Phylogenetic Analyses: Comparing Species to Infer Adaptations and Physiological Mechanisms

Article *in* Comprehensive Physiology · January 2012

DOI: 10.1002/cphy.c100079 · Source: PubMed

2 authors:

Enrico L Rezende
Universidad Andrés Bello
107 PUBLICATIONS **3,283** CITATIONS
SEE PROFILE

Jose Alexandre Felizola Diniz-Filho
Universidade Federal de Goiás
388 PUBLICATIONS **12,450** CITATIONS
SEE PROFILE

Phylogenetic Analyses: Comparing Species to Infer Adaptations and Physiological Mechanisms

Enrico L. Rezende[*1] *and José Alexandre F. Diniz-Filho*[2]

ABSTRACT

Comparisons among species have been a standard tool in animal physiology to understand how organisms function and adapt to their surrounding environment. During the last two decades, conceptual and methodological advances from different fields, including evolutionary biology and systematics, have revolutionized the way comparative analyses are performed, resulting in the advent of modern phylogenetic statistical methods. This development stems from the realization that conventional analytical methods assume that observations are statistically independent, which is not the case for comparative data because species often resemble each other due to shared ancestry. By taking evolutionary history explicitly into consideration, phylogenetic statistical methods can account for the confounding effects of shared ancestry in interspecific comparisons, improving the reliability of standard approaches such as regressions or correlations in comparative analyses. Importantly, these methods have also enabled researchers to address entirely new evolutionary questions, such as the historical sequence of events that resulted in current patterns of form and function, which can only be studied with a phylogenetic perspective. Here, we provide an overview of phylogenetic approaches and their importance for studying the evolution of physiological processes and mechanisms. We discuss the conceptual framework underlying these methods, and explain when and how phylogenetic information should be employed. We then outline the difficulties and limitations inherent to comparative approaches and discuss potential problems researchers may encounter when designing a comparative study. These issues are illustrated with examples from the literature in which the incorporation of phylogenetic information has been useful, or even crucial, for inferences on how species evolve and adapt to their surrounding environment. © 2012 American Physiological Society. *Compr Physiol* 2:639-674, 2012.

Introduction

Explaining the diversity of form and function of living organisms remains the ultimate goal of evolutionary biology, and the comparative method one of its most powerful research tools. Comparisons among organisms, focusing on their similarities and differences, are fundamental to systematization and classification: they enable researchers to contrast what is unknown against well-established information, to identify regularities in seemingly different biological processes and mechanisms, to determine which attributes are unique to some groups and ultimately to understand how this biological complexity has evolved. Patterns of phenotypic similarity among species led Darwin to propose his evolutionary theory of descent with modification, laying the conceptual foundation for fields such as systematics, paleontology, or biogeography. In animal physiology, central concepts such as Krogh's principle and the choice of appropriate animal models for biomedical studies are inherently comparative (143, 165, 251, 271, 281), and many of the accomplishments in the field during the last century stem from a comparative framework. Interspecific comparisons were crucial to understand how organisms work and evolve in response to the ecological circumstances in which they live, as summarized by Schmidt-Nielsen (274) in his classic textbook on the subject:

> *"The understanding of how organisms function is helped enormously by using a comparative approach. By comparing different animals and examining how each has solved its problem of living within the constraints of the available environment, we gain insight into the general principles that otherwise might remain obscure. No animal exists, or can exist, independently of an environment, and the animal that utilizes the resources of the environment must also be able to cope with the difficulties it presents. Thus, a comparative and environmental approach provides deeper insight into physiology."*

Implicit in this view is a close correspondence between phenotypic and environmental variation, a pattern that emerges from adaptive evolution where natural selection is the primary process behind evolutionary changes (note that many other processes such as genetic drift, bottlenecks, and founder effects also contribute for phenotypic evolution, but these processes are fundamentally stochastic and

[*]Correspondence to enrico.rezende@uab.cat

[1]Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain

[2]Departamento de Ecologia, Universidade Federal de Goiás, Goiás, Brazil

Published online, January 2012 (*comprehensivephysiology.com*)

DOI: 10.1002/cphy.c100079

nonadaptive). Accordingly, in comparative studies the general *modus operandum* to study evolutionary adaptations—i.e., responses occurring at the population level that involve multiple generations and changes in gene frequencies, which should not be confounded with phenotypic adaptations encompassing plastic responses in an individual (see 20)—has involved identifying associations between phenotypic characters and environmental variables, as if species phenotypes have resulted from a long-term natural experiment. As discussed previously in the literature (75, 101, 150), this analogy entails serious misconceptions because researchers only have access to the final output of this so-called experiment, and virtually no historical information on the "experimental conditions" (the conditions lineages encountered during their evolution or the contribution of alternative processes to observed patterns of phenotypic variation).

To illustrate this problem, consider the difference between groups in a drug trial under controlled conditions and inferences based on a comparison between species to test adaptive hypotheses. Whereas in the first case differences between groups support a treatment effect because all other relevant variables are (presumably) controlled, in the second case species may simply differ because of their separate evolutionary paths, irrespective of the trait, or traits, under study (100). Similarly, adaptive inferences based on comparisons between species inhabiting contrasting environments may be impaired if these species also differ in taxonomic affiliation (e.g., suppose they belong to different genera). Observed differences among species may have been inherited from their ancestors, hence it is not possible to determine if they reflect recent evolutionary adaptations to contrasting environmental conditions or result from shared ancestry. Importantly, adaptive explanations have not been refuted in either of the examples, neither have alternative scenarios. Ultimately, these competing hypotheses cannot be appropriately addressed because analyses involve at least one confounding factor that obscures comparisons.

The realization that interspecific comparisons must explicitly consider species evolutionary history resulted in profound changes in comparative studies during the last decades (82, 101, 137, 226). Conceptual advances borrowed from evolutionary biology are widely applied in current analyses, which incorporate rigorous evolutionary perspectives and phylogenetic information. These methods have greatly increased the types and qualities of inferences that can be drawn from comparative data, and opened new venues of research in physiology that are still unfolding. From our own experience, one of the primary challenges for researchers with a physiological background interested in comparative analyses consists in understanding the evolutionary framework inherent to phylogenetic statistical methods, hence here we focus primarily on the conceptual basis that underlies the study of physiological evolution from a comparative historical perspective. We consider the drawbacks of employing conventional statistical techniques to analyze comparative datasets and how phylogenetic statistical methods can circumvent many of these issues (whereas others, such as low statistical power associated with small sample sizes, remain general to virtually all statistical approaches), and discuss the inherent limitations of a comparative framework, its correlational nature and the sources of phenotypic variation that physiologists should take into consideration—and, hopefully, control for—during a comparative study. Subsequently, we explain how to employ phylogenetic information in comparative analyses, discussing in more detail different analytical methods that have been applied to many morphological, physiological, biochemical, and behavioral investigations.

Adaptation in A Historical Context

Living organisms are related to each other in a hierarchical fashion, and these relationships are often represented by a branching tree structure (Fig. 1). A phylogenetic tree depicts the hypothesized evolutionary history of a clade—that is, a group consisting of one ancestor and all its descendants—starting from a single common ancestor in the past (root) diversifying into different lineages that ultimately give rise to extant species and/or populations (tips) (note that a treelike structure is only adequate for multiple populations within species when migration and gene-flow among them are negligible; see 84, 295). Intermediate nodes represent speciation events where one ancestral lineage splits into two or more descendent lineages, and the length of each branch roughly represents the expected amount of time that each lineage had to evolve, which is often inferred from variation in DNA sequences or other molecular data. Given that phylogenies are employed to estimate (and often control for) patterns of resemblance in comparative data due to shared ancestry, researchers should always keep in mind two important considerations that will be discussed in more detail subsequently: (i) phylogenies depict hypothesized patterns of relatedness among species, hence they are prone to error and uncertainty (8, 9, 85) and (ii) even though branch lengths, in principle, reflect elapsed time and these estimates are perfectly suitable for several analyses, branch length transformations are often employed for practical purposes to emulate different evolutionary processes underlying trait variation (below). These points are crucial because the reliability of analyses depends fundamentally on whether the phylogeny employed reflects with some degree of accuracy the phylogenetic structure of the phenotypic data.

Because lineages in different branches evolve independently from one another, the overall phenotypic variation observed between them tends to accumulate (hence increase) with divergence time. Consequently, closely related species will tend to resemble each other for historical reasons and this will happen independently of any deterministic process, such as natural selection, driving trait variation. Even though this trend is pervasive in the tree of life and occurs at all levels of biological organization, notable exceptions often result from adaptive evolution. The intrinsic association between

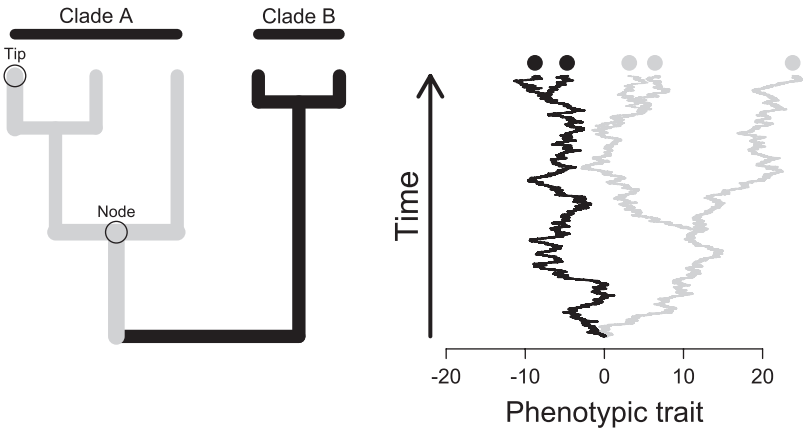


Figure 1 A hypothetical phylogeny representing the evolutionary relationships among five species, and its consequences at the level of phenotypic variation. The *tips* of the phylogenetic tree represent extant species and *nodes* depict the most recent common ancestor of a clade, that is, a hierarchically arranged, monophyletic group of species. For illustrative purposes, only the separation between clades A and B are shown, but note that these two clades belong to a larger clade that encompasses all species in the phylogeny, with a common ancestor known as the *root node* of the tree. Given the hierarchical patterns of relatedness among species, phenotypic data in comparative studies may not necessarily provide independent sources of information, as shown for the two pairs of closely related species that are phenotypically very similar. Consequently, patterns of phenotypic resemblance may be interpreted as evidence of evolutionary convergence (adaptation) when in fact they reflect common ancestry. For this particular example, phenotypic evolution proceeded as a random walk (i.e., a Brownian motion model of evolution).

form and function can constrain the number of potential solutions for any given problem and, consequently, different lineages may adopt similar strategies to cope with similar challenges (for instance, insects, pterosaurs, birds, and bats employ wings for flying). An intuitive way to visualize alternative evolutionary scenarios in this context involves contrasting the degree of phenotypic resemblance between species against their relatedness (Table 1). Whereas similarity between relatives is expected from simple descent with modification, natural selection might explain pronounced differences between closely related species (divergence) or phenotypic similarity between distant lineages (convergence). Changes in osmotic tolerance during freshwater invasion by the copepod *Eurytemora affinis* provide an illustrative example of both processes (172), where the high tolerance to saltwater in ancestral populations shifted repeatedly to freshwater tolerance during separate colonization events (Fig. 2).

Table 1 Comparative Patterns and The Origin and Diversification of Phenotypic Traits

	Taxa	
Traits	**Related**	**Not related**
Similar	Shared ancestry**	Convergent evolution
Not similar	Divergent evolution	Uninformative

Multiple evolutionary scenarios can account for differences and similarities across taxa; hence consideration of how these taxa are related is crucial to discriminate between adaptive explanations and common descent.
** Note that the maintenance of a trait by natural selection may also account for this pattern. Modified, with permission, from Brooks and McLennan (33).

Thus, although a general correspondence between phenotypic and environmental variation is still necessary for adaptive inferences, phylogenetic analyses can reveal which characters are ancestral and consequently not adaptive in the strict sense because, from a historical evolutionary perspective, adaptations encompass derived characters that evolved in response to natural selection (20, 51, 100, 123, 137, 140). This definition focuses primarily on the origin and diversification of phenotypic traits, which tends to be the primary interest of comparative physiologists, and emphasizes the importance of a historical approach to falsify adaptive hypotheses. For example, the hemoglobin of South American camelids has a high affinity for O_2 that was presumed to have been favored during the colonization of the Andes, though comparisons with the distantly related lowland dromedary suggest that this condition was ancestral in this lineage (245). Therefore, the historical approach suggests that high O_2 affinity is an exaptation in camelids (123), because it may have enabled them to invade high altitudes but its origin preceded the colonization of these habitats. This highlights that, even though this historical definition of adaptation may be adequate for most comparative analyses, some phenotypic traits may be maintained by natural selection and are consequently adaptive in a broader sense. The same applies to the sutures of the human skull, which may be considered adaptive for allowing

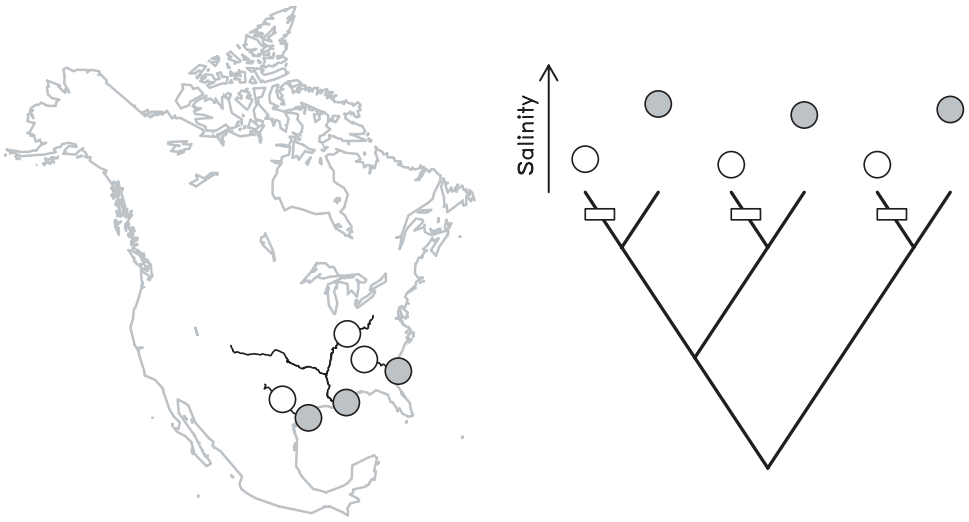


Figure 2 Evolution of osmotic tolerance across populations of *Eurytemora affinis*, a copepod that invaded freshwater environments repeated times. Phylogenetic information suggests that tolerance to low osmotic pressures has evolved at least three different times from a marine ancestral lineage, and this adaptive response resulted in divergence between close relatives and convergence across distantly related species. Adapted from Lee et al. (172) with permission of the University of Chicago Press.

the birth of large-brained offspring even when its origin preceded the evolution of an enlarged brain (102). Regardless of the controversies involving the operational definition of evolutionary adaptation, there is no doubt that a historical perspective is crucial to understand how and when different traits have originated, and may shed light on the evolutionary processes underlying the origin and maintenance of these traits (36,100,102,107,130,132,137,138,149,217,226,233). For instance, recent phylogenetic studies that suggest a marine origin of snakes (61, 174) and, consequently, that limb loss might be associated with aquatic locomotion, illustrate how the selective pressures involved in the origin of some traits may only be inferred with a comparative approach. Additionally, as noted by Darwin, the existence of vestigial structures with presumably no function—for example, flight motor neurons in flightless grasshoppers (76), the pelvic girdle in whales or the appendix in humans—can only be understood from a historical perspective.

Comparisons between ancestral and derived traits are inherent to the historical definition of adaptation. Analyses comparing ancestral and derived characters are often referred to as "directional comparisons" because the direction and rate of evolution along the phylogeny can be inferred (137, 138). In physiological systems, directional studies have been employed to study the historical sequence underlying protein evolution and molecular adaptation (97, 120, 208, 245), thermal coadaptation between behavioral preferences and physiological performance (106, 151), and the origin of evolutionary novelties such as heat exchangers (27) or the swimming bladder in fishes (22, 23) (Fig. 3). The main drawback of directional comparisons is that ancestral states are rarely known with accuracy; hence they are generally inferred from current data assuming a given model of evolution (see below). Alternatively, "nondirectional comparisons" between contemporary species are generally employed to analyze patterns of covariation between phenotypic characters or the fit between these characters and environmental variables, which has been the traditional analytical approach in comparative physiology. Conventional and phylogenetic regressions and correlations, for example, have been frequently employed to study the interspecific variation of continuously varying physiological traits such as metabolic rates, thermal tolerance, diving performance, costs of transport, and so on (e.g., 31,39,50,140,152-154,162,164,212,224,274,275,303,315). These statistical methods have now been expanded to consider explicitly the role played by phylogeny; hence, alternative evolutionary hypotheses can be rigorously tested employing a comparative approach.

Pattern Versus Process

One criticism regarding comparative analyses that has been routinely discussed in the literature involves the difference between pattern and process, or correlation and causation (75, 101, 149, 176, 320). Comparative approaches in general, and phylogenetic analyses in particular, are primarily correlational because it is virtually impossible to know with certainty the conditions in which the phenotypic traits of study have evolved and, consequently, inferring causality and ruling out alternative mechanisms that might explain observed patterns in comparative data may be extremely difficult (182). Historically, any given association between environmental traits and phenotypic variation was interpreted as evidence of adaptive evolution driven by natural selection, and this adaptationist program was severely criticized before and during the advent

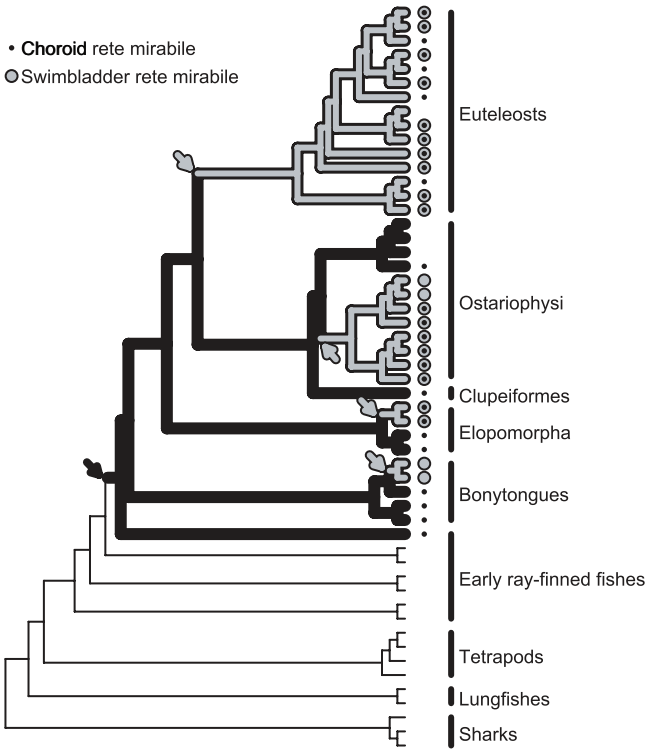


Figure 3 The evolutionary steps behind the origin of the swim bladder can be traced onto the phylogeny of jawed vertebrates. Some fishes present complex vascular counter-current systems known as retia mirabilia that are, among other functions, involved in the secretion of gases by blood acidification (Root effect). Phylogenetic analyses support a single origin for the choroid rete mirabile, suggesting that the physiology behind oxygen secretion first evolved within the ray-finned fishes to maintain a metabolically active retina. This preceded the evolution of the swimbladder rete mirabile, which occurred in four independent lineages (gray arrows) and enabled them to control buoyancy by physiological means. Symbols at the tips of the phylogeny indicate the presence of the choroid or the swimbladder retia mirabilia in extant species, and different branches illustrate the hypothesized state of ancestral lineages according to parsimony (evolutionary losses are not shown in the phylogeny for clarity, but can be inferred from the tip data). Modified, with permission, from Berenbrink (22).

of modern phylogenetic statistical methods because virtually every pattern had an adaptive explanation provided *a posteriori* (102, 122). In many instances, these *a posteriori* explanations entirely ignore the historical component involved in the evolution of a trait, incorrectly interpreting associations between form and function in a derived condition as evidence supporting its adaptive origin; e.g., the observation that flying animals have wings should not be interpreted as evidence that wings evolved as adaptations to fly (see 64-66 for studies on the evolution of avian flight, and 338 on the evolution of winged flight in insects). Certainly, wing morphology is likely associated with flying performance in these animals and may be under selection, but the evolutionary processes underlying the maintenance of any given phenotypic character should not be confounded with those behind the origin of this character.

Even when evolutionary history is taken into account, as in the phylogenetic analyses discussed below, inferences regarding adaptive evolution should be performed with caution. It is often the case that many different processes can give rise to the same pattern, and diverse evolutionary mechanisms, such as founder effects, migration, drift, constraints, genetic hitchhiking, and correlated evolution, among others, can result in significant associations between genetic/phenotypic and environmental variation (244). Comparative biologists must therefore always consider which alternative nonadaptive explanations might account for any observed pattern. For example, two different hemoglobin haplotypes have been described in the deer mouse *Peromyscus maniculatus*, their frequency in the population being highly correlated with the altitude of the collection site (282). This correlation may have emerged in response to selection related with altitude or, alternatively, it may have resulted from stochastic biogeographic processes because altitude trends and geographic trends are confounded (one haplotype is found in high frequency in central Colorado and declines in all directions from this region). Discriminating between adaptive and nonadaptive explanations behind correlational patterns is not simple and must rely on a variety of evidence. With regards to the hemoglobin polymorphism observed in *P. maniculatus*, subsequent studies showed that the hemoglobin haplotypes differed in their binding affinity for O_2, resulted in enhanced aerobic performance during strenuous exercise and cold exposure at their native altitude (i.e., performance in individuals with the "high-altitude" haplotype was significantly higher at high altitudes and lower at low altitudes) and showed contrasting patterns of geographic distribution when compared with other genetic markers (43, 44, 284). Taken together, this information suggests that the altitudinal gradient in hemoglobin haplotype frequencies has emerged for adaptive reasons (see also 297, 298).

Even though at this level some readers may be tempted to dismiss comparative analyses to study adaptation, it is important to recall that a lack of association between hemoglobin haplotypes in deer mice and altitude would challenge any speculation on the role of these haplotypes in altitude adaptation. Similarly, as discussed in the introduction, the high hemoglobin affinity for O_2 in camelids across different continents is at odds with the hypothesized adaptive origin of this trait during the colonization of the Andes. As eloquently stated by Losos and Miles (183), "historical studies can unequivocally falsify adaptive hypotheses, but, because experiments are not possible, support for an adaptive hypothesis requires building up a convincing case by integrating disparate lines of evidence" (p. 66). In this context, the so-called revolution associated with the advent of phylogenetic comparative methods lies in the development of analytical tools that have enabled researchers to study, with unprecedented accuracy, the emergence of biological patterns in a historical perspective. Nonetheless, the correlational nature of phylogenetic analyses, and its inherent limitations, should always be kept in mind (182). While a phylogenetic perspective provides a more elaborate and detailed account of the historical patterns that have resulted in the observed distribution of phenotypic traits across species or other taxa, inferences on underlying

evolutionary processes must always rely on complementary sources of information (75, 149, 176). Consequently, mechanistic research on how organisms work and theoretical approaches such as optimization models remains of paramount importance to study physiological evolution (140, 212, 274), and these studies can (and should) be complemented with approaches borrowed from evolutionary biology such as experimental evolution, population and quantitative genetic analyses, measurements of selection acting in natural populations, studies of individual variation in physiological traits, and so on (20, 75, 81, 99, 101, 102, 149, 176, 183, 212, 244, 284).

Phylogenetic Statistical Approaches

From a statistical point of view, a fundamental problem in comparative analysis is that species do not provide independent sources of information on how phenotypic traits have originated and evolved (Fig. 4). Closely related species have had less time to diverge than distantly related species, therefore conventional statistical methods treating species as independent risk pseudoreplication, which can inflate the type I error rate of significance tests (the null hypothesis will be rejected too often), affect the statistical power to detect relationships, and result in inaccurate estimates of correlations, slopes, and/or group differences (101,202,203,228,302) (Fig. 5). Modern phylogenetic statistical methods circumvent this problem by taking phylogenetic information into consideration. They acknowledge that "the closer two species are in the phylogeny, the more likely that they resemble each other" and control for expected patterns of phenotypic covariation due to shared ancestry. These methods differ from previous analyses that incorporate taxonomic levels—e.g., nested designs have been employed prior to the development of phylogenetic methods to deal, at least to some extent, with this issue (80, 137)—in that they are able to incorporate more complete information on the evolutionary history of the clade, such as the shape of the phylogenetic tree and divergence time estimates, as well as make explicit assumptions on the rates of phenotypic (or genetic) evolution and the model of character evolution (101).

