summarized in several recent reviews (b0260 Nishikawa, 2002; b0395 Zakon, 2002). These include: electric communication in mormyriform (African) and gymnotiform (South American) electric fishes (b0085 Albert and Crampton, 2005), prey capture among frogs (b0255 Nishikawa, 1999), sound localization among owls (b0150 Grothe *et al.*, 2005), and thermoreception in snakes (b0165 Hartline, 1988; b0245 Molenaar, 1992).

p0050 Reversals are among the most common forms of homoplasy, and are often the most difficult to detect even in the context of a resolved phylogenetic hypothesis of relationships (b0070 Cunningham, 1999). The reason for this is the phenotypes of some reversals may be quite literally identical, as in the case of

ELSEVIER FIRST PROOF

convergent loss of structures (e.g., the derived loss of paired limbs in snakes and limbless lizards).

s0025 **1.03.2.3 Character State Polarity**

p0055 A central task of ancestral character state reconstruction is determining the direction or polarity of evolutionary change between alternative states of a character. The ancestral state is referred to as plesiomorphic or primitive, and the descendent state is referred to as apomorphic or derived. Establishing the polarity of a character state transformation is critical to understanding the functional significance of that event. Phenotypes determined to be primitive simply mean they precede the derived state in time and are not necessarily functionally inferior. It is often although my no means always the case that characters evolve from more simple to more complex states, or from the absence of a particular state to the presence of that state.

p0060 There are several methods in use to determine character state polarity. The most widely used method is the so-called outgroup criterion, which employs conditions observed in members of clades other than the clade in which the derived state is present. The basic idea of the outgroup criterion is that for a given character with two or more states within a group, the state occurring in related groups is assumed to represent the plesiomorphic state. In other words, the outgroup criterion states that if one character is found in both ingroup and outgroup, this character is then postulated to be the ancestral state (plesiomorphic). Of course, it is always possible that a given outgroup exhibits an independently derived state of a given character, which is why the condition in several outgroup taxa is regarded as a more reliable test of the plesiomorphic condition.

s0030 **1.03.2.4 Character or Trait Data**

p0065 Methods for estimating ancestral character states and analyzing phenotypic evolution may treat trait data either as continuous (quantitative) or discrete (qualitative) (Zelditch *et al*., 1995; Rohlf, 1998; Wiens, 2001). Continuously distributed trait values have no easily distinguished boundaries between phenotypes. Examples of continuous traits include the sizes of brains and brain regions (e.g., nuclei), the number of cells in a brain region, pigment intensity, amplitude or timing of communication signals, and the amount of gene expression in a tissue. Continuous phenotypic variation typically reflects the additive effects of alleles at multiple loci and is frequently also influenced by environmental factors. Patterns of intraspecific (within species) continuous variation are often analyzed using parametric statistics, including such devices as the population mean and standard deviation. Methods for the analysis of interspecific (between species) continuous traits are useful for assessing the quantitative relationships among variables to address questions regarding, for example, the trade-offs and constraints among correlated traits.

p0070 Discontinuous traits have only a few distinct phenotypes. In many cases alternative alleles generate phenotypes that differ from each other in discrete steps, such that each phenotype can be clearly distinguished from the others. Many classes of phenotypic data are inherently discrete, such as meristic counts (e.g., number of body segments, rhombomeres, and cortical visual maps), and genetic polymorphisms (e.g., left- vs. right-handedness). Nucleotide bases at a locus are discrete states of a character. The presence (or absence) of derived traits on a phylogenetic tree also constitutes a class of discrete phenotypes. Such derived traits that underlie or explain subsequent evolutionary events are referred to as key innovations. Some widely cited examples of putative key innovations in the comparative neurosciences include arthropod cephalic tagmosis (Strausfeld, 1998), cephalopod eyes (Hanlon and Messenger, 1996), craniate neural crest (Northcutt and Gans, 1983), and ray-finned fish genome duplication (Taylor *et al*, 2003; Postlethwait *et al*., 2004). Each of these novelties is thought to have been critical in the diversification of the taxon in which it originated.

1.03.2.5 Adaptation s0035

p0075 One of the most widely applied uses of ancestral character state reconstruction is in the study of adaptation. The word adaptation is derived from the Latin *ad* (to, towards) and *aptus* (a fit), and is used to imply a feature or phenotype that evolved to serve a particular function or purpose. For example the function or purpose of an animal central nervous system is to coordinate sensory information and motor output patterns; that is to say, a centralized brain is an adaptation for sensory-motor coordination. Adaptation is therefore used both as a noun to describe the features that arose because of natural selection, and as a verb, the process of natural selection through which the features originated. In an evolutionary context, an adaptation is not only a static description of the match between form and function, but is also an explanation for the origin of that relationship (Russell, 1916).

p0080 It is important to distinguish among several distinct uses of the word adaptation in the biological sciences. A physiological adaptation is an

ELSEVIER FIRST PROOF

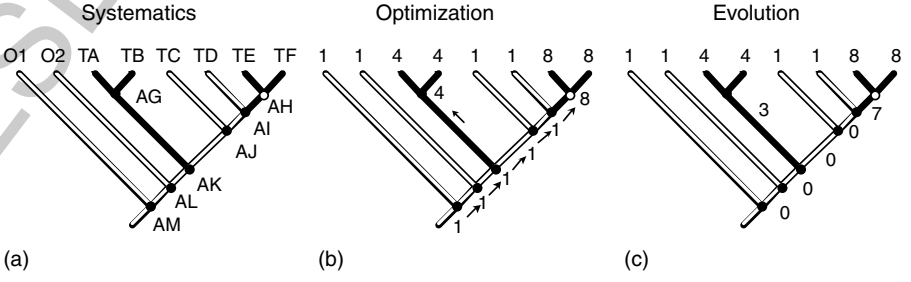
organismal response to a particular stress: if you heat up from the sun you may respond by moving into the shade (a behavioral adaptation), or you may respond by sweating (a physiological adaptation). In an evolutionary context, adaptation is also a change in response to a certain problem, but the change is genetic. Evolutionary adaptations that result from the process of natural selection usually take place over periods of time considerably longer than physiological time scales. Traits are referred to as adaptations only when they evolved as the solutions for a specific problem; that is for a particular function or purpose. A physiological response can itself be an adaptation in the evolutionary sense.

