

Fitting macroevolutionary models to phylogenies: an example using vertebrate body sizes

Arne Ø. Mooers^{1,2} & Dolph Schluter¹

¹Department of Zoology and Centre for Biodiversity Research, University of British Columbia, Vancouver, Canada V6T 1Z4. E-mail: mooers@bio.uva.nl; ²Institute for Systematics and Population Biology, University of Amsterdam, P.O. Box 94766, 1090 GT Amsterdam, The Netherlands

Keywords: Brownian motion, macroevolution, maximum likelihood, phylogenies, vertebrate body size, evolution

Abstract

How do traits change through time and with speciation? We present a simple and generally applicable method for comparing various models of the macroevolution of traits within a maximum likelihood framework. We illustrate four such models: 1) variance among species accumulates in direct proportion to time separating them (gradual model); 2) variation accumulates with the number of speciation events separating them (speciational model); 3) differences between species are unrelated to phylogenetic relatedness (pitchfork model); and 4) a free model where the trait evolves at its own idiosyncratic rate among lineages. Using species-specific body size, we compare the four models across two data sets: twenty-one clades of vertebrate species, and two clades of bird families. For the twenty-one vertebrate trees, the pitchfork model is most successful, though not significantly, and the most successful by far for the youngest clades. The speciational model seems to be preferred for older clades. For both clades of bird families, the speciational model offers the best fit to family-level body size evolution. However, the pitchfork model does much worse for one clade than for the other, suggesting a difference in the relationship between diversification and body-size evolution in the two groups. These examples highlight some possibilities afforded by this simple approach.

Contents

Introduction	3
Methods	4
The models	4
Data collection	6
Analysis	8
Results	9
Species-level vertebrate trees	9
Families of birds	9
Discussion	9
Choice of scenarios and models	9
Body-size evolution	11
Conclusion	15
Acknowledgements	15
References	15

Introduction

How can we best model the evolution of morphological traits among species? There are several reasons to investigate the fit between traits and *a priori* models. First, a successful fit to an *a priori* model would quantify the phylogenetic component to the evolution of the trait, revealing historical constraints on evolution. Second, the degree of success among models might point to instances of niche-shift, whereby closely related species are divergent in the trait of interest. This would bear directly on the forces associated with speciation. Third, results from contrasting models will lead us to a clearer assessment of how traits change with time, information critical to the success of the modern comparative method (Harvey & Pagel, 1991).

Most current models of trait evolution include parameters such as the time available for evolution along lineages, the form and strength of selection, and degrees of relatedness among groups. The impetus for many of these models comes originally from systematic studies: an evolutionary model of trait evolution offers criteria for the choice of a preferred tree from among several possible by looking for the tree that best fits the known character states of the terminal taxa. These same models can be used to investigate the evolution of characters on trees. For example, maximum parsimony, which searches for the minimum number of transitions among states, performs best under a model of evolution wherein the rate of change of characters is low relative to the number of speciation events and not excessively unequal across lineages

(Felsenstein, 1983). This is the model assumed by default when using maximum parsimony for both tree reconstruction and character reconstructions on trees (see e.g. Brooks & McLennan, 1991). A popular class of models is based on random walks – the Markov process (see Maddison, 1991; Pagel, 1994, 1997; Schluter et al., 1997) for discrete traits and Brownian motion for continuous traits (Edwards & Cavalli-Sforza, 1964; Maddison, 1995; Felsenstein, 1988; Pagel, 1997; see also Schluter et al., 1997). Brownian motion is also at the centre of a widely-used approach to comparative analysis – Felsenstein's (1985) phylogenetic independent contrasts method. The models are simple and tractable, having very few parameters and known statistical properties.

Here, we illustrate how such a model-based approach can be applied to macroevolutionary hypotheses. We test four scenarios representing four different views of the evolution of a particular quantitative trait, body size. Body size is tightly linked to many aspects of species-specific ecology (Peters, 1983; Brown, 1995) and is often the focus of macroevolutionary studies. Using stochastic Brownian motion as our core assumption, we test (i) the gradual model, where (squared) differences in body size among species are expected to be directly proportional to the time available for change; (ii) a speciation model, where these differences in body size are proportional to the number of speciation events separating species; (iii) a pitchfork model, where differences in body size are independent of both time and number of speciation events, such that differences among species are expected to be the same for any pair of species, regardless of phylogeny; (iv) a null model, wherein the trait can evolve at a different rate on each branch of the phylogeny: the resulting tree describes the observed differences among taxa. Other, more complex models can be envisioned, both within a Brownian motion framework, and using different approaches (e.g. the Ornstein-Uhlenbeck process – see Lande, 1976; Felsenstein, 1988; Hansen, 1997).

We test these four models with two complementary data sets. The first data set is an exhaustive sample of complete phylogenetic trees for vertebrate clades constructed from molecular data.

They must be complete for our purposes because the speciation model needs information on all speciation events. Because they are complete, the trees tend to be small and sample events of the relatively recent past.

The second data set consists of two complete clades (the Ciconiiformes and Passeriformes) of avian families. Here we consider the evolution of body size at the family level and above, which is the same level in the tapestry of the birds (Sibley & Ahlquist, 1990) considered by Nee et al. (1992) in their studies of macroevolution. Both clades are parts of accelerated radiations within the phylogeny of birds (Nee et al., 1992). They are fairly large (with 28 and 31 taxa, respectively) and sample events of the more distant past than do those in the first data set.

Because a common process of character change (Brownian motion; Felsenstein, 1985) was used as the basis in all four evolutionary models, we can compare them using the concept of maximum likelihood (Edwards, 1992): the better the fit between model and data, the higher the likelihood returned.

