Adaptation, Constraint, and the Comparative Method: Phylogenetic Issues and Methods
Jonathan B. Losos and Donald B. Miles

INTRODUCTION

Ecomorphological analyses generally seek to link the structure and function of organisms with relevant features of the environment. The scope of ecomorphological studies is far reaching and includes the adaptive significance of morphological design, convergence, evolution of function, morphological evolution, structure-function correlations, and community organization. Two schools of ecological morphology have independently flourished (see Ricklefs and Miles, chap. 2, this volume). One approach, characterized by functional morphologists, focuses on the effects of design and function on ecology (e.g., habitat choice or diet breadth). Another, championed by ecologists, uses morphology as an index to ecology and, through the analysis of morphological variation within an assemblage of species, derives inferences about ecological processes. Both schools share a common assumption that variation in morphology bears a strong relationship to variation in a suite of ecological characteristics.

Regardless of the goals of a specific project, many studies assume that the patterns revealed in an ecomorphological analysis reflect adaptation to prevailing selective pressures. In the absence of historical information, however, such interpretations are problematic. For example, suppose that one postulated a relationship between the presence of lamella-bearing subdigital pads and invariant clutch size in lizards. At least 89% of all lizard species with invariant clutch size have pads (from Shine and Greer, 1991)—seemingly a strong relationship. But, when one takes a phylogenetic perspective, a different picture emerges. The relationship exists because two clades, the geckos (>850 species) and the anoles (>300 species), evolved both pads and invariant clutch sizes. However, invariant clutch size has evolved at least twenty times in taxa without pads (Shine and Greer, 1991). Thus, the historical view indicates that the evolution of invariant clutch size is not linked to the presence of subdigital pads, contrary to what the distribution of the traits among extant taxa might suggest.

As this example illustrates, ecomorphological studies are beset with problems attributable to the shared ancestry of species. Taxa which share a trait due to common ancestry cannot be viewed as independent points for statistical analysis; further, in the absence of historical information, evolutionary rate and direction cannot be deduced (Felsenstein, 1985; Harvey and Pagel, 1991). Conversely, inclusion of phylogenetic information in ecomorphological and other studies presents several advantages. First, hypotheses regarding the rates and direction of evolutionary change are possible. When conducted with ecological data simultaneously, tests of the correlated evolution of ecomorphological relationships are also possible. Second, limits to the expression of morphological variation in design or function may be recognized and evaluated. Third, the statistical dilemma of non-independence of species as data points is resolved. Fourth, phylogenetic analyses may suggest hypotheses that can be experimentally tested.

In this chapter, we discuss the importance of including phylogenetic information in ecomorphological analyses. Although many researchers recognize that the form of a trait is an admixture of genealogy and recent adaptive variation (e.g., Darwin, 1859), until recently the analytical methods necessary to account for phylogeny in comparative analyses were not available. In particular, we discuss the various aspects of ecomorphology that have benefited from the inclusion of a phylogenetic approach; how knowledge of phylogenies can refine an analysis; the methods for including phylogenetic information into ecomorphological analyses; and the new comparative methods that explicitly include phylogenetic information in the analyses. The first part of the paper deals with various areas to which a phylogenetic perspective is important. The second part, by reviewing and critiquing the rapidly burgeoning field of statistical comparative methods, highlights the advantages and assumptions of each method and indicates under which circumstances each method is appropriate.

Pertinence of Phylogeny to Ecological Morphology

Reconstruction of Evolutionary Pathways. The fossil record can often tell us a great deal about evolutionary pattern and process (e.g., Van Valkenburgh, chap. 7, this volume). In the absence of a sufficient fossil record (and even in its presence), the derivation of a phylogenetic hypothesis allows inferences to be drawn about the character states of hypothetical ancestral taxa. Comparisons between extant and inferred ancestral taxa enable tests of the number of times a trait has evolved independently and the existence of directional trends of evolution within a clade. Phylogenetic information is crucial for studying a variety of questions ranging from patterns of morphological evolution to the adaptive signifi-
cance of functional trait complexes. Below we discuss how phylogenetic approaches can refine ecomorphological analyses.

The presence of a given character in more than one species in a clade may be attributable to either shared ancestry or independent evolution of the trait. By mapping the character states on a phylogeny, these possibilities can often be distinguished. Phylogenetic studies can be critical in assessing how plastic a trait is evolutionarily (i.e., how frequently evolutionary change in the trait may have occurred) and can be used in testing scenarios about the evolution of a trait. For example, Losos (1990a) concluded that body size has evolved infrequently among northern Lesser Antillean Anolis lizards (see below), whereas Luke (1986) concluded that toe fringes had evolved in lizards from seven families no fewer than twenty-six times!

One initial procedure in an evolutionary analysis is the determination of character polarity: Does character A resemble an ancestral condition or has it changed since divergence? Character transformations can be determined by referring to a cladogram and using cladistic techniques (see Wiley et al., 1991; Maddison and Maddison, 1992, and references therein). Only with such information can one meaningfully attempt to understand the evolution of a character and test scenarios concerning its evolution. For example, a number of authors have questioned the selective features leading to the evolution of leaf retention in the common oak, Quercus robur, and the beech, Fagus sylvatica. Wanntorp (1983), however, noted that leaf retention, rather than shedding, appears to be the ancestral condition of the Fagaceae. Thus, one should look at characteristics of the family as a whole, rather than at attributes of particular species, to understand why leaf retention evolved (unless leaf retention is characteristic of an even larger group). More proximate explanations would be appropriate in understanding those taxa that have evolved leaf-shedding (Wanntorp, 1983; for another example see Doddington, 1986).

Further, one can ask why taxa are similar. Taxa may share the same condition of a trait either through common ancestry or because they have evolved the same condition independently. These possibilities often require radically different scenarios. In the absence of a detailed fossil chronology, the determination of character transformations can only be accomplished with reference to the phylogenetic relationships of the constituent taxa (Mickey and Johnson, 1976; Wake, 1991). However, caution must be exercised because different methodologies for reconstructing ancestral states may tend to under- or overestimate the occurrence of parallel evolution (see below). Russell's (1976, 1979) studies of the evolution of pad structure in geckos elegantly demonstrates the utility of a phylogenetic perspective, in this case distinguishing when taxa are padless due to the inheritance of the primitive padless condition from when padlessness has resulted from convergent secondary loss of the pads (fig. 4.1).

Phylogenetic analyses can also be used to investigate proposed evolutionary sequences and scenarios. For example, plethodontid salamanders display a bewildering array of shapes of the premaxilla bone. Mapping variation in premaxillary shape along a phylogeny revealed that a paired premaxillary structure in adults represents the primitive condition, but a single, fused bone has evolved repeatedly. Further, consideration of ontogenetic data in a phylogenetic context suggested that the fusion has resulted from paedomorphosis in several cases, but peramorphosis in others (Larson, 1984; Lombard and Wake, 1986; Wake and

![Figure 4.1](image-url) The evolution of subdigital pad structure in two gecko subfamilies. Comparison of trends for the two subfamilies indicates parallel patterns of pad evolution and subsequent secondary reduction (from Russell, 1979, with permission).
Larson, 1987; Wake, 1991). In a similar manner, Liebherr and Hajek (1990), testing predictions of the taxon pulse and taxon cycle theories concerning the evolution of habitat use in New World carabid beetles, found no trends in evolutionary direction relative to a null model of random change in habitat use. Dorit (1990) examined a proposed trend in the evolution of neurocranial shape in a clade of cichlid fishes. Rather than the proposed series of incremental changes in which each morphotype is derived from the next less specialized form, Dorit found that the most and least generalized species were sister taxa, and the intermediate species formed a separate monophyletic group.

Understanding the evolution of structurally complex phenotypic features, such as the avian wing or the vertebrate eye, has been a difficult problem for evolutionary biologists. But by examining the nested pattern of appearance of features in a phylogenetic hierarchy, it is possible to reconstruct how a complex structure has evolved through time. The recognition of intermediate states can facilitate investigation concerning the long-standing question about how such forms may have arisen (e.g., Mivart, 1871). For example, many authors have discussed under what selective circumstances the flying apparatus of modern birds evolved. However, the flight apparatus did not evolve in one step, but rather step-by-step, involving the sequential evolution of many characters over a fifty million-year period (fig. 4.2; Cracraft, 1990). Thus, one cannot simplistically talk about the evolution of flight capacity as if it were a single element. Similarly, King (1991) suggested that the cardia organ in the digestive tract of muscoid flies evolved gradually through a series of intermediate forms.

