

Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods

S. P. BLOMBERG & T. GARLAND, JR

Department of Biology, University of California, Riverside, CA, USA

Keywords:

adaptation;
comparative methods;
constraint;
independent contrasts;
natural selection;
phylogenetic inertia;
statistics.

Abstract

Before the Evolutionary Synthesis, 'phylogenetic inertia' was associated with theories of orthogenesis, which claimed that organisms possessed an endogenous perfecting principle. The concept in the modern literature dates to Simpson (1944), who used 'evolutionary inertia' as a description of pattern in the fossil record. Wilson (1975) used 'phylogenetic inertia' to describe population-level or organismal properties that can affect the course of evolution in response to selection. Many current authors now view phylogenetic inertia as an alternative hypothesis to adaptation by natural selection when attempting to explain interspecific variation, covariation or lack thereof in phenotypic traits. Some phylogenetic comparative methods have been claimed to allow quantification and testing of phylogenetic inertia. Although some existing methods do allow valid tests of whether related species tend to resemble each other, which we term 'phylogenetic signal', this is simply pattern recognition and does not imply any underlying process. Moreover, comparative data sets generally do not include information that would allow rigorous inferences concerning causal processes underlying such patterns. The concept of phylogenetic inertia needs to be defined and studied with as much care as 'adaptation'.

Introduction

We review the concept of 'phylogenetic inertia' and consider if and how it can be studied by modern comparative methods. Before the Evolutionary Synthesis, phylogenetic inertia was associated with theories of orthogenesis, which, according to Mayr (1982, pp. 529–530), claimed that organisms possessed an endogenous 'perfecting principle'. The term 'inertia' was used by some proponents of orthogenesis as a direct analogy with inertia in physics: once organisms begin to evolve in a particular direction, they tend to keep evolving in the same direction (cf. Burt, 2001). In his refutation of orthogenetic arguments, Simpson (1944) acknowledged that the fossil record did sometimes exhibit patterns that suggested evolution was proceeding in a particular

direction (e.g. during the evolution of the toes and teeth of horses). However, he argued that: (1) the patterns showed many irregularities; (2) no evidence of an endogenous mechanism had ever been found; and (3) invocation of such a mechanism was unnecessary because patterns in the fossil record could be explained by natural selection (Simpson, 1944, pp. 161–163). Thus, for Simpson, phylogenetic inertia (Simpson uses the term 'evolutionary inertia') represented a *pattern* of 'rectilinear' evolution, by which he meant linear directional trends in the fossil record, generated by the *process* of natural selection. Rectilinear trends were part of the dynamics (tempo) of evolution, and not associated with phyletic stasis. This is in stark contrast with more recent definitions of the concept.

Since the Synthesis, the concept of phylogenetic inertia has undergone considerable revision. Instead of just describing patterns in the fossil record, the term has become associated with various factors, other than natural and sexual selection, that may affect phenotypic evolution. Ridley (1983) claims that the term was

Correspondence: Theodore Garland, Jr, Department of Biology, University of California, Riverside, Riverside, CA 92521, USA.
Tel.: (909) 787-3524; fax: (909) 787-4286;
e-mail: tgarland@citrus.ucr.edu

introduced by Wilson (1975), who writes (p. 32): 'Phylogenetic inertia...consists of the deeper properties of the population that determine the extent to which its evolution can be deflected in one direction or another, as well as the amount by which its rate can be speeded or slowed'. Wilson lists four mechanisms that create phylogenetic inertia in the social behaviour of animals.

1. *Genetic variability*. Organisms can only respond to selection in proportion to the genetic component of phenotypic variability.
2. *Antisocial factors*. Wilson notes that various idiosyncratic, lineage-specific effects can affect the direction of evolution.
3. *The complexity of the social behaviour*. The more complicated the behavioural phenotype and the supporting physiological machinery, the greater the inertia.
4. *The effect of the evolution on other traits*. Inertia of behavioural phenotypes is increased if they are correlated with other traits that may affect fitness. Ridley (1983) notes that Cain (1964) lists similar processes under the term 'genetic inertia', which apparently is attributable to Darlington & Mather (1949). In any case, Wilson proposes various processes (mechanisms) that can affect the course of evolution. All of these would appear to involve well-accepted (if sometimes obscure) biological principles, not any endogenous 'perfecting principle'.

Clearly, Wilson's (1975) conception of phylogenetic inertia (factors that tend to resist selection) is the virtual opposite of Simpson's (1944) (patterns that have resulted from selection). However, Wilson (1975) also says that phylogenetic inertia includes pre-adaptation, the concept of traits that evolve for one function and later get co-opted as an adaptation for another, different function. As noted by Gittleman *et al.* (1996a), after discussing phylogenetic inertia on page 32, Wilson (1975) emphasized adaptive explanations for behavioural traits, and phylogenetic inertia is mentioned in only one other place in the entire book.

Ridley (1983) defines phylogenetic inertia simply as a character that is shared among related species, although he admits that this definition is unclear. Berger (1988) offers a similar definition. However, many modern researchers have followed Wilson's (1975) sentiment, at least in part, and now equate phylogenetic inertia with nonadaptive (or maladaptive) phenotypic stasis. Gould & Lewontin (1979), for example, recognize phylogenetic inertia in the fact that humans are not optimally designed for upright posture, because much of our *Bauplan* evolved originally for quadrupedal locomotion. Edwards & Naeem (1993) provide a particularly good discussion in the context of the evolution of cooperative breeding in birds. They define phylogenetic inertia as occurring '...when traits persist in lineages after the cessation of selective forces thought to have produced or maintained them or through episodes of selectively important environmental oscillations'. They say their definition implies a tendency for traits to resist change, despite environmental perturbations. Edwards & Naeem (1993) provide

five possible causes for phylogenetic inertia of cooperative breeding:

1. *Limited genetic or phenotypic variation* (similar to Wilson's 1., above).
2. *Pleiotropy and genetic covariance*. If the same genes underlie two or more traits, then evolution of one of the traits (e.g. cooperative breeding) may be constrained by selection on the other(s) (similar to Wilson's 4., above).
3. *High correlations with metric traits*. This seems to be a special case of 2.
4. *Functional interdependency of components of a trait*. If juvenile survival is linked to cooperative breeding, then alternative strategies may find it impossible to 'invade'. This appears to be a special case of selection on the components of cooperative breeding.
5. *Behavioural plasticity*. If individuals can modify their behaviour, then they can limit the effect of selection on traits (see also Garland *et al.*, 1990).

A recent attempt to clarify the term phylogenetic inertia (and related terms) has been undertaken by Burt (2001). He argues that phylogenetic inertia should be treated as a phenomenological *pattern* description of traits among species (fossil or extant). For Burt, phylogenetic inertia is defined in analogy to inertia in physics, '...a character with an unchanged character state will remain unchanged and a character experiencing consistent directional change will maintain that evolutionary pattern between generations of a lineage unless an external resultant force acts on it'. Burt points out that in most situations we are unable to observe each generation, and so offers an alternative operational definition, '...a character with an unchanged character state will remain unchanged and a character experiencing consistent directional change will maintain that evolutionary pattern between branches of a phylogenetic tree unless an external resultant force acts on it'.