Methods

The models

Underlying all four models is the assumption that changes in body size can be modeled by a continuous random walk (Brownian motion). Under this process, change is continuous, reversals in direction are frequent, the expected rate of change is the same at all stages and in all lineages, and evolution is unbounded (Felsenstein, 1985). Different theories of biological evolution assume different relationships between the amount of time that passes and the amount of evolutionary change that occurs. The gradual model (Fig. 1A) is the only one that conforms precisely to Brownian motion: the squared differences in body size between any two species should be proportional to the time separating them. For example, closely-related species should have similar sizes. This is the pattern expected under conditions of evolution by genetic drift (Lande, 1976; Lynch, 1990). However, natural selection

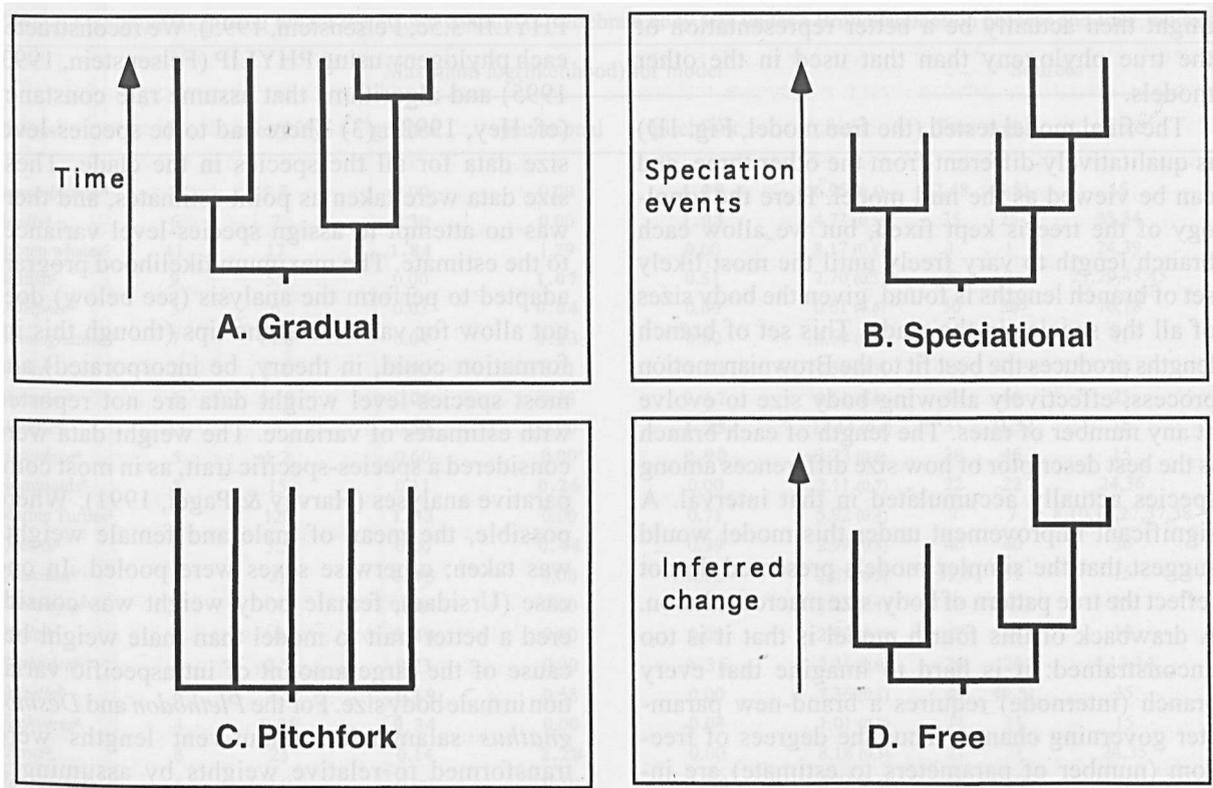


Fig. 1. Characterization of four models for the evolution of quantitative traits. Branch lengths represent the expected amount of change occurring along that lineage. (A) Gradual model, where change is correlated with time. This mirrors the actual phylogenetic tree. (B) Speciational model, where change is correlated with speciation events. (C) Pitchfork model, where there is no phylogenetic component to trait evolution, but each tip is equally divergent from all others. (D) Free model, where each branch is free to vary, and the tree represents the set of branch lengths which best fits the Brownian motion process to the trait data.

can lead to a similar outcome if selection pressures vary unpredictably (Felsenstein, 1981).

The other three models are made to correspond to a Brownian motion process by adjusting relative branch lengths on the phylogenetic tree, essentially manipulating time. Under the speciational model (Fig. 1B), the squared differences between species-specific body size should be proportional to the number of speciation events that separate them, irrespective of the time elapsed. Under this view (for examples, see Rohlf et al., 1990; Harvey & Purvis, 1991: box 2), change in trait values occurs rapidly at, or shortly after, speciation, and stasis follows until the next speciation event. This pattern is associated with the view that speciation is somehow necessary for change to occur (Eldredge & Gould, 1972; Gould & Eldredge, 1993) and may arise for many reasons – for example, if speciation

is associated with niche shifts, which may happen commonly in adaptive radiations. This model is a common default setting for trees used in comparative studies when there is no branch length information.

Under the third model (the pitchfork model, Fig. 1C) squared differences are unrelated to time or speciation events. Closely-related species are no more likely to be similar in size than any two species picked at random. We can test it using the Brownian motion process by setting all internode distances to zero, and setting all terminal branches to unit length, creating a star, or pitchfork phylogeny. Success of this scenario would suggest that phylogenetic history is of no importance to the evolution of body size. Another interpretation of this model is that estimated topology used is very wrong: removing any internal structure (making a pitchfork)

might then actually be a better representation of the true phylogeny than that used in the other models.

The final model tested (the free model, Fig. 1D) is qualitatively different from the other three, and can be viewed as the null model. Here the topology of the tree is kept fixed, but we allow each branch length to vary freely until the most likely set of branch lengths is found, given the body sizes of all the species in the clade. This set of branch lengths produces the best fit to the Brownian motion process, effectively allowing body size to evolve at any number of rates. The length of each branch is the best descriptor of how size differences among species actually accumulated in that interval. A significant improvement under this model would suggest that the simpler models presented do not reflect the true pattern of body-size macroevolution. A drawback of this fourth model is that it is too unconstrained: it is hard to imagine that every branch (internode) requires a brand-new parameter governing change. Thus the degrees of freedom (number of parameters to estimate) are inflated, making simpler models harder to reject.