![Figure 4.2](image)

**Figure 4.2** Evolution of the flight apparatus in birds. The fifteen characters of the flight apparatus did not evolve simultaneously, but, rather, constitute derived characters of four hypothetical ancestral taxa spanning more than fifty million years. Characters 1–3: elongated forelimbs, bowed metacarpal III, sternum; 4–7: feathers, loss of scapulocoracoid fusion, rotation of the forelimb, enlarged brain; 8–11: pygostyle, strutlike coracoid, procoracoid process, shift in scapulocoracoid articulation; 12–15: keel, carpometacarpus, external condyle of the ulna developed as a semilunate ridge, and modified humerus (modified from Cracraft, 1990, with permission).

This phylogenetic approach can be expanded to simultaneously consider more than one trait. Evolutionary biologists often wonder whether the evolution of various organismal traits is related. By examining the evolution of traits in the context of a phylogeny, one can determine whether the traits have evolved independently of each other. For example, Huey and Bennett (1987; Garland et al., 1991) used phylogenetic methods to investigate whether preferred body temperature and optimal temperature for maximal sprinting speed evolved in a correlated manner among scincid lizards. As a second example, Sillén-Tullberg (1988) found that gregariousness evolved fifteen times in butterflies subsequent to the evolution of aposematism, but never before it (though several lineages which evolved gregariousness never evolved aposematism). However, demonstration of an evolutionary correlation is only the first step in an adaptive analysis. Inferences concerning a causal explanation, be it adaptive or developmental, requires mechanistic analysis in a phylogenetic context (see below).

Phylogenetic reconstructions of character evolution can also test hypotheses suggested by studies of extant taxa. For example, by analyzing functional data phylogenetically, one can test whether structural and functional evolution are necessarily coupled (Lauder, 1989, 1990; Losos, 1990; Schaefer and Lauder, 1986). Data on extant taxa, for example, might suggest that a particular structure is causally related to a particular functional capability. This hypothesis is testable using a phylogenetic hypothesis, for if structure and function are necessarily correlated, then they should evolve concordantly (e.g., Wainwright and Lauder, 1993). Conversely, phylogenetic studies might suggest that traits evolve concordantly; subsequent functional studies could look for a mechanistic basis for such a relationship.

**Adaptation and the Comparative Method.** Adaptation has been a contentious topic among evolutionary biologists for more than a century. The validity of current approaches to the study of adaptation, and in some cases the concept itself, has been questioned on a number of grounds—for example, that it is untestable, relies on plausibility criteria, fosters the atomization of organisms into component parts, and ignores nonadaptive alternatives (e.g., Gould and Lewontin, 1979; Gould, 1980; Ho and Saunders, 1984; Reif et al., 1985). These criticisms have been ably countered by Mayr (1983), Fisher (1985), and Baum and Larson (1991), among others, and will not be discussed here.

Here, we follow Gould and Vrba (1982, p. 6) and define an adaptation as "any feature that promotes fitness and was built by selection for its current role." Features that currently promote fitness (i.e., have "current utility"), but arose for some other reason (either as an adaptation for something else or for nonadaptive reasons), are termed "exaptations," and features which currently promote fitness,
but whose historical genesis is unknown, are simply called "aptations." A historical approach is obviously important for the study of adaptation. By contrast, other methods for the study of aptation, such as the optimality approach, which judges a feature by how closely it conforms to the optimal state based on functional criteria (e.g., engineering or foraging economics theory), or the microevolutionary approach, which assesses whether natural selection currently favors a trait, usually focus on current character states and ongoing evolutionary processes. In some cases these ahistorical methods can make a strong case for adaptation even in the absence of historical information (e.g., Schluter et al., 1985), but in most cases these approaches can more profitably be combined with historical analyses to shed light on the origin and evolution of a trait (see Baum and Larson, 1991, for a recent discussion).

Because the action of natural selection rarely can be directly demonstrated in the past, adaptive hypotheses must be evaluated in the same manner as hypotheses in the other historical sciences (Cracraft, 1981). That is, historical studies can unequivocally falsify adaptive hypotheses, but, because experiments are not possible, support for an adaptive hypothesis requires building a convincing case by integrating disparate lines of evidence. Several examples illustrate how historical data can negate an adaptive hypothesis. The hypothesis that coleoid cephalopods never developed hearing ability because whales would be able to stun them with blasts of intense sound was refuted by the observation that earless coleoids evolved 160 million years prior to the evolution of whales (Reid et al., 1986). Even in the absence of chronological data, phylogenetic information can negate hypotheses. In considering scenarios concerning the evolution of flight, Gauthier and Padian (1985) demonstrated that most of the "avian" characters of Archaeopteryx are shared by other coelurosaurs (the clade of dinosaurs to which birds belong) and consequently did not arise to facilitate flying, because other coelurosaurs were flightless.

Historical studies can make a convincing case for adaptive hypotheses by considering organismal analyses of function and selection in a phylogenetic context. Lauder (1981), Ridley (1983), Wanntorp (1983), and Greene (1986) were among the first to recognize that one must investigate the cause of the evolution of a trait at the phylogenetic level at which it arose. If a trait evolved as an adaptation to some problem posed by the environment, then that feature must have arisen in a lineage experiencing that selective regime. Further, a trait can confer enhanced fitness only when it leads to an increase in functional capabilities, which, in turn, affect reproductive success and survival (Alexander, 1990; Arnold, 1983; Garland and Losos, chap. 10, this volume; Huey and Stevenson, 1979). Consequently, if a trait evolved as an adaptation for a given selective regime, then the trait and its associated "performance advantage" should arise coincidentally (fig. 4.3; Coddington, 1990; Greene, 1986; Losos, 1990c; Wainwright and Lauder, in press). The historical approach to the study of adaptation thus has three components:

1. Identification of selectively important features in the environment and where, on a phylogeny, ancestral taxa experienced new selective conditions. Baum and Larson (1991) termed such features "selective regimes," which they defined as (p. 4) "the aggregate of . . . environmental and organismic factors
that combine to determine how natural selection will act upon character variation.” Ideally, the biology of the organism would be well enough understood to predict how, in a selective regime, natural selection would direct the evolution of a trait (Baum and Larson, 1991).

2. Identification of which lineages in a phylogeny have experienced evolutionary change in the trait of interest to determine if evolution of the trait was concordant with entrance into the new selective regime.

3. Identification of which lineages have experienced change in functional capability and whether change in the trait of interest is associated with this change in capabilities. Performance advantages can be deduced by biomechanical theory or modelling as well as by direct measurement (Baum and Larson, 1991).

Few studies to date have fully employed this methodology. In a study of several taxa of sunfishes, Wainwright and Lauder (1993) found that evolutionary change in mouth size and gill raker structure evolved simultaneously with differences in foraging ability and diet. Similarly, limb proportions, jumping and running ability, and habitat use evolved concordantly among West Indian Anolis lizards (Losos, 1990c). Greene (1992) presented an example in which the evolution of a performance advantage was not concordant with evolutionary change in either morphology or selective regime. The thermally sensitive pit of pit-vipers has often been suggested to be an adaptation for detecting and capturing warm-blooded prey. However, no evidence to date indicates that the pit offers a performance advantage in prey-capture: non-pit-bearing vipers can detect thermal cues as well as pit-vipers. Further, there is no obvious shift in selective regime; no derived aspect of their dietary biology distinguishes pit-vipers from their near relatives (fig. 4.4; Greene, 1992).