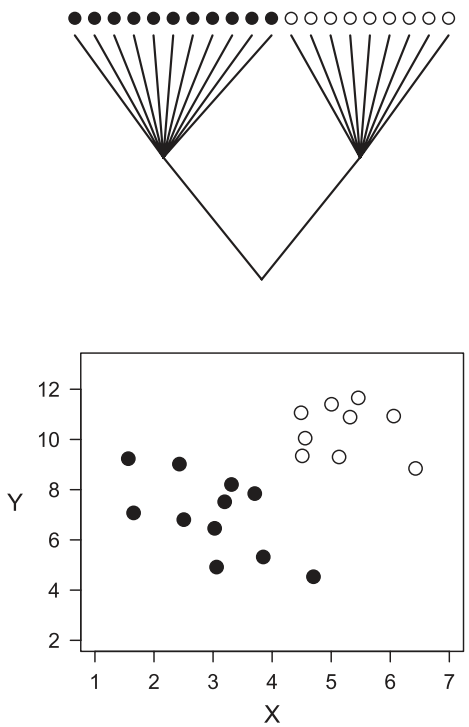


Figure 4 The problem of analyzing phylogenetically structured data with conventional statistical methods. Ignoring phylogeny, one would conclude that X and Y are positively correlated (Pearson $r = 0.48$, 2-tailed $P = 0.034$), when in fact this relationship emerges primarily from the high divergence in X and Y between the two clades at the root of the phylogeny. Modified from Felsenstein (86), with permission of the University of Chicago Press.

Different sources of information such as gene sequences and morphological traits can be employed to build a phylogeny, as long as these characters are independent from the phenotypic traits analyzed (101). Nowadays, the topology of the phylogeny is generally obtained from DNA sequences, and the degree of divergence across sequences then provides an estimate of divergence times between species. For instance, Sibley and Ahlquist (279) compiled the phylogeny of birds around the world employing DNA-DNA hybridization and distance methods to estimate the proximity between certain nodes, which was subsequently used in many comparative studies (260, 262, 323). More sophisticated methods to statistically infer phylogenies from gene sequences are currently available (85), and several clades of Sibley and Ahlquist's classification have been disputed (10, 17). This emphasizes that, in a comparative study, phylogenies are working hypotheses on the true evolutionary relation among species and/or populations; hence, they are subject to error (68, 69, 145, 147, 207, 226, 248, 301). Although some approaches have been developed to deal with phylogenetic uncertainty (145, 147), most comparative analyses currently employ one working phylogeny.

Phylogenetic information can be employed to estimate and statistically remove the effects of nonindependence in comparative data, and for most phylogenetic comparative methods this involves knowledge of divergence times and the incorporation of explicit models of character evolution. Conceptually, it is important to distinguish between the evolutionary history of a clade, described by a phylogeny with its topology and divergence times, and the evolutionary processes that may account for its phenotypic diversity. This distinction is not entirely obvious during the implementation of some phylogenetic methods because they are ultimately concerned about patterns of phenotypic resemblance, which are not always associated with the degree of relatedness among lineages. This is why several comparative analyses have employed branch lengths in units of expected character change (generally referred to as "expected variance of character change") that may or may not be proportional to time (29, 82, 87, 98, 104, 105, 107, 109, 132, 146, 220, 226, 236, 267).

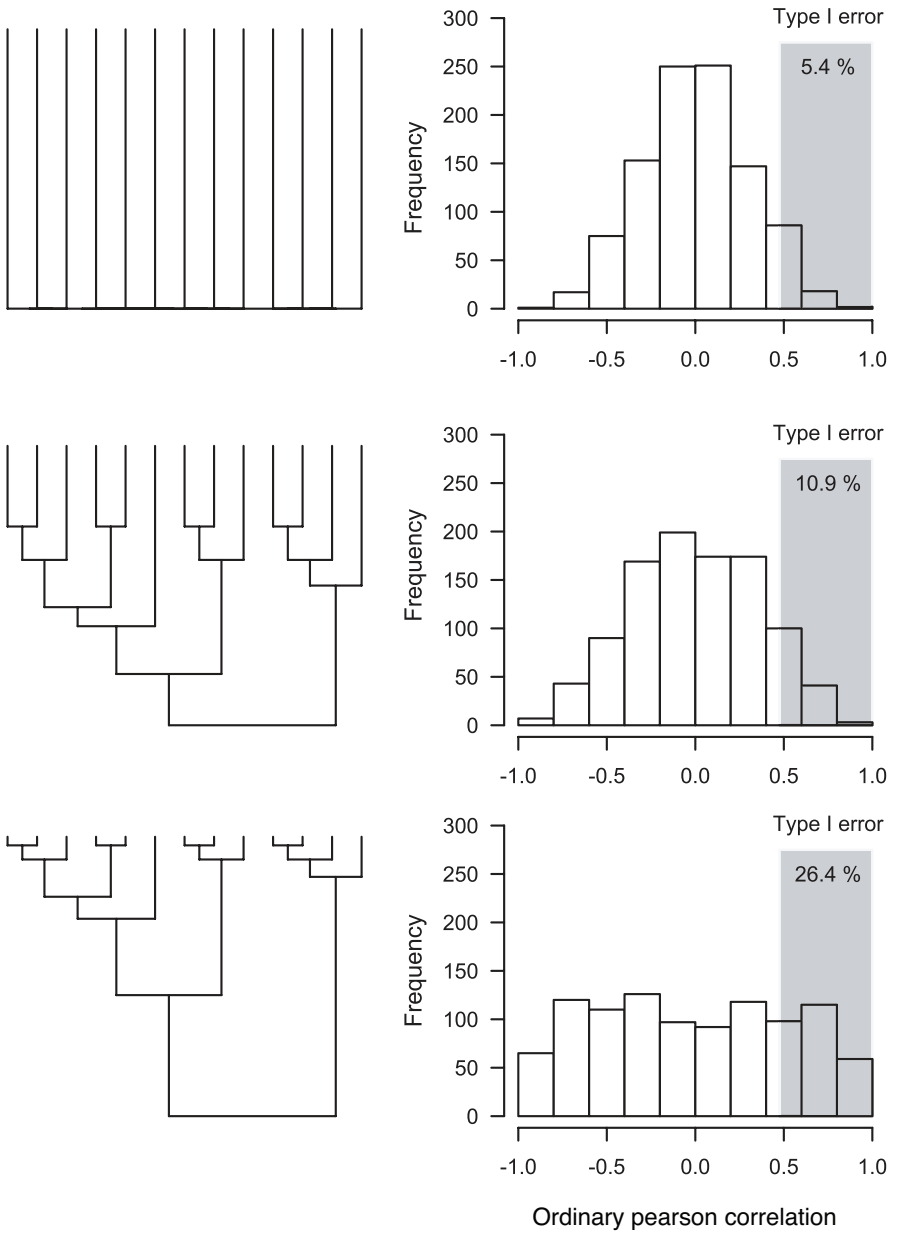


Figure 5 Increased type I error rates of conventional statistics in analyses of interspecific data. When two traits evolve independently along a phylogeny according to Brownian motion, the probability of rejecting the null hypothesis of no correlation (type I error) increases with the amount of phylogenetic structure of the data. The shaded area represents simulations where the resulting ordinary Pearson coefficient falls above the tabular critical value of +0.476 (11 degrees of freedom), which would incorrectly suggest that the two traits are correlated. Simulations with a star phylogeny result in the error rates of 5%, which is the expected type I error rate if conventional (nonphylogenetic) analyses are used. Type I error rates can be higher than 25% if the data shows a strong phylogenetic structure (for one obvious example where the correlation between two traits is incorrectly inferred, see Fig. 4). Modified, with permission, from Garland et al. (101).

Even though intuitively one would expect that the degree of phenotypic divergence increases with time, this is only true for gradual evolutionary models where phenotypic changes accumulate in direct proportion to elapsed time. A random walk or Brownian motion model of phenotypic evolution (see Fig. 1) emulates processes such as random genetic drift or stochastic fluctuating selection (82, 101) that, in the absence of any additional process or boundaries for phenotypic evolution, results in a linear association between divergence times and the expected phenotypic variance (Fig. 6). Other gradual models, such as the Ornstein-Uhlenbeck model (OU) that mimic the action of stabilizing selection toward a "phenotypic optimum," or the ACDC model in which evolutionary rates accelerate or decelerate in time, result in a nonlinear

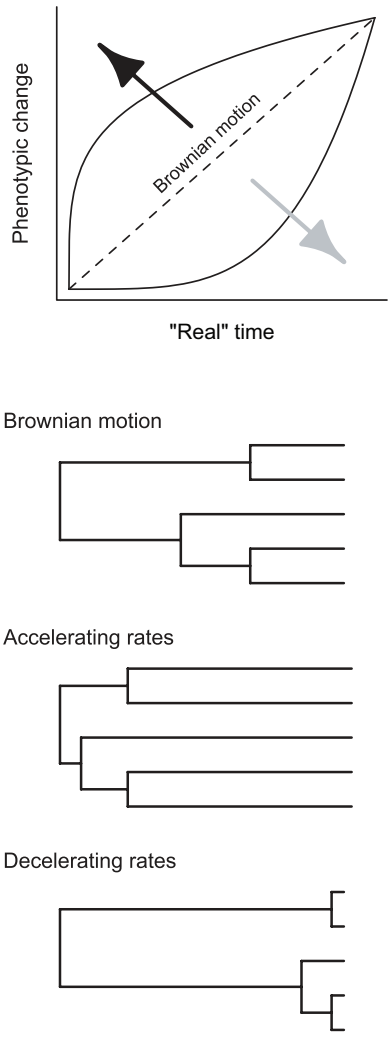


Figure 6 Branch lengths in comparative analyses. The branches of a phylogeny indicate the elapsed time between speciation events, but the degree of phenotypic similarity expected among species (which is the main concern in a comparative dataset) will depend on the elapsed time and on the evolutionary model of character evolution. Under an evolutionary model of Brownian motion, the "expected variance of character change" is proportional to elapsed time. Alternatively, the expected variance is not directly proportional to time when rates of character change accelerate (*gray arrow*) or decelerate in time (*black arrow*), which affects the relative contribution of recent versus past evolution to the overall distribution of phenotypes. Accelerating evolutionary rates results in less phylogenetic signal than expected from Brownian motion because most phenotypic variation reflects recent evolutionary history, whereas decelerating rates generate the opposite pattern. A star phylogeny depicts the extreme case in which all effects of shared ancestry and past evolution have been blurred by recent phenotypic evolution (at the tips of the phylogeny). Many evolutionary processes, such as divergent or stabilizing selection, can account for a nonlinear association between character change and time. Modified, with permission, from Diniz-Filho (71).

relationship between evolutionary time and expected phenotypic variance among species (29,36). Conversely, in models such as the speciational model, in which phenotypic evolution occurs only during speciation and the remaining periods are characterized by evolutionary stasis, character evolution proceeds in jumps and is independent of divergence times (236). In practice, a speciational model can be emulated by setting all branches to a constant and assuming Brownian motion, which illustrates how branch length manipulation can be employed to mimic particular evolutionary processes (105). Processes such as competition between close relatives, character displacement following speciation and adaptive evolution as species invade new niches may justify the saltational behavior of these models. To summarize, researchers must keep in mind that divergence times may not adequately describe patterns of phenotypic resemblance in comparative data (e.g., in a recent allometric analysis of mammalian basal metabolic rates, regressions employing arbitrary branch lengths provided a better fit to the data than divergence times; see 324), the primary reason being that phenotypic traits may evolve in multiple ways and at varying rates. Another possible reason is measurement error, which may obscure and distort phylogenetic patterns present in the data (see below). These alternatives are not mutually exclusive, however, and the potential impact of measurement error must always be taken into consideration when proposing evolutionary explanations for observed patterns of phenotypic distribution.

By now, the intricacies associated with phylogenetic statistical methods may be off-putting for those physiologists interested primarily on analyzing their data. However, the fundamental take-home message is that phylogenetic comparative methods lie on a robust conceptual basis that encompasses (i) an understanding of the processes underlying phenotypic evolution and (ii) how these processes translate into interspecific variation. Such a framework is fundamental for reliable statistical inferences based on comparative data and has enabled comparative biologists to rigorously test their hypotheses in an evolutionary framework. While some of the criticisms of these methods stem from the fact that we often do not know—an may never know—in detail how phenotypic evolution has occurred in the past, it is important to keep in mind that drawing inferences based on comparative data with a phylogenetic structure is primarily a statistical problem (which conventional analyses are not able to circumvent). For instance, recent comparative analyses can test the adequacy of several alternative models to determine which branch lengths best fit the analyzed dataset (53, 110, 171, 299, 324, 329). This approach is particularly powerful because researchers can compare a wide variety of evolutionary scenarios (i.e., whether evolution occurs gradually or in jumps, accelerates or decelerates in time and so forth), and their adequacy will depend primarily on whether they are good descriptors of past phenotypic evolution of the clade under study (92).

Phylogenetic signal

The primary concept behind phylogenetic comparative analyses is that of phylogenetic signal, or the tendency of closely related species to resemble each other. Phylogenetic signal results from the simplest evolutionary models that include hierarchical patterns of relatedness (28), hence its presence should be pervasive across interspecific data.

Accordingly, in comparative studies with large samples where low statistical power to detect signal is not a major issue ($n > 20$ species), more than 90% of the datasets analyzed exhibit significant phylogenetic signal (29); see also (5, 14, 55, 67, 93, 170, 223, 261, 265, 269, 336). These analyses include behavioral, physiological, morphological, life history, and ecological traits measured in very disparate taxa (e.g., birds, insects, and plants), emphasizing the generality of this pattern. In addition, significant phylogenetic signal has been also detected in intraspecific datasets (14, 54), showing that phenotypic variation across populations of a single species can also exhibit a phylogenetic structure (importantly, if migration and gene flow occurs across populations, standard phylogenies may not be appropriate because patterns of relatedness differ from a treelike genealogy; see 84, 295). Nonetheless, different factors such as the existence of evolutionary constraints (including boundaries to phenotypic evolution) or adaptive evolution in response to natural selection may decrease and even disrupt patterns of resemblance (29, 265). For instance, as discussed previously, close relatives may diverge rapidly due to local adaptation and distantly related species may converge for the same reason (Fig. 2). In addition, other factors such as measurement error and phenotypic plasticity can obscure any phylogenetic signal that might exist. Given that similarities due to common ancestry are probably the primary reason accounting for the nonindependence of comparative data, testing for the presence of phylogenetic signal in a comparative dataset is a crucial to diagnose if phylogenetic correction is necessary during analyses (1, 28, 29, 93, 265). Similarly, significant signal in the residual error of regression models (not in the traits themselves) would also indicate that phylogenetic analyses are more appropriate (256).

Even though the presence or absence of signal has little to say about the evolutionary processes that underlie observed patterns (257), it can be very informative on how the phenotypic variation is distributed across species and how it varies over different time scales in the phylogenetic tree. This sort of information is very important to understand, for example, how different clades may respond to environmental stress, illustrating the potential of comparative physiological studies to improve the efficacy of biomonitoring and conservation programs (34, 90, 331). Testing for phylogenetic signal can be accomplished in different yet complementary ways, with randomization procedures or by model fitting comparing phylogenies with varying degrees of hierarchy (1, 28, 29, 93, 196). Even though the Mantel test has been commonly used to test for phylogenetic signal (by testing if a matrix of pairwise distances between trait values are correlated with a matrix of phylogenetic distances; e.g., 7, 55, 59), it will not be discussed here given its poor statistical performance, that is, low statistical power and inflated type I error rates (133).

In randomization analyses testing for phylogenetic signal, the amount of phenotypic divergence in the actual dataset is compared against a null distribution obtained from datasets in which phenotypic values are randomly shuffled across species, hence destroying any signal in the data (29, 196). Because indexes of phenotypic divergence—for example, the minimum number of transitions for categorical data or the mean squared error in phylogenetic linear models for continuous traits—are expected to be low when close relatives resemble each other, one concludes that the studied dataset exhibits significant phylogenetic signal with $P < 0.05$ when the index of divergence calculated on the original dataset is lower than 95% of the values estimated across multiple randomizations (Fig. 7).

Alternative methods have been proposed to estimate the optimal branch lengths that maximize the explained variance of the character of study, providing a direct assessment of its distribution and association with the phylogenetic structure. Hierarchical trees with long basal branches indicate that the phenotypic variation of the descending clades tends to be clustered, whereas trees with long branches leading to the tips indicate that most of the variation occurs across species rather than among groups at higher levels. The branch-length-transformation approach to test for phylogenetic signal consists in comparing the model fitness of the best phylogenetic tree against that of a star phylogeny, where all species derive from a single common ancestor without intermediate steps (29, 93, 171, 323). One can test if the phenotypic data presents a significant hierarchical structure by contrasting the fit of alternative models (35, 127, 158, 173, 185, 309, 310), which can be quantified with measures of goodness of fit such as the Akaike information criterion (AIC). This measure estimates the relative support of each candidate model, rewarding overall goodness of fit given by its maximum likelihood L and penalizing for its number of estimated parameters k ($\mathrm{AIC} = 2k - 2\ln L$): the smaller the AIC value, the better the relative fit of the model (35, 309, 310). Multiple model parameters such as λ (93, 233) or d (29) can be employed during optimization procedures to estimate the degree of phylogenetic dependence of the comparative data and, even though they differ in their underlying assumptions (while the parameter λ is not based on an explicit model of evolution, the parameter d emulates the effects of a OU process in the phylogenetic structure), the overall rationale is very similar. Values close to 1 indicate that the phylogeny with optimal branch lengths resembles the original (presumably hierarchical) phylogeny, while values close to zero indicate little hierarchy and, consequently, low phylogenetic signal (Fig. 8). For an expansion of the OU branch-length transformation method to test for signal in groups of traits, see Zheng et al. (339).

The star phylogeny ($\lambda = 0$ and $d = 0$) describes the extreme scenario where no resemblance exists among relatives (Fig. 9); hence, they provide independent sources of information on the evolution of the trait of study. In this case, trait will display no phylogenetic signal and results from phylogenetic analyses employing a star phylogeny converge to those of conventional statistics because the mathematical assumptions of these models are the same: independence and homogeneity of variances. In the phylogenetic jargon, this respectively implies that no covariation exists due to shared evolutionary history (i.e., no intermediate nodes and branches) and that branch

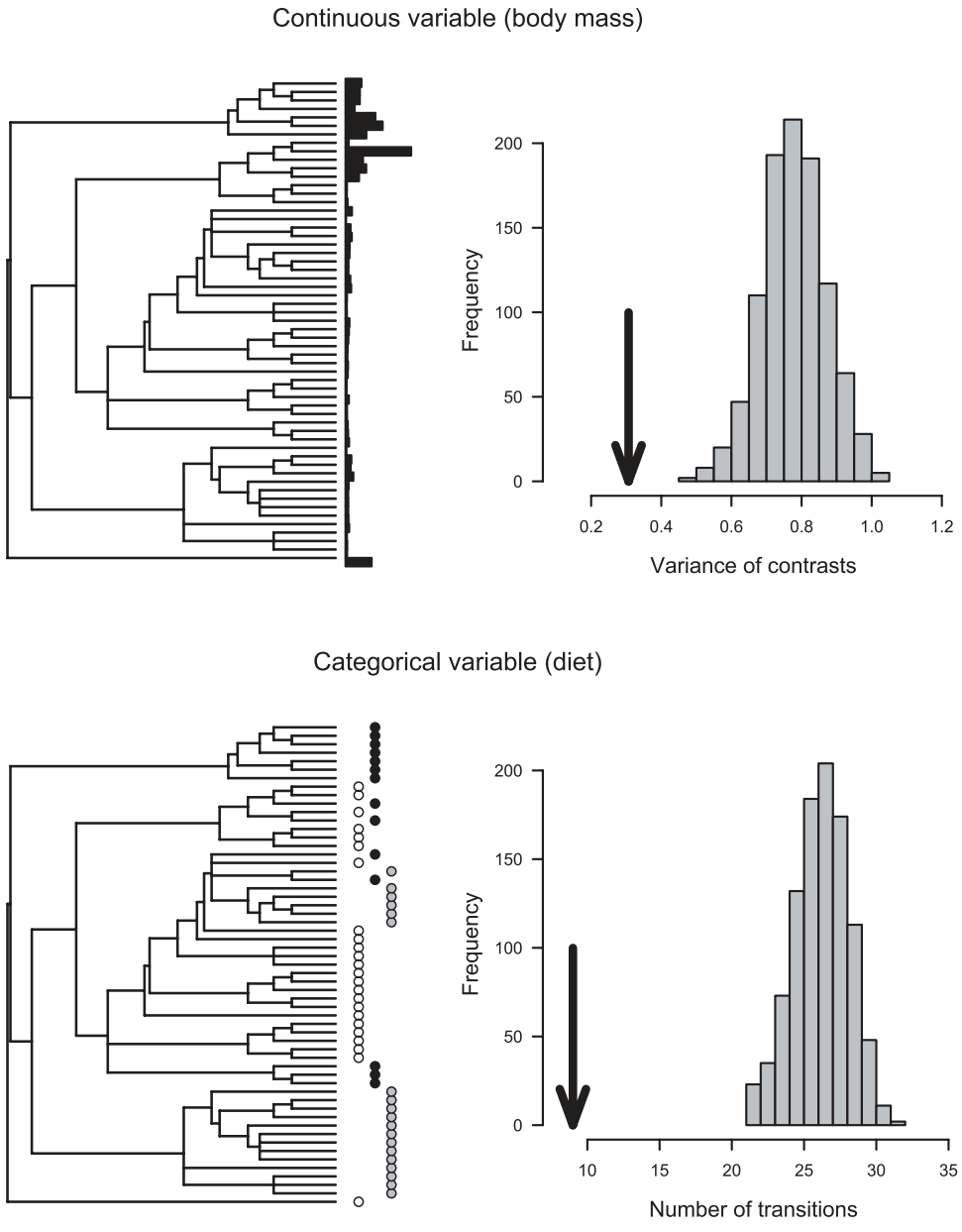


Figure 7 Randomization analysis to test for significant phylogenetic signal in continuous and categorical traits, illustrated with the distribution of body mass and diet categories (*white* = omnivore, *gray* = granivore, *black* = herbivore) of rodent species. Phylogenetic similarity was estimated as the variance of phylogenetic independent contrasts for the continuous trait (log_{10}-transformed body mass) (29) and as the minimum number of transitions according to unrestrained parsimony (i.e., all transitions are possible) for the categorical trait diet (196). These indexes are then calculated after shuffling the phenotypic characters across the tips of the phylogeny, breaking any phylogenetic structure and providing a null random distribution of *n* replicates ($n = 999$ in this example) where phylogenetic signal is absent. The histograms illustrate how both indexes calculated in the real dataset (represented by the arrows) fall consistently below randomizations, hence one can conclude that the distribution of body mass and diet shows significant phylogenetic signal across rodent species ($P < 0.001$ in both cases). Data and phylogeny, with permission, from Rezende et al. (261).

lengths from the common ancestor to the tips are identical across all species (101, 248) (see *top panel* in Fig. 5). More importantly, converging results as the degree of hierarchy decreases and eventually the phylogeny collapses into a star phylogeny emphasize that conventional statistical methods make implicit assumptions on the phylogenetic structure of the data, and actually correspond to a subset of more generalized methods where these assumptions are relaxed (29, 101, 124, 171). Consequently, whether phylogenetic or conventional statistical methods are more appropriate in a comparative study depends primarily on the amount of phylogenetic signal of the trait (or traits) involved, and by now the long-standing