Burt's (2001) analogy with physics is historically interesting, in that at first sight he appears to retrace the steps of the pre-Synthesis orthogeneticists in drawing a close parallel with physical inertia; moreover, he ignores subsequent attempts in the biological literature to give phylogenetic inertia more biological meaning (e.g. Simpson, 1944; Wilson, 1975; Edwards & Naeem, 1993). Burt (2001) did not discuss how the various terms he defines might actually be studied with comparative data. We can ask, however, 'What are the forces?' Clearly, they are not to be interpreted as physical forces, but simply the usual evolutionary processes of natural and sexual selection, mutation, genetic drift, and gene flow. Here the analogy with physics breaks down for various reasons. Unless overridden by selection, random mutation and genetic drift will cause genetic and hence phenotypic evolution in any finite population. Therefore, (absolute) stasis cannot generally be assumed as the null hypothesis for trait evolution. For quantitative phenotypic traits, selection can produce no change in the population mean (stabilizing selection), directional change or even disruptive change (selection against the mean), yet the process is fundamentally the same in all

cases: differential survival and reproduction of individuals determined by their genetically heritable traits.

Phylogenetic 'inertia' and 'constraint'

The various meanings and applications of 'constraint' in evolutionary biology have been discussed in great depth (e.g. Maynard Smith *et al.*, 1985; Antonovics & van Tienderen, 1991; Janson, 1992; McKittrick, 1993; Schwenk, 1995), yet the relationship between constraint and phylogenetic inertia is unclear. Our own view is that modern biologists often use the two terms in an interchangeable fashion, and often casually. We surveyed five popular textbooks of evolutionary biology in order to gauge the 'received view' of constraints as well as phylogenetic inertia (Price, 1996; Ridley, 1996; Futuyma, 1998; Strickberger, 2000; Freeman & Herron, 2001). Surprisingly, none of the examined texts list 'phylogenetic inertia' in the table of contents, glossary or index, and inspection of each failed to find mention of the topic. However, all had some discussion of the role of constraints in evolution. A consensus textbook definition of 'constraint' might be: 'A property of a trait that, although possibly adaptive in the environment in which it originally evolved, acts to place limits on the production of new phenotypic variants'. Similarly, Derrickson & Ricklefs (1988, p. 418) defined 'phylogenetic constraints broadly as differentiation of the evolutionary responsiveness of the phenotype, which may result from intrinsic factors (the genetic covariation patterns) or extrinsic factors (the array of selective pressures impinging on diverse members of a clade)'. Thus, developmental and genetic processes can restrict the types of phenotypes that arise in the future and, given that the environment (and hence selective regime) changes over time, yesterday's adaptation may become tomorrow's constraint.

Ridley (1996) provides a good discussion of developmental and historical constraints; Futuyma (1998) describes several kinds of constraints (e.g. physical, genetic, developmental), and discusses their consequences for evolution. Among the consequences listed by Futuyma (1998) are the absence of adaptive characters (organisms may be constrained from evolving adaptive traits because of a lack of variation in the required direction), presence of directional trends (trends may be observed because developmental pathways make some variants more likely than others), and low rates of evolution (because of limited genetic variation). Note that all three of these consequences have been used as descriptions of phylogenetic inertia, from various perspectives. For example, Simpson (1944) describes inertia as directional change, and several other authors describe phylogenetic inertia as little to no evolutionary change even when selection should be favouring change (e.g. Wilson, 1975; Edwards & Naeem, 1993).

Examples of empirical studies that invoke phylogenetic inertia

Modern empirical studies, especially of behaviour, often invoke 'phylogenetic inertia'. Shapiro (1981), for instance, analysed egg deposition in pierid butterflies. He suggested phylogenetic inertia as a reason why some butterflies disperse their eggs on their new hosts in an inappropriate manner (adaptive on previous hosts, now nonadaptive on the new, current hosts), resulting in a shortage of laying sites. Peterson (1991) attributed the occurrence of delayed maturation in the soft-part colour of some New World jays to phylogenetic inertia, as a hypothesized strong relationship between delayed maturation and plural breeding (groups that have two or more breeding pairs) was not observed. Peterson suggested that a loss of plural breeding in some species may not have coincided with selection against delayed maturation.

Bon *et al.* (1995) described the behaviour of a population of mouflon (wild sheep, *Ovis gmelini*) in the absence of predators. Ewes with older lambs isolate themselves in safe ranges, even when predators are absent, an effect attributed to phylogenetic inertia. The behaviour remains although the selection (predation by foxes) no longer exists. Chu (1994) studied the evolution of delayed plumage maturation in shorebirds and concluded that it is an incidental consequence of the phylogenetic inertia (retention) of molts in this group. Prey handling in *Eumeces gilberti* lizards was studied by de Queiroz & de Queiroz (1987), who concluded that headfirst ingestion of prey was ancestral in tetrapods. Thus, this behaviour could not necessarily be considered adaptive in *E. gilberti*, and hence was attributable to phylogenetic inertia. Strike-induced chemosensory searching (chemosensory searching following an attempted predation event) in anguid lizards was attributed to phylogenetic inertia and not adaptation by Cooper (1995), who found that the phylogenetic evidence favoured an ancient origin for this stereotypic behaviour. Inflexibility in the social structure of hamadryas baboons (compared with common baboons) was attributed to phylogenetic inertia by Barton *et al.* (1996), who noted that the typical baboon social structure persists in captivity in hamadryas baboons, whereas in common baboons other social structures appear readily.

Sih *et al.* (2000) studied antipredator behaviour in salamanders (the closely related *Ambystoma barbouri* and *A. texanum*). They concluded that the behaviour of *A. barbouri* in the presence of predatory sunfish could be attributed to phylogenetic inertia because *A. barbouri* showed ineffective antipredator behaviour, similar to *A. texanum*. *A. barbouri* evolved from an *A. texanum*-like ancestor, and because *A. texanum* lives in fishless ephemeral ponds, *A. barbouri* probably inherited its antipredator behaviour from an ancestor that also lived in

fishless ponds. *A. barbouri's* antipredator behaviour appears not to have been moulded by selection in the new predator-rich environment.

The role of phylogenetic inertia in evolution has been discussed in depth with reference to sexual dimorphism in body size and canine size in primates. Cheverud *et al.* (1985) and Lucas *et al.* (1986) argued for the role of phylogenetic inertia in the evolution of body size and canine size, respectively. The results of Cheverud *et al.* (1985) have been heavily criticized by Ely & Kurland (1989), and their implementation of the autocorrelation comparative method (see below) apparently included a mathematical error (Rohlf, 2001). Various other studies have also purported to find evidence for a phylogenetic 'effect' on dimorphism in body size (e.g. Smith & Cheverud, 2002). However, other authors have claimed no role for phylogenetic inertia in canine tooth size of primates (Plavcan & van Schaik, 1992).

What these examples show is not whether phylogenetic inertia was confirmed or falsified in each case – however, defined by the researcher – but that phylogenetic inertia is a widely used explanation for the existence of some biological phenomena. Apparently, the concept continues to be of value whether or not the definition is clear or the criteria for establishing its existence are adequate.

Phylogenetic inertia and adaptation: alternative hypotheses?