p0085 In reconstructing ancestral phenotypes it is important to bear in mind the primitive condition may be more or less variable than the conditions observed in living species. In some cases physiological or developmental plasticity is itself an evolutionary (genetic) specialization that permits organisms to adapt physiologically or behaviorally. For example, many species are characterized as eurytopic, or tolerant of a wide variety of habitats. Other species are stenotopic, or adapted to a narrow range of habitats. Similarly, individual characters may be more or less variable within a species, and this variability may itself be subject to evolutionary change. Flexible phenotypes may be more adaptive in a variable environment and stereotyped phenotypes more adaptive in a stable environment (van b0370 Buskirk, 2002).

s0040 1.03.2.6 Phylogenetic Trees

p0090 Implicit in all phylogenetic methods for studying character evolution is a tree-shaped branching diagram, alternatively called a dendrogram, cladogram, phenogram, or tree, depending on the methods used to construct the diagram, and the information content it is intended to convey. It is important to note that each of the many alterative methods for building trees that are currently available was designed to communicate different kinds of information. The methods grouped formally as 'phylogenetic systematics' (cladistics) exclusively use derived similarities (synapomorphies) to hypothesize genealogical relationships. This is to be contrasted with phenetic methods which use measures of overall similarity to group taxa, including both primitive and derived aspects of similarity. Cladistic methods generate branched diagrams referred to as cladograms, which should be viewed as summary diagrams depicting the branching pattern most consistent with a given data set (morphological or molecular). It is important to distinguish raw cladograms from phylogenetic trees; there is no time dimension to a cladogram *per se,* and the branch lengths are simply proportional to the minimum number of steps required to map all the character states onto that tree. A robust phylogenetic tree is usually the result of several or many phylogenetic analyses. The geological time frames associated with branching events are usually estimated from external paleontological, molecular, and biogeographic sources of information.

p0095 Figure 1 provides a conceptual overview for how phylogenetic trees may be used to study phenotypic evolution. All comparative approaches begin by assuming (or building) a hypothesis of genealogical interrelationships among the taxa of interest. There are many methods, even whole philosophies, of tree building, and the reader is referred to Page and b0285 Holmes (1998) for an introduction to this literature. Phylogenetic methods are then used to optimize character states at internal nodes of the tree; these nodes or branching points are hypothesized speciation events. Comparisons of trait values at ancestral



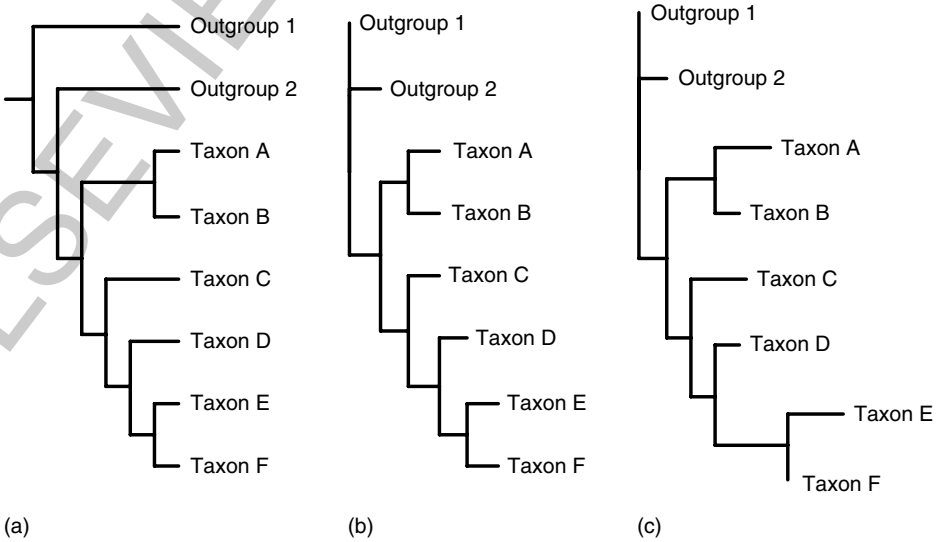
f0005 **Figure 1** Summary of the comparative approach for inferring phenotypic evolution. (a), Phylogenetic systematics (i.e., tree building): reconstruction of genealogical interrelationships among taxa (extant and/or fossil) using morphological and/or molecular sequence data. Taxa are species or clades (monophyletic groups of species): phylogeny includes six ingroup terminal taxa (TA-TF) and two outgroup taxa (O1-O2). (b), Character state optimization at internal nodes (branching points or hypothesized speciation events). Observed trait values at tips of the tree. Seven internal tree nodes represented by ancestral taxa (AG-AM) with trait values estimated by linear parsimony. (c), Evolution: tracing the history of phenotypic changes along branches of the tree. Numbers indicate absolute amount of trait change on the branch.

ELSEVIER FIRST PROOF

and descendant nodes of the tree allow the history of phenotypic changes to be traced. The distribution of these phenotypic changes (also known as steps or transformations) can then be assessed, qualitatively or quantitatively, depending on the types of data examined and the analytical methods employed.