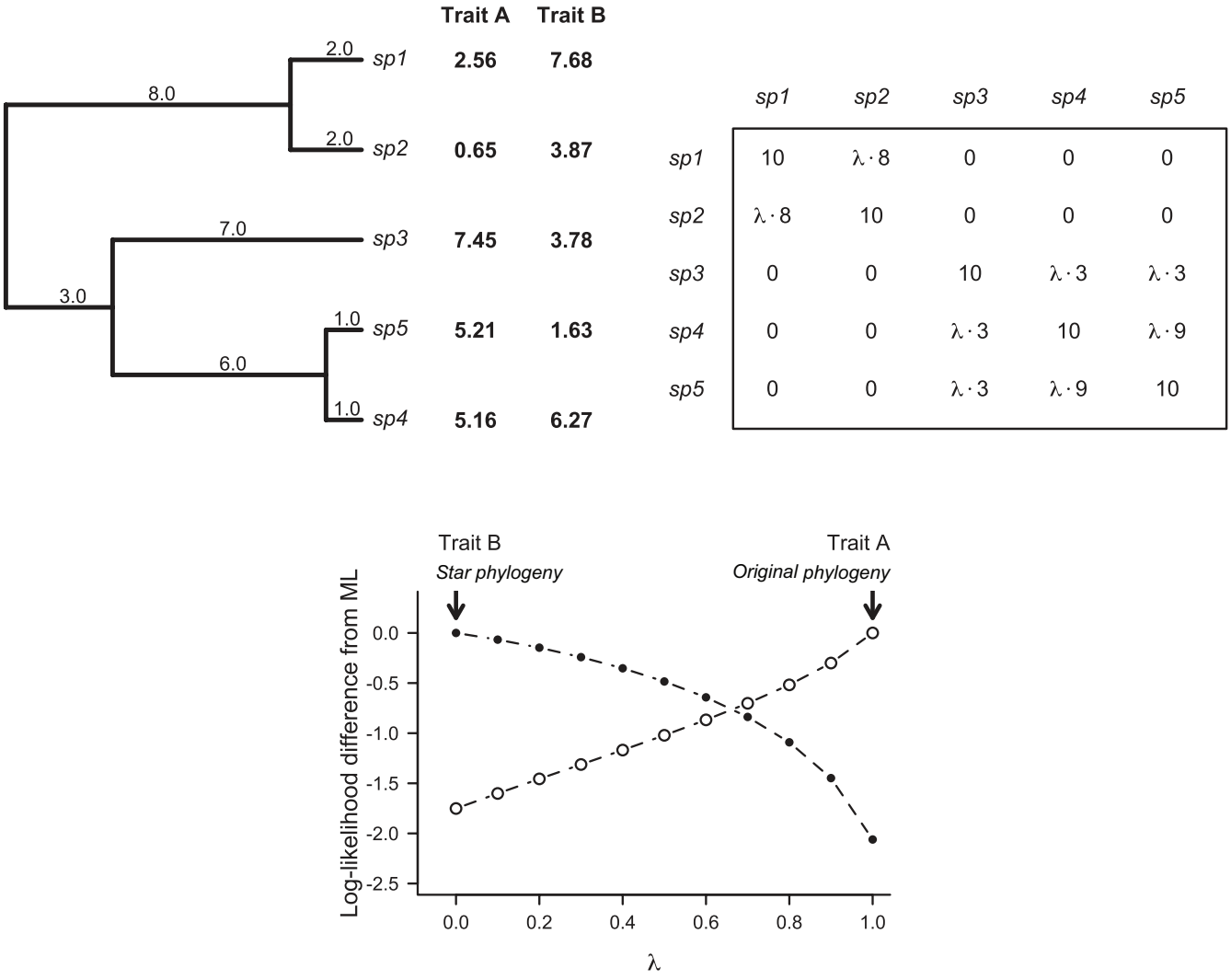


Figure 8 Branch length transformation analyses to test for the presence of phylogenetic signal attempt to find the phylogenetic structure that best fits the data. A phylogeny can be described as a matrix of variance-covariance describing the expected residual distribution of the comparative data. The diagonals depict the expected phenotypic variance (i.e., how species are expected to differ from the overall mean) that, under Brownian motion, corresponds to the total distance from the root to the tips. The off-diagonals provide the expected phenotypic covariance among species (or how they are expected to resemble each other due to shared ancestry), which corresponds to the total amount of evolutionary history shared by each pair of species. Because the amount of phylogenetic signal essentially encapsulates the amount of phenotypic covariance among species, parameters such as λ that affect the degree of hierarchy of a phylogeny by manipulating the length of the internal branches can be employed to estimate which phylogeny best fits the data (see Fig. 9). In this hypothetical example, the phylogeny that best fits trait A is very hierarchical ($\lambda = 1$) suggesting that close relatives tend to resemble each other for this trait. Conversely, the phylogeny that best fits the distribution of trait B shows no hierarchy ($\lambda = 0$), suggesting that signal in trait B is negligible. A log-likelihood ratio test comparing the likelihoods of models when $\lambda = 0$ versus $\lambda = 1$ can be employed for significance testing (not shown). Adapted from Freckleton et al. (93) (see also 256), with permission of the University of Chicago Press.

debate on this subject has shifted from a philosophical problem to a strictly statistical one. Importantly, the power to detect phylogenetic signal drops dramatically at small sample sizes ($n < 20$ species) (29,60), therefore caution is warranted when signal is not detected in small datasets. When it is impossible to differentiate between these alternative models, it is advisable to run both conventional and phylogenetic analyses and ensure that studied patterns are qualitatively robust. Otherwise, contrasting results should be reported and conclusions tempered accordingly. Given the conceptual insights resulting from an explicitly historical perspective and analytical rigorousness, we ultimately encourage the use of phylogenetic information in comparative analyses whenever it is possible (but see 182).

Phylogenetically independent contrasts

Since its introduction (82), the method of independent contrasts has been widely employed to analyze and control for phylogenetic effects in comparative studies. It is considered the first explicit phylogenetically based statistical method because it takes into consideration the branching pattern of the

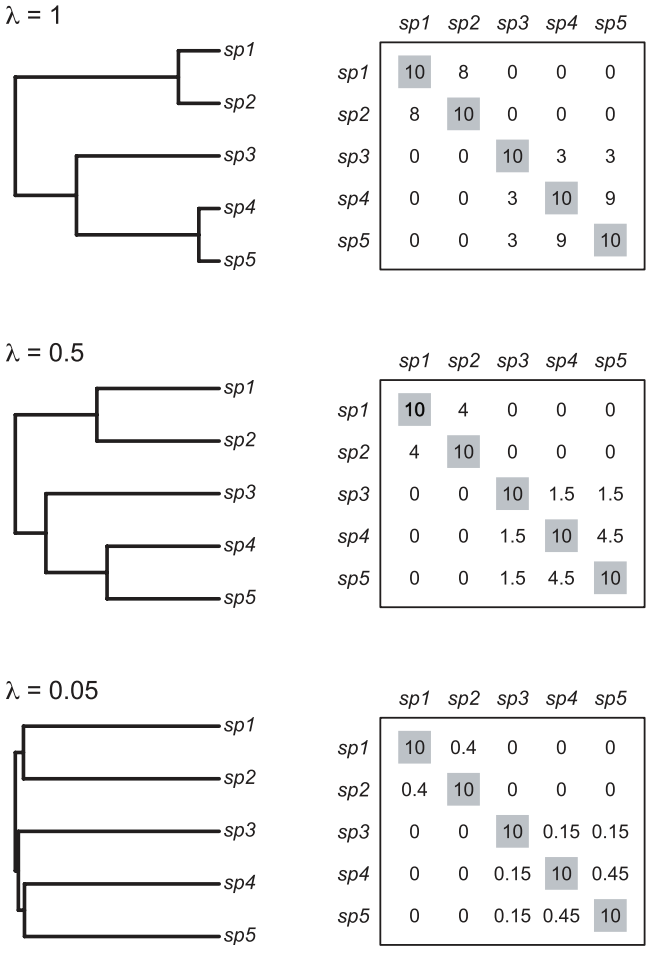
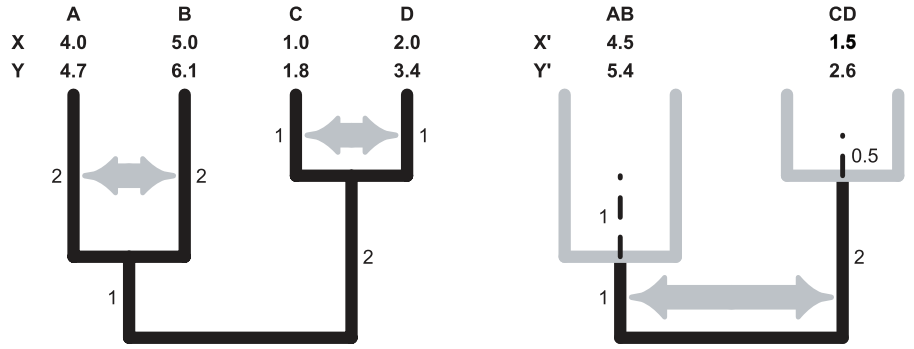


Figure 9 Varying degrees of hierarchy in phylogenetic trees expressed as matrices of phenotypic variance-covariance among species, illustrated for the phylogeny in Figure 8. The expected amount of phenotypic similarity due to shared ancestry decreases as λ approaches zero, and as when $\lambda = 0$ the matrix of phenotypic variance-covariance converges to the identity matrix (i.e., the expected residual distribution of conventional statistical analyses). In other words, comparative studies employing conventional analyses inherently assume that the data does not exhibit phylogenetic signal and species provide independent sources of information, which may or may not be true (compare traits A and B in Fig. 8).

phylogenetic tree, divergence times, and an explicit model of character evolution (Brownian motion). Succinctly, this method converts the original data, with its phylogenetic structure, into a series of phenotypic divergences that are independent from each other (82, 98, 101, 105, 107, 137, 203, 231). To calculate independent contrasts, the phenotypic divergence between two related nodes is standardized by the "time" they had to diverge (note that in the strict sense the branch lengths do not necessarily reflect time, as discussed above). Starting from the tips of a bifurcating phylogeny, this algorithm transforms the N original measurements into N–1 contrasts between sister lineages until it reaches the root node (Fig. 10). Importantly, as phenotypic divergences are estimated deeper in the phylogeny, the uncertainty associated with these estimations increases, resulting in an "extra" amount of expected variance in more ancestral nodes. In practice, this involves an additional step in the computation of contrasts that has been forgotten in several worked examples in the literature (see 101), which consists on adding extra variance during the standardization of contrasts—i.e., lengthening the branches by a certain amount—that can be calculated assuming a Brownian motion evolutionary model (Fig. 10). As demonstrated by Garland and Ives (107) and Rohlf (267), independent contrasts can be understood with the generalized least-squares regressions framework (below), ensuring strong theoretical statistical foundations for the method and making its solution well understood in this context. Several computer programs readily available can calculate independent contrasts, and may be employed in conjunction with any conventional statistical package. Assuming that the topology and branch lengths are correct, the resulting contrasts meet the requirements of independence and homogeneity of variances of conventional parametric statistics, and can be analyzed with these tools. Additional worked examples of how to appropriately calculate contrasts are available in other references (95, 100, 101).

One major advantage of independent contrasts is that they retain basic notions of phylogenetic thinking, because their distribution ultimately represents a series of phenotypic divergences over evolutionary time. The representation of evolutionary patterns employing contrasts immediately follows from the distribution of the raw data, but differs considerably from bivariate plots in the original dimensions (Fig. 10). These plots provide complementary information on the distribution of the original data and the evolutionary trends that underlie this distribution. For instance, in mammals, the size and shape of the ear bones are involved in stabilization and balance during locomotion (289, 290). Allometric plots clearly show that variation in size of the semicircular canal system is associated with body mass (Fig. 11), and the clustered distribution of the data suggests that both body mass and the size of the semicircular canal exhibit phylogenetic signal (see Fig. 1 in 289). Ignoring this fact, the allometry of the semicircular canal obtained with a conventional linear regression employing $\log_{10}$-transformed data equals $1.025 \times Mass^{0.067}$ (95% CI for the exponent lying between 0.049 and 0.085). However, it is clear that a grade shift in the size of the semicircular canal has occurred in cetaceans, presumably as an adaptation associated with the colonization of aquatic environments (289), resulting in a major impact on the scaling relationship. When cetaceans and all other mammals are analyzed separately, the allometric relationships equal $0.134 \times Mass^{0.167}$ (95% CI of 0.131 and 0.202 for the exponent) and $0.640 \times Mass^{0.137}$ (95% CI of 0.127 and 0.148), resulting in over a 2-fold increase in the calculated scaling exponent. Knowledge of patterns of relatedness among species provides a more accurate historical perspective, with the analysis with contrasts being able to control for phylogenetic effects both across and within taxonomic groups and also indicating where the grade-shift has occurred (Fig. 11). The resulting allometric



Variable	Contrast name	Contrast value	Raw contrast	Standard deviation	Standardized contrast
X	A - B	4.0 - 5.0	-1.0	$\sqrt{2+2}$	-0.5
	D - C	2.0 - 1.0	1.0	$\sqrt{1+1}$	0.71
	AB - CD	4.5 - 1.5	3.0	$\sqrt{2+2.5}$	1.41
Y	A - B	4.7 - 6.1	-1.4	$\sqrt{2+2}$	-0.7
	D - C	3.4 - 1.8	1.6	$\sqrt{1+1}$	1.13
	AB - CD	5.4 - 2.6	2.8	$\sqrt{2+2.5}$	1.32

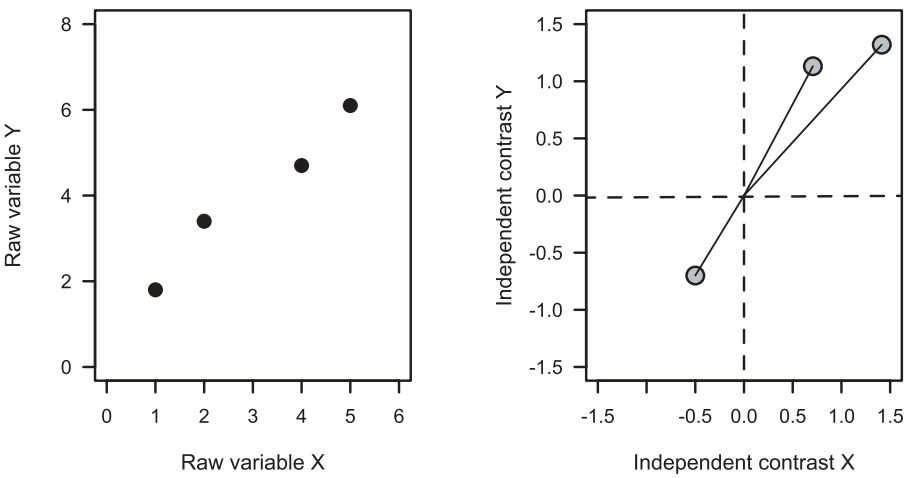


Figure 10 Calculation of phylogenetic independent contrasts for two hypothetical variables X and Y. Contrasts estimate the amount of phenotypic divergence across sister lineages standardized by the amount of time they had to diverge (the square root of the sum of the two branches). The algorithm runs iteratively from the tips to the root of the phylogeny, transforming *n* phenotypic measurements that are not independent in *n*–1 contrast that are statistically independent. Because phenotypic estimates at intermediate nodes (X′ and Y′) are not measured, but inferred from the tip data, divergence times employed to calculate these contrasts include an additional component of variance that reflects the uncertainty associated with these estimates. In practice, this involves lengthening the branches (dashed lines) by an amount that, assuming Brownian motion, can be calculated as (*daughter branch length 1* × *daughter branch length 2*) / (*daughter branch length 1* + *daughter branch length 2*). As a result, the association between the hypothetical phenotypic variables X and Y analyzed employing conventional statistics and independent contrasts may seem remarkably different, as shown in the bottom panels. Because independent contrasts estimate phenotypic divergence after speciation and are expressed as deviations from zero (i.e., the daughter lineages were initially phenotypically identical), correlation and regression analyses employing contrasts do not include an intercept term and must be always calculated through the origin (82, 105, 175). Note that the sign of each contrast is arbitrary; hence many studies have adopted the convention to give a positive sign to contrasts in the x-axis and invert the sign of the contrast in the y-axis accordingly (this procedure does not affect regression or correlation analyses through the origin; see 105). Even though the classic algorithm to calculate contrasts neglects important sources of uncertainty such as individual variation and measurement error, recent methods can account for these sources of error (87, 157, 206).

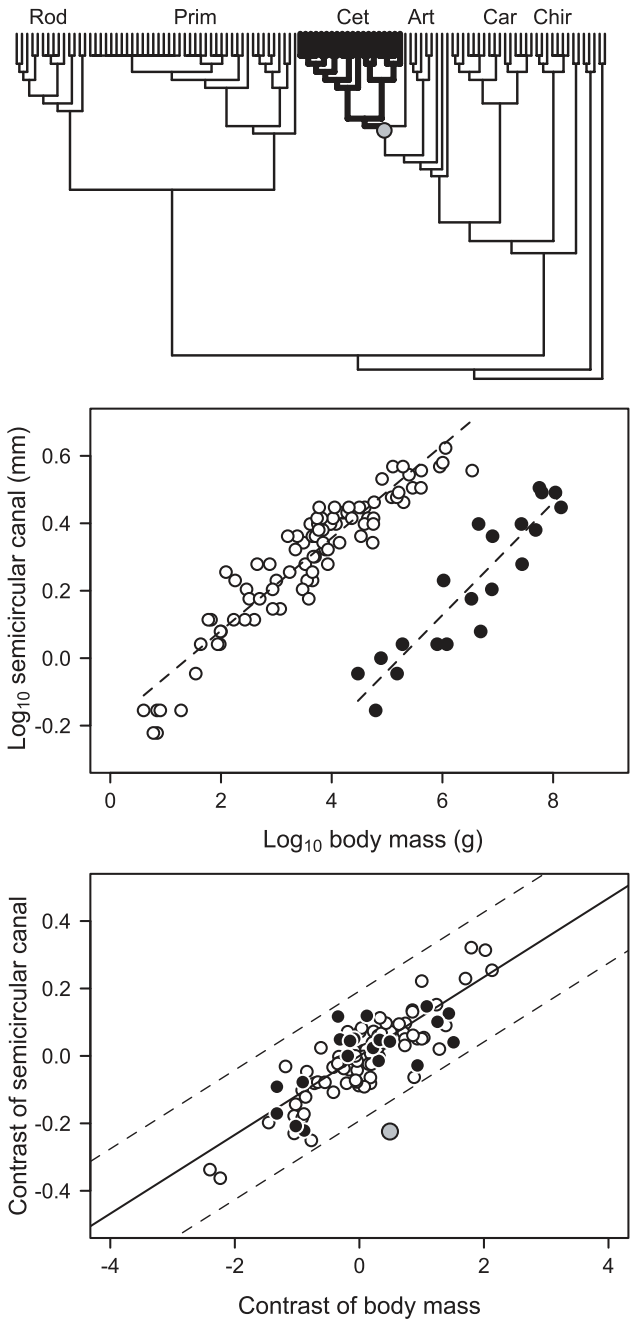


Figure 11 Correlated evolution and grade shifts in comparative data, plotted in raw dimensions and employing independent contrasts. The semicircular canal system in mammals (Rod = rodentia, Prim = primates, Cet = cetacea, Art = artiodactyla, Car = carnivora, Chir = chiroptera) contributes to stabilization and balances during locomotion and varies positively with body size. The highly derived system of cetaceans (close symbols) seems to have evolved as an adaptation to an aquatic environment, and not taking this fact into account (i.e., pooling all species during analyses) would result in an underestimation of the allometric slope of the semicircular canal system across mammals. A regression with independent contrasts controls for this problem during scaling analyses and also extracts the region of the phylogeny where the grade shift has occurred (gray symbol), since the node separating cetaceans from their artiodactyl sister lineage falls outside the 99% prediction interval for a new observation when it is removed from the analysis. Modified, with permission, from Spoor et al. (289), the phylogeny was built employing the mammalian supertree reconstructed by Beck et al. (19).

exponent with contrasts equals $Mass^{0.117}$ (with 95% CI between 0.101 and 0.132), emphasizing that the effects of the grade shift were appropriately removed in this analysis (for other studies discussing grade shifts and differences among clades, see 3, 99, 104, 107, 197, 209, 227, 262, 277).

Data transformed into independent contrasts can be employed in tests of correlation, regression, and other linear models. Because contrasts ultimately correspond to estimates of divergence of continuous phenotypic traits, many statistical properties require some discussion. First, the sign of each contrast is arbitrary, depending solely of how sister nodes are ordered during calculations, yet it is imperative that the ordering of species remains identical during calculations of contrasts for separate traits. Otherwise, the sign of one association between two traits may change its sign if the contrast for one trait is calculated between species $A - B$ while the contrast for the other trait is computed as $B - A$. Second, correlation and regression analyses must be always calculated though the origin (Fig. 10), because the expected average phenotypic divergence equals zero under Brownian motion (failing to calculate correlations and regressions through the origin when working with contrasts is a relatively frequent mistake of many published studies) (82, 105, 175). As a result, hypothesis testing with contrast involve $N - 1$ data and no intercept, ultimately resulting in the same degrees of freedom of conventional statistics (100). Even though analyses with contrasts do not include an intercept term, intercepts can be calculated by remapping the origin to the ancestral state estimated for the root node (107), which mathematically corresponds to the phylogenetically corrected mean of x and y from phylogenetic generalized linear models (see below). Importantly, multivariate analyses with contrasts can include categorical predictor variables (104), hence standard analysis of variance (ANOVA) and covariance (ANCOVA) may also be performed with this method. In these analyses, categorical variables with n categories are included in the regression model employing $n - 1$ dummy variables of $0 - 1$, whose effect is tested together in the model (121,261). This is mathematically equivalent to a phylogenetic ANOVA or ANCOVA, hence the $n - 1$ degrees of freedom in the numerator of the F statistic of these analyses.

As with any other statistical model, analyses based on contrasts and statistical inferences will be affected by violations of underlying assumptions (115, 203, 207, 249). Importantly, some analyses are more sensitive than others to violations of underlying assumptions; hence researchers should always take into account the adequacy and reliability of these methods in the context of the study (i.e., the application being used and the question being investigated, see 160). For instance, inferred ancestral states calculated with contrasts for the root node (see below) will be affected by directional trends, such as the increase in body size during the evolution of horses, yet independent contrasts may still adequately account for phylogenetic effects and result in data that are statistically independent if the primary goal is to study the correlation between characters during the evolution of a clade

(230). Computer simulations suggest that tests of correlated evolution employing contrasts are relatively robust to deviations of Brownian motion and equivocal branch lengths (68, 69), particularly when diagnostic and remedial measures are taken prior to analyses (3, 29, 74, 93, 98, 105, 107, 109). Before we discuss these diagnostic measures, we recall that performing tests for the presence of phylogenetic signal in the comparative data and in the residual error of linear models is highly recommended (incidentally, when phylogenetic signal is absent, analyses with independent contrasts can be performed on a star phylogeny and the slope and the P-value will be mathematically identical to those of conventional analyses; see 101, 248).

When phylogenetic correction is necessary, the resulting contrasts must be normally distributed and appropriately standardized, which may be accomplished with a variety of transformations on both the phenotypic traits and/or on the branch lengths (105). For instance, body mass is log-normally distributed and such a distribution cannot be obtained from Brownian motion in the original scale, hence many comparative analyses employ $\log_{10}$-transformed estimates of body mass (Fig. 11). In addition, adequately standardized estimates of phenotypic divergence should be homogeneously distributed across short and long branches. Plotting and correlating the absolute contrasts against their standard deviation (the square root of the sum of their branch lengths) is often employed to diagnose if this is the case, and the standardization is considered appropriate if no significant correlation is observed (105). Significant negative (or positive) correlations in this diagnostic imply that contrasts calculated between closely related sister lineages are consistently higher (or lower) than those calculated for distantly related sister lineages, hence it would not be possible to determine in subsequent analyses if observed patterns reflect a biological phenomenon or an artifact resulting from inadequate standardization. Additionally, studying the distribution of the absolute contrasts is important for reasons discussed in more detail below: outliers and potentially influential points can indicate regions of the phylogeny where evolutionary divergences were particularly large, while differences in distribution of contrasts across clades would suggest that the rates of phenotypic evolution differ between them (98, 107).

Phylogenetic regression

The phylogenetic regression is a generalization of standard regression techniques developed by Grafen (124) to deal with the non-independence of comparative data. A detailed historical account of the development of phylogenetic generalized least-squares methods, which is now a routine comparative tool (e.g., see 29, 38, 74, 93, 107, 146, 206, 267, 324), is provided in Appendix A of Lavin et al. (171). While the standard regression assumes that the residual errors for each species are equal and there is no covariation between them, the generalized model includes a matrix of expected variance-covariance between residuals that is based on phylogenetic distances (Fig. 9). Data in the diagonal of this matrix provide the expected residual phenotypic variance for each species, which roughly corresponds to the total amount of evolutionary time available for each species from the root node to the tip of the tree. The degree of resemblance or covariance among species, described in the off-diagonals, corresponds to the amount of time these species evolved together as a single lineage (Fig. 9). While close relatives are expected to be very similar due to shared ancestry, distantly related species are more likely to differ given the increase in divergence time. By accounting for the expected variance-covariance between residuals, parameter estimation, and hypothesis testing in these regressions take into consideration the evolutionary relations between species (107, 124, 125, 267, 268, 299, 324). These generalized linear methods are very versatile and can be employed to perform the phylogenetically correct version of many standard comparative techniques, such as ANOVA and ANCOVA.

Importantly, under a Brownian motion model of character evolution, phylogenetic regression analyses yield parameter estimates that are numerically identical to those obtained by independent contrasts (e.g., compare Table 2 with analyses listed in Section "Phylogenetically independent contrasts"). However, in spite of being functionally equivalent for most purposes, both approaches can be nonetheless complementary (101, 107, 267). Phylogenetic thinking is retained with independent contrasts, which facilitates identification of clades with contrasting rates of phenotypic evolution (e.g., 30, 107) and the calculation of ancestral phenotypes or predicted phenotypes for unmeasured species by manipulating (rerooting) the phylogeny (107). Additionally, outliers in contrast

Table 2 Allometry of the Semicircular Canal in Mammals, Comparing Models with Different Phylogenetic Structures (Star Versus Hierarchical Phylogeny)

Evolutionary model	Intercept ± SE	Slope ± SE	AIC	w_i
Nonphylogenetic	0.011 ± 0.040	0.067 ± 0.009	−73.42	0.000
Phylogenetic	**−0.167 ± 0.083**	**0.117 ± 0.008**	**−277.78**	**1.000**

Results from phylogenetic generalized linear regressions employing $\log_{10}$-transformed data. The best-fitting model was determined with the Akaike information criterion (AIC), and corresponds to the one with the smallest AIC value shown in **bold**. Akaike weights (w_i) suggest that the model with a hierarchical structure has 100% support *given the two models tested*, hence one can conclude that the radius of the semicircular canal exhibits high phylogenetic signal and that the phylogenetic regression is more reliable. Data, with permission, from Spoor et al. (289).