Data collection

The first data set comprised a group of species-level molecular phylogenies of vertebrates obtained from the literature. Candidate phylogenies had to meet three criteria: (1) They had to include at least $N-1$ of the N known species of the ingroup (“complete phylogenies”, sensu Mooers, 1995). This requirement greatly restricts the pool of available trees but is necessary when considering the speciation model of morphological evolution, where morphological change is concentrated at speciation events. We consider the effects of nonrandom extinction in the discussion. (2) Data sufficient to reconstruct the author’s tree using their algorithm had to be included. Most phylogenies were reconstructed from distance data (genetic distances from allozymes [Nei, 1978; Rogers, 1972], p-values based on RFLP data [Nei & Li, 1979], DNA-DNA hybridization distances [see Sibley & Ahlquist, 1990]), or aligned gene sequence data and commonly used models of base substitution (cf.

PHYLP 3.5c; Felsenstein, 1995). We reconstructed each phylogeny using PHYLP (Felsenstein, 1993; 1995) and algorithms that assume rate constancy (cf. Hey, 1992). (3) There had to be species-level size data for all the species in the clade. These size data were taken as point estimates, and there was no attempt to assign species-level variances to the estimate. The maximum likelihood program adapted to perform the analysis (see below) does not allow for variance at the tips (though this information could, in theory, be incorporated) and most species-level weight data are not reported with estimates of variance. The weight data were considered a species-specific trait, as in most comparative analyses (Harvey & Pagel, 1991). Where possible, the mean of male and female weights was taken; otherwise sexes were pooled. In one case (Ursidae), female body weight was considered a better trait to model than male weight because of the large amount of intraspecific variation in male body size. For the *Plethodon* and *Desmognathus* salamanders snout-vent lengths were transformed to relative weights by assuming a constant allometry among species. For the baleen whales, marine turtles, and the kodkod (an ocelot) species-specific allometric relationships were used (see Table I for references) to estimate body size. All body weights were logarithmically transformed prior to analysis, such that we studied changes in proportional rather than absolute body size. Twenty-one trees from the literature met the criteria for inclusion. The clades ranged from three to thirteen species, including ten groups of birds, six of mammals, three of reptiles, and two of amphibians. We restricted ourselves to molecular phylogenies. We do not feel they are inherently superior, but only molecular phylogenies allow us to assign tentative branch lengths to the resulting trees, using the assumption of the molecular clock.

In addition to body mass, we recorded the age of the group, estimated as the time of the earliest split, the number of species in the group, and the class of molecular data (allozymes, restriction fragments or DNA sequences). The ages were estimates, and were made using a combination of fossil dates, biogeographic information, and molecular calibrations taken from the original papers. For allozyme frequency data, Roger’s D (Rogers, 1972)

Table 1. Maximum likelihood fits for the macroevolution of vertebrate body size under a Brownian motion process and four models.

Taxon ²	N	age (10 ⁶ y)	Maximum log(likelihood) for model ¹				Sources ⁴		
			gradual	speciational	pitchfork	free ³	Tree Age		Size
<i>Ammodramus</i> ^a	8	5.5	0.00	0.09	1.28	6.95 (0.5)	2,48	48,51	15
<i>Anolis</i> ^a	6	7	0.30	0.00	1.03	4.73 (0.5)	25	25	33,34
Baleen whales ^c	11	17	1.84	1.29	0.00	8.17 (0.8)	1	1	24,39
Billfish ^c	9	55	0.00	1.07	0.51	7.70 (0.5)	6	3	27,29,43,45
<i>Didelphis</i> ^d	3	6	0.03	0.04	0.00	0.61 (0.6)	30	30	16,18
<i>Desmognathus</i> ^c	7	14.6	0.04	0.23	0.00	4.60 (0.6)	42	9	41
<i>Equus</i> ^b	3	2.8	0.00	0.01	0.22	1.36 (0.3)	19	51,19	24
<i>Geospiza</i> ^a	6	0.18	0.00	0.33	0.67	4.61 (0.4)	46	46	23
Gruinae ^c	13	4	0.59	0.00	1.25	13.14 (0.4)	31	31,32	15
<i>Heliodoxa</i> ^a	5	1.2	0.60	0.00	0.90	3.23 (0.6)	20	46	15
Hominoids ^c	5	15	0.11	0.26	0.00	2.11 (0.7)	22	22	24,36
Marine Turtles ^c	7	120	1.33	0.00	0.71	5.92 (0.4)	7	7	8,10,17,27,37,38
Ocelots ^a	6	3.5	0.00	0.44	0.39	2.99 (0.8)	40	40	36
<i>Quiscalus</i> ^b	4	21	0.45	0.00	0.02	2.67 (0.3)	51,5	51	15
<i>Phalaropus</i> ^b	3	4	0.004	0.00	0.13	2.27 (0.1)	12	51,12	15
<i>Pipilo</i> ^b	4	4	0.44	0.00	0.09	2.92 (0.3)	47	47	15
<i>Plethodon</i> ^a	6	2.5	0.23	0.00	0.36	2.31 (0.9)	26	26	4,13,14
<i>Spizella</i> ^b	6	7.5	1.19	0.58	0.00	7.36 (0.1)	49	49,51	15
<i>Tachyeres</i> ^a	4	0.06	0.34	0.00	0.04	1.01 (0.9)	11	11	15
<i>Ursus</i> ^a	8	23	0.14	2.28	0.00	8.18 (0.5)	21	21	24,35
<i>Zonotrichia</i> ^b	5	2.2	0.00	0.19	0.50	3.74 (0.4)	50	50	15
SUM			7.67	6.81	8.10	96.58 (0.4)			

¹ Poorest-fit model within a row is given a score of 0.00. The bold entry is the best-fit model.

² a: Allozyme data; b: RFLP data; c: sequence data; d: DNA-DNA hybridization data.