p0100 A tree-shaped branching diagram conveys two kinds of information (weather they are intended or not): the tree topology, or the sequential order in which the taxa branch from one another, and the lengths of the individual branches (Figure 2). These two aspects of a tree correspond to the cladogenesis and the anagenesis of Rensch (1959). The tree topology (branching order) is reconstructed from distribution of shared–derived traits among taxa. The traits examined may be morphological novelties or nucleotide substitutions. Branch lengths may be reconstructed from one or more sources of information, including alternative models (or modes) of character evolution, or from empirical data (Losos, 1989). Under models of constant (or near constant) evolution (e.g., molecular clocks), all terminal taxa are treated as equidistant from the root (or base) of the tree. Terminal taxa are those at the tips of the tree, as opposed to ancestral taxa at internal nodes (branching points) within the tree. Under models of punctuated equilibrium, all (or most) character evolution occurs at branching points (nodes), and all branches are therefore of equal (or almost equal) length. Branch lengths derived from empirical data sets may be treated as proportional to the amount of character state change on that particular tree topology, or from stochastic models of evolution assuming that DNA nucleotide substitutions occur at an equal rate (Sanderson, 2002). The constant evolution and punctuated equilibrium models represent extremes of branch-length heterogeneity, between which branch lengths derived from empirical data sets usually fall. Branch lengths for clades with known fossilized members can also be estimated from the geological age of these fossils (Benton *et al.*, 2000; Near and Sanderson, 2004). Calibrations based on molecular sequence divergence or fossil data can take one of two forms: assignment of a fixed age to a node, or enforcement of a minimum or maximum age constraint on a node. The later option is generally a better reflection of the information content of fossil evidence.

p0105 It is important to recognize an analytical difference in the two kinds of information represented in a phylogeny: whereas the tree topology is transitive the branch lengths are not. In the language of formal logic, 'transitive' means that a relationship necessarily holds across (i.e., it transcends) the particularity of data sets. In the case of phylogenetic trees, the branching order derived from analysis of one data set is expected to predict the branching order of independent data sets (e.g., those derived from different genes, genes and morphology, osteology and neurology). Branch lengths, however, are intransitive, meaning the branch length values derived from one data set are not expected to predict those of other data sets. The reason for this is that we believe there has been a single phylogenetic history of life; a



f0010 **Figure 2** Alternative branch length models. (a), Molecular clock: all terminal taxa equidistant from root to from an ultrametric tree. (b), Equal branch lengths: all character evolution (anagenesis) occurs at branching events, as in punctuated equilibrium. (c), Empirical: branch lengths proportional to amount of character evolution and/or geological ages determined from fossils. Note: tree topology is transitive; branch lengths are not.

ELSEVIER FIRST PROOF

unique sequence of speciation events that gave rise to the species richness of the modern world. This single history underlies the evolution of all aspects of organismal phenotypes. There are, however, no such expectations of homogeneity in the rates of phenotypic (or gene sequence) evolution; in fact, the differential effects of directional and stabilizing selection on different phenotypes may be expected to result in longer or shorter branches for some traits than others.

1.03.3 Methods

1.03.3.1 Parsimony Optimization of Discrete Traits

The principle of parsimony (i.e., Occam's Razor) is widely used in the natural sciences as a method for selecting from among numerous alternative hypotheses. The principle of parsimony underlies all scientific modeling and theory building. The basic idea is that one should not increase, beyond what is necessary, the number of entities required to explain anything. In this context, parsimony means that simpler hypotheses are preferable to more complicated ones. It is not generally meant to imply that Nature itself is simple, but rather that we as observers should prefer the most simple explanations.

Maximum parsimony (MP) is a character-based method used in phylogenetic systematics to reconstruct phylogenetic trees by minimizing the total number of evolutionary transformations (steps) required to explain a given set of data. In other words, MP minimizes the total tree length. The steps may be nucleotide base or amino-acid substitutions for sequence data, or gain and loss events for restriction site and morphological data. MP may also be used to infer ancestral states of a character within a phylogenetic tree (this is discussed in the following).

1.03.3.2 Binary and Multistate Characters

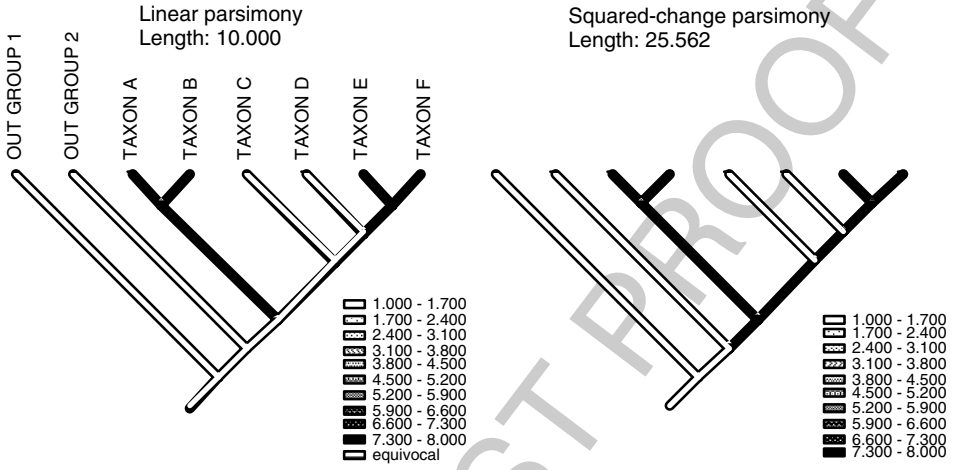
Discrete characters may be characterized as either binary (coded into two mutually exclusive alternative states) or as multistate (a transformation series of three or more discrete states). The alternative states of a binary character are generally (although not necessarily) explicit hypotheses of the primitive and derived (advanced) states of a single evolutionary transformation event, such as the origin (or loss) of a novel feature. A multistate character is a more complex intellectual device with many more interpretations of meaning. Multistate characters may be presented as many stages of a long-term phylogenetic trend (e.g., larger relative brain size, larger body size) or as independent alternative trends from a common ancestral plan (e.g., large brains evolving from enlargement of the cerebellum in chondrichthyans vs. the telencephalon in mammals). An ordered transformation series models a preconceived phylogenetic sequence of changes, such that in the series 1-2-3, state 3 is only permitted to be derived from state 2. In an unordered transformation series, state 3 may be derived from either of states 1 or 2. Following a similar logic, reversals (e.g., from 2 to 1) may be allowed, penalized, or prohibited, depending on the preconceptions of the investigator. Of course, building *a priori* conceptions of order or reversibility into an analysis of character state change precludes the use of that analysis as an independent test of those assumptions. To summarize this section, treating all characters as unpolarized and unordered means that all transitions among states are regarded as equally probable.