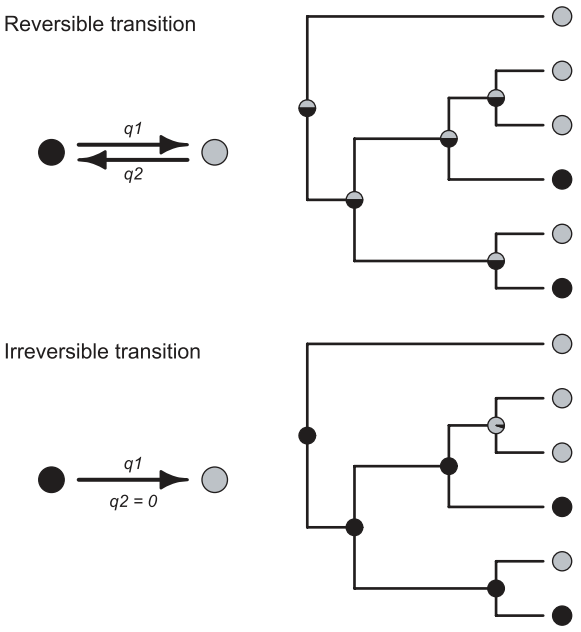


Figure 12 The general approach to study the evolution of categorical traits. Different evolutionary models can be emulated by varying the rates of transitions *q1* and *q2* between categorical states: evolution is reversible when transitions in both directions are possible (i.e., the probability that a transition occurs is different from zero for *q1* and *q2*), and irreversible when the probability of a transition in one of the directions is constrained. Different set of rules about transformations between states can have a major impact on the analytical results, as shown for the ancestral reconstruction performed with maximum likelihood assuming a Markov model of evolution (see text). In the first case, transition rates are assumed to be identical ($q1 = q2$), resulting in equiprobable ancestral states in all nodes, which contrasts dramatically with the outcome of the same analysis assuming that evolution is irreversible.

analyses may point to regions of the phylogeny where evolutionary divergence is seemingly fast or in a different direction from the general trend (Fig. 12; see also references 67, 100 [Fig. 4], 108 [Fig. 4c], 167 [Fig. 7]). Conversely, phylogenetic regressions may easily accommodate a variety of alternative evolutionary scenarios and compare their adequacy employing maximum likelihood or any other criterion for model choice (e.g., 171, 299, 323, 324).

Different traits may contain contrasting amounts of phylogenetic signal, that is, the best-fitting phylogenies would have different hierarchical structures; hence testing for the presence of phylogenetic signal in the residual variation of the regression model has been recently recommended as a diagnostic tool to assess the adequacy of the model and ensure appropriate phylogenetic correction (256). As with independent contrasts, branch length transformations can be done from a purely statistical perspective to ensure that phylogenetically structured data provide independent sources of information (68, 69, 93, 105, 124). Regression models may include independent categorical factors (e.g., diet or taxonomic group), continuous variables and phylogenies with different levels of hierarchy, including star phylogenies. Consequently, alternative regression models may be compared while also testing for the presence phylogenetic signal in the dataset (171, 299, 323). Computer programs can currently determine the optimal branch-lengths that fit the data, assuming different models of phenotypic evolution such as the OU process—a model with a stochastic component such as a random walk and a deterministic term that mimics the action of selection—and other processes that include evolutionary rates accelerating or decelerating in time and speciational models of evolutionary change (29, 36, 68, 83, 104, 131, 202, 206). This flexibility has made phylogenetic regressions particularly compelling for comparative studies of scaling, that is, estimating the parameters of allometric scaling relationships (129, 261, 262, 280, 324), comparing these estimates against theoretical predictions (280, 288, 324), or testing whether a single allometry fits better the data than separate estimations for different clades (e.g., 129, 171, 280, 299, 324).

Partition approaches

An entirely different set of statistical approaches to comparative analyses was developed based on quantitative genetics and spatial autocorrelation techniques for decoupling variation into environmental and phylogenetic components (46, 72, 114, 115, 146, 187). Importantly, these approaches differ from the previous methods in that they do not assume any *a priori* evolutionary model (such as Brownian motion or OU) and are based instead on the ability to empirically fit trait values to phylogenetic relatedness contained in the phylogeny. Although, some researchers have criticized these methods on this ground (91, 204, 267) (see also 137, p. 137), it is important to recall that phylogenetic linear models and independent contrast analyses employing branch length that are transformed primarily for statistical purposes—that is, for appropriate standardization—have the same problem since they do not necessarily reflect any known evolutionary model. For instance, Reynolds and Lee (260) raised the branch lengths estimated from DNA hybridization to the power −0.2 to obtain independent contrasts of basal metabolic rates in birds that were statistically adequate according to diagnostic tests (see Section "Phylogenetically independent contrasts"). This protocol transforms the longest branches in the shortest ones, which is difficult to reconcile with any known model of phenotypic evolution (107). Similarly, several authors have developed algorithms to assign arbitrary branch lengths into phylogenies in which divergence times are not known (124, 237, 247), which may be adequate for statistical purposes even when no evolutionary model was envisioned. Consequently, even though we emphasize that an explicit consideration of the model of character evolution behind comparative analyses is certainly beneficial (and crucial in some instances), statistical methods based on variance partitioning may be useful when analyzing real datasets due to the inherent complexity of phenotypic evolution (202).

Methods based on variance partitioning are mathematically different from independent contrasts and phylogenetic linear models (115, 267), and they usually do not outperform

model-based methods in simulations (in which evolutionary models that generated the data are known). In these analyses, which can be collectively called "partition" methods, the main idea is that trait values can be decomposed into a phylogenetic (inherited) component and a component resulting from independent evolution unique for each tip of the phylogeny (the specific component). Several techniques can be employed to achieve this, and the first approach was the autoregressive method proposed by Cheverud et al. (46) (see also 45). Employing maximum likelihood to estimate the autoregression coefficient ρ, this method maximizes the correlation between the studied phenotypic trait and a matrix representing expected phenotypic similarities predicted from phylogenetic distances. This matrix differs from the variance/covariance matrix discussed previously (Fig. 9) and is generally calculated as $1 - D$ or $1/D$, where D is the phylogenetic pairwise distance between species (scaled between 0 and 1; diagonals equal 1 given that D is by definition zero in this case). The autoregressive coefficient ρ expresses the fraction of the observed variation in the phenotypic trait explained by the expectation based solely on phylogenetic relationships, while the model fit given by the coefficient r^2 can be used to estimate phylogenetic signal. If the model fits the data well, the predicted value of the model represents the expected value of the trait under the pure historical divergence of the trait along the phylogeny (for improvements that allow taking into account more complex models, see 114, 239).

Conversely, residuals from the model (the specific component) express the deviations from phylogenetic expectations. As with independent contrasts, the specific component of different phenotypic traits can be employed to study correlated evolution, the idea being that a significant correlation between different components—that are presumably statistically independent of phylogeny—indicates that when one trait deviates from the ancestor in a given direction, the other trait tends to deviate in the same direction. Several simulation studies showed that the correlation among specific components is an estimate of the input correlation (74, 200, 202), which is also estimated by the correlation employing Felsenstein's (82) independent contrasts. Although one simulation study indicates that autocorrelation residuals often perform less well than contrasts for estimating and testing correlation between characters (248), results from the two methods are generally quite similar (104, 198, 240, 319).

The autoregressive model does not allow much flexibility because it is based on a direct relationship between the phenotypic trait studied and the phylogeny (expressed in the matrix of phylogenetic similarity), mediated exclusively by the autoregressive coefficient ρ. More complex and flexible methods have been developed (63, 146, 187), and they all ultimately attempt to decompose the variation into a phylogenetic component and a residual variation independent of phylogeny. As has been recently recommended for the residuals of phylogenetic regressions (256), assessing if residuals from the model show significant phylogenetic signal is a useful diagnostic to determine if phylogenetic effects were effectively controlled during analyses (70, 72, 114). Importantly, one proposed method, the phylogenetic eigenvector regression (72), has been correctly criticized on the grounds that it may confound variation due to phylogeny with the species-specific component (91, 267). Succinctly, the method partitions the phenotypic variation into a set of eigenvectors and eigenvalues (which reflect the nested relations in the phylogeny, from root to tips), yet how many eigenvectors are necessary to extract all phylogenetic effects remains uncertain (73). Consequently, the coefficient r^2 may not be as informative on the phylogenetic structure of the data as its relative change following the sequential introduction of eigenvectors in the analysis. Nonetheless, in spite of the mathematical shortcomings of the method, residuals may be employed for comparative analyses after they are properly tested for phylogenetic independence (see 73).

In spite of the widespread use of partition methods in areas such as conservation biology and ecology, few examples are available in the physiological literature (ironically, this seems to be the case primarily for historical reasons). On one notable exception, Withers et al. (336) employed partition methods, independent contrasts, and phylogenetic regression to analyze the allometry and environmental correlates of physiological variation in marsupials, and showed that these analyses generally—but not always—provide qualitatively similar results (see also 202, 240). Another study employed more complex variation partition analysis (described in 63) to determine to what extent phylogeny accounted for the observed variation of body mass, bone-growth rate, and resting metabolism across 13 species representing different groups of amniotes (mammals, birds, lizards, and turtles) (218). These authors showed that phylogeny accounted for a large fraction of the variation in the data (i.e., high phylogenetic signal); hence they subsequently employed independent contrasts to test for correlated evolution between these phenotypic traits (importantly, a more appropriate procedure employed nowadays involves determining the error structure that best fits the data during analyses; see 256).

Categorical and Continuous Traits

Phenotypic characters employed in phylogenetic analyses can be discrete, such as the presence or absence of a given structure, or continuous, such as body size and most morphological and physiological traits (e.g., organ size, blood pressure, enzyme activity, or metabolic rate). Even though all phylogenetic statistical methods ultimately share common concepts of phylogenetic thinking, whether characters are continuous or categorical will partly determine what kind of analytical approach is more appropriate (82, 126, 156, 191, 194, 203, 232, 234). For instance, randomization analyses to test for the presence of phylogenetic signal in categorical (196) and continuous traits (29) employ different algorithms and test statistics to estimate the amount of phenotypic evolution in each replicate—e.g., ordered or unordered maximum

parsimony may be employed to infer the minimum number of evolutionary transitions in categorical data, while the variance of independent contrasts or the mean squared error of generalized least squares can be used for continuous traits (Fig. 7). Nonetheless, these methods are conceptually very similar. Note that both analyses work primarily in the same way: the statistic of the real dataset is compared against a null distribution obtained from multiple replicates where tips have been randomly reshuffled; hence both methods ultimately test if phenotypic evolution inferred in the real dataset is significantly lower than what is observed from randomized tip data.

Similarly, for obvious reasons, models of character evolution for categorical traits differ from evolutionary models employed to analyze continuous phenotypic traits. However, all models ultimately attempt to mimic some pattern of evolutionary change. For instance, the Markov process employed to model transitions between discrete states can simulate random walks in continuous time (177, 232, 234, 273), being analogous to Brownian motion for continuous traits (see 273 for a brief discussion on the similarities between both models). This model works with the probability that, at a given time interval, a transition between character states can occur (Fig. 12), being the categorical equivalent of the evolutionary rates discussed previously for continuous phenotypic traits. Multiple variants of the Markov model can be obtained by defining different set of rules about transformations between states (190,232,234,273)—for example, symmetrical character state change can be simulated by assuming equiprobable transitions from a character state 0 to 1 or vice-versa, different probabilities will result in asymmetrical transition rates, and setting one of the probabilities of transition to zero actually mimics irreversible evolution (Fig. 12). Additionally, accelerating (or decelerating) rates of evolution for categorical transitions can be obtained by increasing (or decreasing) the probability of a transition within any given time interval. In summary, even though many of the algorithms employed to analyze comparative categorical or continuous data differ, the fundamental premises behind these different methods are quite similar. And as with continuous variables, results, and inferences may drastically change depending on the evolutionary model employed during analyses with categorical data (Fig. 12), hence a fundamental take-home message is to always keep in mind (and understand) the underlying assumptions of each statistical procedure.

Analyses with continuous and discrete variables have different strengths and limitations that should be kept in mind when designing a comparative study. For example, categorical traits that ultimately describe continuous variation in one (or many) environmental or physiological character are common practice in comparative studies (5, 121, 184, 261), even though this procedure can potentially affect the outcome and interpretation of analyses (176, 330). Categorizations tend to decrease the statistical power to detect associations as they ignore quantitative differences within categories and often involve ambiguous criteria to place species in different categories. This procedure may also result in high collinearity between phylogeny and the independent variable of study, as observed in many comparative studies that condense the variation in diet across species into categorical variables as, for instance, herbivores, omnivores, or carnivores (see 223, 261 for a detailed discussion on the subject). Furthermore, many categories reflect the interaction of multiple variables with effects that are virtually impossible to discern. For instance, differences between species from "arid" and "mesic" environments may reflect the combined effects of contrasting environmental temperatures, precipitation, and primary productivity (67, 308).

Although in these instances, it is preferable to employ quantitative estimates of phenotypic and environmental variation (49, 67, 223), qualitative differences across species may prevent the use of continuously varying characters in other occasions. For example, in studies focusing on the evolution of complex evolutionary novelties, analyses of the evolutionary sequence underlying the origin of highly derived conditions may require approaches based on categorical variables. Some interesting examples in this context involve the evolution of the swimming bladder (22, 23), heat-production organs (27) and the placenta in fishes (216, 264), the venom system in lizards and snakes (97), baleen in mysticete whales (62), electric organs in electric fishes, ballistic tongue protraction in frogs, sound localizations in owls (225), and so forth. Some studies also analyze categorical data as continuous when categories can be traced to a single phenotypic dimension (see 330). For example, Birdsey et al. (25) employed a code from 1 to 5 to classify diet of different mammalian species as carnivorous, carnivorous/omnivorous, omnivorous, omnivorous/herbivorous, and herbivorous, which described with some degree of accuracy the variation in diet across these species, and treated this variable as continuous in subsequent analyses. The same approach was employed to study the association between agility and locomotor performance in primate species and the size of the semicircular canal (a structure in the ear associated with balance in mammals), which employed categorical variables from 1 to 6 to describe species as fast, medium/fast, medium, medium/slow, slow and extra slow, and subsequently employed this variable as a continuous trait in the regression model (290).

Another important area of research that deals primarily with categorical data involves molecular evolution, since gene sequences or amino acid composition vary qualitatively. Molecular and biochemical evolution can also be analyzed with comparative approaches (120, 306), providing important insights on the evolution of physiological systems, cooptation and the origin of evolutionary novelties. In essence, changes in amino acid composition and their phenotypic effects at other levels of organization can be studied as any other categorical variable, as has been done in research of hemoglobin evolution during the colonization of high altitudes (208, 245). It is important to note that these comparative studies go one step beyond merely describing the stepwise evolution of amino acid composition along the phylogeny, primarily because the physiological basis underlying hemoglobin function is

becoming better understood. Consequently, a direct association between observed interspecific patterns and the evolutionary processes underlying these patterns (e.g., adaptation to hypoxic environments) may eventually be established (see 140,245,296). A similar approach has been employed to study the association between amino acid sequences and the evolution of hemoglobin buffer function and the Root effect in fishes (23, see Fig. 3). Other studies of protein evolution often involve methodological tools borrowed from comparative genomics and systematics, such as maximum parsimony, maximum likelihood and bootstrap analyses that are beyond the scope of this chapter (for a detailed discussion regarding these methods, see 85). Nonetheless, these methods can provide very interesting insight on the origin of evolutionary novelties at the biochemical level. Phylogenetic analyses of venom evolution suggest, for instance, a single evolutionary origin of the venom system in squamate reptiles (97) and striking similarities in protein composition of several types of venom families across a variety of taxa [for a review on the subject, see Fry et al. (96); for a phylogenetic analysis of the evolutionary origin of fangs in snakes, see Vonk et al. (312)]. Importantly, comparative studies focusing on molecular evolution share the same conceptual basis and "tree-thinking" of more traditional phylogenetic approaches such as independent contrasts, in spite of the differing analytical approaches and areas of expertise required to study biochemical evolution.

Many phylogenetic statistical methods—including independent contrasts, generalized linear models and phylogenetic eigenvector regressions—can accommodate both categorical and continuous independent variables; hence their concomitant effects may be studied in a single statistical model or by means of model comparisons (e.g., 5,171,184,238,261,299). Categorical data may be included as factors in phylogenetic regressions, or transformed into binary (0 or 1) dummy variables in analyses with contrasts (e.g., 121, 261). Some categorical or discrete traits reflect continuous variation underlying them—for example, carnivores and herbivores differ on the relative contribution of meat to their diets, which is continuous—and can be analyzed with a phylogenetic extension of the threshold model borrowed from quantitative genetics (86). This is advantageous because lineages can be polymorphic and have more than one state. Recently developed phylogenetic logistic regression (156) allows researchers to perform comparative analyses of binary dependent variables controlling for species phylogenetic affiliation, and also compare the adequacy of alternative models, that is, different set of branch lengths, independent factors and covariables included in the model, and so forth (see also 238). Analyses such as principal components for phylogenetically structured data (255) can be employed to study the evolution of quantitative traits in multivariate space, as in morphometric analyses (48, 134), while a phylogenetic paired *t*-test can be employed to study the effects of sex (males vs. females) or treatment (control vs. treatment) within each species in a phylogenetic comparative framework (179). Computer simulations provide an additional powerful and quite flexible tool to obtain phylogenetically correct null phenotypic distributions of categorical and continuous data and to study the performance of different statistical methods (e.g., see 68, 69, 104, 115, 135, 196, 202, 203, 228). In summary, a wide range of phylogenetic methodological approaches is available for hypothesis testing and evolutionary inferences based on different sorts of phenotypic data.

Designing a Comparative Study

What kind of characters can be studied with phylogenetically based statistical methods? In principle, any measurable environmental and physiological trait. Even though phylogenetic analyses implicitly assume some sort of inheritance between ancestral to descendent lineages, phenotypic characters do not have to be inherited in a conventional (genetic) sense. Environmental characteristics such as ambient temperature, humidity, soil, or water pH, are often inherited from parents to offspring (105, 181, 331, 332) and cultural traits such as language may be passed on from generation to generation by means of imitation (189, 226). Accordingly, many studies have reported significant phylogenetic signal in environmental conditions (67, 261, 333), suggesting that statistical analyses including these variables should take into account the phylogenetic relation among species. Regardless of the nature of the process behind phenotypic inheritance across generations, the fundamental point is that such inheritance may eventually result in a phylogenetically structured dataset where species data do not provide independent sources of information (28, 257). Although the adequacy of phylogenetic comparative analyses have been historically questioned given their implicit assumption that phenotypic variation partly corresponds to patterns of relatedness among species (1,265), the general *modus operandum* nowadays consists in determining which explanatory model bests fits the data while simultaneously estimating and controlling for phylogenetic signal (see 171, 256).

As occurs with any analysis, the accuracy of inferences from phylogenetic comparisons across species and/or populations will depend primarily on the quality of the comparative data. Three main aspects must be considered in this context: (i) how reliable estimates of trait values are and the potential impact of uncontrolled factors, such as measurement error, phenotypic plasticity in response to environmental conditions or sampling variation, (ii) the adequacy of the sampled species to perform hypothesis testing and to discriminate between alternative evolutionary scenarios, and (iii) the accuracy of the phylogenetic hypothesis (topology and branch lengths) employed to perform analyses. These factors will have an impact on the outcome of statistical tests and potentially on the evolutionary inferences made from the data (e.g., for contrasting inferences based on conventional versus phylogenetically informed approaches, see 4, 121, 261, 329), and their relative contribution in this context will depend on the phenotypic traits of study, the source of the data, the

information employed to infer phylogenetic relations and so forth. For instance, measurement variation in morphological traits is probably lower than for physiological or behavioral traits (29), and potential effects of experimental protocol should be more pronounced in comparative studies that have compiled their datasets from multiple sources of information (205).

Sources of phenotypic variation

Comparisons among species inherently assume that observed differences have a genetic basis, even though a multitude of other factors contribute to the observed phenotypic distribution across species. Environmental effects and measurement error can provide an important source of variation in comparative studies, as illustrated by recent analyses of interspecific variation in metabolism of birds and mammals. Comparative studies have reported significant effects of thermal acclimation on maximum thermogenic capacity of rodents (327) and of measurement protocol and captivity on avian basal metabolism (209, 210), emphasizing that researchers must consider the confounding role of environmental effects during comparative analyses. To remove potential environmental effects, it is imperative that species and/or populations are maintained under common conditions, even though some factors such as diet or age can be difficult or virtually impossible to standardize across distantly related species (100, 101). Additionally, confounding effects can be partly accounted for statistically, by including factors such as details of measurement or calculation methods that differ among studies as nuisance variables in the model (e.g., 261,323,337). Importantly, even though environmentally induced phenotypic variation is often considered a confounding factor in comparative studies, the evolution of phenotypic plastic responses across lineages is a relevant problem in itself that can be studied with a comparative approach (128, 178, 188, 222, 241, 311).

Individual variation provides another important source of phenotypic variability within a species or population. This effect has generally been dismissed in most empirical comparative studies, which consequently work with the assumption that within-species variation is negligible (for more detailed discussions on the subject, see 87, 101, 206, 266, 267). Differences among individuals and environmental effects can have an important impact on phenotypic estimates and affect the performance of statistical phylogenetic methods (see 87, 135, 157). Proposed remedies to account for this problem consist of including estimates of within-species variation (such as the variance of the mean) in analyses with phylogenetic regression and independent contrasts (87, 157). This may be also accomplished by including individuals at the tips of the phylogeny separated from each other by branches of length zero, which assume that they are independent samples from each population (87). In practice, these approaches ultimately involve lengthening the branches leading to the species/populations at the tip to compensate the inflated variance in phenotypic data due to within-species variation and sampling effects (87,157). Similarly, branch-length optimization procedures can partly circumvent the confounding effects of intraspecific variation when no data is available in this context (Figs. 8 and 9), but working with some measure of within-species variation is always preferable given the information it conveys on the quality of the comparative dataset and reliability of statistical analyses.

In summary, multiple factors account for the observed phenotypic distribution across populations and species. Phenotypic variation across individuals, differences in sample sizes, environmental conditions, and methodology across studies are a general problem that comparative biologists must deal with. In physiological studies, this problem is exacerbated because physiological traits are inherently plastic (see 161, 241 for expanded discussions on the subject), with some traits exhibiting seasonal or circadian rhythms and others being very sensitive to experimental conditions (although plasticity is the norm in physiological systems, plastic responses can contribute to overall measurement uncertainty when their effects are ignored in a comparative study). For instance, many species show a dramatic increase in free corticosterone levels during an acute stress response, presumably to mobilize glucose and react to avoid chronic stress. However, estimating and comparing acute stress responses among species is difficult because animals exhibit a sharp increase in free corticosterone levels during handling, which may result in a several-fold difference between blood samples taken within minutes in the same individuals (32). Similarly, measurement protocol can have a dramatic impact on estimates of critical thermal limits in ectothermic species (47,221,304), which has impaired broad-scale comparisons aimed at studying thermal adaptation in these lineages (see also 11, 263). The impact of uncontrolled factors is particularly relevant with the increasing amount of information available for large-scale datasets, as in recent analyses of basal and field metabolic rates of birds and mammals (40, 184, 213, 280, 287, 324). To improve the accuracy of evolutionary inferences from the data, these comparative studies must control for as many confounding variables as possible, which can be partly accomplished with more rigorous criteria for data inclusion based on the quality of phenotypic estimates—for example, estimates based on $n > 3$ individuals (210)—by including additional factors and covariables in the statistical model, and employing branch length optimization methods to remove confounding phylogenetic effects and maximize the residual variation independent from phylogeny.

Selecting species

Studies in evolutionary physiology combine different fields of knowledge, including organism physiology and its underlying functional principles with phylogenetic thinking and a historical perspective. In spite of the new and more advanced statistical methods and increased computing power, reliable evolutionary inferences ultimately depend on the adequacy of the data compiled to test alternative hypotheses. Comparative

analyses involve measurements encompassing at least a few species and can eventually comprise hundreds of species and are, needless to say, inherently limited by the range of species that can ultimately be sampled—for example, researchers may rely on data from the literature or have problems measuring species that are rare or inhabit remote areas, some experimental methods are only applicable to organisms within a limited phenotypic range, such as a small body size or fast generational times.