³ P-values in parentheses are for the improvement of the best-fit model over other models using a log-likelihood ratio test with 2N-4 degrees of freedom.

⁴ References:

1-Arnason & Gullberg, 1994; 2-Avise, pers. comm.; 3-Benton, 1993; 4-Bishop, 1943; 5-Bjorklund, 1991; 6-Block et al., 1993; 7-Bowen et al., 1993; 8-Bustard, 1973; 9-Caccone et al., 1994; 10-Carr, 1952; 11-Corbin et al., 1988; 12-Dittman & Zink, 1991; 13-Duncan & Highton, 1979; 14-Dunn & Heinze, 1933; 15-Dunning, 1993; 16-Eisenberg, 1989; 17-Ernst et al., 1994; 18-Gardner, 1982; 19-George & Ryder, 1986; 20-Gerwin & Zink, 1989; 21-Goldman et al., 1989; 22-Goodman et al., 1994; 23-Grant et al., 1985; 24-Grzimek, 1990; 25-Hedges & Burnell, 1990; 26-Highton & Larson, 1979; 27-Hirth, 1982; 29-Joseph et al., 1980; 30-Kirsch et al., 1993; 31-Krajewski & King, 1996; 32-Krajewski & Fetzner, 1994; 33-Losos, 1990; 34-Losos, pers. comm.; 35-Morrow, 1964; 36-Nowak, 1991; 37-Pritchard, 1967; 38-Pritchard, 1971; 39-Ridgeway & Harrison, 1985; 40-Slattery et al., 1994; 41-Tilley & Bernardo, 1993; 42-Titus & Larson, 1996; 43-Wilson et al., 1991; 44-Wilson et al., 1985; 45-Witzell, 1989; 46-Yang & Patton, 1981; 47-Zink, 1988; 48-Zink & Avise, 1990; 49-Zink & Dittman, 1993; 50-Zink et al., 1991a; 51-Zink et al., 1991b.

distances were converted to Nei's D (Nei, 1978) distances using an empirical calibration supplied by N. Grabovac (pers. comm.) before tree construction. While necessarily crude, this allowed ultrametric trees to be constructed for these data.

The second data set is the higher level phylogeny for two clades of birds (the Ciconiiformes and the Passeriformes, Sibley & Ahlquist, 1990). The two bird clades are fairly large (with 28 and 31 tips, respectively), but no raw data are available to

reconstruct the trees for ourselves. We constrained ourselves to the same level in the tree as Nee et al. (1992), roughly the family level, where we can be fairly confident of a complete tree (no missing lineages). The UPGMA tapestry was used, with the branch lengths (in ΔT_{50H} units; Sibley & Ahlquist, 1990) taken to be linearly related to time. Estimates for representative body sizes of the taxa were reconstructed by hierarchical weighting such that speciose taxa do not bias the estimate (Harvey

& Mace, 1982), using weights from Blackburn & Gaston (1994). Species were first averaged within genera, and then genera were averaged within tribes, tribes within subfamilies and subfamilies within families. Under the gradual scenario, the family-level representative body sizes were placed at the highest split within the family, following the conventions of Mooers et al. (1994). This means that body-size estimates made from families that radiated soon after the $10 \Delta T_{50}H$ unit cut-off will be found on short terminal branches, while estimates from late-radiating families will be found on the ends of longer branches. This allows more time for change in groups whose family-level estimates sample less elapsed time. This procedure conforms with the underlying Brownian motion process (see below) and does not bias the results towards preferring one model over another.

Analysis

The goal of the analysis was to determine, for each clade, which model of character change best described the evolution of species-specific body size. To do this we calculated the goodness of fit of each model to the body sizes of the taxa at the tips of the tree. Goodness of fit was measured as the logarithm of the likelihood of each model given the data (body sizes) under the Brownian motion process of character change. Better fits between the data and a model for body-size evolution will have higher log(likelihood) scores.

The fits of alternative models were compared using the differences in their corresponding log(likelihoods). The first three models tested have the same number of estimated parameters. For the gradual model, there is only one rate. For the speciation and pitchfork models, we assume that the traits evolve at different rates on different branches (e.g. at rate = 0 for the zero length branches of the pitchfork model and at different rates for the unit length branches). However, because we set the branch lengths *a priori* (ignoring the actual inferred time between nodes), only one rate parameter is estimated. The other rates are proportional to it: we could estimate the actual rate for any branch on the tree by dividing the single estimated rate

by the branch length taken from the gradual tree. This rate is uninteresting, however: for the speciation model, it would be the average of the fast rate at speciation and zero for the duration of that lineage's existence. Given that these three models are equally parameterized, any two are deemed significantly different when the difference in their log(likelihoods) is 2.0, corresponding to approximately a seven-fold difference in their likelihoods (Edwards, 1992: 180ff). Risch (1992) would prefer a difference of 3.0 before discriminating among models (corresponding to a 20-fold difference in likelihoods).

The fourth model is qualitatively different. The species' body sizes are used to optimize the branch lengths, which are unconstrained by *a priori* hypotheses, and so the model has many more parameters than the first three. The number of extra parameters is $2N-4$ for a tree having N species (the number of branches in an unrooted tree -1). This allows us to judge whether the free model is significantly better than any of the alternatives by comparing twice the improvement in the log(likelihood) to a chi-square distribution with $2N-4$ degrees of freedom (see Goldman, 1993).

Given that the actual topology and number of tips affects the fit of any given data set to a tree, we cannot compare the values of the same model for different clades; for example we cannot state that the gradual model for one group of birds fits n times better than it does for some other group (Goldman, 1993). However, individual trees are independent and the log(likelihoods) for a given model may be summed across the full data set of twenty-one trees to yield an overall measure of the goodness of fit of each model. Because the differences between the fits of different models are relevant, but not their absolute value, we scaled the log(likelihoods) so that the worst-fit model for any data set was designated zero.