Many current researchers view phylogenetic inertia and adaptation by natural selection as alternative hypotheses for the presence (or absence) of a character in a taxon. Most often, phylogenetic inertia is viewed as a null hypothesis against which to test hypotheses of adaptation (e.g. Edwards & Naeem, 1993). However, this is not always a clear distinction. Indeed, phylogenetic inertia has been called a 'last explanatory resort' by Shapiro (1981), and there is some concern that the two hypotheses (adaptation and phylogenetic inertia) actually occupy different 'levels of analysis' (*sensu* Sherman, 1988; see Edwards & Naeem, 1993; Reeve & Sherman, 2001). Phylogenetic inertia may be an explanation of the *origin* of a trait, whereas if stabilizing selection is acting, adaptation by natural selection may be involved in the *maintenance* of the trait. Thus, traits are a product both of their evolutionary history and natural selection in the recent and current environment. Alternatively, a trait may have evolved originally by natural selection, experienced such strong selection that genetic variation was eliminated, and then persisted because of the lack of genetic variation [point 1 of Wilson (1975) and Edwards & Naeem (1993)], even in the face of altered environmental conditions and hence a changed selective regime. Moreover, as many workers have noted, if related species tend to share environmental characteristics, and hence selective regimes, then we would expect them also to

share traits that are adaptive for those regimes. In any case, casting phylogenetic inertia and adaptation by selection as alternative hypotheses may be inappropriate.

Orzack & Sober (2001) argue that both phylogenetic inertia and adaptation can contribute to trait values, and that they can be considered as orthogonal factors in an evolutionary, statistical analysis of trait values (see also Reeve & Sherman, 2001). Orzack and Sober define phylogenetic inertia as the influence of the initial state of a character on its end state, an interpretation that is at least compatible with Wilson (1975). This definition implies that phyletic stasis (absence of evolutionary change in a trait) is neither sufficient nor necessary evidence for phylogenetic inertia, in contrast to many recent uses of the term in the empirical literature (see previous section; Reeve & Sherman, 2001). Lack of change in a trait is not sufficient because the trait may be under stabilizing selection (Griffiths, 1996). It is also not necessary because daughter species always inherit at least some trait values from their ancestors. It then becomes a question of how much of each trait is attributable to adaptation to the current environment, and how much to ancestry (Edwards & Naeem, 1993).

Phylogenetic inertia and analytical methods for comparative data

Following publication of two seminal papers in 1985 (Cheverud *et al.*, 1985; Felsenstein, 1985), quantitative and statistical aspects of 'the comparative method' (*sensu* Harvey & Pagel, 1991) have advanced tremendously. These advances have followed from at least six ideas: (1) adaptation by natural selection should not be inferred casually from comparative data; (2) independent phylogenetic information can greatly increase the types and quality of inferences that can be drawn from comparative data (e.g. estimation of character states for hypothetical ancestors); (3) among-species data cannot be assumed to represent independent and identically distributed samples from a 'population' for purposes of statistical analyses; (4) assumptions about the way characters have evolved (such as by a process like Brownian motion or a more complicated model (e.g. see Garland *et al.*, 1993; Hansen, 1997; Orzack & Sober, 2001; Freckleton *et al.*, 2002; Martins *et al.*, 2002) are required for statistical inferences; (5) choice of species to include in a comparative study should be guided by knowledge of their phylogenetic relationships; (6) most comparative studies are purely correlational, so the ability to draw causal inferences from them (e.g. about the importance of natural selection in shaping biological diversity) can be greatly enhanced by additional types of information, such as can be obtained from experimental studies of selection acting within present-day populations or mechanistic studies of how organisms work (e.g. Coddington, 1988; Baum & Larson, 1991; Brooks & McLennan, 1991; Harvey & Pagel, 1991; Lynch, 1991; Eggleton &

Vane-Wright, 1994; Garland & Adolph, 1994; Garland & Carter, 1994; Leroi *et al.*, 1994; Doughty, 1996; Rose & Lauder, 1996; Autumn *et al.*, 2002).

These ideas have led to refinement of the analysis of adaptation through comparative studies and the development of various phylogenetically based statistical methods. For continuous-valued characters, four major statistical methods have emerged (recent reviews in Garland *et al.*, 1999; Rohlf, 2001): phylogenetically independent contrasts (Felsenstein, 1985; Garland *et al.*, 1992), phylogenetic autocorrelation (Cheverud & Dow, 1985; Cheverud *et al.*, 1985; Gittleman & Kot, 1990; Gittleman & Luh, 1994), generalized least-squares approaches (Grafen, 1989; Martins & Hansen, 1997; Garland & Ives, 2000), and Monte Carlo simulations (Martins & Garland, 1991; Garland *et al.*, 1993). Some of these approaches have been linked, directly or indirectly, to the study of 'phylogenetic inertia'. In this section, we briefly review these and some related methods in order to consider whether they can in fact provide measures of phylogenetic inertia in comparative data sets. A thorough discussion of methods is beyond the scope of this paper, but some additional discussion can be found in Blomberg *et al.* (in press).

Biologists have long recognized that distinct evolutionary lineages (clades) may show quantitative differences in such traits as body size, brain size or metabolic rate (the latter two after correction for correlations with body size). These differences have often been referred to as 'grade shifts' (Huxley, 1958; Simpson, 1961). As reviewed in Harvey & Pagel (1991), long before 1985 nested ANOVAS and ANCOVAs were used to partition variance among taxonomic levels (implicitly assumed to represent clades), and differences that were identified were sometimes discussed in terms of phylogenetic inertia or constraint. More recently, it has been recognized that the incorporation of phylogenetic information into clade comparisons greatly reduces the probability of finding statistically significant differences (Garland *et al.*, 1993). Moreover, even if clade differences are identified, attributing their origin to any particular clade-specific feature is problematic because most clades exhibit many synapomorphies, i.e. shared, derived features that are unique to themselves. Thus, although modern phylogenetically based statistical methods, such as Monte Carlo simulations, can protect us from inflated type I error rates during clade comparisons, they do not allow us to infer phylogenetic inertia if we do find significant differences among clades (see also Derrickson & Ricklefs, 1988).

Several authors have implied that the use of phylogenetically independent contrasts (Felsenstein, 1985) allows insight concerning 'phylogenetic inertia'. For example, Manning & Chamberlain (1993), using an independent contrast method, found significant associations between fluctuating asymmetry and canine tooth dimorphism, canine length, mass dimorphism and competition type. They conclude that 'Phylogenetic

inertia did not account for the association between fluctuating asymmetry and sexual selection [in primates]'. Similarly, Manning & Chamberlain (1994), after conducting an analysis using phylogenetically independent contrasts, conclude '...it is unlikely that phylogenetic inertia can explain the relation between gametic redundancy and haploid chromosome number [in primates]'. Manning and Chamberlain therefore interpret phylogenetic comparative analyses as falsifying the hypothesis of phylogenetic inertia if phylogenetically 'correct' statistics show significant results, because in using these methods the likely lack of independence of trait values among related species has been eliminated (at least in principle). Similarly, Hosken *et al.* (2001), after conducting an analysis using independent contrasts conclude, '...after using independent contrasts to control for phylogenetic inertia in these data, baculum length [in bats] was not significantly associated with mating system, testis mass or body mass', and Iwaniuk *et al.* (1999) performed an independent contrasts analysis on brain size and forelimb dexterity in carnivores, 'to account for confounding effects of phylogenetic inertia...'