Adaptive hypotheses can lead to the formulation of testable predictions by combining form-function-fitness studies on extant taxa (Arnold, 1983, 1986; Huey and Bennett, 1986; see Garland and Losos, chap. 10, this volume) with phylogenetic analyses (e.g., Schluter, 1989). For example, historical analysis might indicate that a particular trait evolved via natural selection in a given selective regime. If sufficient natural variation exists for the trait or can be created experimentally (e.g., Carothers, 1986; Sinervo, 1990), and if the selective regime in which a trait presumably arose still exists naturally or can be created artificially, then one can test whether the trait actually does increase fitness, relative to its antecedent condition, in that ecological context (Baum and Larson, 1991; Mitter and Brooks, 1983). Of course, this test could lead to false rejection of the adaptation hypothesis if subsequent evolutionary change has nullified the advantage of the trait or if the ecological context as it exists or is recreated is not the same as that experienced when the trait evolved.

Two schools of thought exist on how to interpret a situation in which a trait is inferred to have evolved after a lineage enters a new selective regime (that is, trait evolution is reconstructed on a more terminal branch of the tree than is change in selective regime). Many authors (e.g., Baum and Larson, 1991; Ridley, 1983) consider that such a scenario would be consistent with a hypothesis of adaptation, but others indicate that one would expect adaptive change to occur rapidly relative to speciation. Thus, if trait evolution is adaptive, both changes (the trait and the new selective regime) should occur on the same branch of the phylogenetic tree (Björklund, 1991; Greene, 1986; Huey and Bennett, 1987; Losos, 1990c).

An adaptive scenario can be devised to explain any single evolutionary event (de Pinna and Salles, 1990). However, the morphology-performance-ecology approach to the study of adaptation also leads to the prediction that the same selective regime should lead to similar evolutionary responses in other lineages. This comparative approach to the study of adaptation has been a staple of biology for centuries (see Harvey and Pagel, 1991). As traditionally applied, the comparative method suggests that when distantly related taxa that experience the same selective regime evolve the same feature, then that feature must be an adaptation to that selective regime. Everyone is familiar with examples such as the convergent body shape of fast-moving aquatic organisms (e.g., sharks, tunas, dolphins, ichthyosaurs). Implicit in the notion of a comparative method is the idea that taxa
independently evolve the same feature, but evolutionary independence is not always as obvious as in this example. In a phylogenetic framework, however, one can explicitly test whether the evolution of a trait has repeatedly been associated with a particular selective regime, as in Sillén-Tullberg's (1988) study of the evolution of gregariousness in butterflies (discussed above). Ideally, functional studies should be included in comparative analyses of adaptation rather than assumed. In the butterfly study, for example, the link between aposmatism and unpalatability is assumed.

A potential difficulty with the comparative method is the assumption that there is a one-to-one match between selective regimes and organismal traits. To the extent that there are several adaptive "solutions" to a selective regime, or that one feature can evolve adaptively in several situations, then the comparative method may fail to identify adaptations (Bock, 1977, 1980; Bock and Miller, 1959; Harvey and Pagel, 1991). Several examples illustrate multiple adaptive solutions to a selective regime: some species of pocket gophers burrow with their forelimbs, others with their teeth (Lessa, 1989); diverse methods and structures for filter-feeding exist in taxa as different as whales, flamingos, and tadpoles (Sanderson and Wassersug, 1993). Conversely, some features can evolve in response to diverse selective regimes: for example, long legs can function to increase running ability in open habitats, as in antelopes, or jumping ability at habitat interfaces, as in frogs. Again, it is crucial to understand the functional basis for the relationship between morphology and environment (Bock and von Wahlert, 1965; Bock, 1977). For example, hawks, shrikes, and Australian shrike-tits have convergently evolved similar beaks, the first two for predation, but the shrike-tit for tearing bark off trees to capture insects. However, all three use their beaks for the same purpose, to seize and rend, even if the context is quite different (Simpson, 1978). In cases such as this, a detailed functional analysis can identify adaptation even when a less-detailed analysis fails to identify correlations between form and environment.

Quantitative traits can also be accommodated within the phylogenetic framework. If one were examining whether a quantitative trait evolved adaptively with respect to a quantitative selective regime (e.g., body size and prey size), then one would predict that evolutionary change in the two variables would be correlated. When considering quantitative characters and categorical selective regimes, one would predict that if the trait evolved adaptively, it would change more when a lineage enters a new selective regime than when the lineage remains in the same regime.

Constraints. Constraints on evolutionary change have been a popular topic for the past fifteen years. In response to the "adaptationist programme," Gould and Lewontin (1979) presented alternative explanations for evolutionary patterns. Certain traits may show a consistency of form within a clade for a variety of reasons, such as developmental canalization, intrinsic design, structural limitations, or stabilizing selection. These factors either prevent evolutionary change or channel it in particular directions. While a number of methods have been proposed to study constraints (e.g., Cheverud et al., 1985; Derrickson and Ricklefs, 1988; Stearns, 1983), explicitly phylogenetic methods allow direct examination of lineage specific effects on evolution of a trait (Carrier, 1991; Donoghue, 1989; Lauder, 1981; Lauder and Liem, 1989; Maddison, 1990; Miles and Dunham, 1992; Schluter, 1989; Wantorp et al., 1990).

Hypotheses of constraint can be tested in a manner identical to hypotheses of adaptation: the evolution of a putatively constraining feature can be plotted on a phylogeny, allowing a comparison of the monophyletic clade containing that constraint with its sister taxon that does not have the constraint. If the putative constraint has actually constrained subsequent evolution, then the two clades should differ in their subsequent evolutionary history. As with studies of adaptation, mechanistic studies should be conducted as well to investigate whether the feature actually is the constraining element rather than a correlated trait (Carrier, 1991; Funk and Brooks, 1990).

The generality of the results could be tested by identifying multiple instances of the evolution of the constraint, with comparison in each case to sister clades. Emerson (1988) provided an example of this comparative method for studying constraints. She tested the hypothesis that evolutionary reduction in the number of elements in a system, in this case the pectoral girdle of anurans, leads to a reduction in morphological diversity among descendant species (following Lauder, 1981). Fusion of the epicoracoids has occurred in eight clades. The shape of the girdle in each clade was compared to the shape exhibited by the probable sister taxon. Emerson then tested the hypothesis that the evolution of epicoracoid fusion had constrained girdle shape to evolve in certain directions. As a generality, similar changes in shape occurred in the lineages experiencing fusion, which would seem unlikely as a null expectation given the morphological diversity and rampant evolution of girdle shape displayed by frogs (for another recent example, see Janson, 1992).

The recent emphasis on constraints and on historical influences has led to the notion that history can serve as a constraint. "Phylogenetic constraint" is often invoked to explain why a particular taxon has not evolved some feature. Along these lines, a number of authors have, either implicitly or explicitly, considered adaptation and history to be alternative explanations for why a particular taxon displays a particular feature (e.g., Derrickson, 1989; Lessios, 1990). However, a dichotomy between adaptation and history is fundamentally flawed. Whereas ad-
aptation is a mechanistic process, as are developmental and other types of constraints, history is no more than a pattern. As such, history (or phylogeny) cannot constrain anything. The term “phylogenetic constraint” is shorthand for saying “the constraints that have previously evolved in and now characterize a lineage.” For example, Gould and Lewontin (1979, p. 594) “invoke phyletic constraint in explaining why no molluscs fly in air and no insects are as large as elephants.” Being a mollusc does not, in itself, preclude the evolution of flying ability. Rather, it is the developmental system and the materials of which molluscs are made that preclude the evolution of wings. Similarly, the surface area-dependent breathing mechanism of insects probably limits their maximum size to something less than elephantine. The point is, Gould and Lewontin do not really claim that phylogeny is a constraint; rather, the constraints exhibited by a particular clade, which were derived historically, are what limit evolutionary pathways.

This is more than mere semantics because adaptations, as well as constraints, may be inherited phylogenetically (Altaba, 1991; Clutton-Brock and Harvey, 1984; Fisher, 1985; Greene, 1986; Harvey and Pagel, 1991; Ridley, 1983)—historical and adaptational analyses are overlapping, rather than exclusive, perspectives. For example, enlarged and sharp canine teeth presumably arose at the base of the radiation of the order Carnivora as an adaptation for carnivory. Consequently, the presence of pointy canines in carnivores can be explained as a result of history. Nonetheless, they arose and, in most taxa, still function as adaptations—in this case adaptation and history are completely confounded.