1.03.3.3 Squared-Change and Linear Parsimony

There are two general types of MP widely used in tracing the evolution of continuous traits; squared-change parsimony and linear parsimony. Squared-change algorithms (Rogers, 1984) seek to minimize the amount of squared change along each branch across the entire tree simultaneously, using a formula in which the cost of a change from state x to y is $(x-y)^2$. Squared-change parsimony assigns a single ancestral value to each internal node to minimize the sum of squares change over the tree (Maddison, 1991).When using squared-change parsimony, the absolute amount of evolution over the whole is not necessarily minimized, and some degree of change is forced along most branches. Linear parsimony reconstructs ancestral node values by minimizing total changes (Figure 3). Linear-parsimony algorithms (Kluge and Farris, 1969) seek to minimize the total amount of evolution and consider only the three nearest nodes when calculating the ancestral character states. In linear parsimony the cost of a change from x to y is $|x-y|$. The result of this local optimization is that changes are inferred on very few or single branches. Linear parsimony therefore permits the accurate reconstruction of discontinuous events, or of large changes in trait values on a tree. Although evolutionary change is often thought of as gradual, large changes on a tree may result from a variety of real biological processes, not the least of which is the extinction of taxa with intermediate trait values (Butler and Losos, 1997).

ELSEVIER FIRST PROOF

Taxon	*Character*
Outgroup 1	1
Outgroup 2	1
Taxon A	4
Taxon B	4
Taxon C	1
Taxon D	1
Taxon E	8
Taxon F	8



AU11

f0015 **Figure 3** Alternative methods for estimating ancestral character states. (a), Linear parsimony. (b), Squared-change parsimony. Character state data by taxon reported in the table.

ELSEVIER FIRST PROOF

s0065 **1.03.3.4 Maximum Likelihood and Bayesian Optimization**

p0130 Maximum likelihood (ML) methods for tracing character evolution select ancestral trait values with highest likelihood on a given phylogenetic hypothesis given a model of trait evolution (defined by user). Bayesian analysis (BA) selects the ancestral trait value with the highest posterior probability, given the probabilities of priors (external evidence) and assumptions of trait evolution (defined by user). Because they are model-based approaches, ML and BA optimization methods are more commonly used in the analysis of gene sequence data, using explicit models of changes between nucleotide bases (Liò and Goldman, 1998; Sullivan *et al*., 1999). ML has been used in the analysis of continuous character evolution where the models may vary from very simple (e.g., Brownian motion) to quite complex; there is a large literature regarding methods to test the validity of using particular models (Díaz-Uriarte and Garland. 1996; Oakley, 2003).

s0070 **1.03.3.5 Which Optimization Approach to Use?**

p0135 Empirical studies using simulated data sets and those derived from evolution in a test tube have concluded that model driven approached like ML and BA give more accurate results than MP when the modeled parameters (i.e., likelihood or probability of nucleotide substitutions) are known, but can be positively misleading when the parameters are unknown (Hillis *et al*., 1992; Oakley and Cunningham, 2000). MP often provides less resolution (more interior tree nodes reconstructed with ambiguous states), than ML or BA methods, which usually give very precise estimates with high confidence levels even under circumstances in which available data are insufficient to the task. In this regard, MP methods are regarded as more conservative, with lower risk of committing type I errors or false positives (Webster and Purvis, 2002).

p0140 Most studies on the evolution of neural characters use MP approaches because, unlike molecular sequence data, it is not straightforward how to pose or parametrize models on the evolution of complex phenotypes. Continuously varying aspects of neural features, like the size or shape of structures, have been modeled as simple Brownian motion or random walk processes, under the assumptions that the trait has not experienced selection and that there are no constraints on variance through time (Butler and King, 2004). Whether or not the assumptions of Brownian motion or any other specific model are satisfied by real neural or behavioral data is almost completely unknown.

p0145 A general conclusion reached by a number of review studies is that, under most circumstances faced by comparative morphologists, linear

parsimony is the most conservative method for reconstructing ancestral trait values (Losos, 1999). Unlike squared-change parsimony, linear parsimony does not average out change over the interior nodes of a tree, but rather permits discontinuous changes along a branch. This has the advantageous effect of not forcing gradual trait evolution on the tree, and also of not forcing unnecessary trait reversals (Figure 3). A methodological advantage of linear over squared-change parsimony is that it permits the reconstruction of ambiguous ancestral character state reconstructions (Figure 4). This is a desirable property in cases where the available data are in fact insufficient to resolve the trait value at a specified internal nodes (Cunningham, 1999). A methodological disadvantage of linear parsimony is that, computationally, it requires a completely resolved tree topology in which all branching events are divided into only two daughter clades. Unfortunately, fully resolved trees are unusual in most studies with many (>30) species. By contrast, squared-change parsimony can be calculated on a tree with unresolved multichotomies (also called polytomies), and therefore often becomes the method of choice by default. One alternative to using squared-change parsimony when faced with an incompletely resolved tree is to use linear parsimony on numerous (100, 1000) arbitrarily resolved trees, and then report statistics (e.g., minimum and maximum) of the trait values obtained. Software for this procedure is available in the freely available Mesquite software package (Maddison and Maddison, 2005).

1.03.3.6 Correlative Comparative Methods

Ordinary least-squares regression allows one to investigate relationships between two variables in order to ask if change in one of these variables is associated with change in the other. One may ask, for example, how is variation in brain size related to body size, ecological role (predator vs. prey), climate, life history mode, or locomotion (Albert *et al.*, 2000; Safi and Dechmann, 2005)? The least-squares fitting procedure is commonly used in data analysis in comparative studies, and conventional regression analysis has been one of the main tools available to comparative neurobiology and ecological physiology to study of form–function relationships and adaptation (Garland and Carter, 1994). However, it is now widely recognized that interspecific observations generally do not comprise independent and identically distributed data points, thus violating fundamental assumptions of conventional parametric statistics (Felsenstein, 1985, 1988; Pagel and Harvey, 1989; Harvey and Pagel, 1991).