Besides this obvious limitation, many other aspects should be taken into consideration when choosing species for a comparative study. For example, the statistical power of phylogenetic analyses increases with sample size; therefore inferences based on datasets involving many species are generally more reliable. Although a minimum of two species/populations is required for a comparative study, any evolutionary inference based on this sort of comparison will be severely limited (100, 102, 149). Different species are expected to eventually diverge simply because they have followed separate evolutionary paths, and virtually no observation can reject this null hypothesis in the absence of additional information. For the same reason, there is a 50% probability that the divergence between two species occurs in the expected direction when testing directional hypotheses, and consequently sign tests based on multiple independent paired comparisons may be employed to test adaptive hypotheses or trait associations (82, 100, 137). For example, pairwise comparisons between closely related species of birds, mammals, turtles, and fishes suggest that hypoxia tolerance is partly explained by increased hemoglobin affinity for oxygen (144), which is also supported by comparisons among deer mice populations (283). Even though paired comparisons have appropriate type I error rates, this approach ignores divergence times among species and focuses on only a subset of possible comparisons, which entails a considerable loss of statistical power and incorrect estimates of trait associations (3, 126). In contrast, phylogenetic statistical methods such as independent contrasts impinge no loss of information.

The nonindependence of comparative data due to shared evolutionary history should be considered in every step of a comparative study, including during the choice of species (3,12,101,149,150,226). Even though increased sample sizes may improve the statistical power to detect associations, how much information is gained with the inclusion of any additional species and/or population may vary considerably depending on its phylogenetic affiliation (12, 226). In principle, evolutionary inferences based on datasets where the independent variable of study and phylogeny are orthogonal are substantially more robust than those datasets where these factors are confounded (320). Consider the differences between two comparative studies (Fig. 13), the first one testing for the effects of temperature on thermal conductance (276) and the second one for effects of diet on home-range area (104). In both studies, effects are estimated by comparing two groups of mammals, arctic versus tropical and carnivore versus herbivore species. However, hypothesized environmental effects under investigation are partly independent from phylogenetic history in the first study, while in the second any potential effect of diet will be confounded with phylogeny (Fig. 13). According to maximum parsimony, at least seven transitions between arctic and tropical environments are necessary to result in the observed pattern, corresponding to 21% of the 33 nodes of the working phylogeny. Conversely, differences in diet between carnivores and artiodactyls are highly collinear with their phylogenetic affiliation, with a single transition (2% of the 48 nodes) being enough to account for the observed distribution. Consequently, it may be difficult or even impossible to discriminate between phylogenetic affiliation and the effects of diet, in spite of the large sample of 49 species. Although some methods can improve inferential power in this type of scenario and statistically significant differences can eventually be detected (e.g., see analyses testing for diet effects in this dataset in 101), inferences should be performed with caution given the overwhelming influence of a single evolutionary event (75, 82, 105, 168, 169, 176).

On the same basis, an ideal comparative scenario to increase analytical power would include groups of organisms that are closely related (low phylogenetic variation) and experience a wide range of environmental conditions (a high environmental variation) (e.g., 25, 57, 67, 121, 172, 208, 245, 261, 276, 285, 308). Working with closely related taxa is preferred because patterns of phenotypic distribution are more likely to reflect recent evolutionary events, while analyses involving distantly related species include a longer temporal window for divergence at many different (and uncontrolled) phenotypic levels (101). This comparative design increases the statistical power (i.e., lower probability of a type II error) to detect associations between phenotypic and environmental variation, and decreases the likelihood that observed associations are spurious and reflect the effect of other factors that have not been considered in the analysis (low type I error). For example, recent analyses suggest that the allometry of basal metabolic rates varies significantly among lineages within mammals (40, 280, 324), hence scaling relationships based on distantly related species and evolutionary inferences stemming from the observed residual variation may eventually be incorrect. It is important to keep in mind that, because environmental variation and phylogeny history are often intrinsically associated, in practice it may be difficult to effectively select species with this criterion (and this is why analyses that take into account the potential confounding effects of evolutionary history are important).

Phylogenetic hypotheses

Working phylogenetic hypotheses may be obtained from a variety of different sources, which include DNA sequences, morphological or taxonomic information, the only requirements being that the information employed to build the phylogenies should be unrelated to (independent of) the phenotypic traits of study to avoid circularity and, hopefully, that it recapitulates with some degree of accuracy the

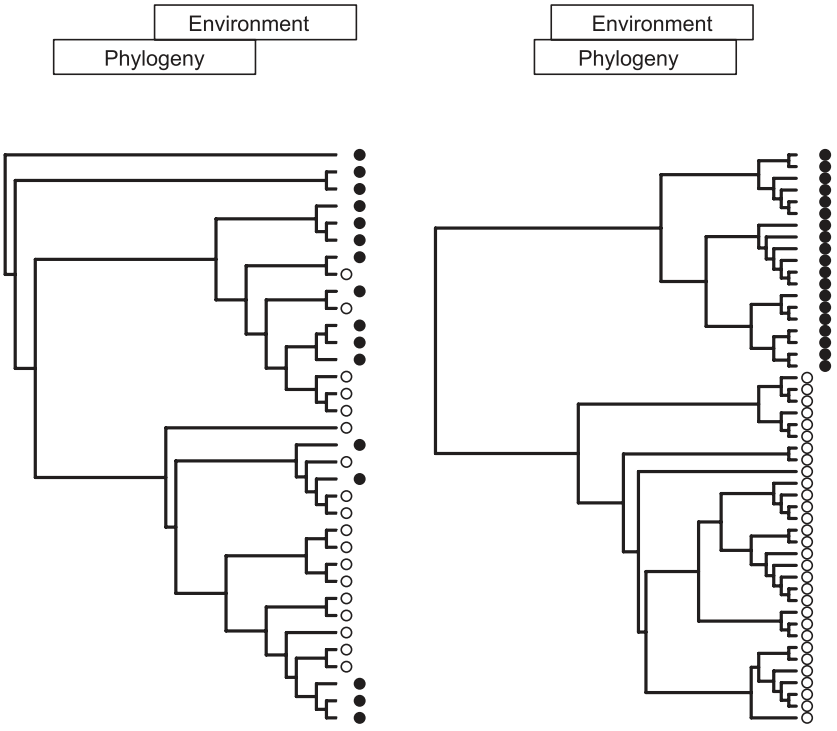


Figure 13 The statistical power to detect associations will depend on the degree of overlap between the independent variable of study and phylogeny relations among species. *Left panel*. Hypothesized phylogenetic relationships between arctic (open symbols) and tropical (black symbols) mammalian species (276), where it is relatively simple to discriminate between phylogenetic and environmental effects. *Right panel*. The worst-case scenario for a comparative analysis, because all carnivore species (black symbols) are clustered within Carnivora while all herbivores (open symbols) are ungulates (98) (adapted, with permission, from the University of Chicago Press). In this case, any potential effect of diet will be highly confounded with phylogeny.

true branching patterns during the evolution of the clade being studied. As emphasized by many authors, phylogenetic comparative analyses may perform relatively poorly when their underlying assumptions are severely violated (68,69,74,93,103,139,146,202,203,246,294,301), and virtually all phylogenetic statistical methods share the same basic assumptions that the phylogenetic topology used for analyses is correct. In addition, methods such as independent contrasts also assume that the branch lengths are known in units proportional to expected variance of character evolution. Hypothesized patterns of relationships among taxa involve an additional source of error that may affect the outcome of a comparative analysis and, as discussed above, conventional statistical methods provide no alternative to this issue given that they implicitly assume that relationships among species resemble a star phylogeny (101). Consequently, any evolutionary inference from a comparative analysis is susceptible to modification or falsification if the phylogenetic hypothesis is eventually revised.

Inaccuracies in the phylogeny fall in three broad categories: topological errors, incorrect branch lengths (i.e., inaccurate estimates of divergence and/or the assumed model of character evolution), and unresolved nodes. The impact of these factors is highly variable, as well as the strategies proposed to circumvent these sources of uncertainty. While methods to estimate optimal branch lengths—for example, by estimating λ so that the resulting branches more accurately reflect expected amounts of evolutionary change along them (Figs. 8 and 9)—have relaxed the requirements of accurate *a priori* information on divergence times among species (29, 68, 69, 93, 124, 139, 146, 202, 233, 294), gross errors in topologies clearly must be minimized. Topological errors will obviously have an impact on the outcome of analyses, yet its magnitude will depend on the nature of the error (e.g., whether disruptions are located closer to the root or the tips of the phylogeny, whether they involves multiple taxa, and so forth), the distribution of the data and the comparative method of choice (207, 294, 301). This is particularly important when taxonomic classification is employed as surrogate of phylogeny because, even within a well-studied clade such as mammals, several taxonomic groups are seemingly not monophyletic (291, 292). Another drawback of employing taxonomic information is that, depending on the information employed, the resulting phylogeny will have many unresolved nodes and branch lengths that do not necessarily correspond to divergence times. Phylogenetic information

inferred from DNA sequences should in principle be more reliable than taxonomy, hence the general strategy employed by most researchers is to build a composite tree based on phylogenies inferred from sequences reported in the literature (with the increased availability of sequencing technology and computer programs, some physiologists are building phylogenies *ad hoc* for their comparative dataset; for example, 27, 49, 159, 308). Additionally, statistical methods to infer the phylogenetic relations across large sets of species from the combination of several phylogenies (19, 24, 41) or disparate information such as gene sequences and morphological traits (61) are in ongoing development, and may be particularly useful for comparative analyses.

Despite the increasing amount of phylogenetic information available, hypothesized patterns of relatedness contain many nodes with little statistical support, and inferred relationships based on different sequences may also differ. Possible solutions depend on the degree of discrepancies between alternative topologies, and involve working with multiple topologies and analyzing the robustness of results under contrasting evolutionary scenarios (18, 141, 302, 321), including phylogenetic uncertainty in the analyses (145, 147, 148, 186, 234, 235) or leaving unresolved relations as soft polytomies in the phylogeny (103, 248). The last approach has been adopted in most comparative studies in the physiological literature, hence some practical considerations must be considered. Polytomies are nodes that lead to more than two descendent lineages, which may reflect the true evolutionary history of a clade (hard polytomy) or simply our lack of knowledge of how species of populations within this clade are actually related (soft polytomy) (124, 191, 248), and can be included in a phylogenetic hypotheses as a series of bifurcations with some branch lengths equal to zero (82, 87, 103, 248). The underlying nature of hard and soft polytomies differs, and so does their statistical repercussions. While no corrections are necessary for hard polytomies because they reflect true patterns of relatedness among species, the uncertainty associated with soft polytomies impinge on the same statistical problems of conventional statistics—to a lesser degree, though, because some phylogenetic relationships are actually known (or strictly speaking, assumed to be true) (294). Different approaches have been suggested to account for their potential impact in comparative analyses (124, 125, 145, 183, 200). For example, one possibility consists in testing the comparative hypothesis across the set of possible phylogenies that are fully resolved and weighting the results by the probability distribution of these candidate phylogenies, which can be estimated with randomization or Bayesian analyses (145, 148, 186, 234, 235). Another proposed solution involves reducing the degrees of freedom employed for hypothesis testing by an amount that corresponds to the number of unresolved relationships (i.e., internal branch lengths equal to zero; 103, 248). This approach can be very conservative, and the power to detect an association is expected to decrease when the phylogenetic hypothesis contains many soft polytomies, therefore the number of soft polytomies should be minimized (particularly in comparative analyses involving only a few species). In conclusion, although several approaches have been proposed to accommodate phylogenetic uncertainty in comparative analyses, the primary recommendation before performing a comparative study is to spend some time and effort building moderately well resolved, accurate phylogenetic hypotheses.

Comparative Analyses and Evolutionary Inferences

Phylogenetic statistical methods can be employed in a wide range of analyses, such as correlation, multiple regression, analysis of variance and covariance, canonical correlations, principal components analysis, and meta-analysis (4, 82, 104, 105, 111, 124, 166, 255, 258). Compared with conventional methods, phylogenetic analyses include a temporal component that can increase the inferential accuracy and scope of evolutionary questions studied with comparative data. For instance, analyses with contrasts allow for the objective recognition of divergent subsets in the data (e.g., the allometric grades in Fig. 11), reconstructing phenotypic attributes of hypothetical ancestors, comparing evolutionary rates across different clades and/or contrasting alternative evolutionary scenarios against the inferred temporal sequence of events. Nonetheless, as discussed previously, the evidence provided by comparative analyses is primarily correlational, hence inferences on adaptive evolution based on comparative analyses must be repeatedly contrasted against—and complemented with—information stemming from other approaches such as optimality models, experimental evolution and so forth (see Section "Pattern Versus Process"). In addition, researchers should always keep in mind how nonadaptive evolutionary processes, such as founder effects and migration, may contribute to observed trends in comparative data. Below, we discuss some of the most frequent evolutionary questions addressed with phylogenetic statistical methods, and review some comparative studies in the literature where the use of phylogenetic information has contributed to our understanding of physiological evolution.

Evolutionary adaptation

The concept of adaptation inherently involves a functional correspondence between phenotype and environment, implying that adaptive variation across taxa can be identified by seeking associations between particular traits, or set of traits, and environmental characteristics (33, 75, 77, 149). To address whether putative functional traits vary consistently with a selective regime, statistical comparative tests analyze if observed patterns of association are likely to have been produced by chance or reflect a nonrandom distribution consistent with the hypothesis of adaptation. For example, the hypothesis of increased thermogenic capacities in endothermic species inhabiting cold environments has been tested in mammals and

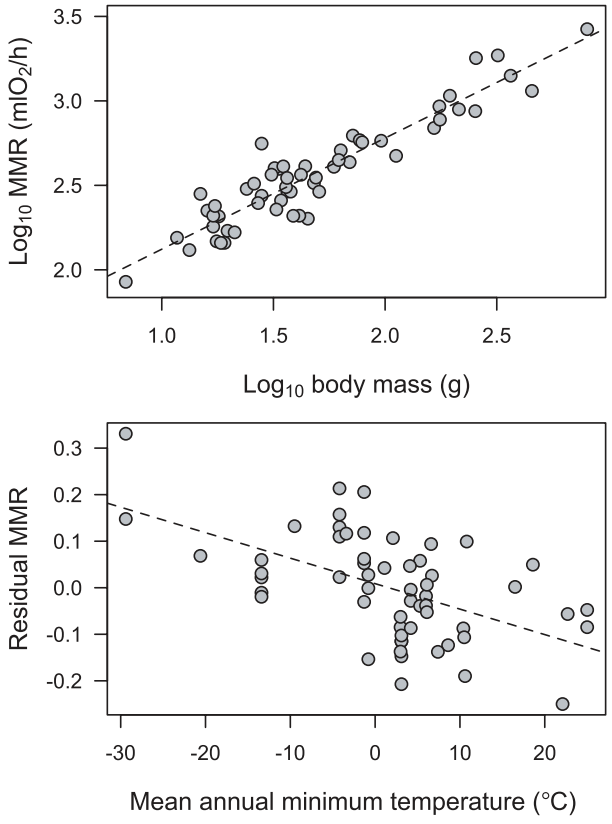


Figure 14 Relationship between maximum metabolic rates during thermogenesis (MMR), body mass, and environmental temperature for rodents across the world. The residual variation in MMR—obtained from a regular regression because this trait did not exhibit phylogenetic signal—is significantly correlated with ambient temperature, suggesting that interspecific variation in this trait may be partly explained by thermal adaptation. Modified, with permission, from Rezende et al. (261).

birds, and is supported by the significant negative association between mass-independent maximum metabolic rates and environmental temperatures in both groups (261,299,335). Consequently, cold-adapted endothermic species seem to have evolved higher thermogenic capacities than their counterparts from warmer regions (Fig. 14), which is expected from a functional perspective because heat production must compensate for increased heat loss in the cold (assuming that body temperature is similar across species and that variation in insulation cannot fully compensate for differences across thermal environments). A comparative approach can also be employed to study the evolutionary origin of presumably adaptive traits, as recently demonstrated for the evolution of defensive mechanisms in poison frogs (57). There was a recurrent association, observed in different regions of the dendrobatid phylogeny, between containing defensive alkaloids in the skin and being a diet specialist consuming primarily ants. This pattern strongly suggests that these alkaloids are exogenous, and have been coopted by many ant-eating species as a defensive mechanism in the course of their evolution.

The recent incorporation of model comparisons in phylogenetic analytical methods has provided a valuable tool to weigh the evidence supporting alternative adaptive hypotheses. The model comparison approach consists in fitting alternative models that might explain the phenotypic variation observed in a comparative dataset, and assessing which of these scenarios best fits the data (35, 127, 158, 160, 173, 185, 309, 310). Evolutionary models incorporating multiple environmental variables can be employed to weigh the evidence in favor of different adaptive scenarios (as well as different phylogenetic hypotheses to test for the presence of phylogenetic signal, see above). Consider, for example, the lower basal metabolic rates of birds and mammals inhabiting hot deserts (42, 211, 214, 307, 308, 321), which may have evolved as a response to different selective pressures: lower metabolic rates may be favored in these environments to reduce the risk of hyperthermia, or to minimize water loss. Previous comparative analyses among lark species strongly support an adaptive variation in basal metabolism associated with aridity (308), but discriminating between these hypotheses was virtually impossible given the strong correlation between maximum ambient temperatures and precipitation ($r = –0.93$, $P < 0.001$) (see discussion in 261). Employing a model comparison approach, White et al. (323) has shown that the model including maximum temperature fits the data substantially better than the model with precipitation, as suggested by its lower AIC values (Table 3). One might be tempted to conclude that low metabolic rates have evolved in desert environments as an adaptation to reduce both energy expenditure and endogenous heat load, while the reduction in water loss probably emerged as a correlated (and very convenient) response. Nonetheless, two caveats are worth emphasizing regarding this conclusion. First, these alternative hypotheses are not mutually exclusive; hence the observed interspecific variation in basal metabolism may reflect selection associated with both temperature and precipitation (in fact, the intraannual coefficient of variation of precipitation remained statistically significant when included in the model with temperature; 323). Second, inferences based on a given comparative study are limited to the dataset analyzed, and extrapolations to other taxa must be always performed with caution. Another illustrative study of this sort investigated the repeated evolution of discontinuous respiratory gas-exchange cycles (DGC) in insects (325).

Contrasting with these examples where the primary evidence supporting an adaptive explanation stems from a correlation between continuously varying physiological attributes and ecological data, other patterns may also be indicative of adaptive evolution. For example, based on some prior information suggesting that exceptional evolution might have occurred in that particular region of the phylogeny, researchers may be interested on determining whether a single species or clade has diverged from the allometric pattern observed among their close relatives (254). Single evolutionary events of disproportionally large effects may result from rapid adaptive evolution following an ecological transition in traits such as habitat or diet (215, 293), and different

Table 3 Environmental and Phylogenetic Effects on Avian Basal Metabolism, Comparing The Fit of Models With Different Independent Variables And Phylogenetic Structures (Star Versus Hierarchical Phylogeny)

Statistical model	Evolutionary model	AIC[§]	w_i
Mass	Nonphylogenetic	−120.7	< 0.001
	Phylogenetic	−150.42	< 0.001
Mass + productivity	Nonphylogenetic	−125.12	< 0.001
	Phylogenetic	−149.68	< 0.001
Mass + mean temperature	Nonphylogenetic	−137.32	< 0.001
	Phylogenetic	**−169.52**	**0.999**
Mass + temperature range	Nonphylogenetic	−120.69	< 0.001
	Phylogenetic	−150.2	< 0.001
Mass + precipitation	Nonphylogenetic	−121.67	< 0.001
	Phylogenetic	−148.89	< 0.001
Mass + precipitation variability	Nonphylogenetic	−120.79	< 0.001
	Phylogenetic	−148.53	< 0.001

Results from separate phylogenetic generalized linear regressions including $\log_{10}$-transformed basal metabolic rate of 92 populations (representing 90 species) of birds as the dependent variable. The best model to fit the data was inferred with the Akaike information criterion (AIC) and is shown in **bold**. Lower AIC values indicate models with a better fit, and the support to each model *given all models analyzed* is computed with Akaike weights (w_i). Modified, with permission, from White et al. (323).

methods can be employed to analyze whether this is the case (100, 107, 215, 229, 254, 305). These methods can address if the rates of phenotypic divergences observed during hypothesized transitions correspond to outliers falling outside the intervals observed from the remaining dataset (which provide the null distribution of how much lineages can diverge in the absence of transitions), and are applicable to hypotheses concerning the evolutionary rate along a single or multiple isolated branches in a phylogeny. The large divergence in the size and shape of the semicircular canal in cetaceans, that coincides with the colonization of aquatic environments and falls outside the prediction interval of contrasts observed for the remaining mammalian clades (Fig. 11), illustrates how this approach can be employed to test an adaptive hypothesis. A similar approach has been employed to show that the desert ringtail *Bassariscus astutus* has evolved significantly lower basal metabolic rates than predicted for its body mass, presumably due to the colonization of deserts (100, 107).

Because single or isolated evolutionary events are being compared against the general null distribution inferred from the comparative data, the branches in which exceptional divergences are hypothesized must be defined *a priori* and based on independent sources of information (215,254). In this context, the adaptive explanation proposed for the evolution of the semicircular canal stems primarily from a functional knowledge of how this structure works and its association with ecologically relevant traits such as balance during locomotion, which might be under selection during the evolutionary transition to an aquatic environment (289). The analysis with contrasts is congruent with this possibility, providing additional evidence supporting an adaptive response underlying the observed pattern. The opposite approach—that is, searching *a posteriori* for adaptive explanations to any outlier observed in the distribution of contrasts—should be avoided because, in a statistical sense, no hypothesis defined *a priori* has been tested.

Correlated evolution and scaling

Phenotypic traits do not involve independently from one another, and studying patterns of association among traits can be very important to understand the evolutionary processes underlying phenotypic variation. Correlated evolution may result in general associations between physiological and environmental traits when these traits were in fact not subject to selection (75, 169, 176). For example, the aerobic capacity model for the evolution of endothermy proposes that increased metabolic rates (and heat generation) during resting conditions have evolved in mammals and birds as a byproduct of selection on increased maximum aerobic performance (21). Even though testing this hypothesis in the ancestral lineages leading to mammals and birds may be virtually impossible, interspecific comparative analyses suggest that basal metabolic rates are associated with maximum metabolic rates during cold exposure in birds and mammals (79, 261, 262) and during sustained activity in birds (335). From a strictly physiological perspective, these results suggest a functional association between basal and maximum levels of metabolism (that may also be studied with a comparative framework; for example, see 112, 152-154). From an ecological and evolutionary perspective, analyses indicate that these traits may evolve in a correlated fashion in response to selection. Consequently, adaptive explanations based on the association between any of these traits and environmental characteristics (184, 214, 307, 321, 336) must acknowledge the possibility that observed patterns reflect correlated responses to selection (261) and, if possible, analyze the adequacy of these

alternative scenarios (for a comparative phylogenetic method that explicitly incorporates information on how traits are genetically correlated, see 142).

Another pervasive example of correlated evolution involves body size effects and allometric scaling. Many physiological, ecological and life-history traits vary predictably with body size (39, 129, 253, 275), and the form of these allometric relationships have been a subject of intense research in the comparative literature for decades. Allometric exponents obtained from comparative data have been employed to test the adequacy of different mechanistic hypotheses attempting to explain universal trends (78, 116, 117, 272, 288), to support functional associations between physiological traits at different levels of organization (26, 58, 112, 152, 153, 314-316) or to analyze potential differences in interspecific scaling among clades (40, 104, 171, 262, 280, 324). Analyses of many relationships employing phylogenetic comparative methods often differ from those of conventional analyses (107, 262), hence allometric estimates obtained with regular regressions should be reanalyzed with phylogenetic statistical tools and inferences revised accordingly (see also 227). For example, several theoretical models have postulated alternative functional explanations for the 3/4 power scaling relationship observed across many taxa (15, 16, 317, 318). In endotherms, geometric expectations based on surface area suggest that the 2/3 exponent may be more adequate (324, 326), whereas recent calculations of maximum heat dissipation capacities predict a 1/2 exponent (288). Several empirical studies have focused on the scaling exponent of basal or field metabolic rates with body mass to test the predictions of these theoretical models, which generally involves estimating the exponent from empirical data and determining if the theoretical expectation (exponents of 0.75, 0.67, or 0.50) falls within the confidence interval of these estimates (e.g., 78, 280, 288, 324, 326). Results of these studies are by no means conclusive, but it is currently evident that the phylogenetic hypotheses and the taxa employed in these analyses have a major impact on estimates of allometric relationships (280, 302, 324).

Importantly, a framework for model comparison can shed light on which evolutionary model best fits the data and, consequently, on the adequacy of different allometric exponents when testing alternative hypotheses (e.g., 6, 78, 89, 110, 171, 299, 322, 324). For instance, a comparison of AIC values obtained with ordinary least squares versus a phylogenetic regression to calculate how the semicircular canal system in mammals scales with body mass strongly suggests that results from the phylogenetic regression are more reliable (Table 2). A model comparison approach is particularly powerful because multiple scenarios can be tested simultaneously, which includes testing if the data exhibits significant phylogenetic signal (as seems to be the case with the semicircular canal system). More flexible model-fitting methods, in which the strength of phylogenetic signal in the residual variation is estimated simultaneously with the regression coefficients, go one step further and allow researchers to study allometric relationships and find the optimal branch lengths in a single analysis (171). In situations in which no single model is clearly superior to some of the other models tested, a model averaging approach can be employed to reduce the effects of model selection bias on parameter estimates (35). For example, if AIC values did not allow to discriminate which of the models in Table 2 is estimated to be best, an allometric equation employing a weighted average estimate of the slope and intercept, weighting these parameters by the Akaike weights (w_i) of each model, would provide a more robust scaling relationship.