All analyses were performed on a modified version of the CONTML program of PHYLIP (version 3.0, Felsenstein, 1993). The modification allowed CONTML to return a log(likelihood) for a data set that was independent of the scale of measurement of body size and total tree length. The modification is available from the authors upon request. The Brownian motion model does not require direc-

tionality (i.e., rootedness) in the tree, and in order to use CONTML, the rooted trees in Fig. 1 were made unrooted before analysis. This does not bias the results in favour of any specific model.

Results

Species-level vertebrate trees

Results for the twenty-one vertebrate trees are listed in Table I. We will denote the log(likelihood) of a model as 'L(model)' in the text. Ursidae is the only group where the best-fit scenario offers a significantly better fit than the alternatives: L(speciational) is at least eight times better than its nearest rival, L(gradual).

There was no significant correlation between the age of the group and the number of species included ($r=0.10$, $p=0.16$, $N=21$, on logarithmically transformed age data). The number of species included in the group did not correlate with the ranking of the models (for all three possible comparisons of models: Mann-Whitney tests for clade size, with groups $L(\text{model } 1) > L(\text{model } 2)$ and $L(\text{model } 2) < L(\text{model } 1)$, $p > 0.5$, $N=21$ for all tests).

By inspection, no one model of evolution was found to be associated with a particular method of tree reconstruction. For none of the clades did the free model, allowing different rates along each branch of the tree, offer a significantly better fit to the data than the simpler models (Table I). Therefore the simpler models cannot be rejected as wholly inadequate descriptions of body-size evolution in these groups.

The pitchfork model was the most successful model overall, fitting the data 1.75 times better than did the runner-up, gradual model (evaluated as $\exp[7.49-6.93]$ from Table I). It returned the highest likelihood in 9 of 21 cases. Furthermore, the pitchfork model seemed to best fit the younger trees (Mann-Whitney test of differences in age of clades, between groups $L(\text{pitchfork}) > L(\text{others})$ and $L(\text{pitchfork}) < L(\text{others})$, $p=0.008$, $N=21$; this is marginally significant with the Bonferroni correction [Rice, 1989] for six independent tests [corrected $\alpha = 0.008$]). By contrast, the speciation model may fit the older trees better (Mann-Whitney test of differences in age of clades, between groups

Table II. Maximum likelihood fits for the macroevolution of body size for higher clades of birds (trees from Sibley & Ahlquist, 1990) under a Brownian motion process and four models.

Taxon	N	Maximum likelihood for model ¹			
		gradual	speciational	pitchfork	best-fit ²
Ciconiiformes	28	6.43	9.66	0.00	42.97 (0.08)
Passeriformes	31	0.00	5.11	3.28	32.65 (0.6)

¹ Poorest-fit model within a row is given a score of 0.00.

² P-values in parentheses are for the improvement of best-fit over other models using a log-likelihood ratio test with $2N-4$ degrees of freedom.

$L(\text{speciational}) > L(\text{others})$ and $L(\text{speciational}) < L(\text{others})$, $p=0.07$, $N=21$).

Families of birds

Table II lists the fits of the four models of body-size change to the two family-level trees. For both groups, the speciation model of evolution offered the best fit to the data. The clades were not equally well fit by the other two models: the pitchfork model did relatively better for the Passeriformes than for the Ciconiiformes. For neither clade did the fourth model, whereby each branch has its own rate of evolution, fit significantly better than the best of the other three models, though this model did relatively better for the Ciconiiformes. The ages of the major groups of extant birds extends into the Cretaceous (Cooper & Penny, 1997), and calibrations presented by Sibley & Ahlquist (1990) for DNA-DNA hybridization data suggests both orders originated between 40-80 million years ago.

Discussion

Choice of scenarios and models

The macroevolutionary scenarios presented here are but a sample of those conceivable. Díaz-Uriarte & Garland (1996) simulated comparative data using fifteen different scenarios; not all are tractable. One entire class of theirs which cannot be tested in our framework is a random model with a trend: with only end points, it would be impossible to estimate the rate at which the mean value for a

trait changed through time. Another class presented by Díaz-Uriarte & Garland (1996) allowed for random motion within preset boundaries. Boundaries on trait evolution are intuitively appealing (trait evolution must certainly be bounded, certainly on a multiplicative scale) and easy to simulate, but difficult to model analytically: estimating boundaries from data would require arbitrary decisions. In Díaz-Uriarte & Garland's simulation study, the boundaries were set arbitrarily to maximize differences between this class and others.

A third class of scenario and one which warrants more attention is true punctuated evolution (*sensu* Eldredge & Gould, 1972): given a model of peripatric speciation, change is restricted to the peripheral isolate, and is uncorrelated with time. Graphically, at every node in the tree, one daughter lineage would be represented with a branch of some length, and the other would have its branch length set to zero. There is no reason that all peripheral isolates should undergo the same expected amount of change (Grafen & Ridley, 1997), but this could serve as a reasonable first approximation. Díaz-Uriarte & Garland (1996) state that this scenario assumes no extinction: this can be relaxed if we assume that extinction is random across lineages, in which case extinction should obscure but not destroy the pattern. This holds for the speciation scenario tested here as well. However, if extinction and speciation rates are positively correlated (Rosenzweig, 1995), then the speciation hypothesis should fare well (Ferris et al., 1979). If there is a bias in future extinction and speciation probabilities between ancestral species and new peripheral isolates (*cf.* Losos & Adler, 1995), then the speciation hypothesis will be hampered. As with random motion within boundaries, true punctuated evolution is easy to simulate, but difficult to model analytically: it is akin to the free scenario, where the data are used to estimate the branch lengths, but with peculiar and discrete constraints: branches can only be of zero or unit length, and one of each must occur at each node.