We argue that, in general, phylogenetically independent contrasts are ill-suited to the study of phylogenetic inertia because the mathematical definition of independent contrasts attempts to remove all effects of ancestry from the calculation (see also Orzack & Sober, 2001). [The same arguments would apply to existing generalised least squares (GLS) methods because independent contrasts are a special case of them (Garland & Ives, 2000; Rohlf, 2001).] A simple example will illustrate this fact. Suppose that an ancestral species A (with trait value a) undergoes speciation to form two daughter species, D_1 and D_2 , with trait values d_1 and d_2 , respectively. The evolution of species A to D_1 has caused the trait to evolve from state a to state $d_1 + a$ (i.e. by an amount equal to d_1). Similarly, evolution of A to D_2 has caused the trait value to evolve to $d_2 + a$. The total amount of evolution is therefore $(d_2 + a) - (d_1 + a) = d_2 - d_1$, which is independent of a , and hence independent of A (and all other ancestors). [The actual calculation of contrasts involves division of the differences between species by the square root of the sum of the branch lengths leading from their ancestor (see Felsenstein, 1985), which does not affect the present argument.] As independent contrasts attempt to make orthogonal all previous evolution in the calculation, they cannot be used to measure phylogenetic inertia in any simple way, because inertia is related in some way to previous evolutionary history.

Nevertheless, Burt (1989) proposes a test based on independent contrasts for what he describes as phylogenetic inertia. Burt (1989) argues that if one trait changes faster than another, then there should be a negative relationship between the degree of change within each contrast for the two traits (the within-contrast 'slope', calculated as the contrast in the X variable divided by the contrast in the Y variable), and

divergence time. The reasoning behind this is that if changes in trait X cause changes in the selective regime experienced by trait Y, but trait Y has not responded in time to reach its optimum, then the faster X changes, the more Y will 'lag' behind its optimum, leading to an association between the within-contrast 'slope' and time since divergence. We are unaware of any applications of Burt's method, aside from his own analysis of the data of Sessions & Larson (1987). Deaneer & Nunn (1999) use a similar method, except they examine the relationship of the residuals of the contrasts (of one trait regressed on the other) with divergence times. Under the 'lag' hypothesis, small residuals will be associated with shorter divergence times.

The definition of phylogenetic inertia according to Burt (1989) is the same as the 'evolutionary lag' of Deaneer & Nunn (1999) (although the latter authors do not cite the former). Evolutionary lag occurs when changes in one trait occur later in evolution than changes in a second trait. Lag can thus be caused by different strengths of selection on each trait, or by phylogenetic inertia (*sensu* Wilson, 1975; Edwards & Naem, 1993) owing to insufficient genetic variation in the lagging trait. [According to Simpson (1944), the origin of the term evolutionary lag is probably Darlington (1939).] The Burt and Deaneer-Nunn methods for measuring evolutionary lag are conceptually simple, but they do not correspond to most definitions of phylogenetic inertia as usually expressed. Lag can *cause* a pattern that may be recognized as phylogenetic inertia, but lag alone is not the same as phylogenetic inertia.

In one sense, phylogenetic inertia may correspond to some type of lag: lag of a trait in tracking environmental optima set by a particular selection regime. To quote Simpson (1944, p. 179), 'Response to selection pressure is not instantaneous, and inertia, in the sense of lag in following a shifting optimum, is an important element in evolution'. This definition has added to the confusion surrounding the terminology of phylogenetic inertia. Simpson's final likening of inertia to lag, after earlier defining evolutionary inertia as rectilinear evolution, illustrates one of the problems of the definition of phylogenetic inertia. Some authors use it as a synonym for phylogenetic or evolutionary lag. For example, Sih *et al.* (2000) use phylogenetic inertia in the sense of Simpson (1944), when describing the lack of adaptive response of the behaviour of salamanders (*Ambystoma barbouri*) to predation by sunfish (the selective regime).

The phylogenetic autocorrelation method introduced by Cheverud & Dow (1985) and Cheverud *et al.* (1985) would appear to show more promise for quantifying phylogenetic inertia. Indeed, the latter paper is titled 'The quantitative assessment of phylogenetic constraints in comparative analyses...' The method works by fitting a model which partitions trait values (mean values for a series of species or populations) into (1) shared phylogenetic and (2) independent

components: $\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{e}$, where \mathbf{y} is the vector of trait values, \mathbf{W} is a weighting matrix that expresses the phylogenetic relationships of the species under study, \mathbf{e} is a vector representing values unique to each terminal taxon, and ρ is a scalar autocorrelation coefficient. It was suggested that \mathbf{e} could be used to test hypotheses about adaptation among terminal taxa. This has been criticized, however, because (subject to limitations of estimation), \mathbf{e} is intended to consist only of values unique to the terminal taxa and does not contain information about higher taxa. It is not clear why an hypothesis of adaptation should be tested only with variation that has arisen since the most recent furcations in a phylogeny (Harvey & Pagel, 1991; Garland *et al.*, 1999).

Use of the autocorrelation method for testing hypotheses of phylogenetic inertia centres on interpretation of both the autocorrelation coefficient, ρ , and the 'true R^2 ' [both of which may be affected by transformation of the \mathbf{W} matrix (Gittleman & Kot, 1990; Garland *et al.*, 1999)]. Positive and statistically significant ρ -values are thought to indicate significant phylogenetic inertia, and one can compare ρ -values among traits to establish which traits exhibit more or less phylogenetic inertia (e.g. Morales, 2000). Additionally, one can calculate the true R^2 , which provides a measure of the proportion of the total phenotypic variance explained by phylogeny and can be used to judge the fit of the model (Cheverud *et al.*, 1985; Martins & Hansen, 1996). R^2 has also been used as a measure of the phylogenetic 'effect' in its own right (Gittleman & Kot, 1990; Gittleman *et al.*, 1996a). However, it is clear from the statistical model that ρ only represents the degree of similarity among species, given their phylogeny. This is similar to Ridley's 1983 definition of phylogenetic inertia (shared traits among species), but it lacks any explanatory power. As with other statistics that have been derived solely from comparative data (and a phylogenetic tree), ρ does not provide us with any information on the cause of species resemblance. Similar criticisms can be made of related techniques that claim to measure phylogenetic inertia, including phylogenetic correlograms (Gittleman & Kot, 1990; Gittleman *et al.*, 1996a,b; Diniz-Filho, 2001) and phylogenetic eigenvector regression (Diniz-Filho *et al.*, 1998). Rohlf (2001) discusses the above methods and concludes that phylogenetic autocorrelation cannot be made equivalent to independent contrasts or generalized least-squares methods under any known model of evolution; thus, it is difficult to interpret biologically the results of an autocorrelation analysis. Rohlf (2001) also points out a mathematical error in all previous calculations using the autocorrelation method, which complicates interpretation of published results.

An important phylogenetically based statistical method was introduced by Lynch (1991). By analogy with methods from quantitative genetics, Lynch used a mixed-model formulation and maximum likelihood methods to estimate 'phylogeny-wide' mean values for

characters, the variance–covariance structure of the components of the taxon-specific means, and the estimated mean phenotypes for ancestral characters. Methods to test hypotheses of correlations among characters were also presented. Lynch (1991) argues that the variance–covariance structure of phylogenetic effects in his model can be used to describe phylogeny-wide macroevolutionary patterns, whereas the variance–covariance structure of the residual effects may be used to describe microevolutionary patterns in the data, once measurement error is taken into account. In large part because of computational difficulties, the method has rarely been applied (but see Christman *et al.*, 1997). However, recent work has provided usable algorithms for estimating parameters in the statistical model (although this will generally require large sample sizes) and begun to clarify how it relates to other comparative methods (Freckleton *et al.*, 2002; Martins *et al.*, 2002; E. A. Housworth and M. Lynch, pers. comm.). It is important to note that Lynch (1991) presented his method as a way to describe patterns in the data and not to make causal inferences, and he concludes with remarks that empirical work in ecology and genetics is necessary to shed light on the causal factors that underlie patterns that occur in comparative data.