Phylogenetic comparative analyses can also be employed to study how complex phenotypes involving multiple traits at different levels of organization evolve in response to selection. Natural selection is thought to act most directly on ecologically relevant performance abilities, with behavioral variation ultimately determining which aspects of performance may be favored (13, 102). For example, the choice to fight or flight in the presence of a predator will result in different selective regimes associated with these behaviors and potentially divergent evolutionary responses to the same selective pressure (i.e., predation). Because behavior and performance are ultimately determined by variation in morphological and physiological traits at lower levels (e.g., neuronal circuits, cardiac performance, muscle fiber types, and so on), selection should result in concerted responses and ultimately coadaptation at these lower levels. Interspecific comparative analyses in ectothermic species partly support coadaptation between thermal physiology and thermal preferences (18, 149, 151). Similar correlational analyses have shown that interspecific variation in jumping performance in anurans can be partly explained by differences in hind limb length. Pronounced differences in jumping performance and morphology across habitats suggest that these traits evolved in a correlated fashion during ecological adaptation (121), illustrating how correlated evolution at different levels can be studied from a comparative perspective to draw inferences on how complex phenotypes evolve and adapt in response to natural selection.

Rates of evolution

Phylogenetic statistical methods can also be employed to study how rates of evolutionary change differ across phenotypic traits, biological levels of organization, and/or taxonomic groups. For instance, comparisons between quantitative estimates of phylogenetic signal across traits mapped onto the same phylogeny can suggest that some traits evolve faster than others (as long as other potential sources of variation such as measurement error are not involved). Intuition leads us to expect that higher rates of evolution result in less signal, and a study that compared signal in different traits shows that body size and morphological traits, which are thought to evolve more slowly, tend to have more phylogenetic signal than physiological traits such as metabolic rates or behavioral traits such as home range (29). The same pattern is apparent in Rezende et al. (261), that reported high phylogenetic signal for body size across rodent species and no signal for

mass-independent maximum metabolic rate. However, these results must be interpreted with caution for two reasons. First, it is plausible that physiological and behavioral traits are intrinsically more variable and/or prone to higher measurement error (see above), hence differences in phylogenetic signal *per se* may not conclusively indicate that traits at different levels of organization evolve at different rates (note that these alternatives are not mutually exclusive). And second, the relationship between phylogenetic signal and evolutionary rates is complex and depends on the model of character evolution—for example, in some models such as random walk, the amount of phylogenetic signal is independent of the rates of evolutionary change (Fig. 6), while in other models phylogenetic signal decreases with increased evolutionary rates (2, 257). Another aspect to keep in mind is that phenotypic evolution may resemble a random walk within some limits—for example, hummingbirds and shrews are probably close to the smallest viable body size for a continuously endothermic animal because endothermy may not be sustained at substantially smaller sizes (243)—which would be reached more often by fast evolving traits (the result being that the evolutionary model would deviate from strict Brownian motion). Thus, even though one might intuitively expect that traits evolving faster would consistently show low phylogenetic signal, this is not always the case.

Statistical analyses of evolutionary rates can also be employed to study how character evolution proceeded in time, supporting different scenarios of adaptive phenotypic evolution (36, 199). One example involves contrasting the magnitude of phenotypic divergence across sister lineages to determine if evolution proceeded unusually fast in specific regions of the phylogeny. This can be readily done with phylogenetic independent contrasts (98,215) or using likelihood (229,305). Fast evolutionary rates may reflect adaptive responses to directional selection, as explained above for the desert ringtail population and the evolution of the semicircular canal in cetaceans. Similarly, in adaptive radiations, phenotypic evolution may proceed very fast in early stages and decrease in subsequent stages, as the available niches become occupied (29, 118). This matches a decelerating model of phenotypic evolution, where rates of evolution decrease toward the present. Employing phylogenetic methods and model comparison analytical tools, a recent study has shown that the observed interspecific variation in body size in different taxa (squamates, birds, fishes, insects, mammals, and amphibians) provides little support to the early burst model of adaptive radiation (136). Instead, analyses show that body size evolution across these groups was more or less continuous through time, supporting evolutionary models that assume a random walk and/or a selective optimal (OU model of stabilizing selection). This study provides a good example of the degree of flexibility of modern comparative approaches, and the scope of new venues of research that are only possible with a phylogenetic perspective (see also 53). Similar analyses may eventually shed light on how the origin of morphological and physiological novelties, such as wings, lungs, or eyes, might be associated with the occupation of new niches, and consequently with the radiation and evolutionary success of certain lineages (e.g., bats or birds). For example, a comparative analysis of phenotypic divergence during the radiation of dusky salamanders (genus *Desmognathus*) shows that the evolution of species with aquatic larvae was followed by a high rate of lineage accumulation (163), suggesting an association between phenotypic variation and diversification in this group (see also 52, 94, 155, 195, 250, 252, 286).

Another level of inquiry involves comparing rates of phenotypic evolution across taxa (98, 229). These analyses are on one hand necessary to ensure the appropriate standardization of the data—otherwise, phylogenetic effects may not be appropriately removed, entailing the statistical problems described above (105, 107)—while on the other hand they are informative of the distribution of the phenotypic data across the studied dataset (which may be employed to test alternative evolutionary hypotheses 229). One well-known example involves the differences in evolutionary rates in body mass observed between passerine and nonpasserine birds (107, 210, 262), which result in a more homogeneous distribution of body sizes across passerines than nonpasserines (Fig. 15). The difference between groups is evident when absolute standardized independent contrasts within passerines and nonpasserines are compared, with the former group having significantly lower evolutionary rates (as explained above, phylogenetic independent contrasts ultimately express rates of phenotypic divergence through time). It is important to notice that differences in evolutionary rates between groups ultimately result in differences in *dispersion* of phenotypic values, but not necessarily in *mean values* across these groups. Passerines are on average smaller than nonpasserines according to a phylogenetic ANCOVA (2-tailed $P = 0.037$), but this pattern resulted primarily from the large basal divergence between passerines and their nonpasserine sister group, which should be mathematically independent from the other contrasts if the phylogenetic correction is adequate. Also note that, this interpretation relies on the assumption that the branch-length estimates obtained from DNA hybridization (279) are accurate estimations of divergence times, which may not necessarily be the case because rates of DNA evolution can vary significantly across clades. Regardless of the evolutionary processes underlying this pattern, diagnostic analyses testing whether evolutionary rates differ between clades is advisable to ensure that effects within each clade is appropriately weighted in subsequent analyses.

Reconstruction of ancestral states

Ancestral reconstruction is another aspect of comparative biology and physiology that—by definition—requires a phylogenetic perspective, which can subsequently be used to study a wide variety of evolutionary problems (e.g., quantifying and testing for character change along the branches; see 151,259). A large number of methods are available for reconstructing ancestral attributes, and generally fall within two broad

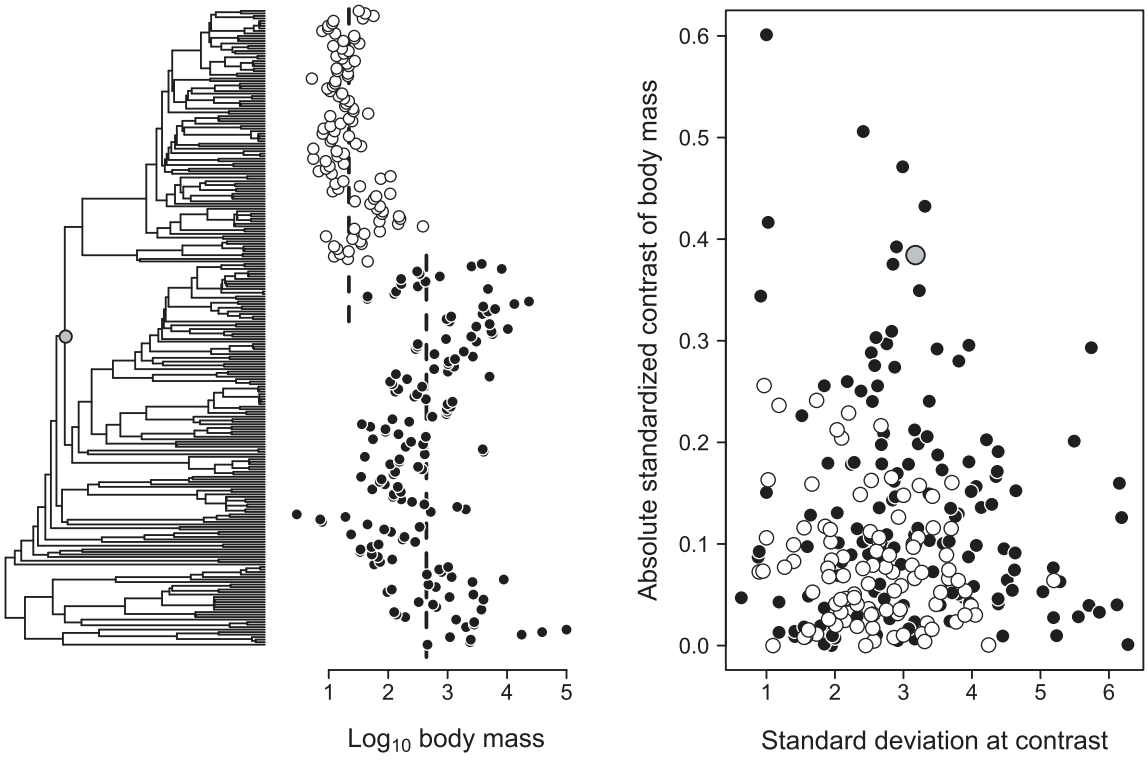


Figure 15 Phylogenetic analyses can detect differences in rates of phenotypic evolution between clades, as illustrated for body mass in passerines (open symbols) and nonpasserines (black symbols). Passerines show a more homogeneous distribution (i.e., lower variance) of body mass than nonpasserines, which is supported by statistical comparisons between absolute standardized contrasts ($P < 0.001$) after excluding the contrast between passerines and their nonpasserine sister clade (gray symbol). Assuming that branch lengths accurately reflect elapsed times, this pattern supports a lower rate of evolution of body mass in passerines. Modified from Garland and Ives (107), with permission of the University of Chicago Press, original data from Reynolds and Lee (260).

categories, parsimony, and model-based. These methods can be employed to reconstruct ancestral phenotypes for both continuous and categorical traits. However, these approaches are fundamentally different, they differ greatly in their assumptions and can lead to contrasting reconstructions. Parsimony analyses attempt to minimize the amount of evolutionary change mapped onto the phylogeny of the studied clade, based on the principle of minimum evolution (or maximum parsimony), searching for the combination of ancestral values that minimizes statistics such as the sum of squared evolutionary changes across branches for continuous traits, or the number of transitions for categorical traits (37, 56, 190, 192, 193, 300). Conversely, model-based methods calculate the likelihood of alternative ancestral character states, estimated as the probability that each distribution would have been generated by a specific evolutionary model on the tree (190, 219, 273). Even though the rationale behind parsimony and maximum likelihood is considerably different, in some circumstances these analyses yield very similar results. For example, for quantitative phenotypic traits, ancestral character estimates obtained with weighted squared changes parsimony converge to values obtained with maximum likelihood under a Brownian motion model of evolution (192, 201, 206). One major advantage of maximum likelihood, however, is that this method also provides estimates of the standard errors of those ancestral states, such as confidence intervals when traits are continuous or the likelihood of each character state in analyses with categorical traits (e.g., see Fig. 12).

In addition to these two general approaches, independent contrasts and phylogenetic generalized least square regressions can also be employed to estimate ancestral values of continuously varying traits, resulting in estimates that are highly correlated to those with weighted parsimony and maximum likelihood when the evolutionary model resembles Brownian motion (107, 109, 201, 206). The ancestral trait value for the root (basal) node of the phylogeny calculated with independent contrasts is identical to the phylogenetic mean calculated with generalized least-square methods (107, 206) and also corresponds to the most parsimonious (optimal) value of the common ancestor of the clade studied, according to weighted squared changes parsimony (109, 192). Ancestral trait values for the remaining nodes, on the other hand, are not optimal and merely reflect intermediate steps in the tip-down algorithm employed to calculate contrasts (205).

For this reason, a rerooting procedure—that is, placing the intermediate node of interest in the base of the phylogeny while maintaining the patterns of relatedness among the remaining nodes unchanged—is advisable when performing ancestral reconstruction with contrasts (107, 109). This procedure is adequate from a statistical perspective because phenotypic evolution is assumed to proceed as a random walk and, consequently, the amount of evolutionary change across lineages depends solely on the divergence times among them and is independent of where the phylogeny is rooted. A related problem that can be addressed with this approach is predicting the value of a hypothetical unmeasured species (extant or extinct) with prediction intervals (107). Importantly, neither independent contrasts nor phylogenetic regressions are suitable to estimate ancestral states of categorical traits, even though both methods may include categorical traits in linear analyses and contrasts can even provide estimates of node values. These values are, strictly speaking, statistical averages computed as part of the algorithm, which may be interpreted as ancestral character values in very particular conditions. For example, a phylogenetic regression including a binary character 0 or 1 that codes for diet (herbivory vs. carnivory) as an independent variable corresponds mathematically to an ANOVA between these groups where the data is phylogenetically structured, and yields identical results to contrast analyses calculated for this character (101, 107). However, the trait value at the root node will correspond to the phylogenetic mean and fall between 0 and 1, which is not an estimate of the diet of the hypothetical last common ancestor of the species in the dataset (strictly speaking, in this example the ancestral trait value might be interpreted as a probability of the ancestor being a carnivore or a herbivore, but, in analyses including multiple categories, these estimates are really meaningless).

Given that estimates of ancestral character states roughly correspond to mean values weighted by the phylogenetic distance, it is clear that directional trends will have a major impact on the accuracy of ancestral estimates (88, 109, 230). Computer simulations show that ancestral reconstructions with parsimony, generalized linear methods and maximum likelihood can be fairly accurate when evolutionary models resemble a random walk, but suggest that results may not be so reliable if traits are under directional or stabilizing selection during their evolution (201,230). Even under a Brownian motion model of phenotypic evolution, the confidence intervals provided by methods such as maximum likelihood and independent contrasts often exceed the range of variation observed in extant species (109,242). Researchers should therefore keep in mind that results of multivariate analyses employing contrasts and phylogenetic generalized least squares are substantially more robust to deviations of Brownian motion and sampling effects, being able to adequately account for the nonindependence of the data due to phylogenetic history while failing to provide accurate estimates of ancestral trait values (see 230). Nonetheless, detecting directional trends in comparative data is possible with other phylogenetic approaches (233) and ancestral values can be estimated with increased accuracy incorporating data from the fossil record (88, 170). Even though most physiological and behavioral traits do not leave any traces on fossilized specimens, recent studies have applied the use of proxies commonly preserved in fossils to study physiological and behavioral evolution. For instance, the reduction of the semicircular canal during the evolution of whales and dolphins (Fig. 11) is evident in the fossil record and predates the evolution of the modern cetacean body plan (289). Because the size of the canal is associated with balance in other mammals and is significantly correlated with agility and locomotor performance in extant primates, this trait can be employed to estimate with some degree of accuracy whether ancestral species in the fossil record were agile or not (290). Similarly, patterns of deposition of bone layers, ontogenetic growth trajectories obtained from fossil bones or the presence of nasal turbinates have been proposed as proxies for increased metabolic rates, which may be employed to study the evolution of endothermy in the avian and mammalian lineages (113,218,270). Since this sort of evidence is indirect, comparative studies of physiological evolution employing fossil data must always (i) establish a functional association between the morphological proxies and the physiological/behavioral traits of interest and (ii) estimate the degree of association between these traits and determine if they are indeed significantly related in extant species. Even when these conditions are met, researchers must keep in mind the assumptions and limitations associated with this sort of approach (e.g., that relationships observed in extant species hold for extinct lineages).

To summarize, most methods developed to estimate ancestral trait values or characters can perform relatively well in simple evolutionary scenarios that encompass just a few transitions for categorical traits or when the phenotypic evolution of a quantitative trait actually resembles a random walk (201, 242). The origin of complex traits can often be studied with these approaches, as elegantly demonstrated for the stepwise phenotypic evolution of baleen in mysticete whales from an ancestor with mineralized teeth (62), which suggests that the hypothetical common ancestor of this lineage may have presented a transitional form with the simultaneous occurrence of teeth and baleen (the same applies to the evolution of the swim bladder in teleosts, as shown in Fig. 3). In other evolutionary scenarios that involve multiple evolutionary transitions, reversions and more complex models of phenotypic evolution, it is debatable to what extent ancestral states inferred from phylogenetic analyses are reliable (119,180,182,273,313,328,334). Consequently, results from ancestral reconstruction obtained with these methods should always be contrasted against information from complementary sources, such as the fossil record (62, 170, 242, 289), biogeographical patterns (278) and developmental evidence (62, 312), and importantly, estimates of error on reconstructions should always be provided.

Conclusion

Comparative analyses have been radically restructured during the last decades, with the development of new statistical methods that incorporate phylogenetic information and explicit models of phenotypic evolution into the design and analysis of interspecific data. This revolution in comparative methodology has had repercussions in virtually all areas of biology, with phylogenetic statistical approaches being currently employed on a regular basis in seemingly disparate fields such as molecular biology, physiological and community ecology, animal behavior, conservation biology, and linguistics. Phylogenetic statistical methods currently come in many flavors; they can be implemented in a large variety of statistical packages and are becoming increasingly flexible to deal with different types of data, evolutionary models, phylogenetic uncertainty, and so forth (see Section "Further Reading"). Nonetheless, in spite of the methodological advances resulting from the advent of the phylogenetic comparative methodology, understanding the conceptual framework behind these methods remains a condition *sine qua non* for adequate interpretation of results following a comparative analysis (320).

As outlined by Blomberg and Garland (28), several aspects are general to all phylogenetic analyses and should be considered in every step of a comparative study. First, adaptation by natural selection should not be casually inferred from comparative data, taxa are expected to differ simply because of their separate evolutionary paths hence *a posteriori* adaptive explanations based solely on observed differences should be generally avoided. Second, given the hierarchical patterns of relatedness among organisms due to evolutionary history, comparative data cannot be assumed to be independent as required by conventional statistical methods. Whether phylogenetic or conventional analyses are more appropriate will depend both on the question being investigated and the amount of phylogenetic signal detected in the data. Third, comparative methods allow explicit inferences about evolutionary history that are not possible with traditional statistical analyses (e.g. studying the origin of a phenotypic character or inferring the character state of hypothetical ancestors). Fourth, phylogenetic analyses make explicit assumptions about character evolution, which will have an impact on the outcome of a comparative study. That is, most phylogenetic statistical methods ultimately deal with how and how fast phenotypic traits evolve in time (i.e., evolutionary models and evolutionary rates of phenotypic evolution). Fifth, taxa used in comparative analyses should be chosen in regard to their phylogenetic affinities as well as the area of functional investigation. Finally, phylogenetically based comparisons remain strictly correlational, therefore inferences on the evolutionary processes underlying observed patterns must always be contrasted against information stemming from complementary approaches, that include functional and mechanistic studies, experimental evolution, etc.

On a broader context, we believe the incorporation of an explicit evolutionary perspective into comparative physiological studies has been highly beneficial, not only from a practical perspective but also (primarily) from a conceptual point of view. We encourage colleagues and students concerned about the intricacies of phylogenetically comparative methods to look beyond the technicalities of the evolutionary jargon and the inherent complexity of statistical procedures, the fundamental goal of these analyses is to obtain a more accurate representation of reality: organisms evolve, and all their functional characters are the products of evolution. The diversity of form and function of living organisms is the result of millions of years of continuous change, and the incorporation of phylogenies in the comparative method continually reminds us of this fact.

Acknowledgements

We are very grateful to our colleagues who kindly shared their data for inclusion in this chapter, three anonymous reviewers and Ted Garland for comments and suggestions that have greatly improved the clarity, depth and scope of this chapter. E.L.R. is supported by a Ramón y Cajal contract and by grant BFU2009-07564 from the Ministerio de Ciencia e Innovación (Spain), and J.A.F.D.-F. is supported by several research grants from CNPq and CAPES. Financial support by grant 2009SGR 636 from Generalitat de Catalunya to the Grup de Biologia Evolutiva is also gratefully acknowledged.

References

1. Abouheif E. A method for testing the assumption of phylogenetic independence in comparative data. *Evol Ecol Res* 1: 895-909, 1999.
2. Ackerly D. Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proc Natl Acad Sci U S A* 106: 19699-19706, 2009.
3. Ackerly DD. Taxon sampling, correlated evolution, and independent contrasts. *Evolution* 54: 1480-1492, 2000.
4. Adams DC. Phylogenetic meta-analysis. *Evolution* 62: 567-572, 2008.
5. Al-kahtani MA, Zuleta C, Caviedes-Vidal E, Garland T. Kidney mass and relative medullary thickness of rodents in relation to habitat, body size, and phylogeny. *Physiol Biochem Zool* 77: 346-365, 2004.
6. Alerstam T, Rosen M, Backman J, Ericson PGP, Hellgren O. Flight speeds among bird species: allometric and phylogenetic effects. *Plos Biol* 5: 1656-1662, 2007.
7. Alexander HJ, Taylor JS, Wu SST, Breden F. Parallel evolution and vicariance in the guppy (Poecilia reticulata) over multiple spatial and temporal scales. *Evolution* 60: 2352-2369, 2006.
8. Alfaro ME, Holder MT. The posterior and the prior in Bayesian phylogenetics. *Annu Rev Ecol Evol S* 37: 19-42, 2006.
9. Alfaro ME, Huelsenbeck JP. Comparative performance of Bayesian and AIC-based measures of phylogenetic model uncertainty. *Syst Biol* 55: 89-96, 2006.
10. Alstrom P, Ericson PGP, Olsson U, Sundberg P. Phylogeny and classification of the avian superfamily Sylvioidea. *Mol Phylogenet Evol* 38: 381-397, 2006.
11. Angilletta MJ. *Thermal Adaptation: A Theoretical and Empirical Synthesis* New York: Oxford University Press, 2009.
12. Arnold C, Nunn CL. Phylogenetic targeting of research effort in evolutionary biology. *Am Nat* 176: 601-612, 2010.
13. Arnold SJ. Morphology, performance and fitness. *Am Zool* 23: 347-361, 1983.
14. Ashton KG. Comparing phylogenetic signal in intraspecific and interspecific body size datasets. *J Evol Biol* 17: 1157-1161, 2004.
15. Banavar JR, Damuth J, Maritan A, Rinaldo A. Ontogenetic growth: modelling universality and scaling. *Nature* 420: 626-626, 2002.
16. Banavar JR, Moses ME, Brown JH, Damuth J, Rinaldo A, Sibly RM, Maritan A. A general basis for quarter-power scaling in animals. *Proc Natl Acad Sci U S A* 107: 15816-15820, 2010.

17. Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J. Phylogeny and diversification of the largest avian radiation. *Proc Natl Acad Sci U S A* 101: 11040-11045, 2004.
18. Bauwens D, Garland T, Castilla AM, Vandamme R. Evolution of sprint speed in lacertid lizards: morphological, physiological, and behavioral covariation. *Evolution* 49: 848-863, 1995.
19. Beck RMD, Bininda-Emonds ORP, Cardillo M, Liu FGR, Purvis A. A higher-level MRP supertree of placental mammals. *BMC Evol Biol* 6: 93, 2006.
20. Bennett AF. Adaptation and the evolution of physiological characters. In: Dantzler WH, editor. *Handbook of Physiology: Comparative Physiology*, New York: Oxford University Press, 1997, sect. 13, pp. 3-16.
21. Bennett AF, Ruben JA. Endothermy and activity in vertebrates. *Science* 206: 649-654, 1979.
22. Berenbrink M. Historical reconstructions of evolving physiological complexity: O_2 secretion in the eye and swimbladder of fishes. *J Exp Biol* 210: 1641-1652, 2007.
23. Berenbrink M, Koldkjaer P, Kepp O, Cossins AR. Evolution of oxygen secretion in fishes and the emergence of a complex physiological system. *Science* 307: 1752-1757, 2005.
24. Bininda-Emonds ORP, Gittleman JL, Purvis A. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biol Rev* 74: 143-175, 1999.
25. Birdsey GM, Lewin J, Holbrook JD, Simpson VR, Cunningham AA, Danpure CJ. A comparative analysis of the evolutionary relationship between diet and enzyme targeting in bats, marsupials and other mammals. *Proc R Soc B* 272: 833-840, 2005.
26. Bishop CM. The maximum oxygen consumption and aerobic scope of birds and mammals: getting to the heart of the matter. *Proc R Soc B* 266: 2275-2281, 1999.
27. Block BA, Finnerty JR, Stewart AFR, Kidd J. Evolution of endothermy in fish: mapping physiological traits on a molecular phylogeny. *Science* 260: 210-214, 1993.
28. Blomberg SP, Garland T. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J Evol Biol* 15: 899-910, 2002.
29. Blomberg SP, Garland T, Ives AR. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57: 717-745, 2003.
30. Bonine KE, Gleeson TT, Garland T. Comparative analysis of fiber-type composition in the lliofibularis muscle of Phrynosomatid lizards (Squamata). *J Morphol* 250: 265-280, 2001.
31. Bozinovic F, Rosenmann M. Maximum metabolic-rate of rodents: physiological and ecological consequences on distributional limits. *Funct Ecol* 3: 173-181, 1989.
32. Breuner CW, Lynn SE, Julian GE, Cornelius JM, Heidinger BJ, Love OP, Sprague RS, Wada H, Whitman BA. Plasma-binding globulins and acute stress response. *Horm Metab Res* 38: 260-268, 2006.
33. Brooks DR, McLennan DA. *Phylogeny, Ecology and Behavior: A Research Program in Comparative Biology*. Chicago: The University of Chicago Press, 1991.
34. Buchwalter DB, Cain DJ, Martin CA, Xie L, Luoma SN, Garland T. Aquatic insect ecophysiological traits reveal phylogenetically based differences in dissolved cadmium susceptibility. *Proc Natl Acad Sci U S A* 105: 8321-8326, 2008.
35. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* New York: Springer, 2002.
36. Butler MA, King AA. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am Nat* 164: 683-695, 2004.
37. Butler MA, Losos JB. 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: 1623-1635, 1997.
38. Butler MA, Schoener TW, Losos JB. The relationship between sexual size dimorphism and habitat use in Greater Antillean *Anolis* Lizards. *Evolution* 54: 259-272, 2000.
39. Calder WA. *Size, Function, and Life-History*. Cambridge: Harvard University press, 1996.
40. Capellini I, Venditti C, Barton RA. Phylogeny and metabolic scaling in mammals. *Ecology* 91: 2783-2793, 2010.
41. Cardillo M, Bininda-Emonds ORP, Boakes E, Purvis A. A species-level phylogenetic supertree of marsupials. *J Zool* 264: 11-31, 2004.
42. Careau V, Morand-Ferron J, Thomas D. Basal metabolic rate of canidae from hot deserts to cold arctic climates. *J Mammal* 88: 394-400, 2007.
43. Chappell MA, Hayes JP, Snyder LRG. Hemoglobin polymorphisms in deer mice (*Peromyscus maniculatus*): physiology of beta-globin variants and alpha-globin recombinants. *Evolution* 42: 681-688, 1988.
44. Chappell MA, Snyder LRG. Biochemical and physiological correlates of deer mouse alpha-chain hemoglobin polymorphisms. *Proc Natl Acad Sci U S A* 81: 5484-5488, 1984.
45. Cheverud JM, Dow MM. An auto-correlation analysis of genetic variation due to lineal fission in social groups of *Rhesus* macaques. *Am J Phys Anthropol* 67: 113-121, 1985.
46. Cheverud JM, Dow MM, Leutenegger W. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body-weight among primates. *Evolution* 39: 1335-1351, 1985.
47. Chown SL, Jumbam KR, Sorensen JG, Terblanche JS. Phenotypic variance, plasticity and heritability estimates of critical thermal limits depend on methodological context. *Funct Ecol* 23: 133-140, 2009.
48. Claude J, Pritchard PCH, Tong HY, Paradis E, Auffray JC. Ecological correlates and evolutionary divergence in the skull of turtles: a geometric morphometric assessment. *Syst Biol* 53: 933-948, 2004.
49. Clauss M, Hofmann RR, Fickel J, Streich WJ, Hummel J. The intraruminal papillation gradient in wild ruminants of different feeding types: implications for rumen physiology. *J Morphol* 270: 929-942, 2009.
50. Clusella-Trullas S, Terblanche JS, Blackburn TM, Chown SL. Testing the thermal melanism hypothesis: a macrophysiological approach. *Funct Ecol* 22: 232-238, 2008.
51. Coddington JA. Cladistic tests of adaptational hypotheses. *Cladistics* 4: 3-22, 1988.
52. Collar DC, Near TJ, Wainwright PC. Comparative analysis of morphological diversity: does disparity accumulate at the same rate in two lineages of centrarchid fishes? *Evolution* 59: 1783-1794, 2005.
53. Cooper N, Purvis A. Body size evolution in mammals: complexity in tempo and mode. *Am Nat* 175: 727-738, 2010.
54. Corl A, Davis AR, Kuchta SR, Comendant T, Sinervo B. Alternative mating strategies and the evolution of sexual size dimorphism in the side-blotched lizard, *Uta stansburiana*: a population-level comparative analysis. *Evolution* 64: 79-96, 2010.
55. Cubo J, Ponton F, Laurin M, de Margerie E, Castanet J. Phylogenetic signal in bone microstructure of sauropsids. *Syst Biol* 54: 562-574, 2005.
56. Cunningham CW, Omland KE, Oakley TH. Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol Evol* 13: 361-366, 1998.
57. Darst CR, Menendez-Guerrero PA, Coloma LA, Cannatella DC. Evolution of dietary specialization and chemical defense in poison frogs (Dendrobatidae): a comparative analysis. *Am Nat* 165: 56-69, 2005.
58. Darveau CA, Suarez RK, Andrews RD, Hochachka PW. Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* 417: 166-170, 2002.
59. Davis EB. Comparison of climate space and phylogeny of *Marmota* (Mammalia: Rodentia) indicates a connection between evolutionary history and climate preference. *Proc R Soc B* 272: 519-526, 2005.
60. de Queiroz A, Ashton KG. The phylogeny of a species-level tendency: species heritability and possible deep origins of Bergmann's rule in tetrapods. *Evolution* 58: 1674-1684, 2004.
61. de Queiroz A, Gatesy J. The supermatrix approach to systematics. *Trends Ecol Evol* 22: 34-41, 2007.
62. Demere TA, Mcgowen MR, Berta A, Gatesy J. Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales. *Syst Biol* 57: 15-37, 2008.
63. Desdevises Y, Legendre P, Azouzi L, Morand S. Quantifying phylogenetically structured environmental variation. *Evolution* 57: 2647-2652, 2003.
64. Dial KP. Wing-assisted incline running and the evolution of flight. *Science* 299: 402-404, 2003.
65. Dial KP, Jackson BE, Segre P. A fundamental avian wing-stroke provides a new perspective on the evolution of flight. *Nature* 451: U985-U983, 2008.
66. Dial KP, Randall RJ, Dial TR. What use is half a wing in the ecology and evolution of birds? *Bioscience* 56: 437-445, 2006.
67. Diaz GB, Ojeda RA, Rezende EL. Renal morphology, phylogenetic history and desert adaptation of South American hystricognath rodents. *Funct Ecol* 20: 609-620, 2006.
68. Diaz-Uriarte R, Garland T. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Syst Biol* 45: 27-47, 1996.
69. Diaz-Uriarte R, Garland T. Effects of branch length errors on the performance of phylogenetically independent contrasts. *Syst Biol* 47: 654-672, 1998.
70. Diniz-Filho JAF. Phylogenetic autocorrelation under distinct evolutionary processes. *Evolution* 55: 1104-1109, 2001.
71. Diniz-Filho JAF. Phylogenetic diversity and conservation priorities under distinct models of phenotypic evolution. *Conserv Biol* 18: 698-704, 2004.
72. Diniz-Filho JAF, De Sant'ana CER, Bini LM. An eigenvector method for estimating phylogenetic inertia. *Evolution* 52: 1247-1262, 1998.
73. Diniz-Filho JAF, Bini LM, Rangel TF, Morales-Castilla I, Olalla-Tarraga MA, Rodriguez MA, Hawkins BA. On the selection of phylogenetic eigenvectors for ecological analyses. *Ecography* (in press), 2011. DOI: 10.1111/j.1600-0587.2011.06949.x.

74. Diniz-Filho JAF, Torres NM. Phylogenetic comparative methods and the geographic range size: body size relationship in new world terrestrial carnivora. *Evol Ecol* 16: 351-367, 2002.
75. Doughty P. Statistical analysis of natural experiments in evolutionary biology: comments on recent criticisms of the use of comparative methods to study adaptation. *Am Nat* 148: 943-956, 1996.
76. Dumont JPC, Robertson RM. Neuronal circuits: an evolutionary perspective. *Science* 233: 849-853, 1986.
77. Dunbar RIM. Adaptation, fitness, and the evolutionary tautology. In: Sociobiology KsC., editor. *Current Problems in Sociobiology*, Cambridge: Cambridge University Press, 1982, pp. 9-28.
78. Duncan RP, Forsyth DM, Hone J. Testing the metabolic theory of ecology: allometric scaling exponents in mammals. *Ecology* 88: 324-333, 2007.
79. Dutenhoffer MS, Swanson DL. Relationship of basal to summit metabolic rate in passerine birds and the aerobic capacity model for the evolution of endothermy. *Physiol Zool* 69: 1232-1254, 1996.
80. Elgar MA, Harvey PH. Basal metabolic rates in mammals: allometry, phylogeny and ecology. *Funct Ecol* 1: 25-36, 1987.
81. Feder ME, Bennett AF, Huey RB. Evolutionary physiology. *Annu Rev Ecol Syst* 31: 315-341, 2000.
82. Felsenstein J. Phylogenies and the comparative method. *Am Nat* 125: 1-15, 1985.
83. Felsenstein J. Phylogenies and quantitative characters. *Annu Rev Ecol Syst* 19: 445-471, 1988.
84. Felsenstein J. Contrasts for a within-species comparative method. In: Slatkin M, Veuille M, editors. *Modern Developments in Theoretical Population Genetics*. Oxford: Oxford University Press, 2002, pp. 118-129.
85. Felsenstein J. *Inferring Phylogenies*. Sunderland, MA: Sinauer Associates, Inc, 2004.
86. Felsenstein J. Using the quantitative genetic threshold model for inferences between and within species. *Philos Trans R Soc B* 360: 1427-1434, 2005.
87. Felsenstein J. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *Am Nat* 171: 713-725, 2008.
88. Finarelli JA, Flynn JJ. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Syst Biol* 55: 301-313, 2006.
89. Finarelli JA, Flynn JJ. Brain-size evolution and sociality in Carnivora. *Proc Natl Acad Sci U S A* 106: 9345-9349, 2009.
90. Fisher DO, Owens IPF. The comparative method in conservation biology. *Trends Ecol Evol* 19: 391-398, 2004.
91. Freckleton RP, Cooper N, Jetz W. Comparative methods as a statistical fix: the dangers of ignoring an evolutionary model. *Am Nat* 178: E10-E17, 2011.
92. Freckleton RP, Harvey PH. Detecting non-Brownian trait evolution in adaptive radiations. *Plos Biol* 4: 2104-2111, 2006.
93. Freckleton RP, Harvey PH, Pagel M. Phylogenetic analysis and comparative data: a test and review of evidence. *Am Nat* 160: 712-726, 2002.
94. Freckleton RP, Phillimore AB, Pagel M. Relating traits to diversification: a simple test. *Am Nat* 172: 102-115, 2008.
95. Freeman S, Herron JC. *Evolutionary Analysis*. Upper Saddle River, New Jersey: Pearson/Prentice Hall, 2004.
96. Fry BG, Roelants K, Norman JA. Tentacles of venom: toxic protein convergence in the kingdom animalia. *J Mol Evol* 68: 311-321, 2009.
97. Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SFR, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E. Early evolution of the venom system in lizards and snakes. *Nature* 439: 584-588, 2006.
98. Garland T. Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Am Nat* 140: 509-519, 1992.
99. Garland T. Phylogenetic comparison and artificial selection: two approaches in evolutionary physiology. *Adv Exp Med Biol* 502: 107-132, 2001.
100. Garland T, Adolph SC. Why not to do 2-species comparative studies: limitations on inferring adaptation. *Physiol Zool* 67: 797-828, 1994.
101. Garland T, Bennett AF, Rezende EL. Phylogenetic approaches in comparative physiology. *J Exp Biol* 208: 3015-3035, 2005.
102. Garland T, Carter PA. Evolutionary physiology. *Annu Rev Physiol* 56: 579-621, 1994.
103. Garland T, Diaz-Uriarte R. Polytomies and phylogenetically independent contrasts: examination of the bounded degrees of freedom approach. *Syst Biol* 48: 547-558, 1999.
104. Garland T, Dickerman AW, Janis CM, Jones JA. Phylogenetic analysis of covariance by computer simulation. *Syst Biol* 42: 265-292, 1993.
105. Garland T, Harvey PH, Ives AR. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst Biol* 41: 18-32, 1992.
106. Garland T, Huey RB, Bennett AF. Phylogeny and coadaptation of thermal physiology in lizards: a reanalysis. *Evolution* 45: 1969-1975, 1991.
107. Garland T, Ives AR. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *Am Nat* 155: 346-364, 2000.
108. Garland T, Janis CM. Does metatarsal femur ratio predict maximal running speed in cursorial mammals? *J Zool* 229: 133-151, 1993.
109. Garland T, Midford PE, Ives AR. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. *Am Zool* 39: 374-388, 1999.
110. Gartner GEA, Hicks JW, Manzani PR, Andrade DV, Abe AS, Wang T, Secor SM, Garland T. Phylogeny, ecology, and heart position in snakes. *Physiol Biochem Zool* 83: 43-54, 2010.
111. Giannini NP, Goloboff PA. Delayed-response phylogenetic correlation: an optimization-based method to test covariation of continuous characters. *Evolution* 64: 1885-1898, 2010.
112. Gillooly JF, Allen AP. Changes in body temperature influence the scaling of VO_2max and aerobic scope in mammals. *Biol Lett* 3: 99-102, 2007.
113. Gillooly JF, Allen AP, Charnov EL. Dinosaur fossils predict body temperatures. *Plos Biol* 4: 1467-1469, 2006.
114. Gittleman JL, Kot M. Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst Zool* 39: 227-241, 1990.
115. Gittleman JL, Luh HK. On comparing comparative methods. *Annu Rev Ecol Syst* 23: 383-404, 1992.
116. Glazier DS. Beyond the '3/4-power law': variation in the intra- and interspecific scaling of metabolic rate in animals. *Biol Rev* 80: 611-662, 2005.
117. Glazier DS. Effects of metabolic level on the body size scaling of metabolic rate in birds and mammals. *Proc R Soc B* 275: 1405-1410, 2008.
118. Glor RE. Phylogenetic insights on adaptive radiation. *Annu Rev Ecol Syst* 41: 251-270, 2010.
119. Goldberg EE, Igic B. On phylogenetic tests of irreversible evolution. *Evolution* 62: 2727-2741, 2008.
120. Golding GB, Dean AM. The structural basis of molecular adaptation. *Mol Biol Evol* 15: 355-369, 1998.
121. Gomes FR, Rezende EL, Grizante MB, Navas CA. The evolution of jumping performance in anurans: morphological correlates and ecological implications. *J Evol Biol* 22: 1088-1097, 2009.
122. Gould SJ, Lewontin RC. Spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist program. *Proc R Soc B* 205: 581-598, 1979.
123. Gould SJ, Vrba ES. Exaptation: a missing term in the science of form. *Paleobiology* 8: 4-15, 1982.
124. Grafen A. The phylogenetic regression. *Philos Trans R Soc B* 326: 119-157, 1989.
125. Grafen A. The uniqueness of the phylogenetic regression. *J Theor Biol* 156: 405-423, 1992.
126. Grafen A, Ridley M. Statistical tests for discrete cross-species data. *J Theor Biol* 183: 255-267, 1996.
127. Grueber CE, Nakagawa S, Laws RJ, Jamieson IG. Multimodel inference in ecology and evolution: challenges and solutions. *J Evol Biol* 24: 699-711, 2011.
128. Hahn TP, MacDougall-Shackleton SA. Adaptive specialization, conditional plasticity and phylogenetic history in the reproductive cue response systems of birds. *Philos Trans R Soc B* 363: 267-286, 2008.
129. Halsey LG, Butler PJ, Blackburn TM. A phylogenetic analysis of the allometry of diving. *Am Nat* 167: 276-287, 2006.
130. Hansen TF. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51: 1341-1351, 1997.
131. Hansen TF, Martins EP. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50: 1404-1417, 1996.
132. Hansen TF, Pienaar J, Orzack SH. A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62: 1965-1977, 2008.
133. Harmon LJ, Glor RE. Poor statistical performance of the Mantel test in phylogenetic comparative analyses. *Evolution* 64: 2173-2178, 2010.
134. Harmon LJ, Kolbe JJ, Cheverud JM, Losos JB. Convergence and the multidimensional niche. *Evolution* 59: 409-421, 2005.
135. Harmon LJ, Losos JB. The effect of intraspecific sample size on type I and type II error rates in comparative studies. *Evolution* 59: 2705-2710, 2005.
136. Harmon LJ, Losos JB, Davies TJ, Gillespie RG, Gittleman JL, Jennings WB, Kozak KH, McPeek MA, Moreno-Roark F, Near TJ, Purvis A, Ricklefs RE, Schluter D, Schulte JA, Seehausen O, Sidlauskas BL, Torres-Carvajal O, Weir JT, Mooers AO. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64: 2385-2396, 2010.
137. Harvey PH, Pagel MD. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press, 1991.

138. Harvey PH, Purvis A. Comparative methods for explaining adaptations. *Nature* 351: 619-624, 1991.
139. Harvey PH, Rambaut A. Comparative analyses for adaptive radiations. *Philos Trans R Soc B* 355: 1599-1605, 2000.
140. Hochachka PW, Somero GN. *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. New York: Oxford University Press, 2002.
141. Hodges WL. Evolution of viviparity in horned lizards (*Phrynosoma*): testing the cold-climate hypothesis. *J Evol Biol* 17: 1230-1237, 2004.
142. Hohenlohe PA, Arnold SJ. MIPoD: a hypothesis-testing framework for microevolutionary inference from patterns of divergence. *Am Nat* 171: 366-385, 2008.
143. Holmes DJ, Kristan DM. Comparative and alternative approaches and novel animal models for aging research. *Age* 30: 63-73, 2008.
144. Hopkins SR, Powell FL. Common themes of adaptation to hypoxia: insights from comparative physiology. *Adv Exp Med Biol* 502: 153-167, 2001.
145. Housworth EA, Martins EP. Random sampling of constrained phylogenies: conducting phylogenetic analyses when the phylogeny is partially known. *Syst Biol* 50: 628-639, 2001.
146. Housworth EA, Martins EP, Lynch M. The phylogenetic mixed model. *Am Nat* 163: 84-96, 2004.
147. Huelsenbeck JP, Rannala B. Detecting correlation between characters in a comparative analysis with uncertain phylogeny. *Evolution* 57: 1237-1247, 2003.
148. Huelsenbeck JP, Rannala B, Masly JP. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* 288: 2349-2350, 2000.
149. Huey RB. Phylogeny, history, and the comparative method. *New Directions in Ecological Physiology*. New York: Cambridge University Press, 1987, pp. 79-98.
150. Huey RB, Bennett AF. A comparative approach to field and laboratory studies in evolutionary biology. In: *Predator-Prey Relationships: Perspectives and Approaches from the Study of Lower Vertebrates*. Chicago: Chicago University Press, 1986, pp. 82-98.
151. Huey RB, Bennett AF. Phylogenetic studies of coadaptation: preferred temperatures versus optimal performance temperatures of lizards. *Evolution* 41: 1098-1115, 1987.
152. Hulbert AJ, Else PL. Basal metabolic rate: history, composition, regulation, and usefulness. *Physiol Biochem Zool* 77: 869-876, 2004.
153. Hulbert AJ, Else PL. Membranes and the setting of energy demand. *J Exp Biol* 208: 1593-1599, 2005.
154. Hulbert AJ, Faulks S, Buttemer WA, Else PL. Acyl composition of muscle membranes varies with body size in birds. *J Exp Biol* 205: 3561-3569, 2002.
155. Hunter JP, Jernvall J. The hypocone as a key innovation in mammalian evolution. *Proc Natl Acad Sci U S A* 92: 10718-10722, 1995.
156. Ives AR, Garland T. Phylogenetic logistic regression for binary dependent variables. *Syst Biol* 59: 9-26, 2010.
157. Ives AR, Midford PE, Garland T. Within-species variation and measurement error in phylogenetic comparative methods. *Syst Biol* 56: 252-270, 2007.
158. Johnson JB, Omland KS. Model selection in ecology and evolution. *Trends Ecol Evol* 19: 101-108, 2004.
159. Johnston IA, Fernandez DA, Calvo J, Vieira VLA, North AW, Abercromby M, Garland T. Reduction in muscle fibre number during the adaptive radiation of notothenioid fishes: a phylogenetic perspective. *J Exp Biol* 206: 2595-2609, 2003.
160. Kelchner SA, Thomas MA. Model use in phylogenetics: nine key questions. *Trends Ecol Evol* 22: 87-94, 2007.
161. Kelly SA, Panhuis TM, Stoehr AM. Phenotypic plasticity: molecular mechanisms and adaptive significance. *Comp Physiol* (in press), 2012.
162. Koteja P, Weiner J. Mice, voles and hamsters: metabolic rates and adaptive strategies in muroid rodents. *Oikos* 66: 505-514, 1993.
163. Kozak KH, Larson AA, Bonett RM, Harmon LJ. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: *Desmognathus*). *Evolution* 59: 2000-2016, 2005.
164. Kram R, Taylor CR. Energetics of running: a new perspective. *Nature* 346: 265-267, 1990.
165. Krebs HA. Krogh's principle: for many problems there is an animal on which it can be most conveniently studied. *J Exp Zool* 194: 221-226, 1975.
166. Lajeunesse MJ. Meta-analysis and the comparative phylogenetic method. *Am Nat* 174: 369-381, 2009.
167. Lambert AJ, Boysen HM, Buckingham JA, Yang T, Podlutsky A, Austad SN, Kunz TH, Buffenstein R, Brand MD. Low rates of hydrogen peroxide production by isolated heart mitochondria associate with long maximum lifespan in vertebrate homeotherms. *Aging Cell* 6: 607-618, 2007.
168. Lauder GV. Form and function: structural analysis in evolutionary morphology. *Paleobiology* 7: 430-442, 1981.
169. Lauder GV, Leroi AM, Rose MR. Adaptations and history. *Trends Ecol Evol* 8: 294-297, 1993.
170. Laurin M. The evolution of body size, Cope's rule and the origin of amniotes. *Syst Biol* 53: 594-622, 2004.
171. Lavin SR, Karasov WH, Ives AR, Middleton KM, Garland T. Morphometrics of the avian small intestine compared with that of nonflying mammals: a phylogenetic approach. *Physiol Biochem Zool* 81: 526-550, 2008.
172. Lee CE. Rapid and repeated invasions of fresh water by the copepod. *Evolution* 53: 1423-1434, 1999.
173. Lee CH, Blay S, Mooers AO, Singh A, Oakley TH. CoMET: a mesquite package for comparing models of continuous character evolution on phylogenies. *Evol Bioinform* 2: 183-186, 2006.
174. Lee MSY. Molecular evidence and marine snake origins. *Biol Lett* 1: 227-230, 2005.
175. Legendre P, Desdevises Y. Independent contrasts and regression through the origin. *J Theor Biol* 259: 727-743, 2009.
176. Leroi AM, Rose MR, Lauder GV. What does the comparative method reveal about adaptation? *Am Nat* 143: 381-402, 1994.
177. Lewis PO. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst Biol* 50: 913-925, 2001.
178. Liefting M, Hoffmann AA, Ellers J. Plasticity versus environmental canalization: population differences in thermal responses along a latitudinal gradient in *Drosophila serrata*. *Evolution* 63: 1954-1963, 2009.
179. Lindenfors P, Revell LJ, Nunn CL. Sexual dimorphism in primate aerobic capacity: a phylogenetic test. *J Evol Biol* 23: 1183-1194, 2010.
180. Losos JB. Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods. *Anim Behav* 58: 1319-1324, 1999.
181. Losos JB. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol Lett* 11: 995-1003, 2008.
182. Losos JB. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *Am Nat* 177: 709-727, 2011.
183. Losos JB, Miles DB. Adaptation, constraint, and the comparative method: phylogenetic issues and methods. In: Wainwright PC, Reilly SM, editors. *Ecological Morphology*. Chicago: The University of Chicago Press, 1994, pp. 60-98.
184. Lovegrove BG. The zoogeography of mammalian basal metabolic rate. *Am Nat* 156: 201-219, 2000.
185. Ludden TM, Beal SL, Sheiner LB. Comparison of the Akaike information criterion, the Schwarz criterion and the F-test as guides to model selection. *J Pharmacokinet Biop* 22: 431-445, 1994.
186. Lutzoni F, Pagel M, Reeb V. Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 411: 937-940, 2001.
187. Lynch M. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45: 1065-1080, 1991.
188. MacDougall-Shackleton SA, Stevenson TJ, Watts HE, Pereyra ME, Hahn TP. The evolution of photoperiod response systems and seasonal GnRH plasticity in birds. *Integr Comp Biol* 49: 580-589, 2009.
189. Mace R, Pagel M. The comparative method in anthropology. *Curr Anthropol* 35: 549-564, 1994.
190. Maddison DR, Maddison WP. *MacClade Version 4: Analysis of Phylogeny and Character Evolution*. Sunderland, MA: Sinauer Associates, 2000.
191. Maddison WP. A method for testing the correlated evolution of 2 binary characters: are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* 44: 539-557, 1990.
192. Maddison WP. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. *Syst Zool* 40: 304-314, 1991.
193. Maddison WP. Calculating the probability distributions of ancestral states reconstructed by parsimony on phylogenetic trees. *Syst Biol* 44: 474-481, 1995.
194. Maddison WP. Testing character correlation using pairwise comparisons on a phylogeny. *J Theor Biol* 202: 195-204, 2000.
195. Maddison WP. Confounding asymmetries in evolutionary diversification and character change. *Evolution* 60: 1743-1746, 2006.
196. Maddison WP, Slatkin M. Null models for the number of evolutionary steps in a character on a phylogenetic tree. *Evolution* 45: 1184-1197, 1991.
197. Martin RD, Genoud M, Hemelrijk CK. Problems of allometric scaling analysis: examples from mammalian reproductive biology. *J Exp Biol* 208: 1731-1747, 2005.
198. Martins EP. A comparative study of the evolution of *Sceloporus* push-up displays. *Am Nat* 142: 994-1018, 1993.
199. Martins EP. Estimating the rate of phenotypic evolution from comparative data. *Am Nat* 144: 193-209, 1994.
200. Martins EP. Phylogenies, spatial autoregression, and the comparative method: a computer simulation test. *Evolution* 50: 1750-1765, 1996.
201. Martins EP. Estimation of ancestral states of continuous characters: a computer simulation study. *Syst Biol* 48: 642-650, 1999.