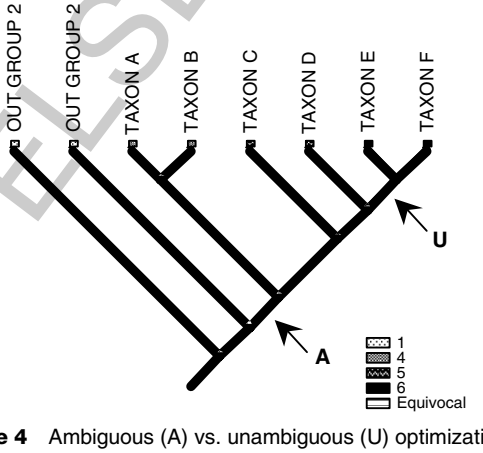
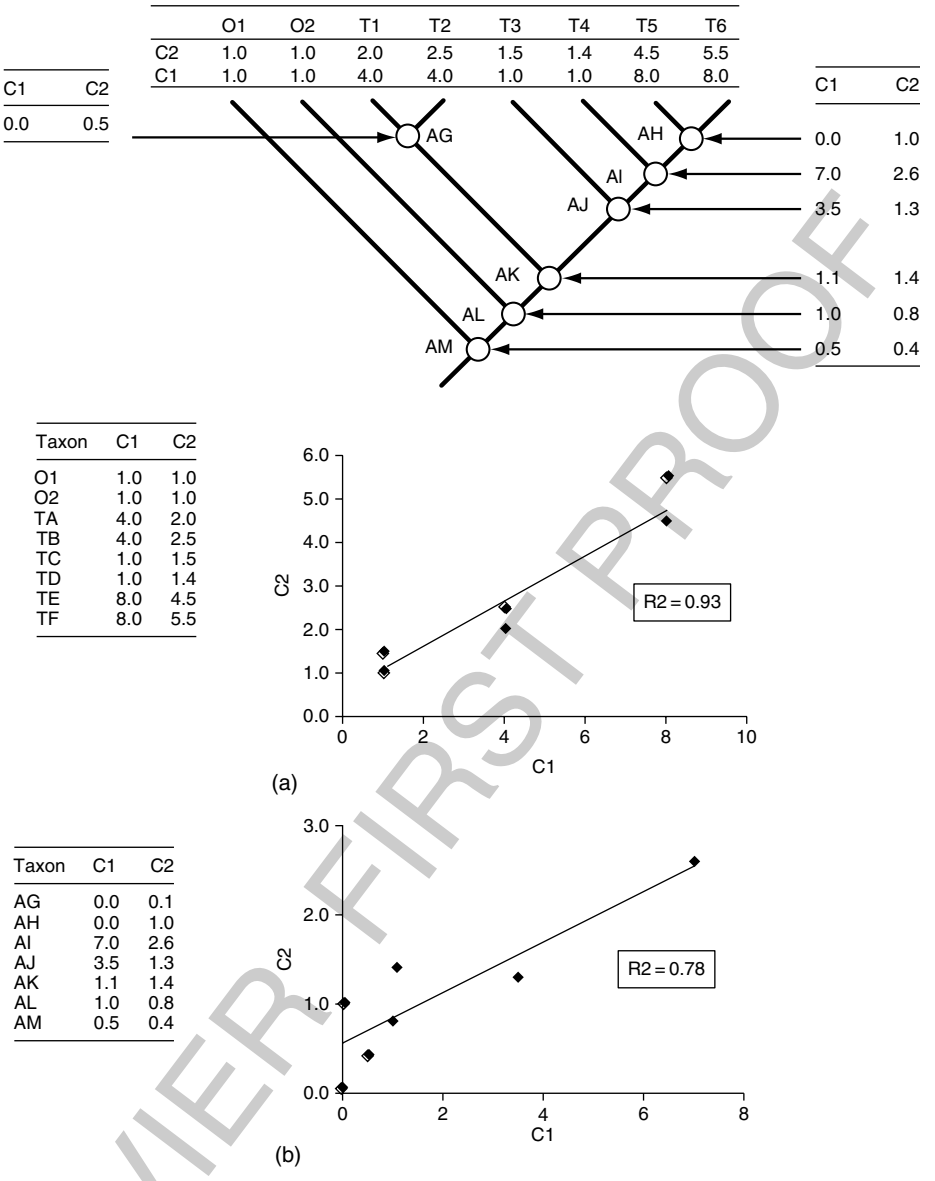


Figure 4 Ambiguous (A) vs. unambiguous (U) optimizations.

Phylogenetically based statistical methods allow traditional topics in comparative neuroanatomy and physiology to be addressed with greater rigor, including the form of allometric relationships among traits and whether phenotypes vary predictably in relation to behavior, ecology, or environmental characteristics (Brooks and McLennan, 1991; Frumhoff and Reeve, 1994; Losos, 1996). In a conventional regression analysis the data points represent terminal taxa. In a phylogenetic regression the data points represent sister-taxon comparisons (Grafen, 1989). These two methods are compared in Figure 5, in which identical data are analyzed using conventional and phylogenetic regression methods. The phylogeny of Figure 5 includes six terminal taxa (TA-TF) and two outgroup taxa (O1-O2), which are represented by two continuously distributed characters (C1 and C2). The tree topology has been determined from data other than characters 1 and 2, and the branch lengths are treated as equal (under a model of punctuated equilibrium). There are seven internal tree nodes represented by ancestral taxa (AG-AM) with trait values estimated by least-square parsimony. By removing psuedorepilcates, the phylogenetic regression compares fewer taxa, has fewer degrees of freedom, and has a lower correlation coefficient (R^2 value) than does the conventional regression. The phylogenetic regression therefore provides a better quantitative measure of correlated evolution between the two traits, and is a more conservative measure of the strength of adaptive pressures.