Thus, a phylogenetic constraint simply refers to the observed pattern that a phenotypic trait does not evolve as expected within a given clade due to historical contingencies experienced by that clade. The term “phylogenetic constraint,” then, should not be considered comparable to other types of constraints. Rather, a phylogenetic constraint results when a constraint (e.g., developmental, functional) evolves and affects subsequent evolution of the descendant clade. Phylogenetic constraints are thus the historical pattern, whereas other types of constraints refer to the actual mechanistic cause. Phylogenetic constraints may thus be recognized as “different evolutionary responses among taxa from different historical backgrounds when they are subjected to the same environmental selection pressure” (Schluter, 1989, p. 82; see also Maynard Smith et al., 1985).

As such, a phylogenetic constraint is a relative concept investigated by comparing clades and seeing how they differ in evolutionary response to the same condition. Once a phylogenetic constraint is recognized, one can then determine, by plotting characters on a phylogeny, what features may be responsible for the constraint and what type of constraint it is (e.g., developmental).

Finally, it is important to distinguish between phylogenetic constraint and phylogenetic effect (Derrickson and Ricklefs, 1988; Miles and Dunham, 1992)—the latter simply indicates that taxa are similar, for whatever reason (including stabilizing selection), due to traits inherited from a common ancestor.

Community Structure. One of the major goals of ecology is the elucidation of the factors responsible for observed patterns of local diversity in biological communities. Characteristics such as species diversity and niche breadth, which species are present, how they interact, and what emergent properties characterize the community as a whole describe community structure. Ecologists attempt to explain differences in local diversity by invoking various mechanisms (e.g., competition, predation) that enhance or diminish the coexistence of species (MacArthur, 1972).

The development of ecological theory assumed that communities and their constituent species were at an equilibrium. Knowledge of contemporary species interactions were considered sufficient to describe processes responsible for maintaining diversity. The theories which were derived to explain the genesis and maintenance of diversity escaped the necessity of historical explanations by assuming that communities were saturated (Ricklefs, 1987). Consequently, the explanation for community-level phenomena is sought by examining conditions currently characterizing that community.

Species, however, are not interchangeable parts to be substituted into theoretical equations. Rather, they have specific adaptations and constraints that define how they can interact with and respond to other taxa and the environment. As with the study of adaptation and constraint, knowledge of the history of a community must be integrated with understanding of ongoing processes to understand patterns (Ricklefs, 1987; Brooks and McLennan, 1991).

Community ecology is contentious because the same pattern often can be produced by several processes, and experimental tests to distinguish alternatives are not always possible. Further, to the extent that observed patterns are the outcome of historical events, an historical perspective can be crucial in deciding between alternative possibilities. For example, in the northern Lesser Antilles, sympatric species of Anolis lizards are more dissimilar in size than would be expected by chance. This pattern could be produced either through in situ evolutionary change (e.g., character displacement) or differential colonization success (i.e., only dissimilar-sized species can successfully colonize the same island). Based on a phylogenetic analysis, Losos (1990a) argued that character displacement was responsible for the evolution of large and small taxa, but it may have occurred only once; the existence of size patterns across numerous islands consequently must be due to differential colonization success. The phylogenetic analysis thus indicates that both processes have operated in the northern Lesser Antilles—a conclusion not possible without phylogenetic information.
Ecological morphologists are often interested in comparing to what extent, if any, the structure of communities differs in diverse regions or habitats. Although in some cases differences may result purely from present-day causes, a historical perspective can often yield considerable insight. For example, carnivore communities in North America (Yellowstone) are more dispersed in morphological space than similar communities elsewhere, due primarily to the presence of bears (Van Valkenburgh, 1985, 1988). This difference in morphological diversity may exist because the environmental circumstances which have elicited the evolution of bear-like morphologies or allowed the immigration of ursids may never have existed in other communities, such as the Serengeti—differences in the use of "morphospace" would thus indicate ecological differences among sites. Alternatively, the difference in diversity may have a historical basis, resulting because the ursid lineage never colonized sub-Saharan Africa and the evolutionary flexibility of carnivore lineages present in Africa is constrained (Cadle and Greene, 1993)—perhaps bear-like morphologies are not possible in felid or viverrid lineages, for example. This latter perspective does not indicate that ecological considerations have been unimportant in the structuring of carnivore communities, but suggests that alternative lines of inquiry are appropriate. Why have bears never reached sub-Saharan Africa? What prevents other carnivore lineages from producing bear-like morphologies in the presence of bear-like ecological conditions? Similar phylogenetic effects are apparent in the differences in ecological structure of neotropical assemblages of colubrine snakes (Cadle and Greene, 1993), the abundance of nocturnal species in desert lizard assemblages around the world (Duellman and Pianka, 1990; Pianka, 1986), and the composition of helminth parasite assemblages of neotropical stingrays (Brooks and McLennan, 1991).

Theoretical analyses suggest that communities in similar environments should converge in structure. Convergence could result by the same ecological and/or morphological types occurring in each community, or it could result at the level of aggregate properties of communities, such as species richness and spacing (Blondel et al., 1984; Orians and Paine, 1983; Schluter, 1986, 1990; Strauss, 1987; Wiens, 1989). Although community convergence is clearly within the realm of historical approaches, its study will prove more difficult than studies involving single taxa and lineages because community structure results from differential colonization success, in situ evolutionary change of species within a community, and in situ production of new species. Schluter (1986) proposed a novel indirect method to study community convergence by comparing the variance between replicate communities in the same habitats on different continents to the variance between communities in different habitats within continents. To the extent that lineages are not shared between continents, replicate communities in the same habitats should be phylogenetically independent (see also Winemiller, 1991).

A more direct method of examining community structure would require phylogenetic information for each lineage in the communities, a forbidding task for all but the smallest communities. Even with such information, it will not be easy to relate to the evolution of aggregate properties of communities, such as niche packing, and few such approaches have appeared to date: for example, Nee et al. (1991) have shown how phylogenetic considerations can shed light on the relationship between abundance and body size in British birds (for another promising approach, see Winemiller, 1991). Perhaps by randomizing both evolutionary change within lineages and species co-occurrence, one could assess whether communities have converged relative to random expectations.

Brooks and McLennan (1991) and Gorman (1993) have devised an alternative approach, focusing more on species-level interactions, that incorporates aspects of vicariance biogeography and adaptational studies to partition the relative importance to community structure of colonization, interspecific interactions, and the interaction between the two. In contrast to Schluter's method, this approach requires that the members of the same lineages occur in the different communities so that changes in associations and traits can be compared and interpreted phylogenetically. In a particularly interesting example, Gorman (1993) found that interspecific interactions among stream fishes could be predicted with historical data—taxa that had a long history of association (as revealed by a combination of phylogenetic and biogeographical data) competed less strongly than species whose association was more recent.

The evolution of community (or assemblage) structure is simpler to comprehend when one is dealing with a monophyletic group, as is often the case on islands. The use of monophyletic groups also facilitates understanding of the sequence of addition and modification that occurs as a community assembles. For example, Williams (1972) used a phylogeny for the Anolis of Puerto Rico, which have diversified morphologically and ecologically, to formulate a hypothesis about the causal factors underlying the radiation (fig. 4.5). Based on the pattern of character evolution, he suggested that interspecific competition first leads to divergence in size followed by habitat partitioning. Subsequently, a phylogeny for Jamaican Anolis allowed factors governing the parallel radiations on the two islands to be compared (Losos, 1992; fig. 4.6). A comparison of the reconstructions in figures 4.5 and 4.6 reveals that the anole communities on the two islands are quite similar and that the sequence of ecomorph types (sensu Williams, 1972) evolved in identical order. For instance, the ecomorph type absent on Jamaica, the grass-bush anole, is the last to evolve on Puerto Rico, which offers a proximate explanation for its absence on the former island (Losos, 1992).
Figure 4.5 Diversification of anole fauna on Puerto Rico. Sequence in the evolution of ecomorph types (a, b, c, d).

Figure 4.6 Diversification of anole fauna on Jamaica. Sequence in the evolution of ecomorph types (a, b, c).