202. Martins EP, Diniz JAF, Housworth EA. Adaptive constraints and the phylogenetic comparative method: a computer simulation test. *Evolution* 56: 1-13, 2002.
203. Martins EP, Garland T. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution* 45: 534-557, 1991.
204. Martins EP, Hansen TF. A microevolutionary link between phylogenies and comparative data. In: Harvey PH, Leigh Brown AJ, Maynard Smith J, Nee S, editors. *New Uses for New Phylogenies*. Oxford: Oxford University Press, 1996, pp. 273-288.
205. Martins EP, Hansen TF. The statistical analysis of interspecific data: a review and evaluation of phylogenetic comparative methods. In: Martins E, editor. *Phylogenies and the Comparative Method in Animal Behavior*. Oxford: Oxford University Press, 1996, pp. 22-75.
206. Martins EP, Hansen TF. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am Nat* 149: 646-667, 1997.
207. Martins EP, Housworth EA. Phylogeny shape and the phylogenetic comparative method. *Syst Biol* 51: 873-880, 2002.
208. McCracken KG, Barger CP, Bulgarella M, Johnson KP, Sonsthagen SA, Trucco J, Valqui TH, Wilson RE, Winker K, Sorenson MD. Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. *Mol Ecol* 18: 3992-4005, 2009.
209. McKechnie AE, Freckleton RP, Jetz W. Phenotypic plasticity in the scaling of avian basal metabolic rate. *Proc R Soc B* 273: 931-937, 2006.
210. McKechnie AE, Wolf BO. The allometry of avian basal metabolic rate: good predictions need good data. *Physiol Biochem Zool* 77: 502-521, 2004.
211. Mcnab BK. Climatic adaptation in the energetics of heteromyid rodents. *Comp Biochem Physiol A* 62: 813-820, 1979.
212. McNab BK. *The Physiological Ecology of Vertebrates: A View from Energetics*. New York: Cornell University Press, 2002.
213. Mcnab BK. Ecological factors affect the level and scaling of avian BMR. *Comp Biochem Physiol A* 152: 22-45, 2009.
214. Mcnab BK, Morrison P. Body temperature and metabolism in subspecies of *Peromyscus* from arid and mesic environments. *Ecol Monogr* 33: 63-82, 1963.
215. McPeek MA. Testing hypotheses about evolutionary change on single branches of a phylogeny using evolutionary contrasts. *Am Nat* 145: 686-703, 1995.
216. Meredith RW, Pires MN, Reznick DN, Springer MS. Molecular phylogenetic relationships and the evolution of the placenta in Poecilia (Micropoecilia) (Poeciliidae: Cyprinodontiformes). *Mol Phylogenet Evol* 55: 631-639, 2010.
217. Miles DB, Dunham AE. Historical perspectives in ecology and evolutionary biology: the use of phylogenetic comparative analyses. *Annu Rev Ecol Syst* 24: 587-619, 1993.
218. Montes L, Le Roy N, Perret M, De Buffrenil V, Castanet J, Cubo J. Relationships between bone growth rate, body mass and resting metabolic rate in growing amniotes: a phylogenetic approach. *Biol J Linn Soc* 92: 63-76, 2007.
219. Mooers AO, Schluter D. Reconstructing ancestor states with maximum likelihood: support for one- and two-rate models. *Syst Biol* 48: 623-633, 1999.
220. Mooers AO, Vamosi SM, Schluter D. Using phylogenies to test macroevolutionary hypotheses of trait evolution in Cranes (Gruinae). *Am Nat* 154: 249-259, 1999.
221. Mora C, Maya MF. Effect of the rate of temperature increase of the dynamic method on the heat tolerance of fishes. *J Therm Biol* 31: 337-341, 2006.
222. Morey S, Reznick D. A comparative analysis of plasticity ln larval development in three species of spadefoot toads. *Ecology* 81: 1736-1749, 2000.
223. Munoz-Garcia A, Williams JB. Basal metabolic rate in carnivores is associated with diet after controlling for phylogeny. *Physiol Biochem Zool* 78: 1039-1056, 2005.
224. Nagy KA. Field bioenergetics of mammals: what determines field metabolic rates? *Aust J Zool* 42: 43-53, 1994.
225. Nishikawa KC. Evolutionary convergence in nervous systems: insights from comparative phylogenetic studies. *Brain Behav Evolut* 59: 240-249, 2002.
226. Nunn CL. *The Comparative Method in Evolutionary Anthropology and Biology.* Chicago: University of Chicago Press, 2011.
227. Nunn CL, Barton RA. Allometric slopes and independent contrasts: a comparative test of Kleiber's law in primate ranging patterns. *Am Nat* 156: 519-533, 2000.
228. O'Connor MP, Agosta SJ, Hansen F, Kemp SJ, Sieg AE, McNair JN, Dunham AE. Phylogeny, regression, and the allometry of physiological traits. *Am Nat* 170: 431-442, 2007.
229. O'Meara BC, Ane C, Sanderson MJ, Wainwright PC. Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60: 922-933, 2006.
230. Oakley TH, Cunningham CW. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution* 54: 397-405, 2000.
231. Pagel M. Seeking the evolutionary regression coefficient: an analysis of what comparative methods measure. *J Theor Biol* 164: 191-205, 1993.
232. Pagel M. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc R Soc B* 255: 37-45, 1994.
233. Pagel M. Inferring the historical patterns of biological evolution. *Nature* 401: 877-884, 1999.
234. Pagel M, Meade A. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am Nat* 167: 808-825, 2006.
235. Pagel M, Meade A, Barker D. Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53: 673-684, 2004.
236. Pagel M, Venditti C, Meade A. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* 314: 119-121, 2006.
237. Pagel MD. A method for the analysis of comparative data. *J Theor Biol* 156: 431-442, 1992.
238. Paradis E, Claude J. Analysis of comparative data using generalized estimating equations. *J Theor Biol* 218: 175-185, 2002.
239. Pavoine S, Ollier S, Pontier D, Chessel D. Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theor Popul Biol* 73: 79-91, 2008.
240. Pierce VA, Crawford DL. Phylogenetic analysis of glycolytic enzyme expression. *Science* 276: 256-259, 1997.
241. Piersma T, van Gils JAM. *The Flexible Phenotype: A Body-Centered Integration of Ecology, Physiology, and Behaviour*. Oxford: Oxford University Press, 2011.
242. Polly PD. Paleontology and the comparative method: ancestral node reconstructions versus observed node values. *Am Nat* 157: 596-609, 2001.
243. Pough FH. Advantages of ectothermy for tetrapods. *Am Nat* 115: 92-112, 1980.
244. Powers DA, Ropson I, Brown DC, Vanbeneden R, Cashon R, Gonzalezvillasenor LI, Dimichele JA. Genetic variation in *Fundulus heteroclitus*: geographic distribution. *Am Zool* 26: 131-144, 1986.
245. Poyart C, Wajcman H, Kister J. Molecular adaptation of hemoglobin function in mammals. *Resp Physiol* 90: 3-17, 1992.
246. Price T. Correlated evolution and independent contrasts. *Philos Trans R Soc B* 352: 519-529, 1997.
247. Purvis A. A composite estimate of primate phylogeny. *Philos Trans R Soc B* 348: 405-421, 1995.
248. Purvis A, Garland T. Polytomies in comparative analyses of continuous characters. *Syst Biol* 42: 569-575, 1993.
249. Purvis A, Gittleman JL, Luh HK. Truth or consequences: effects of phylogenetic accuracy on 2 comparative methods. *J Theor Biol* 167: 293-300, 1994.
250. Rabosky DL, Donnellan SC, Talaba AL, Lovette IJ. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc R Soc B* 274: 2915-2923, 2007.
251. Ramirez JM, Folkow LP, Blix AS. Hypoxia tolerance in mammals and birds: from the wilderness to the clinic. *Annu Rev Physiol* 69: 113-143, 2007.
252. Ree RH. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* 59: 257-265, 2005.
253. Reiss MJ. *The Allometry of Growth and Reproduction*. Cambridge: Cambridge University Press, 1989.
254. Revell LJ. On the analysis of evolutionary change along single branches in a phylogeny. *Am Nat* 172: 140-147, 2008.
255. Revell LJ. Size-correction and principal components for interspecific comparative studies. *Evolution* 63: 3258-3268, 2009.
256. Revell LJ. Phylogenetic signal and linear regression on species data. *Methods Ecol Evol* 1: 319-329, 2010.
257. Revell LJ, Harmon LJ, Collar DC. Phylogenetic signal, evolutionary process, and rate. *Syst Biol* 57: 591-601, 2008.
258. Revell LJ, Harrison AS. PCCA: a program for phylogenetic canonical correlation analysis. *Bioinformatics* 24: 1018-1020, 2008.
259. Revell LJ, Johnson MA, Schulte JA, Kolbe JJ, Losos JB. A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution* 61: 2898-2912, 2007.
260. Reynolds PS, Lee RM. Phylogenetic analysis of avian energetics: passerines and nonpasserines do not differ. *Am Nat* 147: 735-759, 1996.
261. Rezende EL, Bozinovic F, Garland T. Climatic adaptation and the evolution of basal and maximum rates of metabolism in rodents. *Evolution* 58: 1361-1374, 2004.
262. Rezende EL, Swanson DL, Novoa FF, Bozinovic F. Passerines versus nonpasserines: so far, no statistical differences in the scaling of avian energetics. *J Exp Biol* 205: 101-107, 2002.

263. Rezende EL, Tejedo M, Santos M. Estimating the adaptive potential of critical thermal limits: methodological problems and evolutionary implications. *Funct Ecol* 25: 111-121, 2011.
264. Reznick DN, Mateos M, Springer MS. Independent origins and rapid evolution of the placenta in the fish genus *Poeciliopsis*. *Science* 298: 1018-1020, 2002.
265. Rheindt FE, Grafe TU, Abouheif E. Rapidly evolving traits and the comparative method: how important is testing for phylogenetic signal? *Evol Ecol Res* 6: 377-396, 2004.
266. Ricklefs RE, Starck JM. Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* 77: 167-172, 1996.
267. Rohlf FJ. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* 55: 2143-2160, 2001.
268. Rohlf FJ. A comment on phylogenetic correction. *Evolution* 60: 1509-1515, 2006.
269. Ross CF, Henneberg M, Ravosa MJ, Richard S. Curvilinear, geometric and phylogenetic modeling of basicranial flexion: is it adaptive, is it constrained? *J Hum Evol* 46: 185-213, 2004.
270. Ruben J. The evolution of endothermy in mammals and birds: from physiology to fossils. *Annu Rev Physiol* 57: 69-95, 1995.
271. Sanford GM, Lutterschmidt WI, Hutchison VH. The comparative method revisited. *Bioscience* 52: 830-836, 2002.
272. Savage VM, Gillooly JF, Woodruff WH, West GB, Allen AP, Enquist BJ, Brown JH. The predominance of quarter-power scaling in biology. *Funct Ecol* 18: 257-282, 2004.
273. Schluter D, Price T, Mooers AO, Ludwig D. Likelihood of ancestor states in adaptive radiation. *Evolution* 51: 1699-1711, 1997.
274. Schmidt-Nielsen K. *Animal Physiology: Adaptation and Environment*. Cambridge: Cambridge University Press, 1997.
275. Schmidt-Nielsen K. *Scaling: Why is Animal Size So Important?* Cambridge: Cambridge University Press, 1999.
276. Scholander PF, Walters V, Hock R, Irving L. Body insulation of some arctic and tropical mammals and birds. *Biol Bull* 99: 225-236, 1950.
277. Schulz TM, Bowen WD. The evolution of lactation strategies in pinnipeds: a phylogenetic analysis. *Ecol Monogr* 75: 159-177, 2005.
278. Seebacher F. Dinosaur body temperatures: the occurrence of endothermy and ectothermy. *Paleobiology* 29: 105-122, 2003.
279. Sibley CG, Ahlquist JE. *Phylogeny and Classification of Birds: A Study in Molecular Evolution*. New Haven: Yale University Press, 1990.
280. Sieg AE, O'Connor MP, McNair JN, Grant BW, Agosta SJ, Dunham AE. Mammalian metabolic allometry: do intraspecific variation, phylogeny, and regression models matter? *Am Nat* 174: 720-733, 2009.
281. Singer MA. *Comparative Physiology, Natural Animal Models and Clinical Medicine Insights Into Clinical Medicine from Animal Adaptations*. London: Imperial College Press, 2007.
282. Snyder LRG. Deer mouse hemoglobins: is there genetic adaptation to high-altitude? *Bioscience* 31: 299-304, 1981.
283. Snyder LRG. Low P50 in deer mice native to high-altitude. *J Appl Physiol* 58: 193-199, 1985.
284. Snyder LRG, Hayes JP, Chappell MA. Alpha-chain hemoglobin polymorphisms are correlated with altitude in the deer mouse, *Peromyscus maniculatus*. *Evolution* 42: 689-697, 1988.
285. Sol D, Duncan RP, Blackburn TM, Cassey P, Lefebvre L. Big brains, enhanced cognition, and response of birds to novel environments. *Proc Natl Acad Sci U S A* 102: 5460-5465, 2005.
286. Sol D, Price TD. Brain size and the diversification of body size in birds. *Am Nat* 172: 170-177, 2008.
287. Speakman JR. The cost of living: field metabolic rates of small mammals. *Adv Ecol Res* 30: 177-297, 2000.
288. Speakman JR, Krol E. Maximal heat dissipation capacity and hyperthermia risk: neglected key factors in the ecology of endotherms. *J Anim Ecol* 79: 726-746, 2010.
289. Spoor F, Bajpal S, Hussaim ST, Kumar K, Thewissen JGM. Vestibular evidence for the evolution of aquatic behaviour in early cetaceans. *Nature* 417: 163-166, 2002.
290. Spoor F, Garland T, Krovitz G, Ryan TM, Silcox MT, Walker A. The primate semicircular canal system and locomotion. *Proc Natl Acad Sci U S A* 104: 10808-10812, 2007.
291. Springer MS, Cleven GC, Madsen O, deJong WW, Waddell VG, Amrine HM, Stanhope MJ. Endemic African mammals shake the phylogenetic tree. *Nature* 388: 61-64, 1997.
292. Springer MS, Stanhope MJ, Madsen O, de Jong WW. Molecules consolidate the placental mammal tree. *Trends Ecol Evol* 19: 430-438, 2004.
293. Stoks R, McPeek MA. A tale of two diversifications: reciprocal habitat shifts to fill ecological space along the pond permanence gradient. *Am Nat* 168: S50-S72, 2006.
294. Stone EA. Why the phylogenetic regression appears robust to tree misspecification. *Syst Biol* 60: 245-260, 2011.
295. Stone GN, Nee S, Felsenstein J. Controlling for non-independence in comparative analysis of patterns across populations within species. *Phil Trans R Soc B* 366: 1410-1424, 2011.
296. Storz JF. Hemoglobin function and physiological adaptation to hypoxia in high-altitude mammals. *J Mammal* 88: 24-31, 2007.
297. Storz JF, Runck AM, Sabatino SJ, Kelly JK, Ferrand N, Moriyama H, Weber RE, Fago A. Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proc Natl Acad Sci U S A* 106: 14450-14455, 2009.
298. Storz JF, Sabatino SJ, Hoffmann FG, Gering EJ, Moriyama H, Ferrand N, Monteiro B, Nachman MW. The molecular basis of high-altitude adaptation in deer mice. *Plos Genet* 3: 448-459, 2007.
299. Swanson DL, Garland T. The evolution of high summit metabolism and cold tolerance in birds and its impact on present-day distributions. *Evolution* 63: 184-194, 2009.
300. Swofford DL, Maddison WP. Reconstructing ancestral character states under Wagner parsimony. *Math Biosci* 87: 199-229, 1987.
301. Symonds MRE. The effects of topological inaccuracy in evolutionary trees on the phylogenetic comparative method of independent contrasts. *Syst Biol* 51: 541-553, 2002.
302. Symonds MRE, Elgar MA. Phylogeny affects estimation of metabolic scaling in mammals. *Evolution* 56: 2330-2333, 2002.
303. Taylor CR, Heglund NC, Maloiy GMO. Energetics and mechanics of terrestrial locomotion 1. Metabolic energy-consumption as a function of speed and body size in birds and mammals. *J Exp Biol* 97: 1-21, 1982.
304. Terblanche JS, Deere JA, Clusella-Trullas S, Janion C, Chown SL. Critical thermal limits depend on methodological context. *Proc R Soc B* 274: 2935-2942, 2007.
305. Thomas GH, Freckleton RP, Szekely T. Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. *Proc R Soc B* 273: 1619-1624, 2006.
306. Thornton JW, DeSalle R. Gene family evolution and homology: genomics meets phylogenetics. *Annu Rev Genom Hum G* 1: 41-73, 2000.
307. Tieleman BI, Williams JB. The adjustment of avian metabolic rates and water fluxes to desert environments. *Physiol Biochem Zool* 73: 461-479, 2000.
308. Tieleman BI, Williams JB, Bloomer P. Adaptation of metabolism and evaporative water loss along an aridity gradient. *Proc R Soc B* 270: 207-214, 2003.
309. Turkheimer F, Hinz R, Cunningham V. Model averaging using Akaike weights. *Neuroimage* 16: S60-S60, 2002.
310. Turkheimer FE, Hinz R, Cunningham VJ. On the undecidability among kinetic models: from model selection to model averaging. *J Cerebr Blood F Met* 23: 490-498, 2003.
311. Van Buskirk J. A comparative test of the adaptive plasticity hypothesis: relationships between habitat and phenotype in anuran larvae. *Am Nat* 160: 87-102, 2002.
312. Vonk FJ, Admiraal JF, Jackson K, Reshef R, de Bakker MAG, Vanderschoot K, van den Berge I, van Atten M, Burgerhout E, Beck A, Mirtschin PJ, Kochva E, Witte F, Fry BG, Woods AE, Richardson MK. Evolutionary origin and development of snake fangs. *Nature* 454: 630-633, 2008.
313. Webster AJ, Purvis A. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc R Soc B* 269: 143-149, 2002.
314. Weibel ER. Understanding the limitation of O_2 supply through comparative physiology. *Resp Physiol* 118: 85-93, 1999.
315. Weibel ER, Bacigalupe LD, Schmitt B, Hoppeler H. Allometric scaling of maximal metabolic rate in mammals: muscle aerobic capacity as determinant factor. *Resp Physiol Neurobi* 140: 115-132, 2004.
316. Weibel ER, Hoppeler H. Exercise-induced maximal metabolic rate scales with muscle aerobic capacity. *J Exp Biol* 208: 1635-1644, 2005.
317. West GB, Brown JH, Enquist BJ. A general model for the origin of allometric scaling laws in biology. *Science* 276: 122-126, 1997.
318. West GB, Brown JH, Enquist BJ. The fourth dimension of life: fractal geometry and allometric scaling of organisms. *Science* 284: 1677-1679, 1999.
319. Westneat MW. Feeding, function, and phylogeny: analysis of historical biomechanics in labrid fishes using comparative methods. *Syst Biol* 44: 361-383, 1995.
320. Westoby M, Leishman MR, Lord JM. On misinterpreting the phylogenetic correction. *J Ecol* 83: 531-534, 1995.
321. White CR. The influence of foraging mode and arid adaptation on the basal metabolic rates of burrowing mammals. *Physiol Biochem Zool* 76: 122-134, 2003.
322. White CR. Allometric estimation of metabolic rates in animals. *Comp Biochem Phys A* 158: 346-357, 2011.
323. White CR, Blackburn TM, Martin GR, Butler PJ. Basal metabolic rate of birds is associated with habitat temperature and precipitation, not primary productivity. *Proc R Soc B* 274: 287-293, 2007.
324. White CR, Blackburn TM, Seymour RS. Phylogenetically informed analysis of the allometry of mammalian basal metabolic rate supports neither geometric nor quarter-power scaling. *Evolution* 63: 2658-2667, 2009.

325. White CR, Blackburn TM, Terblanche JS, Marais E, Gibernau M, Chown SL. Evolutionary responses of discontinuous gas exchange in insects. *P Natl Acad Sci USA* 104: 8357-8361, 2007.
326. White CR, Seymour RS. Mammalian basal metabolic rate is proportional to body mass$^{(2/3)}$. *P Natl Acad Sci USA* 100: 4046-4049, 2003.
327. White CR, Terblanche JS, Kabat AP, Blackburn TM, Chown SL, Butler PJ. Allometric scaling of maximum metabolic rate: the influence of temperature. *Funct Ecol* 22: 616-623, 2008.
328. Whiting MF, Bradler S, Maxwell T. Loss and recovery of wings in stick insects. *Nature* 421: 264-267, 2003.
329. Whitney KD, Garland T. Did genetic drift drive increases in genome complexity? *Plos Genet* 6, 2010.
330. Wiens JJ. Character analysis in morphological phylogenetics: problems and solutions. *Syst Biol* 50: 689-699, 2001.
331. Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies TJ, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* 13: 1310-1324, 2010.
332. Wiens JJ, Graham CH. Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annu Rev Ecol Evol S* 36: 519-539, 2005.
333. Wiens JJ, Graham CH, Moen DS, Smith SA, Reeder TW. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: Treefrog trees unearth the roots of high tropical diversity. *Am Nat* 168: 579-596, 2006.
334. Wiens JJ, Kuczynski CA, Duellman WE, Reeder TW. Loss and re-evolution of complex life cycles in marsupial frogs: Does ancestral trait reconstruction mislead? *Evolution* 61: 1886-1899, 2007.
335. Wiersma P, Chappell MA, Williams JB. Cold- and exercise-induced peak metabolic rates in tropical birds. *Proc Natl Acad Sci U S A* 104: 20866-20871, 2007.
336. Withers PC, Cooper CE, Larcombe AN. Environmental correlates of physiological variables in marsupials. *Physiol Biochem Zool* 79: 437-453, 2006.
337. Wolf CM, Garland T, Griffith B. Predictors of avian and mammalian translocation success: reanalysis with phylogenetically independent contrasts. *Biol Conserv* 86: 243-255, 1998.
338. Yanoviak SP, Kaspari M, Dudley R. Gliding hexapods and the origins of insect aerial behaviour. *Biol Lett* 5: 510-512, 2009.
339. Zheng L, Ives AR, Garland T, Larget BR, Yu Y, Cao KF. New multivariate tests for phylogenetic signal and trait correlations applied to ecophysiological phenotypes of nine *Manglietia* species. *Funct Ecol* 23: 1059-1069, 2009.

Further Reading

Phylogenetic statistical methods are becoming increasingly flexible and complex, and considerable information regarding these approaches, available softwares and forums of discussion, is currently available of the Internet. Even though these sources are not peer-reviewed, they can be very helpful for researchers and students that are not familiar with the field and look for some guidance. Here we list a few web pages that may be useful in this context.

http://en.wikipedia.org/wiki/Phylogenetic_comparative_methods

Lists of software packages

http://evolution.genetics.washington.edu/phylip/software.html
http://bioinfo.unice.fr/biodiv/Tree_editors.html
http://cran.r-project.org/web/views/Phylogenetics.html
http://mesquiteproject.org/mesquite/mesquite.html

Discussion forums, courses and miscelaneous information

http://bodegaphylo.wikispot.org/
http://informatics.nescent.org/wiki/Main_Page
http://phytools.blogspot.com/
http://www.r-phylo.org/wiki/Main_Page

Mailing lists (R-phylo and Mesquite)

https://stat.ethz.ch/mailman/listinfo/r-sig-phylo
http://mesquiteproject.org/mailman/listinfo/mesquitelist