Relationships between brain size and the volume of frontal and visual cortices in mammals have recently been studied using the methods of phylogenetic regression analysis (Bush and Allman, 2004a, 2004b). These studies found that size has a

ELSEVIER FIRST PROOF



f0025 **Figure 5** Comparison of conventional and phylogenetic regression analyses. Phylogeny of six terminal taxa (TA-TF) and two outgroup taxa (O1-O2), represented by two continuously distributed characters (C1 and C2). Tree topology determined from data other than characters 1 and 2, and branch lengths treated as equal. Seven internal tree nodes represented by ancestral taxa (AG-AM) with ancestral trait values estimated by least-square parsimony. (a), Conventional regression of trait values from terminal taxa. (b), Phylogenetic regression of trait values at internal tree nodes using the method of independent contrasts. Note that by removing psuedorepilcates, the phylogenetic regression compares fewer taxa, has fewer degrees of freedom, and has a lower correlation coefficient (R^2 value) than does the conventional regression. The phylogenetic regression therefore provides a more conservative quantitative measure of correlated evolution between the two traits.

ELSEVIER FIRST PROOF

profound effect on the structure of the brain, and that many brain structures scale allometrically; that is, their relative size changes systematically as a function of brain size. They also conclude that the three-dimensional shape of visual maps in anthropoid primates is significantly longer and narrower than in strepsirrhine primates. Using conventional regression analyses, b0375 von Bonin (1947) showed that frontal cortex hyperscales with brain size, and humans have "precisely the frontal lobe which [we deserve] by virtue of the overall size of [our] brain." These are of course precisely the qualitative conclusions arrived at by Bush and Allman using analysis of phylogenetic regressions. In fact, many studies reviewing the uses of phylogenetic methods for reconstructing ancestral states conclude that all methods will recover a very strong historical signal (b0210 Losos, 1999).

1.03.4 Limitations of Methods

The accuracy of ancestral reconstructions has been investigated by comparisons with known phylogenies (e.g., viruses, computer simulations; Oakley and Cunningham (2000)). It is well known that all phylogenetically based methods perform poorly when taxon sampling is low and when rates of evolution in the character of interest are unequal among branches of the tree (Garland *et al.*, 1993; Sullivan *et al.*, 1999; Hillis *et al.*, 2003). Further, all methods for studying character evolution on a tree make certain assumptions about the capacity of trees to faithfully record the actual history of character change. These include the assumptions that: phenotypic diversification results largely from speciation and that the effects of extinction have not erased the signal; that taxon sampling faithfully represent the history of diversification; and that genealogical history is largely or entirely bifurcating (vs. multifurcating or converging). Of course, all methods assume we know the 'true' (or 'nearly true') tree topology. In addition, each of the optimization methods makes assumptions about critical parameters, including branch lengths, models of character evolution, absolute rates of evolution, homogeneity (vs. heterogeneity) of evolutionary rates, reversibility (or the lack thereof), and the orderedness (or unorderedness) of multistate characters.

The accuracy of ancestral trait reconstruction also depends strongly on parameter estimation (e.g., tree topology, branch lengths, models of trait evolution). ML and BA perform well when model assumptions match real parameters. ML and BA are positively misleading when model assumptions are violated. MP is more conservative, making fewer type II errors than ML and BA when biological parameters are not known. Squared-change parsimony, ML and BA minimize large changes, spreading evolution over the internal tree branches. Linear parsimony permits reconstructions at ancestral nodes with no change, and permits ambiguous reconstructions. 'Independent contrasts' assumes that selection operates in the origin but not maintenance of derived traits.

Both conventional and phylogenetic correlations of interspecific character data make assumptions about critical parameters. These assumptions are often of unknown validity, and in some cases are known to be incorrect. Conventional statistics assume that each terminal taxon (tips of the tree) may be treated as independent sample of the relationship under investigation. This means that the character value (phenotype) observed in that taxon evolved independently (without inheritance) from the values in other taxa in the analysis. In an evolutionary context, this is equivalent to assuming that trait values result primarily from stabilizing selection in each species that acts to maintain trait values, rather than from directional selection at the origin of the trait in an ancestral species (Hansen, 1997). In other words, conventional statistics assume traits to be highly labile and without significant phylogenetic inertia. Phylogenetic correlations make converse assumptions, that trait values are due largely or entirely to directional selection at the origin of a feature and that the influence of stabilizing selection is negligible. Phylogenetic correlations also must make particular assumptions about branch lengths and models of trait evolution.

1.03.5 Conclusions

As in all aspects of historical inquiry, the study of character evolution is exceptionally sensitive to the amount of information that has actually survived up to the present. The reality of neural evolution was in most cases almost certainly very complex, and may be reliably regarded to have included vastly more numbers of independent transformations than has been recorded in the distribution of phenotypes preserved among living species. The signature of many historical events has been overwritten by reversals and convergences, or eliminated altogether by extinctions. Paleontologists estimate that more than 99% of all species that have ever lived are now extinct (Rosenzweig, 1995). This figure, of course, includes higher taxa (e.g., trilobites, placoderms, plesiosaurs) that are now entirely extinct, bringing up the aggregate percentage of extinction for all taxa. The proportion of living species that persists within certain targeted taxa may be much higher (e.g., Lake Victoria cichlid fishes). Nevertheless, in comparative studies of neural, physiological, or behavioral phenotypes, it is rare to have information on all extant species. Whether it is from extinction or incomplete surveys, taxon sampling remains one of the greatest sources of error in phylogenetic estimates of character evolution (Sullivan *et al.*, 1999; Zwickl and Hillis, 2002).

Despite all these reservations, we must continue to estimate ancestral traits in order to study phenotypic evolution. None of the methods reviewed in this article should be regarded as a magic bullet, but rather there are advantages and disadvantages of each method as they are applied under different circumstances. All the methods reviewed here have proved to be useful tools in the phylogenetic toolbox. As in other aspects of science, it is important to make our assumptions explicit, and to use reasonable

ELSEVIER FIRST PROOF

assumptions. Further, as in other aspects of evolutionary biology, critical insights into the evolution of neural characters will come from a better understanding of the biology of the phenotypes themselves, and the organisms in which they have evolved.