Methods for Incorporating Phylogenetic Information into Comparative Studies of Ecological Morphology

In this section, we first discuss problems posed by the phylogeny itself, and then detail the methodology, assumptions, and weaknesses of several of the most widely used methods for incorporating phylogenetic information into comparative analyses.
The Underlying Phylogenetic Hypothesis

Any historical analysis is only accurate if the phylogeny used is correct. Consequently, it is important to realize that phylogenies are hypotheses about patterns of evolutionary descent. Rarely, if ever, will it be possible to demonstrate that a phylogeny is correct. Rather, one can only assess to what extent the evidence corroborates a particular phylogenetic hypothesis. This is an important perspective, because historical analyses such as those discussed above tend to be based on a preferred phylogenetic hypothesis. However, in many systematic analyses, the preferred (or “best”) phylogeny is only marginally better than competing hypotheses. Many cladistic analyses of morphological data, for example, simply choose as the preferred hypothesis the tree which requires the fewest evolutionary transitions in character states. The advent of statistical criteria for choosing between hypotheses will prove helpful in delimiting which trees to consider (e.g., DeBry, 1992; Cunningham and Buss, 1993), but these tests are only valid to the extent that their underlying assumptions are met (e.g., that parsimony criteria are appropriate or that transition state probabilities in maximum likelihood methods are accurate); these assumptions have generally proven difficult to verify.

Thus, future workers should not accept whatever phylogeny they can find or patch together, but instead should evaluate the evidence and include all appropriate phylogenetic hypotheses (Swofford, 1991). Eventually, it will be desirable to develop methods to weight the results of analyses using different phylogenies by the perceived likelihood that each phylogeny is correct. In a less thorough fashion, one could conduct an analysis using the preferred phylogeny, but perform sensitivity analyses to see how vulnerable the results are to changes in the phylogenetic structure (e.g., Richman and Price, 1992).

Recently, it has become common for comparative analyses to acknowledge the importance of phylogenetic considerations, but, because no phylogeny is available for the group in question, to conduct analyses ignoring phylogeny, or to use an available taxonomy. It is clear that ahistorical analyses have a high probability of being inaccurate (Felsenstein, 1985; Grafen, 1989; Martins and Garland, 1991); using taxonomies is perhaps less risky, but many available taxonomies probably poorly reflect phylogeny. An alternative approach (Losos, 1994) would be to conduct the analysis on many different phylogenies and to see if the results differ. Such an analysis might indicate, for example, that a significant relationship exists regardless of which phylogeny is used or that a significant result would obtain only when one used a phylogeny that seemed highly improbable. Rather than ignoring phylogeny, this approach would be more positive, laying a predictive groundwork for future systematic work.

Even when a comparative study focuses only on the preferred phylogenetic hypothesis, the analysis may be hampered by the inability to resolve the tree completely into a network of bifurcations. Polytomies (i.e., ancestral nodes, each producing more than two descendant lineages) may actually represent the history of diversification; population fragmentation could give rise to more than two descendant species. In many cases, however, polytomies reflect either lack of data or disagreement among data. Polytomies also will exist when taxonomies are recast as phylogenetic trees if any rank (e.g., genus) contains more than two taxa (such attempts also make the dangerous assumption that each taxon is monophyletic). Grafen (1989), Maddison (1989), and Harvey and Pagel (1991), among others, have developed methods for interpreting the evolution of a character when faced with a polytomy, but these methods involve assumptions that the state of the character can provide information useful in resolving uncertain phylogenetic relationships; Harvey and Pagel’s (1991; see also Pagel, 1992) method, for instance, assumes that phenotypically more similar species are more closely related. Conducting separate analyses using each possible resolution of the polytomy would be less biased (Maddison, 1989). Similarly, when faced with multiple, equally supported phylogenetic hypotheses the most conservative course would be to conduct the analysis using each one (Harvey and Pagel, 1991; Harvey and Purvis, 1991). Rather than attempting to circumvent uncertainty in our knowledge with assumptions whose accuracy may be questionable, this approach would again put matters in a positive light, indicating to what extent different phylogenetic resolutions alter the outcome of historical analyses and encouraging further studies (Losos, 1994).

Variance Apportioning Methods

In some cases, investigators may wish to determine the correlation of a trait with phylogeny, which would provide an index of a trait’s evolutionary lability. A number of quite distinct methods have been developed that partition variation between recently evolved versus inherited values (we will not discuss hierarchical taxonomic methods, which are not strictly phylogenetic and suffer from a number of problems; Felsenstein, 1988a; Harvey and Pagel, 1991; Losos, 1990c; Maddison and Maddison, 1992; Miles and Dunham, 1992). These techniques determine how much of the variation in a trait is attributable to phylogenetic inheritance and how much is due to evolution since divergence from a common ancestor.

Phylogenetic Autocorrelation Analysis. The elucidation of a phylogenetic constraint first requires information about the covariation between a character and phylogeny (Miles and Dunham, 1992). Indeed, if phenotypic values cannot
be predicted as a function of phylogenetic relationships, then incorporating phylogenetic information into statistical analyses may be unnecessary (Gittleman and Luh, in press). Consequently, investigating whether variation in a trait is related to phylogenetic relationships of the group in question may be important. The difficulty with this approach is determining what one means by a correlation between phylogeny and trait variation.

Using an autocorrelation model, Cheverud et al. (1985) showed that correlation between the phenotype and phylogeny may be described by an autocorrelation coefficient which is the product of a phylogenetic connectivity matrix and a vector of standardized trait values. The phylogenetic connectivity matrix represents the pairwise phylogenetic similarity (or relatedness) of all the extant taxa included in the analysis. For each species, the value of a trait can be partitioned into two components: the phylogenetic value is estimated by the product of the autocorrelation and the matrix of phylogenetic weights and describes the degree of invariance of the trait with tree topology (i.e., the phylogenetic effect); the specific value is estimated by the variation that is independent of phylogenetic effects (Cheverud et al., 1985; Gittleman and Kot, 1990; Miles and Dunham, 1992; see fig. 4.7 for a hypothetical example). In other words, the expected value for each species is predicted by summing the phenotypic values for all other species, with the value for each of the other species weighted by some predetermined amount. The specific value, then, represents the difference between the observed and expected values for a species.

The weighting scheme is thus of crucial importance in determining the specific value. Weights are determined by the phylogenetic connectivity matrix, which in the past has been based on taxonomic rank (e.g., congers weighted more highly than conflationary species, which are weighted more highly than members of the same order, and so on [Cheverud et al., 1985; Gittleman and Kot, 1990]). Alternatively, cladistic methods can also be used to determine the weighting scheme (Miles and Dunham, 1992), either by using the patristic distance (i.e., the number of nodes separating two taxa) or by other cladistically based distances. In an extension of the autocorrelation approach, Gittleman and Kot (1990) included a “variable weighting index,” α, which was estimated by maximum likelihood procedures. The rationale for adding this exponent to the weighting matrix was the finding that the autocorrelation coefficient did not decay in monotonic fashion with taxonomic distance; by including α, pairwise distances among taxa are adjusted in the connectivity matrix to remove phylogenetic correlations at all hierarchical levels (see example in Gittleman and Kot, 1990).

Thus, the autocorrelation statistic is simply a measure of how well species' values can be predicted based on the phylogeny. Several authors have used the specific values to test adaptational hypotheses, arguing that the specific values, being free of phylogenetic influences, represent change due to environmental conditions. These analyses look for a correlation between specific values for two characters, such as clutch size and age of maturity (Miles and Dunham, 1992). Obviously, the key to this method revolves around the accuracy of the phylogenetic matrix and how successfully the values for other taxa can be used to predict a species' value. In addition, it remains to be determined what exactly the specific value measures in evolutionary terms. To some extent, the specific value can be seen as an amalgam of the evolutionary history of a taxon, with recent divergent events weighted more highly than more ancient events. The question, then, is whether a particular specific value could result from several, quite distinct, historical scenarios. If, for example, a small specific value could result either from slight change from a sister taxon or from great change during a more distant cladogenetic event, then correlations between specific values for two variables might be suspect. In simulations based on taxonomic hierarchies, however, Gittleman and Luh (1992) found that autocorrelational studies of this sort work reasonably well, at least under certain circumstances.