Maddison & Slatkin (1991) proposed to test for phylogenetic inertia in discretely valued characters via ancestor reconstruction followed by counting the number of evolutionary transitions on a tree. After ancestor reconstruction, the number of character changes in a clade are tabulated, and fewer changes imply greater phylogenetic inertia. One can then compare multiple traits for a given tree. Maddison & Slatkin (1991), under a heading titled ‘Test for Phylogenetic Inertia’, use permutation of character values on a given phylogeny to establish the null distribution of the minimum number of transitions required by the character. If the observed number of transitions is relatively small, then it can be concluded that the character is evolving slowly enough to retain phylogenetic information (Maddison & Slatkin, 1991). Although this approach tests for phylogenetic structure in the data, like other purely comparative approaches it offers no clues as to why such structure might exist. Moreover, although Maddison & Slatkin (1991) equate phylogenetic inertia with low rates of evolution (see also Garland, 1992; Garland & Ives, 2000), this is not necessarily a property of phylogenetic inertia under any of the definitions discussed above.

We have devised a test similar to that of Maddison & Slatkin (1991), but for continuous-valued characters (Blomberg *et al.*, 2001; in press). However, we do not claim to be measuring ‘phylogenetic inertia’. Instead, we simply hope to detect ‘phylogenetic signal’, which we define as a tendency for related species to resemble each other more than they resemble species drawn at random from the tree (see Fig. 1 for a hypothetical example). [This meaning of the term ‘phylogenetic signal’ is similar

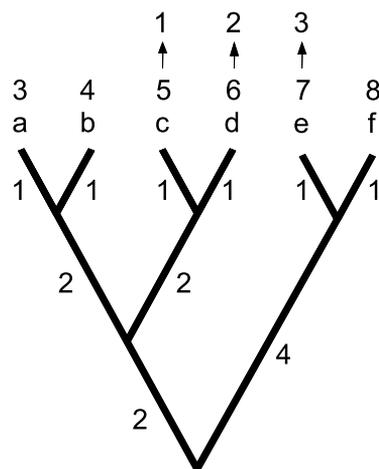


Fig. 1 Hypothetical example illustrating the presence of phylogenetic signal in comparative data, and how adaptation in response to natural selection can reduce phylogenetic signal [numbers adjacent to branches are branch lengths in units of expected variance of character evolution (see Felsenstein, 1985; Martins & Garland, 1991; Garland, 1992; Garland *et al.*, 1992)]. For a hypothetical tree with six species (a–f) and trait values of 3–8, significant phylogenetic signal is detected ($P = 0.047$, based on 1000 randomizations) by the randomization test proposed in Blomberg *et al.* (2001; in press). If species c, d, and e experienced a change in selective regime and evolved to have values as shown (1, 2, 3, respectively), then phylogenetic signal would be obscured (randomization $P = 0.294$). Blomberg *et al.* (in press) propose a statistic, which they term K , that can be used to indicate the amount of phylogenetic signal that is present relative to the amount expected for the given topology, branch lengths, and under a Brownian motion model of character evolution. For these two examples, K is 1.38 and 0.52, respectively. For discussion of other statistics that can quantify the amount of phylogenetic signal, see Freckleton *et al.* (2002) and Blomberg *et al.* (in press).

in spirit to recent usage in systematic biology (e.g. Hillis & Huelsenbeck, 1992), and is also similar to the ‘phylogenetic effect’ of Derrickson & Ricklefs (1988).] An important fact to note is that, on average, for any hierarchical tree, closely related species will tend to resemble each other under such simple evolutionary models as Brownian motion. This resemblance constitutes phylogenetic signal, and its presence, then, clearly does not *require* the invocation of such processes as natural selection (indeed, adaptation via natural selection may often serve to obscure phylogenetic signal, as shown in Fig. 1). That is, random genetic drift alone, occurring along a hierarchical phylogeny, will result in a general tendency for related species to resemble each other. This is why we favour the term phylogenetic signal: it carries no connotation of lack of genetic variation, developmental constraint, character interaction, etc. Whether or not statistically significant phylogenetic signal is detected for any given trait on any given tree will depend on the statistical power of the test as well

as the presence of various factors that may lower the amount of signal, including measurement error in the trait data, errors in the phylogenetic topology, errors in branch lengths, and adaptation or other factors that may cause evolution to deviate from simple Brownian motion. In addition, most comparative studies do not involve data that were gathered for species reared under common environmental conditions, and the possibility of genotype–environment interactions could have unpredictable effects on the amount of phylogenetic signal.

Our method can be implemented with phylogenetically independent contrasts, as follows (Blomberg *et al.*, 2001; in press). First, standardized contrasts are calculated for the original data on the specified topology and branch lengths. Secondly, the variance of these contrasts is calculated. Thirdly, the data are permuted equiprobably across the tips of the tree a large number of times, and again the variance of the contrasts is recorded for each permutation. Permuting the data should, on average, eliminate phylogenetic signal in the data. Finally, the original variance of contrasts is compared with the distribution of variances resulting from the permutation procedure. If the data show phylogenetic signal, then most (e.g. 95%) of the variances resulting from the permuted data should be larger than the original variance. The *P*-value is then calculated as the proportion of variances from the permuted data that are less than the variance of the original contrasts (see example in Fig. 1). [Calculations can be carried out with the Phenotypic Diversity Analysis Programs (PDAP), available from T.G. on request.] Simulations under Brownian motion character evolution indicate that our method has proper type I error rates (it does not falsely claim phylogenetic signal when none exists) and good statistical power for trees with 20 or more species (Blomberg *et al.*, in press).

Most recently, Pagel (1999), Freckleton *et al.* (2002), and Blomberg *et al.* (in press) have emphasized that transformations of branch lengths, such as originally proposed by Grafen (1989), can be used to compare the fit of a continuum of phylogenies, ranging from a star (no hierarchical structure) to a given candidate tree and even to a tree that is more hierarchical than the candidate. Thus, the branch-length transformation parameter can be used as an index of the amount of phylogenetic signal in traits. Such procedures again constitute pattern recognition, but as transformations can be formulated under various explicit models of character evolution, they may allow insight into processes.