References

b0005 Albert, J. S. and Crampton, W. G. R. 2005. Electroreception and electrogenesis. In: The Physiology of Fishes (*eds.* D. H. Evans and J. B. Claiborne) 3rd edn., pp. 431–472.

b0010 Albert, J. S., Lannoo, M. J., and Yuri, T. 1998. Testing hypotheses of neural evolution in gymnotiform electric fishes using phylogenetic character data. *Evolution* 52, 1760–1780.

b0015 Albert, J. S., Froese, R., and Paulay, D. 2000. The brains table. In: FishBase 2000, Concepts, Design and Data Sources (*eds.* R. Froese and D. Paulay), pp. 234–237. ICLARM.

b0020 Benton, M. J. 1993. The Fossil Record 2, p. 845. Chapman & Hall.

b0025 Benton, M. J., Wills, M., and Hitchin, R. 2000. Quality of the fossil record through time. *Nature* 403, 534–538.

b0030 Blackledge, T. A. and Gillespie, R. G. 2004. Convergent evolution of behavior in an adaptive radiation of Hawaiian web-building spiders. *Proc. Nat. Acad. Sci.USA* 101, 16228–16233.

b0035 Blomberg, S. P., Garland, T., and Ives, A. R. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57, 171–745.

b0040 Brooks, D. R. and McLennan, D. A. 1991. Phylogeny, Ecology, and Behavior. University of Chicago Press.

b0045 Bush, E. C. and Allman, J. M. 2004a. Three-dimensional structure and evolution of primate primary visual cortex. *Anat. Rec. Part A* 281A, 1088–1094.

b0050 Bush, E. C. and Allman, J. M. 2004b. The scaling of frontal cortex in primates and carnivores. *Proc. Nat. Acad. Sci. USA* 101(11), 3962–3966.

b0055 Butler, M. A. and Losos, J. B. 1997. Testing for unequal amounts of evolution in a continuous character on different branches of a phylogenetic tree using linear and squared-change parsimony: An example using Lesser Antillean *Anolis* lizards. *Evolution* 51(5), 1623–1635.

b0060 Butler, M. A. and King, A. A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164(6), 683–695.

b0065 Coyne, J. A. and Orr, H. A. 1989. Two rules of speciation. In: Speciation and Its Consequences (*eds.* D. Otte and J. Endler), pp. 180–207. Sinauer Associates.

b0070 Cunningham, C. W. 1999. Some limitations of ancestral character-state testing evolutionary hypotheses. *Syst. Biol.* 48(3), 665–674.

b0075 de Queiroz, A. and Wimberger, P. H. 1993. The usefulness of behavior for phylogeny estimation: levels of homoplasy in behavioral and morphological characters. *Evolution* 47, 46–60.

b0080 Diaz-Uriarte, R and Garland, T. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: Sensitivity to deviations from Brownian motion. *Syst. Biol.* 45(1), 27–47.

b0085 Edinger, T. 1941. The brain of *Pterodactylus*. *Amer. J. Sci.* 239(9), 665–682.

b0090 Eisthen, H. L. 2002. Why are olfactory systems of different animals so similar? *Brain Behav. Evol.* 59, 273–293.

b0095 Felsenstein, J. 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *Am. J. Genet.* 25, 471–492.

Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125, 1–15. b0100

Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* 19, 445–471. b0105

Frumhoff, P. C. and Reeve, H. K. 1994. Using phylogenies to test hypotheses of adaptation: a critique of some current proposals. *Evolution* 48, 172–180. b0110

Garland, T., Jr. and Carter, P. A. 1994. Evolutionary physiology. *Annu. Rev. Physiol.* 56, 579–621. b0115

Garland, T., Jr., Harvey, P. H., and Ives, A. R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* 41, 18–32. b0120

Garland, T., Jr., Dickerman, A. W., Janis, C. M., and Jones, J. A. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* 42, 265–292. b0125

Garland, T., Jr., Midford, P. E., and Ives, A. R. 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. *Am. Zool.* 39, 347–388. b0130

Garstang, W. 1922. The theory of recapitulation: a critical restatement of the biogenetic law. *J. Linn. Soc. Zool.* 35, 81–101. b0135

Gittleman, J.L and Luh, H. K. 1992. On comparing comparative methods. *Annu. Rev. Ecol. Syst.* 23, 383–404. b0140

Grafen, A. 1989. The phylogenetic regression. *Phil. Trans. R. Soc. Lond.* 326, 119–157. b0145

Grothe, B., Carr, C. E., Cassedy, J., Fritzsch, B., and Köppl, C. 2005. Brain pathway evolution and neural processing patterns - parallel evolution? In: Evolution of the Vertebrate Auditory System Springer Handbook of Auditory Research (*eds.* Manley, Popper, and Fay). Springer. b0150

Hanlon, R. T. and Messenger, J. B. 1996. Cephalopod Behaviour, p. 232. Cambridge University Press. b0155

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51, 1341–1351. b0160

Hartline, P. H. 1988. Thermoreception in snakes. *Prog. Brain Res.* 74, 297–312. b0165

Harvey, P. H. and Pagel, M. D. 1991. The Comparative Method in Evolutionary Biology. Oxford University Press. b0170

Hillis, D. M., Bull, J. J., White, M. E., Badgett, M. R., and Molineux, I. J. 1992. Experimental phylogenetics: generation of a known phylogeny. *Science* 255, 589–592. b0175

Hillis, D. M., Pollock, D. D., Mcguire, J. A., and Zwickl, D. J. 2003. Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* 52(1), 124–126. b0180

Jerison, H. J. 1976. Paleoneurology and evolution of mind. *Sci. Am.* 234, 90. b0185

Kluge, A. G. and Farris, S. J. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18, 1–32. b0190

Krebs, H. A. 1975. The August Krogh principle: for many problems there is an animal on which it can be most conveniently studied. *J. Exp. Zool.* 194(1), 221–226. b0195

Liò, P. and Goldman, N. 1998. Models of molecular evolution and phylogeny. *Genome Res.* 8(12), 1233–1244. b0200

Losos, J. B. 1996. Phylogenies and comparative biology, stage II: testing causal hypotheses derived from phylogenies with data drawn from extant taxa. *Syst. Biol.* 45, 259–260. b0205

Losos, J. B. 1999. Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods. *Anim. Behav.* 58, 1319–1324. b0210

MacLeod, N. and Rose, K. D. 1993. Inferring locomotor behavior in Paleogene mammals via eigenshape analysis. *Am. J. Sci.* 293-A, 300–355. b0215