Two cautionary comments need to be made concerning autocorrelation methods. First, Cheverud et al. (1985) argued that correlations based on the specific values may be used in assessing adaptive change independent of change experienced in other lineages. Because many factors may influence this variance component, such as genetic drift, pleiotropic effects, and sampling error, Lynch (1991) suggested caution in interpretation of such correlations and suggested that patterns of variation attributed to phylogeny may be viewed as estimates of mac-

![Figure 4.7 Sample autocorrelation values obtained using the method of Cheverud et al. (1985). a–d, different distributions of trait values. The autocorrelation value (A.C.) ranges from 1.0 (trait values completely predicted by phylogenetic topology) to 0 (no relationship) to -1.0 (closely related taxa dissimilar in trait value). This example, the distance between any two taxa was assigned based on the number of speciation events since their common ancestor, with the provision that all members of a clade are equally distant from taxa outside that clade.](image-url)
roevolutionary effects (see the discussion of Lynch's mixed model below) on a trait, whereas specific values represent patterns of variation due to microevolu-

tionary events.

Second, partitioning variation into a phylogenetic and a specific component is not equivalent to partitioning into an adaptive and a nonadaptive component. As discussed above, adaptations can contribute to phylogenetic effects and phy-
logenetic effects can result from the action of stabilizing selection.

**Lynch's Mixed Model.** Lynch (1991), drawing a parallel between micro-
evolutary studies of selection and heritability and macroevolutionary analyses of adaptation and radiation, suggested extending quantitative genetics methodology to macroevolutionary questions. He suggested that the use of mixed model methods developed in animal breeding may be appropriate for partitioning the variation in a trait. Given $n$ taxa for which $k$ characters have been measured, the model for simultaneously analyzing variation in all traits is given by

$$ z_{ij} = u_i + a_{ij} + e_{ij} $$

where $u$ is the grand mean of the $i$th character for the whole phylogeny, $a$ is the “heritable additive component” of the character for the $j$th taxon (similar to breeding values used in quantitative genetics), and $e$ is the residual from the predicted value ($u + a$). Although similar in spirit to the autocorrelation method of Cheverud et al. (1985), this approach differs in that not all the traits must be included in the analysis to examine patterns of evolutionary change.

The inclusion of phylogenetic effects is accomplished by a matrix of phy-
logenetic relationships, $G$, derived from a phylogenetic tree (including branch lengths). In Lynch’s (1991) example, phylogenetic relationships were estimated by the proportion of evolutionary time, from the root of the tree, shared by any two species. However, as with all methods, other information that captures the phylogenetic relationships among taxa could be used.

The parameters in the mixed model must be estimated directly from data from extant species using an iterative expectation-maximization algorithm. This algorithm produces estimates of the character means (over the whole phylogeny), the additive and residual values (analogous to the phylogenetic and specific values of Cheverud et al., 1985), and the variance-covariance matrices of the additive and residual values.

The variance-covariance matrix for the additive values allows the estimation of phylogenetic heritabilities as well as their associated standard errors, tests of hypotheses regarding the existence of phylogenetic heritability, and the prediction of phenotypes of ancestral species. Unlike other methods for inferring ancestral states, residual deviations (due, for example, to environmental causes or measurement error) do not contribute to the predicted mean of ancestral pheno-
types. Standard regression and correlation procedures on the residual values may be used to describe adaptive patterns among traits. In addition, patterns deduced from the variance-covariance matrix of additive effects may represent macroevolutionary or phylogenetic phenomena, whereas those based on the residual variance-covariance matrix may represent variation specific to taxa since divergence from a common ancestor. To date, this method has been used only on one limited data set (Lynch, 1991).

**Ancestral State Reconstruction and Directional Methods**

A number of methods have been proposed to reconstruct the evolution of a character, given a phylogeny. It is important to recognize that all methods make implicit assumptions and depend on implicit or explicit evolutionary models (Felsenstein, 1988a; Harvey and Pagel, 1991; Harvey and Purvis, 1991; Maddison and Slatkin, 1991; Martins and Garland, 1991; Maddison and Maddison, 1992). Indeed, this is true of phylogenetic statistical methods in general (Harvey and Pagel, 1991; Pagel and Harvey, 1992; Martins, 1993). Qualitative (or categorical) characters have generally been reconstructed using parsimony methods that minimize the number of evolutionary transitions required throughout the tree (Farris, 1970; Swofford and Maddison, 1987). A number of variants exist, de-
pending on particular assumptions (e.g., ordered versus unordered character states, irreversibility of evolution; see Felsenstein, 1983; Maddison and Mad-
dison, 1992). In general, parsimony methods will produce a maximum likelihood estimation of evolutionary pattern when rates of character change have been low relative to cladogenesis or not excessively unequal among lineages (Felsenstein, 1983). To the extent that probability of character change differs on different branches of tree (e.g., due to differences in branch lengths in units of time), then parsimony methods are likely to overestimate the amount of change on some branches and underestimate it on others. Few studies have assessed the degree of error thus introduced: Maddison (1990) used simulations to demonstrate that parsimony did not generally introduce large or systematic errors in his method (discussed below); Hillis et al. (1992) found that parsimony methods were more than 98% accurate in reconstructing ancestral phenotypes in an experimentally derived clade of bacteriophage T7.

A number of methods have been introduced to reconstruct hypothetical ances-
tral states for quantitative characters. Linear parsimony methods, mathematically essentially identical to those for qualitative traits, minimize the change in trait value summed over all branches (Swofford and Maddison, 1987; Losos, 1990c; Maddison and Maddison, 1992). The squared-change parsimony method (Huey and Bennett, 1987; Maddison, 1991) minimizes the square of the amount
of change on each branch summed over all branches. In contrast to conventional parsimony approaches, whose philosophy is to minimize the absolute amount of evolutionary change required, the purpose of minimizing the square of changes is not obvious (Losos, 1990c). The weighted squared-change parsimony method minimizes the square of the amount of change divided by the inverse of branch length (in units of expected variance) summed over all branches (Martins and Garland, 1991). The reconstruction of ancestral character states produced by this method has maximum posterior probability under a Brownian motion model of evolution (Maddison, 1991).

Given the differences in their underlying models of evolutionary change, the methods for reconstructing continuous traits can produce quite different estimates of values for ancestral taxa (Losos, 1990c). In general, linear parsimony tends to concentrate evolutionary changes on relatively few branches of the tree and minimize parallelism and reversal. The other methods spread evolutionary change more homogeneously among branches and exhibit enhanced parallelism and reversal. In addition, linear parsimony methods often yield ambiguities in character reconstruction (e.g., whether a character is interpreted as evolving in parallel in two lineages, or evolving once and subsequently being lost), whereas other methods yield a single estimate for each ancestral taxon. Although the latter methods are thus analytically more tractable, adherents of parsimony would argue that these methods obscure inherent uncertainties in the data (Losos, 1990c). Because there will usually be no way to determine the appropriate model for character evolution, it is important to recognize that choice of methodology will affect the character reconstruction (Harvey and Pagel, 1991; Harvey and Purvis, 1991). To date, little work has been conducted in assessing the extent of error that results when the assumptions of a model are incorrect, though Martins and Garland’s (1991) simulation study suggested that in many cases the effect may be surprisingly small.

Once ancestral character states have been inferred, “directional” comparative methods (sensu Harvey and Pagel, 1991) can be used to investigate whether the evolution of two characters (both organismal or one environmental) is related by comparing the reconstruction of their ancestral states. One of two hypotheses is tested: that the characters evolved simultaneously (in continuous characters, that change in the two is correlated among branches of the tree), or that evolution in one variable is concentrated in clades exhibiting a particular state of the second character. Examples of these approaches have been discussed in the first section of this paper.

Appropriate null models for these tests are more complicated than they initially appear. Maddison (1990) developed a null model to test whether changes in a trait are distributed randomly throughout a tree or are clustered on branches with a particular state of a second trait. For example, although most instances of the evolution of gregariousness in butterflies occurred in aposematic lineages (Sillén-Tullberg, 1988; see above), this might not be surprising because most lineages are aposematic. Indeed, Maddison (1990) found no suggestion that gregariousness had evolved in aposematic lineages more often than expected by chance (for another example, see Donoghue, 1989).