Lynch (1990) suggested that change might be rapid following a speciation event, and subsequently slow down, perhaps at an exponential rate. At the limit, this scenario approaches the speciation hypothesis, but could be better fit with the gradual

hypothesis and an extra parameter governing the rate at which change slows down after the speciation event. Such a transformation of branch lengths is akin to the power functions advocated by students of the comparative method (Grafen, 1989; Gittleman & Kot, 1990; Garland, 1992; Pagel, 1994) for standardizing phylogenetic independent contrasts. This scenario, and others mentioned above, require the fitting of extra parameters, necessitating log-likelihood ratio tests or Monte-Carlo tests (Goldman, 1993) in order to distinguish between them and simpler hypotheses. More work is required to offer guidelines for assessing when data warrant more complex models.

The Brownian motion process is also only one of several possible models of trait evolution. When originally proposed as a model for estimating phylogenies from quantitative character data (Felsenstein, 1981; see Felsenstein, 1988, for a review), recourse was made to genetic drift as the variance-generating mechanism. For traits under natural selection, Brownian motion was deemed "rather arbitrary" (Felsenstein, 1988: 464). Brownian motion can, however, represent change in traits under selection if the selection pressures are multifarious and constantly changing, or if lineages wander randomly from one regularly-spaced adaptive peak to another, both of which may be reasonable representations over long periods of time. Felsenstein (1985) advocated Brownian motion as a possible model with which to investigate correlated evolution of characters under selection (the main class of characters investigated with the comparative method), and this has become the model of choice in this context. In simulations, the model performs well as an approximation even when the true model is quite different (Martins & Garland, 1991; Díaz-Uriarte & Garland, 1996).

Hansen & Martins (1996; Martins & Hansen, 1997; Martins, 1994; Hansen, 1997; see also Garland et al., 1993; Felsenstein, 1988) have advocated the Ornstein-Uhlenbeck process of character change (Lande, 1976), whereby species are randomly perturbed from, and then return towards, some optimum value, with the rate of return increasing with greater perturbation. This model is offered as a theoretical justification for boundaries on trait evolution and so might be seen as a refinement on

Brownian motion. It is more complex than simple random drift, and requires fitting several extra parameters, regardless of the hypothesis tested, which might not be warranted with small data sets. It is hampered by the assumption that all the species in the clade are centered on the same optimum value, which must also be specified or estimated from the data. It also implies that there is no way to recover distantly past events, as evidence of ancient phenotypes will eventually be erased by the pull toward the central point (Felsenstein, 1988). However, the model is grounded in population genetics theory, and links microevolutionary process with macroevolutionary pattern and so may prove to be a valuable and versatile approach. Martins (1994) and Hansen & Martins (1996) discuss ways in which the fit of the O-U model can be compared with strict Brownian motion.

Crucially, our formal approach assumes that the rate of change for the character(s) under question is the same throughout the tree. This will not generally be the case [e.g., changes in discrete characters often appear to be clumped in trees (Grafen & Ridley, 1997)] for large, disparate clades, but might be less of a problem for more homogeneous clades like those considered here. This will add noise rather than bias when comparing among the simple hypotheses, however, and the assumption is partially tested by comparison of these simple hypotheses with the free scenario. The test is not perfect, however: where the assumption of equal rates make the simple scenarios underparameterized, assuming a new rate of change for every branch in the free scenario makes it overparameterized.

Our phylogenetic, model-based approach to macroevolution has antecedents. Ferris et al. (1979) offered a test for comparing the relative importance of speciation events and time on the probability of gene function loss within the catostomid fishes, a discrete character. Their approach was based on the Poisson process (they assumed irreversibility of gene function loss) and they contrasted gradual and speciation models that differed in the number of estimated parameters, necessitating log-likelihood ratio tests. M. Pagel has presented methods for testing the speciation model for discrete (Pagel, 1994) and continuous (Pagel, 1997) characters in a maximum likelihood context. Like Ferris et al.,

Pagel's approach compares models with different numbers of parameters and considers relative fits via a log-likelihood ratio test.

Garland (1992) offers a graphical method for investigating the relationship between differences between sister groups in trait values and their expected variances. This approach, presented in the context of correlated evolution, could be used to distinguish between the speciation and gradual models of evolution. Finally, there have been a number of tests of strict punctuated evolution (reviewed by Gould & Eldredge, 1993), using both paleontological and recent data, some of which make explicit use of trees.

Body-size evolution

Overall, the model which did not incorporate phylogeny (the pitchfork model) offered the best fit to the species-level trees. The difference is not significant, but this may suggest that phylogeny may not be an important factor in body size evolution, at least at the species level. Body size may evolve idiosyncratically with respect to time and speciation. It is known that body size can evolve very rapidly (Brown, 1995).

Overall, differences in fits among the models were small (though for the Ursidae, we might accept the speciation model as significantly the best of the lot). The Brownian motion process is inherently noisy, and so each model might accommodate quite a range of data when the trees are small, such as is the case here (most trees have 6 or fewer tips). Under a Brownian motion process, the test may have low discriminatory power for small trees.

To explore this point, we analysed a data set with a well-established prior expectation. We took the second axis of a principle component analysis of shape (being mostly size-independent beak measures) for members of *Geospiza* and *Zonotrichia* (data available on request) and subjected these measures to the same four macroevolutionary models. Previous work (Schluter & Nagel, 1995) has implicated changes in beak morphology with speciation within *Geospiza*, and so we predicted that the speciation model should best fit the shape data within this clade, with the pitchfork model doing the worst. We had no such expectation for

the sparrows. The results were in accordance with these predictions – the speciation model offered the superior fit in *Geospiza*, performing 125 times better than the pitchfork scenario. Recall that for body size, the pitchfork model offered the best fit in this clade (Table I). Conversely, for *Zonotrichia* the pitchfork model offered the best fit, with a 14-fold difference between it and the worst, gradual model. So for *Geospiza*, body shape does carry a particular phylogenetic signal, consistent with the idea that changes in shape occur in concert with speciation events.

Another factor which might affect our results is tree size. The trees sampled here are so small that single errors (one nonrandom extinction event or one misleading branch length) may have a large effect. This is particularly devastating for the speciation model. Furthermore, if there is phylogenetic signal in body-size differences among species, the relative success of the pitchfork model for the younger trees may be because those trees are the most misinformative, such that the topology is presenting noise rather than signal. Finally, the gradual model must be viewed with some caution for many of the trees listed here because if the rates of change of the molecules deviate strongly from the molecular clock expectation, then the branch lengths may be poor estimates of elapsed time. Reconstructing the trees without assuming constant rates of evolution did not however improve the fit over the speciation or pitchfork models (unpubl. results).