Maddison, W. P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous valued characters on a phylogenetic tree. *Syst. Zool.* 40, 304–314. b0220

ELSEVIER FIRST PROOF

b0225 Maddison, W. P. and Maddison, D. R. 2005. Mesquite: a modular system for evolutionary analysis. Version 1.06 http://mesquiteproject.org

b0230 Martins, E. P. 2000. Adaptation and the comparative method. *Trends Ecol. Evol.* 15, 295–299.

b0235 Mayr, E. 1963. Populations, Species, and Evolution. Harvard University Press.

b0240 McKinney, M. L. and McNamara, K. J. 1990. Heterochrony: The Evolution of Ontegeny. Plenum Press.

b0245 Molenaar, G. J. 1992. Infrared sensitivity of snakes. In: Biology of the Reptilia (*eds.* C. Gans and P. S. Ulinski), Vol. 17, pp. 367–453. University of Chicago Press.

b0250 Near, T. J. and Sanderson, M. J. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Phil. Trans. R. Soc. Lond. B* 359, 1477–1483.

b0255 Nishikawa, K. C. 1999. Neuromuscular control of prey capture in frogs. *Phil. Trans. R. Soc. Lond. B* 354, 941–954.

b0260 Nishikawa, K. C. 2002. Evolutionary convergence in nervous systems: insights from comparative phylogenetic studies. *Brain Behav. Evol.* 59, 240–249.

b0265 Northcutt, R. G. and Gans, C. 1983. The genesis of neural crest and epidermal placodes: a reinterpretation of vertebrate origins. *Q. Rev. Biol.* 58, 1–28.

b0270 Oakley, T. H. 2003. Maximum likelihood models of trait evolution. Comments. *Theor. Biol.* 8, 1–17.

b0275 Oakley, T. H. and Cunningham, C. W. 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution* 54(2), 397–405.

b0280 Packard, A. 1972. Cephalopods and fish - the limits of convergence. *Biol. Rev.* 47, 241–307.

b0285 Page, R. D. M. and Holmes, E. C. 1998. Molecular Evolution: A Phylogenetic Approach. Blackwell Scientific.

b0290 Pagel, M. D. and Harvey, P. H. 1989. Taxonomic differences in the sealing of brain on body weight among mammals. *Science* 244(4912), 1589–1593.

b0295 Patterson, C. 1982. Morphological characters and homology. In: Problems in Phylogenetic Reconstruction (*eds.* K.A Joysey and A. E. Friday), pp. 21–74.

b0300 Postlethwait, J., Amores, A., Cresko, W., Singer, A., and Yan, Y. L. 2004. Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet.* 10, 481–490.

b0305 Rensch, B. 1959. Evolution Above the Species Level., p. 419. Columbia University Press, Methuen & Co. Ltd.

b0310 Rogers, J. S. 1984. Deriving phylogenetic trees from allele frequencies. *Syst. Zool.* 33, 52–63.

b0315 Rogers, S. W. 2005. Reconstructing the behaviors of extinct species: An excursion into comparative paleoneurology. *Am. J. Med. Genet.* 134A(Part A), 349.

Rohlf, F. J. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Syst. Biol.* 47, 147–158. b0320

Rosenzweig, M. L. 1995. Species Diversity in Space and Time. Cambridge University Press. b0325

Russell, E. S. 1916. Form and Function: A Contribution to the History of Animal Morphology. Murray. 1982. Reprint, With a new introduction by G. V. Lauder. University of Chicago Press. b0330

Safi, K. and Dechmann, D. K. N. 2005. Adaptation of brain regions to habitat complexity: A comparative analysis in bats (Chiroptera). *Proc. R. Soc. BBiol. Sci.* 272(1559), 179–186. b0335

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109. b0340

Slowinski, J. B. 1993. Unordered versus ordered characters. *Syst. Biol.* 42, 155–165. b0345 AU8

Stensiö, E. A. 1963. The brain and cranial nerves in fossil, lower craniate vertebrates. *Skrifter Norske Videnskaps-Akademi 1 Oslo I. Mat.-Naturv. Klasse. Ny Serie* 13, 1–120. b0350

Strausfeld, N. J. 1998. Crustacean - insect relationships: The use of brain characters to derive phylogeny amongst segmented invertebrates. *Brain, Behav. Evol.* 52, 186–206. b0355

Sullivan, J., Swofford, D. L., and Naylor, G. J. P. 1999. The effect of taxon sampling on estimating rate heterogeneity parameters of maximum-likelihood models. *Mol. Biol. Evol.* 16(10), 1347–1356. b0360

Taylor, J. S., Braasch, I., Frickey, T., Meyer, A., and Van de Peer, Y. 2003. Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Res.* 13(3), 382–390. b0365

van Buskirk, J. 2002. A comparative test of the adaptive plasticity hypothesis: Relationships between habitat and phenotype in anuran larvae. *Am. Nat.* 160(1), 87–102. b0370 AU9

von Bonin, G. 1947. *Assoc. Res. Nervous Mental Dis.* 27, 67–83. b0375

Webster, A. J. and Purvis, A. 2002. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc. R. Soc. Lond. B.* 269, 143–150. b0380

Wiens, J. J. 2001. Character analysis in morphological phylogenetics: problems and solutions. *Syst. Biol.* 50, 689–699. b0385

Wimberger, P. H. and de Queiroz, A. 1996. Comparing behavioral and morphological characters as indicators of phylogeny. In: Phylogenies and the Comparative Method in Animal Behavior (*ed.* E. P. Martins). Oxford University Press. b0390 AU10

Zakon, H. H. 2002. Convergent evolution on the molecular level. *Brain Behav. Evol.* 59, 250–261. b0395

Zelditch, M. L., Fink, W. L., and Swiderski, D. L. 1995. Morphometrics, homology, and phylogenetics: quantified characters as synapomorphies. *Syst. Biol.* 44, 179–189. b0400

Zwickl, D. J. and Hillis, D. M. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* 51(4), 588–598. b0405

ELSEVIER FIRST PROOF