In the context of diagnosing comparative data to determine whether or not to use phylogenetically based statistical methods, Abouheif (1999) introduced a test that can detect whether traits are ‘significantly correlated to phylogeny’ or possess ‘historical nonindependence’, which we prefer to call phylogenetic signal. Abouheif’s method involves a test for serial independence in comparative data. This test is an analogue of a ‘runs’

test, except for continuous data, and was not originally developed with phylogenetic data in mind. Abouheif incorporates phylogenetic information by conducting the test on the original data, then randomly ‘rotating’ nodes of the tree and calculating the test statistic a large number of times. The mean of this distribution is compared with a null sampling distribution obtained by randomly shuffling taxa on the tips of the tree, and calculating the test statistic again, repeating the process a large number of times. If the original value for the mean test statistic is greater than (say) 95% of the ‘random’ values, then it can be concluded that significant ‘phylogenetic autocorrelation’ exists in the data. For comparative studies, and following earlier suggestions by Gittleman and coworkers, Abouheif (1999) recommends using a phylogenetically based statistical method if and only if such a diagnostic test rejects the assumption of phylogenetic independence. Moreover, if a method such as independent contrasts is applied, he recommends application of his test to the transformed data to verify that the transformation has indeed rendered the data independent. A disadvantage of Abouheif (1999) method is the *ad hoc* way that phylogeny is incorporated into the analysis by the particular randomization method employed. The effects of phylogeny are ‘randomized out’ of the analysis instead of being incorporated in a fundamental way. This is a result of trying to shoe-horn hierarchical data into a linear data structure, for which the test for serial independence was first designed. Also, the test does not use branch length information, so it is unclear how the results would be affected by different evolutionary models. Finally, no analyses of the statistical power to detect phylogenetic autocorrelation have been presented. In any case, Abouheif (1999) method does not allow causal inferences about the source of any phylogenetic signal that might be detected, and accordingly he does not use the terms phylogenetic ‘inertia’ or ‘constraint’.

Attempting to move beyond a simple pattern definition, Orzack & Sober (2001) define phylogenetic inertia as an influence of trait values of ancestors on the trait values of their descendants. They argue that one can, in principle, examine the degree to which both phylogenetic inertia and natural (or sexual) selection contribute to descendent trait values. They propose a test for phylogenetic inertia, based on taking differences between trait values for multiple pairs of species in a phylogeny. The hypothesis that trait values are affected by ancestors at a given depth in the phylogeny is tested by calculating differences (‘controlled comparisons’) between species which share a common ancestor at the phylogenetic level of interest. To test a phylogenetic inertia hypothesis while controlling for effects of selection, one would need to examine a large number of descendants that have been subject to the same selective regime (with respect to the character in question) and whose ancestors had various character states. However, because the character states of ancestors are not given by data for extant species

(which is all that is typically available in a comparative study), and because the validity of using parsimony to reconstruct ancestral character states is questionable, they suggest that the test be modified to employ multiple pairs of extant species (e.g. X1 and X2, Y1 and Y2) in which the first member of each pair has been subject to the same selective regime and the second members exhibit different character states. The question then becomes, does X1 resemble X2 more than X1 resembles Y2, and does Y1 resemble Y2 more than Y1 resembles X2. This method has not yet been applied to real data. We note, however, that Janson (1992) presents a method in which character-state transitions are modelled using a Markov process. Under this process, current character states depend only on the most recent ancestral character state, so it appears that this approach is in the spirit of Orzack & Sober (2001), although it does require inferences about ancestral states. To our knowledge, Janson's method has not been applied other than in his original paper on seed dispersal syndromes.

Conclusions

The concept of phylogenetic inertia has changed during the history of evolutionary biology, and many researchers have applied it in an *ad hoc* way, often without clear definition, similar to the concept of 'constraints' (Antonovics & van Tienderen, 1991). This makes it difficult to determine what phylogenetic inertia really is, and whether it exists in nature. 'Pattern' definitions of phylogenetic inertia, such as Simpson (1944) and (less clearly) Burt (2001), can be subject to quantification by use of phylogenetic comparative methods (e.g. Maddison & Slatkin, 1991; Gittleman *et al.*, 1996a,b; Abouheif, 1999; Blomberg *et al.*, 2001, in press; Freckleton *et al.*, 2002). 'Process' definitions, on the other hand, are not well suited to study by comparative methods alone because comparative data sets typically do not contain information on the (past) genetic architecture of the traits or the selective regimes to which they have been subjected (see also Leroi *et al.*, 1994; Hansen, 1997; Wagner & Schwenk, 2000; Reeve & Sherman, 2001).

Still, comparative analyses may suggest directions for future research that could test underlying causal hypotheses about what drives or impedes evolution within populations. The kinds of studies that are needed are ones that make the possible *causes* of phylogenetic inertia a serious object of investigation, and not just an explanation of last resort (Shapiro, 1981). For example, studies that focus on quantifying the selective regime, heritabilities of traits, and genetic correlations among traits may shed light on whether organisms evolve along the most favourable genetic trajectories (Schluter, 1996) in response to identified selective agents. Such studies should also take into account morphological, physiological, biochemical, developmental, and genetic 'design limitations', because traits that exhibit correlated

responses to selection on other traits may in fact be under some kind of 'constraint' which limits the evolutionary options of the organism (Wake, 1991; Garland & Carter, 1994). In any case, if phylogenetic inertia is to be invoked as an explanation for biological phenomena, then the term should be defined as carefully as 'adaptation' has been in the recent literature.

Moreover, we believe that study of the relative evolutionary lability of traits is a legitimate use of comparative data. For example, Gittleman *et al.* (1996a,b) used autocorrelation methods to compare behavioural, morphological, and life history traits. Although the autocorrelation methods may be viewed as conceptually flawed (see above) and all previous applications have contained a mathematical error (Rohlf, 2001), the basic aims of Gittleman *et al.*'s papers were worthy (we note also that they did not use the term 'phylogenetic inertia'). New analytical methods should allow statistically valid comparisons of different types of traits on a given tree (i.e. for a given set of species) as well as across trees (e.g. Maddison & Slatkin, 1991; Blomberg *et al.*, 2001, in press; Freckleton *et al.*, 2002). For example, de Queiroz & Wimberger (1993) and Wimberger & de Queiroz (1996) compared different types of traits (categorical data) across trees and found that, contrary to some expectations, behavioural characters (at least those chosen by systematists) show no more homoplasy than do morphological characters. However, the behavioural characters chosen for phylogenetic analysis may be unusual in that systematists use prior judgement before including a particular trait in an analysis. Traits chosen for study by behavioural ecologists may exhibit different properties with regard to homoplasy. Nevertheless, such findings are important because although homoplasy is perceived as problematic in cladistic analyses, it has the potential to provide the best evidence for adaptations to common environments (Brooks & McLennan, 1991; Brooks, 1996; but see Wake, 1991). We urge further development of comparative methods for the study of phylogenetic inertia – or at least phylogenetic signal.

Acknowledgments

We thank F. Bashey, C. K. Ghalambor, A. R. Ives, C. L. Nunn, S. H. Orzack, D. N. Reznick, and E. R. Sober for discussions and/or comments on the manuscript. For access to unpublished manuscripts, we thank E. A. Housworth and R. P. Freckleton. This work was supported by NSF grant DEB-0196384 to TG and A. R. Ives, and a Michael Guyer Postdoctoral Fellowship to SPB.

References

- Abouheif, E. 1999. A method to test the assumption of phylogenetic independence in comparative data. *Evol. Ecol. Res* 1: 895–909.