For both groups of bird families, the speciation model of macroevolution offered a better fit to the body-size data than did the other models. The phylogeny used has come under some criticism (see Mooers & Cotgreave, 1994), and it is likely that the branch lengths taken from this tree are not accurate. This handicaps the gradual model. However, the difference in relative fit of the pitchfork model between the two clades suggests that there may be differences in the pattern of body-size evolution in the two groups – incorporating phylogeny causes a much improved fit for the Ciconiiformes versus for the Passeriformes. This is illustrated in Fig. 2, which contrasts the free trees for the two groups. For the Ciconiiformes (Fig. 2A), long branches tend to emanate from long branches (particularly the path leading to the Procellariidae),

illustrating that there is a phylogenetic component to the evolution of body size. For the Passeriformes (Fig. 2B), the free tree looks distinctly pitchfork-like.

This difference may be explained in several ways. The songbird tree may simply be less accurate, so that the topology is presenting so much noise that a pitchfork phylogeny does well in comparison. In the UPGMA tree of the birds, shorter branches are considered less reliable (Sibley & Ahlquist, 1990; Barraclough et al., 1995). However, the average internode length does not differ greatly between the two clades (mean above-family branch lengths: Ciconiiformes = $1.3 \Delta T_{50}H$ units, Passeriformes = $1.1 \Delta T_{50}H$ units; $p=0.14$ based on a t-test of logarithmically transformed data). A more interesting hypothesis is that changes in body size and diversification are more closely linked in the Ciconiiformes than in the songbirds, such that models which do not incorporate this (the pitchfork model) perform poorly. This has intuitive appeal, as there is more variation in body size among families in the former clade (standard deviation of weights: Ciconiiformes = 0.5; Passeriformes = 0.3). In addition, the free model offered a relatively better fit to the Ciconiiformes (Table II). Both of these groups are much older than the average of the smaller species-level vertebrate trees, where the speciation model performed best for the older clades. This trend, if true, suggests an attenuation in the rate of change through time, and is supported by the work of Lynch (1990).

Fig. 2 also illustrates another use of this explicit model-based approach. Long branches under the free model are lineages where much change has occurred (either because there has been much elapsed time or high rates of change). A simple regression of the free branch lengths on the actual (time elapsed) branch lengths can help to identify those lineages of the tree where rates of change have been particularly high or low, directly analogous to looking for outliers in plots of standardized contrasts in comparative analyses (Garland, 1992; see also Martins, 1994). The difference is that our method focusses on single branches rather than differences between reconstructed sister groups. These free model trees offer a graphical representation of tempo and mode in morphological macroevolution.