A further complication is that some reconstruction methods do not consider differences in branch lengths. To the extent that the probability of change along a branch is proportional to time since divergence, one would expect changes to be concentrated on longer branches by chance alone, which could potentially bias Maddison’s (1990) test (Harvey and Pagel, 1991; Maddison, 1990). Harvey and Pagel (1991) have proposed a maximum-likelihood method to estimate the transitional probabilities of change in a categorical character along all branches of a tree. The efficacy of this method is uncertain, however, because the maximum-likelihood estimates are derived from reconstructions of ancestral traits based on parsimony and ignoring branch lengths. Additional problems may arise because parsimony methods usually reconstruct many branches with no observed change, yet maximum-likelihood estimates lead to nonzero expectations for change on all branches. Consequently, if many branches are reconstructed with no change for both characters, standardized scores for the two variables on these branches will be highly correlated (Maddison and Maddison, 1992). Consequently, Harvey and Pagel (1991) suggest considering only branches on which change is reconstructed to have occurred, which necessarily omits some of the available information.

Most analyses that examine the evolution of categorical characters use parsimony methods, but often the existence of alternative, equally parsimonious reconstructions is not mentioned. Such ambiguities can make statistical analysis difficult, but ignoring the uncertainty is not the answer, because it may lead to incorrect conclusions. For example, Altaba (1991) found that resolution of ambiguities affected outcome of an analysis concerning whether evolutionary change in chemical warfare in beetles is related to habitat (see also Donoghue, 1989). Maddison (1990) recommended that when ambiguities in reconstructions exist, one should examine the reconstructions most and least favorable to the hypothesis to suggest how sensitive the analysis is to different reconstructions. Alternatively, one could check all possible combinations of reconstructions for both characters, or a random subset of that universe. Perhaps, if most (95%) reconstructions favored a hypothesis, one could accept it as well supported.

Further, studies should not be limited to the most parsimonious reconstruction of character evolution. As with the decision concerning which phylogenies to consider, one should also conduct sensitivity analyses which consider slightly
less parsimonious character reconstructions to see how they affect evolutionary interpretation (Maddison and Maddison, 1992). Indeed, one underappreciated difficulty with reconstructing ancestral states, regardless of which method is used, is that the reconstructions of the phylogeny and of the ancestral states are conducted sequentially rather than simultaneously, but there is no guarantee that this method produces the overall most parsimonious solution for evolution of the trait in question. For example, trees slightly less parsimonious than the best phylogeny may require considerably fewer changes in the trait in question. An analogous problem occurs in phylogeny reconstruction based on DNA data, when sequence alignment and phylogenetic analysis are conducted sequentially (Felsenstein, 1988b, and references therein).

**Independent Contrast Methods**

Felsenstein (1985) proposed an alternative to directional methods for studies of the evolution of continuous characters which does not require the inference of ancestral traits. By reference to a phylogeny, one can construct judiciously chosen comparisons that resolve the statistical and evolutionary dependence of species as data points. This method uses the information on tree topology and branch lengths to construct "contrasts," that is, a set of sequential comparisons between pairs of taxa (or nodes). Here, "contrasts" are defined as the difference of a trait between one extant species (or node) and its sister species (or node)—evolutionary change in each contrast is independent of change occurring elsewhere in the phylogeny. One continues down the tree, making comparisons from the tips to the root. For a sample of \( N \) taxa there will be \( N - 1 \) possible comparisons (fig. 4.8). Statistical analyses can then be conducted on contrasts generated for two or more characters (see Garland et al., 1992).

Branch lengths, in units of expected variance of evolutionary change, are important to the contrast method in two ways. First, contrasts are based on the difference in character value between sister taxa. When one (or both) of the taxa are internal nodes of the tree, these nodes must be assigned values (these values are not equivalent to reconstruction of ancestral states). The assigned value is an average of the values of its two descendants, inversely weighted by branch length; further, internal branch lengths are lengthened by the addition of an error term, which is itself a function of branch lengths (Felsenstein, 1985; and Garland et al., 1992, provide details for calculations). Second, to meet the assumptions of parametric statistics, contrasts must be standardized by their expected variance. If one assumes that character change can be modeled as a Brownian motion process, then the cumulative variance will be a function of time since divergence. Therefore, previous workers have recommended that each contrast be scaled to a common expected variance through division by the standard deviation. Of course, the Brownian motion model of evolutionary change was originally chosen more for its analytical tractability than for its biological realism (Felsenstein, 1985, 1988a; Harvey and Purvis, 1991). Other models are possible, the most obvious being punctuated equilibrium, in which expected variance on all branches would be equal (Felsenstein, 1988a; Harvey and Pagel, 1991; Martins and Garland, 1991). There are many sources that could provide branch lengths (e.g., genetic distance values or the number of systematic character changes inferred to have occurred along a branch; Garland, 1992; Garland et al., 1992); one only has to be willing to assume that expected variance in trait evolution should be proportional to branch length (one could also propose models in which the relationship between branch length and expected variance varied temporally). Harvey and Pagel (1991) and Garland et al. (1992) provide methods for verifying that contrasts have been adequately standardized.

The incorporation of branch length information is both the principal strength and weakness of the independent contrasts method. To the extent that ancestral nodes are correctly estimated, contrasts reflect nonoverlapping episodes of evolutionary divergence and are statistically independent, thus avoiding the problem of nonindependence inherent in ancestor reconstruction approaches (Felsenstein, 1985; Harvey and Pagel, 1991; Garland et al., 1992). But to the extent that the presumed model of change (e.g., Brownian motion) is inaccurate, then the contrasts lose their claim to statistical independence (Felsenstein, 1985, 1988a). The methods of Harvey and Pagel (1991) and Garland et al. (1992) can detect instances in which the underlying model for contrast standardization is inappropriate, but these methods probably will have low statistical power in many applications, particularly when the true evolutionary model is more complicated.

![Figure 4.8](image_url)
and nonlinear than the standardization model. This is currently an area of active research, with new developments appearing rapidly. Contrasts also may be non-independent because of correlated error (Harvey and Pagel, 1991). Because internal node values are calculated from the values of their descendants, an error in measurement of a character value for an extant taxon will affect all contrasts involving nodes ancestral to that taxon.

The need to standardize contrasts also imposes a cost: standardized contrasts no longer measure absolute change, but rather change per unit standardization (Garland, 1992; Losos, 1990b). This difference, though subtle, might be important in many contexts. For example, the exponent by which metabolic rate scales with body size has long been of interest to physiologists (reviewed in Harvey and Pagel, 1991). In a phylogenetic context, the question is: How does evolutionary change in metabolic rate scale relative to evolutionary change in body size? But when contrasts are scaled by branch length in units of time, the contrasts measure rate of change in metabolic rate relative to body size. It is not obvious that these two values should be the same. Matters become even more complicated when it becomes necessary to standardize by some quantity other than branch lengths in units of time. For example, the number of systematic characters inferred to change on a branch might be used for standardization (Garland, 1992; Garland et al., 1992). But if one is comparing rates of evolution in limb length between two clades (e.g., Garland, 1992), what would it mean that rates of change per unit cladistic character differ between clades? Certainly, this is not the same as saying that rates per unit time differ.

One solution to these difficulties in using independent contrasts, and a means of obviating the need for information on branch lengths, is to ignore branch lengths and not standardize the contrasts. By adopting this course, however, one cannot make comparisons involving internal nodes and is limited to comparisons among pairs of extant species, with no species used more than once and no species pair sharing a branch of the phylogeny (fig. 4.8; Felsenstein, 1985; Burt, 1989; for examples, see Voss, 1988; Müller and Birkhead, 1992). Thus, the numbers of degrees of freedom is cut in half, which may pose a problem for many studies. Further, unless one is willing to assume that the expected amount of change is equal in all comparisons (as would be the case if one used sister species and assumed that evolution was punctuational), then only a sign-test (not ranked) is valid. Although one might expect a sign-test to have low statistical power, reanalysis of two studies that used paired comparisons but failed to use the sign-test (Losos and Miles, unpubl.) indicates only a minor change in significance levels when the sign-test is employed (from P = .026 to .049 in Oakes, 1992, and from P = .017 to .035 in Balmford et al., 1993). The advantages of using such paired-species contrasts that ignore branch lengths and internal nodes is that they are unquestionably independent statistically, avoid the potential difficulties discussed above, and have the added benefit that information on branch lengths is unnecessary.