- Antonovics, J. & van Tienderen, P.H. 1991. Ontoecogenophyloconstraints? The chaos of constraint terminology. *Trends Ecol. Evol.* **6**: 166–168.
- Autumn, K., Ryan, M.J. & Wake, D.B. 2002. Integrating historical and mechanistic biology enhances the study of adaptation. *Quart. Rev. Biol.* **77** (in press).
- Barton, R.A., Byrne, R.W. & Whiten, A. 1996. Ecology, feeding competition and social structure in baboons. *Behav. Ecol. Sociobiol.* **38**: 321–329.
- Baum, D.A. & Larson, A. 1991. Adaptation reviewed: a phylogenetic methodology for studying character macroevolution. *Syst. Zool.* **40**: 1–18.
- Berger, J. 1988. Social systems, resources, and phylogenetic inertia: an experimental test and its limitations. In: *The Ecology of Social Behavior* (C. N. Slobodchikoff, ed.), pp. 157–186. Academic Press, San Diego, CA.
- Blomberg, S.P., Ives, A.R. & Garland, T. Jr 2001. Detecting phylogenetic signal in comparative data. *Am. Zool.* **41**: 1395.
- Blomberg, S.P., Garland, T. Jr & Ives, A.R. (in press). Testing for phylogenetic signal in comparative data. *Evolution*.
- Bon, R., Joachim, J. & Maublanc, M.L. 1995. Do lambs affect feeding habitat use by lactating female mouflons in spring areas free of predators? *J. Zool., Lond.* **235**: 43–51.
- Brooks, D.R. 1996. Explanations of homoplasy at different levels of biological organization. In: *Homoplasy. The Recurrence of Similarity in Evolution* (M. J. Sanderson & L. Hufford, eds), pp. 3–36. Academic Press, San Diego, CA.
- Brooks, D.R. & McLennan, D.A. 1991. *Phylogeny, Ecology, and Behavior*. The University of Chicago Press, Chicago, IL.
- Burt, A. 1989. Comparative methods using phylogenetically independent contrasts. In: *Oxford Surveys in Evolutionary Biology* (P. H. Harvey & L. Partridge, eds), pp. 33–53. Oxford University Press, Oxford.
- Burt, B.B. 2001. Evolutionary stasis, constraint and other terminology describing evolutionary patterns. *Biol. J. Linn. Soc.* **72**: 509–517.
- Cain, A.J. 1964. The perfection of animals. In: *Viewpoints in biology* (J. D. Carthy & C. L. Duddington, eds), pp. 36–63. Butterworth, London.
- Cheverud, J.M. & Dow, M.M. 1985. An autocorrelation analysis of genetic variation due to lineal fission in social groups of rhesus macaques. *Am. J. Phys. Anthr.* **67**: 113–121.
- Cheverud, J.M., Dow, M.M. & Leutenegger, W. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. *Evolution* **39**: 1335–1351.
- Christman, M.C., Jernigan, R.W. & Culver, D. 1997. A comparison of two models for estimating phylogenetic effect on trait variation. *Evolution* **51**: 262–266.
- Chu, P.C. 1994. Historical examination of delayed plumage maturation in the shorebirds (Aves: Charadriiformes). *Evolution* **48**: 327–350.
- Coddington, J.A. 1988. Cladistic tests of adaptational hypotheses. *Cladistics* **4**: 3–22.
- Cooper, W.E. Jr 1995. Strike-induced chemosensory searching by the anguid lizard *Elgaria coerulea*. *Amphib.-Reptilia* **16**: 147–156.
- Darlington, C.D. 1939. *The Evolution of Genetic Systems*. Cambridge University Press, Cambridge.
- Darlington, C.D. & Mather, K. 1949. *Elements of Genetics*. George Allen and Unwin, London.
- Deaner, R.O. & Nunn, C.L. 1999. How quickly do brains catch up with bodies? A comparative method for detecting evolutionary lag. *Proc. Roy. Soc. Lond. B* **266**: 687–694.
- Derrickson, E.M. & Ricklefs, R.E. 1988. Taxon-dependent diversification of life-history traits and the perception of phylogenetic constraints. *Funct. Ecol.* **2**: 417–423.
- Diniz-Filho, J.A.F. 2001. Phylogenetic autocorrelation under distinct evolutionary processes. *Evolution* **55**: 1104–1109.
- Diniz-Filho, J.A., de Sant'Ana, C.E.R.D. & Bini, L.M. 1998. An eigenvector method for estimating phylogenetic inertia. *Evolution* **52**: 1247–1262.
- Doughty, P. 1996. 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.
- Edwards, S.V. & Naem, S. 1993. The phylogenetic component of cooperative breeding in perching birds. *Am. Nat.* **141**: 754–789.
- Eggleton, P. & Vane-Wright, R.I. (eds) 1994. *Phylogenetics and Ecology. Linnean Society Symposium Series*. Academic Press, London.
- Ely, J. & Kurland, J.A. 1989. Spatial autocorrelation, phylogenetic constraints and the causes of sexual dimorphism in primates. *Int. J. Primatol.* **10**: 157–171.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.
- Freckleton, R.P., Harvey, P.H. & Pagel, M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* (in press).
- Freeman, S. & Herron, J.C. 2001. *Evolutionary Analysis*. 2nd edn, 704 pp. Prentice Hall, Upper Saddle River, NJ.
- Futuyma, D.J. 1998. *Evolutionary Biology*. 2nd edn, 763 pp. Sinauer and Associates, Sunderland, MA.
- Garland, T. Jr 1992. Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Am. Nat.* **140**: 509–519.
- Garland, T. Jr & Adolph, S.C. 1994. Why not to do two-species comparative studies: limitations on inferring adaptation. *Physiol. Zool.* **67**: 797–828.
- Garland, T. Jr & Carter, P.A. 1994. Evolutionary physiology. *Annu. Rev. Physiol.* **56**: 579–621.
- Garland, T. Jr & Ives, A.R. 2000. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *Am. Nat.* **155**: 346–364.
- Garland, T. Jr, Bennett, A.F. & Daniels, C.B. 1990. Heritability of locomotor performance and its correlates in a natural population of vertebrates. *Experientia* **46**: 530–533.
- Garland, T. Jr, Harvey, P.H. & Ives, A.R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* **41**: 18–32.
- Garland, T. Jr, Dickerman, A.W., Janis, C.M. & Jones, J.A. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* **42**: 265–292.
- Garland, T. Jr, Midford, P.E. & Ives, A.R. 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. *Am. Zool.* **39**: 374–388.
- Gittleman, J.L. & Kot, M. 1990. Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* **39**: 227–241.

- Gittleman, J.L. & Luh, H.-K. 1994. Phylogeny, evolutionary models and comparative methods: a simulation study. In: *Phylogenetics and Ecology* (P. Eggleton & D. Vane-Wright, eds), pp. 103–122. Academic Press, London.
- Gittleman, J.L., Anderson, C.G., Kot, M. & Luh, H.-K. 1996a. Phylogenetic lability and rates of evolution: a comparison of behavioral, morphological and life history traits. In: *Phylogenies and the Comparative Method in Animal Behavior* (E. P. Martins, ed.), pp. 166–205. Oxford University Press, Oxford.
- Gittleman, J.L., Anderson, C.G., Kot, M. & Luh, H.-K. 1996b. Comparative tests of evolutionary lability and rates using molecular phylogenies. In: *New Uses for New Phylogenies* (P. H. Harvey & A. J. Leigh Brown, J. Maynard Smith & S. Nee, eds), pp. 289–307. Oxford University Press, Oxford.
- Gould, S.J. & Lewontin, R.C. 1979. The spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist programme. *Proc. Roy. Soc. Lond. B* **205**: 581–598.
- Grafen, A. 1989. The phylogenetic regression. *Proc. Roy. Soc. Lond. B* **326**: 119–156.
- Griffiths, P.E. 1996. The historical turn in the study of adaptation. *Br. J. Phil. Sci.* **47**: 511–532.
- Hansen, T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**: 1341–1351.
- Harvey, P.H. & Pagel, M.D. 1991. *The Comparative Method in Evolutionary Biology: Oxford Series in Ecology and Evolution*. Oxford University Press, Oxford.
- Hillis, D.M. & Huelsenbeck, J.P. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *J. Heredity* **83**: 189–195.
- Hosken, D.J., Jones, K.E., Chipperfield, L. & Dixson, A. 2001. Is the bat os penis sexually selected? *Behav. Ecol. Sociobiol.* **50**: 450–460.
- Huxley, J.S. 1958. Evolutionary processes and taxonomy with special reference to grades. *Uppsala University Arssks.* **6**: 21–39.
- Iwaniuk, A.N., Pellis, S.M. & Whishaw, I.Q. 1999. Brain size is not correlated with forelimb dexterity in fissiped carnivores (Carnivora): a comparative test of the principle of proper mass. *Brain, Behav. Evol.* **54**: 167–180.
- Janson, C.H. 1992. Measuring evolutionary constraints: a Markov model for phylogenetic transitions among seed dispersal syndromes. *Evolution* **46**: 136–158.
- Leroi, A.M., Rose, M.R. & Lauder, G.V. 1994. What does the comparative method reveal about adaptation? *Am. Nat.* **143**: 381–402.
- Lucas, P.W., Corlett, R.T. & Luke, D.A. 1986. Sexual dimorphism and tooth size in anthropoids. *Human Evol.* **1**: 23–39.
- Lynch, M. 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* **45**: 1065–1080.
- Maddison, W.P. & Slatkin, M. 1991. Null models for the number of evolutionary steps in a character on a phylogenetic tree. *Evolution* **45**: 1184–1197.
- Manning, J.T. & Chamberlain, A.T. 1993. Fluctuating asymmetry, sexual selection and canine teeth in primates. *Proc. Roy. Soc. Lond. B* **251**: 83–87.
- Manning, J.T. & Chamberlain, A.T. 1994. Sib competition and sperm competitiveness: an answer to 'Why so many sperms?' and the recombination / sperm number correlation. *Proc. Roy. Soc. Lond. B* **256**: 177–182.
- Martins, E.P. & Garland, T. Jr 1991. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution* **45**: 534–557.
- Martins, E.P. & Hansen, T.F. 1996. The statistical analysis of interspecific data: a review and evaluation of phylogenetic comparative methods. In: *Phylogenies and the Comparative Method in Animal Behaviour* (E. P. Martins, ed.), pp. 22–75. Oxford University Press, Oxford.
- Martins, E.P. & Hansen, T.F. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* **149**: 646–667.
- Martins, E.P., Diniz-Filho, J.A.F. & Housworth, E.A. 2002. Adaptive constraints and the phylogenetic comparative method: a computer simulation test. *Evolution* **56**: 1–13.
- Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D. & Wolpert, L. 1985. Developmental constraints and evolution. *Quart. Rev. Biol.* **60**: 265–287.
- Mayr, E. 1982. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Belknap Press, Cambridge, MA.
- McKittrick, M.C. 1993. Phylogenetic constraint in evolutionary theory – has it any explanatory power? *Ann. Rev. Ecol. Syst.* **24**: 307–330.
- Morales, E. 2000. Estimating phylogenetic inertia in *Tithonia* (Asteraceae): a comparative approach. *Evolution* **54**: 475–484.
- Orzack, S.H. & Sober, E. 2001. Adaptation, phylogenetic inertia, and the method of controlled comparisons. In: *Adaptationism and Optimality* (S. H. Orzack & E. Sober, eds), pp. 45–63. Cambridge University Press, Cambridge.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* **401**: 877–884.
- Peterson, A.T. 1991. Sociality and ontogeny of coloration in the blue-and-black jays. *Wilson Bull.* **103**: 59–67.
- Plavcan, J.M. & van Schaik, C.P. 1992. Intrasexual competition and canine dimorphism in anthropoid primates. *Am. J. Phys. Anthropol.* **87**: 461–477.
- Price, P.W. 1996. *Biological Evolution*, 429 pp. Saunders College Publishing, New York, NY.
- de Queiroz, A. & de Queiroz, K. 1987. Prey handling behavior of *Eumeces gilberti* with comments on headfirst ingestion in squamates. *J. Herp.* **21**: 57–63.
- de Queiroz, A. & Wimberger, P.H. 1993. The usefulness of behavior for phylogeny estimation: levels of homoplasy in behavioral and morphological characters. *Evolution* **47**: 46–60.
- Reeve, H.K. & Sherman, P.W. 2001. Optimality and phylogeny. A critique of current thought. In: *Adaptationism and Optimality* (S. H. Orzack & E. Sober, eds), pp. 64–113. Cambridge University Press, Cambridge.
- Ridley, M. 1983. *The Explanation of Organic Diversity: the Comparative Method and Adaptations for Mating*. Oxford University Press, Oxford.
- Ridley, M. 1996. *Evolution*, 2nd edn. Blackwell Science, Cambridge, MA.
- Rohlf, F.J. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* **55**: 2143–2160.
- Rose, M.R. & Lauder, G.V. (eds) 1996. *Adaptation*, 511 pp. Academic Press, San Diego, CA.
- Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**: 1766–1774.
- Schwenk, K. 1995. A utilitarian approach to constraint. *Zoology* **98**: 251–262.

- Sessions, S.K. & Larson, A. 1987. Developmental correlates of genome size in plethodontid salamanders and their implications for genome evolution. *Evolution* **41**: 1239–1251.
- Shapiro, A.M. 1981. The pierid red–egg syndrome. *Am. Nat.* **117**: 276–294.
- Sherman, P.W. 1988. The levels of analysis. *Anim. Behav.* **36**: 616–619.
- Sih, A., Kats, L.B. & Maurer, E.F. 2000. Does phylogenetic inertia explain the evolution of ineffective antipredator behavior in a sunfish-salamander system? *Behav. Ecol. Sociobiol.* **49**: 48–56.
- Simpson, G.G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.
- Simpson, G.G. 1961. *Principles of Animal Taxonomy*. Columbia University Press, New York.
- Smith, R.J. & Cheverud, J.M. 2002. Scaling of sexual size dimorphism in body mass: a phylogenetic analysis of Rensch's rule in primates. *Int. J. Primatol.* **23**: 1095–1135.
- Strickberger, M.W. 2000. *Evolution*, 721 pp. Jones and Bartlett Publishers, Sudbury, MA.
- Wagner, G.P. & Schwenk, K. 2000. Evolutionarily stable configurations: functional integration and the evolution of phenotypic stability. *Evol. Biol.* **31**: 155–217.
- Wake, D. 1991. Homoplasy – the result of natural selection, or evidence of design limitations? *Am. Nat.* **138**: 543–567.
- Wilson, E.O. 1975. *Sociobiology. The New Synthesis*. Harvard University Press, Cambridge.
- Wimberger, P.H. & de Queiroz, A. 1996. Comparing behavioral and morphological characters as indicators of phylogeny. In: *Phylogenies and the Comparative Method in Animal Behavior*. (E. P. Martins, ed.), pp. 206–233. Oxford University Press, Oxford.

Received 22 March 2002; accepted 10 July 2002