Most phylogenies will contain unresolved nodes, that is, polytomies, which prevent the calculation of independent comparisons. One method for resolving ambiguous nodes and constructing comparisons involves the calculation of contrast coefficients commonly used in ANOVA models (Harvey and Pagel, 1991; Pagel and Harvey, 1988). This entails the calculation of a weighted difference score that derives a single score from three or more taxa.

Grafen (1989) proposed an alternative method. The phylogenetic regression is a three-stage analysis for deriving contrasts from unresolved phylogenies. The first phase involves calculating an average value for each node. The average of each node is then expressed as a deviation from its parent node; this procedure results in a "long data set." Second, the values in the long data set are adjusted to account for differences in branch lengths. Grafen (1989, 1992) estimates a branch "height," which is described as one less than the number of taxa below a given node. Then, branch lengths are calculated as the difference between the heights of successive nodes. Grafen assumes that the expected amount of change will be proportional to the length of the branch (i.e., "Brownian motion"). The third step, after adjusting the nodal values by the branch lengths, is to construct a set of linear contrast vectors to remove (or "annihilate") the influence of phylogeny at the nodes with unresolved polytomies. In the phylogenetic regression, any number of independent variables may be annihilated (Grafen, 1992). Strengths and weaknesses of the phylogenetic regression have been discussed by Grafen (1992) and by Pagel and Harvey (1992).

Of course, as an alternative to these methods for dealing with polytomies, one could conduct analyses on all possible resolutions of the polytomy, as discussed above.

Phylogenetically independent contrasts have great potential in ecological morphology but have been used sparingly to date. Although we have elaborated several concerns, it is not clear how serious the biases they introduce may be, and all of these methods still are clearly superior to disregarding phylogenetic information entirely (Felsenstein, 1985), as Martins and Garland’s (1991) simulation study clearly revealed.

Simulations

Randomization and phylogenetic simulation methods are other promising avenues to get around the statistical non-independence problem (Harvey and Pagel, 1991; Martins and Garland, 1991). Martins and Garland (1991) proposed another method for testing the significance of correlated evolution in two continuous
characters given a phylogeny. They simulated the independent evolutionary change in each character through a phylogeny using either a punctuational or Brownian motion model of evolutionary change—the latter leads to greater expected change on longer branches. The correlations of change in two characters from 1000 simulations were then used as a null hypothesis in comparison to the observed correlation between two characters of interest. Garland et al. (1993) subsequently expanded this simulation method to consider questions normally investigated using ANOVA or ANCOVA approaches, such as whether home range size in carnivorous mammals is larger than in herbivorous mammals. In these simulations, the identity of the species (e.g., either ungulate or carnivore) remains the same, but character values for each species are assigned by simulating evolution of the character using the phylogeny for the group. A number of additional modes of character evolution were used in the simulations in addition to the two employed by Martins and Garland (1991). The surprising result of the Garland et al. (1993) study was that differences as extreme as those actually observed between ungulates and carnivores occur relatively often in the simulations.

Comparisons among Methods

Despite the development of numerous analytical procedures whose purpose is to incorporate phylogenetic information, few studies have compared the efficacy of these techniques (e.g., Grafen, 1989; Martins and Garland, 1991; Martins, 1993; Garland et al., 1993). Each of the statistical methods available for assessing the influence of phylogeny in ecomorphological analyses makes specific assumptions or requires varying details about the topology of the phylogenetic tree. For example, the independent contrasts method probably is more statistically justifiable than parsimony-based directional methods (but see concerns discussed above) but also requires more detailed phylogenetic information. Perhaps the best course of action would be to employ several methods; to the extent that they all produce similar results, one might be confident in reporting the outcome.

The applicability of the various methods also depends on the type of question involved. Some procedures compare ancestral and descendant taxa to estimate the direction and magnitude of character change, whereas nondirectional comparisons estimate the phylogenetic concordance between two traits, but do not provide information on direction. Other methods estimate how well knowledge of phylogenetic relationships predicts character values, that is, how frequently evolutionary change has occurred. Researchers should take care to choose the appropriate method for the question at hand.

A trend to date has been the unquestioned use of a particular technique with-
in many (perhaps most?) cases, the key to understanding present character states will be an understanding of the sequence, timing, and cause of evolutionary events in the (sometimes distant) past. Consequently, explicitly directing investigations to this historical component is important.

The phylogenetic approach is also desirable because it can suggest research questions that otherwise might not have been obvious. For example, phylogenetic analyses might indicate that a structure and function evolve concordantly, a hypothesis that experimental or theoretical functional studies can directly test (Lauder, 1989, 1990), and can even suggest the importance of previously overlooked structures. For example, Lauder (1989) demonstrated that the hypochordal longitudinalis muscle arose prior to the evolution of an externally symmetrical tail in ray-finned fishes. Subsequent laboratory studies revealed that the presence of this muscle considerably alters the functional capabilities of the tail and may have played an important role in the evolution of external tail morphology (see also Futuyma and McCaffery, 1990; McLennan, 1991).

The new comparative methods hold great promise for application in ecomorphological studies. Although many studies recognize the importance of including phylogeny in ecomorphological analyses (e.g., James, 1982; Strauss, 1987; Losos, 1990c; Winemiller, 1991; Douglas and Matthews, 1992; Richman and Price, 1992), the use of these techniques is sparse. For the effects of phylogeny to be adequately recovered, future studies must include explicit phylogenetic hypotheses as a central analytical framework (Donoghue, 1989). On a positive note, reference to recent issues of most major journals indicates that the importance of phylogeny is being increasingly appreciated. With rapid developments in the field of comparative biology and its enhanced use, the next few years should prove exciting for systematists and the full spectrum of organismal biologists.

Acknowledgments

We thank T. Garland, H. Greene, A. Larson, E. Martins, T. Price, S. Reilly, D. Stern, P. Wainwright, and two anonymous reviewers for helpful comments; B. Dial, T. Garland, J. Gittleman, O. Gorman, H. Greene, M. Lynch, W. Maddison, T. Price, L. Sanderson, and D. Wake for providing papers prior to publication; and the National Science Foundation (grant BSR 8617688 to DBM) for support. JBL was supported by a postdoctoral fellowship from the Center for Population Biology, University of California, Davis, while this paper was being written.

References


Evaluating the Adaptive Role of Morphological Plasticity

Joseph Travis

INTRODUCTION

The size and shape of morphological characters are rarely constant within a species. Although some characters typically exhibit low levels of phenotypic variation, others vary extensively among individuals from different populations or among individuals from different generations of the same population. Much of this variation is attributable to differences among such groups in the state of some environmental variable experienced during sensitive periods in development. A simple view of such environmental effects is that they represent the necessary and predictable results of biochemical processes that are carried out under different physical and chemical conditions.

In this chapter, I explore another interpretation of these environmental effects, that in many cases they represent an interaction between the genetic system and the environment that has been designed by natural selection to produce different average phenotypes under different conditions. The capacity of a single genotype to produce a variety of phenotypes is called phenotypic plasticity, and the thesis I explore is that some patterns of plasticity in morphological traits are adaptive ones.

This thesis is an attractive one from an empirical perspective and can be developed from two sets of empirical observations. First, the average values of morphological traits that are seen in a species or a population do appear to have been molded by natural selection, and different ecological demands appear to have produced different sizes and shapes in the same morphological traits. Several lines of argument lead to this conclusion: the matching of physical principles of design considerations with the observed organism performance (Denny, chap. 8, this volume); the observations and experiments on the ecological roles of morphological features (Wainwright, chap. 3, this volume); and comparisons and correlations across a range of morphological variation and ecological milieux (this volume: Ricklefs and Miles, chap. 2; Emerson et al., chap. 6; Norberg,