The results of this study have implications for

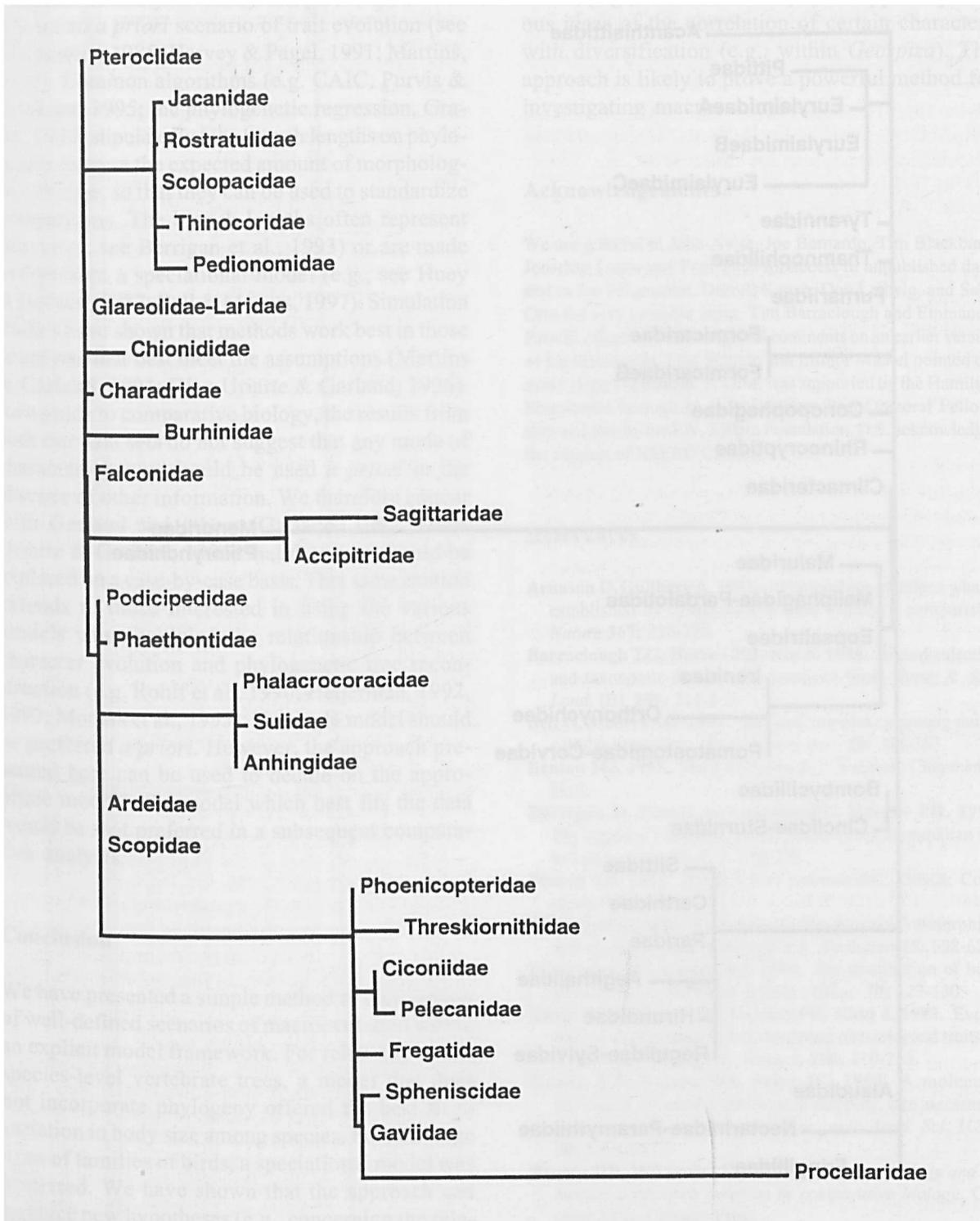


Fig. 2A. The free model for the Ciconiiformes (based on Sibley & Ahlquist, 1990), where the tree represents the set of branch lengths which best explain the body-size data. The tree retains some of the original topology.

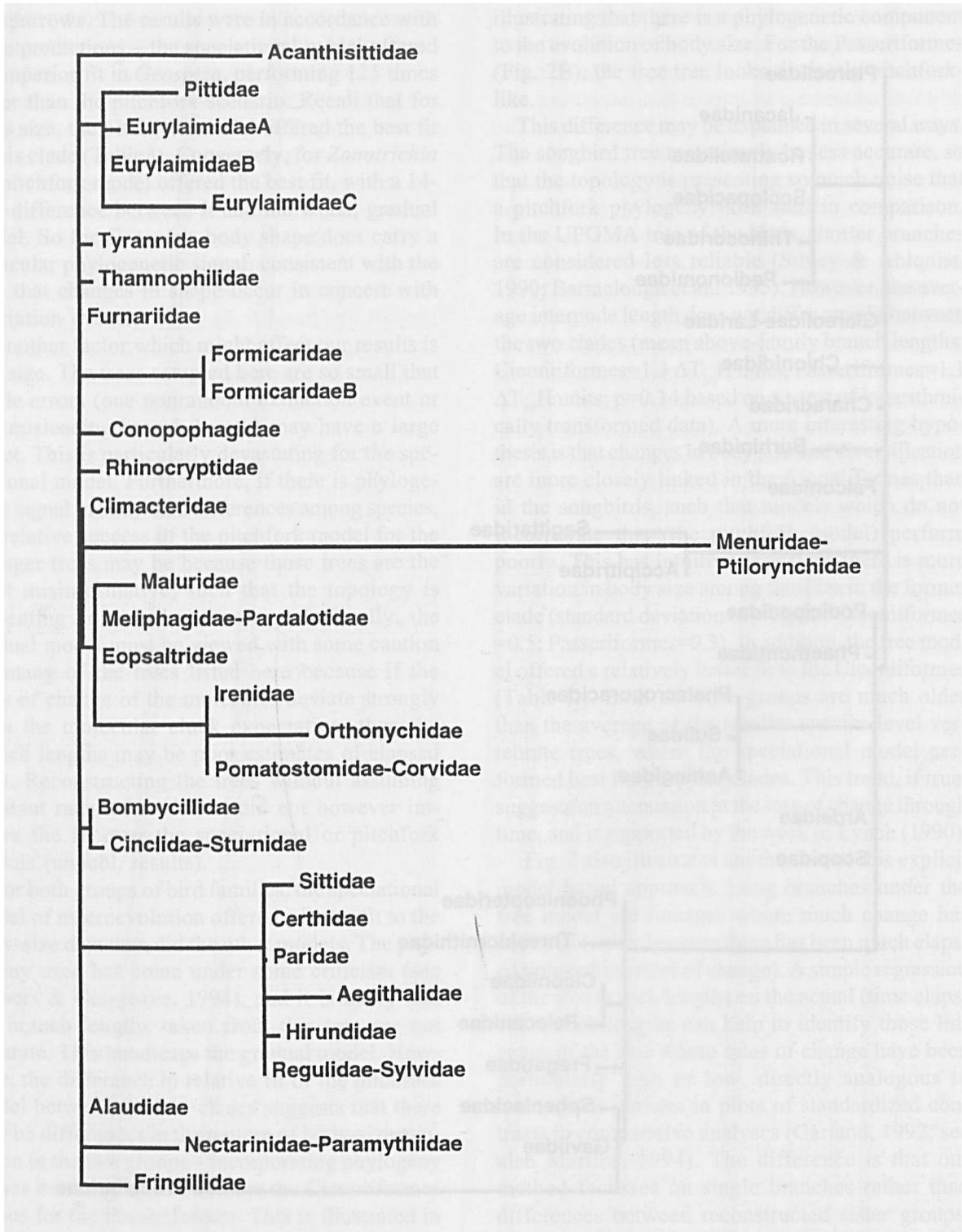


Fig. 2B. The free model for the Passeriformes. The tree resembles a pitchfork phylogeny. The tips are the lineages that arise at $10\Delta T_{50}H$ units; tips with hyphenated names comprise all those families subtending that branch (see Mooers et al., 1994 for full explanation).

comparative analyses. Many comparative methods rely on an *a priori* scenario of trait evolution (see Felsenstein, 1985; Harvey & Pagel, 1991; Martins, 1996). Common algorithms (e.g. CAIC, Purvis & Rambaut, 1995; the phylogenetic regression, Grafen, 1989) stipulate that the branch lengths on phylogenies express the expected amount of morphological change, so that they can be used to standardize comparisons. The branch lengths often represent time (e.g., see Berrigan et al., 1993) or are made to represent a speciation model (e.g., see Huey & Bennett, 1987; Bell & Mooers, 1997). Simulation studies have shown that methods work best in those situations that best meet the assumptions (Martins & Garland, 1991; Díaz-Uriarte & Garland, 1996). As a guide to comparative biology, the results from both our data sets do not suggest that any mode of character change should be used *a priori* in the absence of other information. We therefore concur with Garland and others (Garland, 1992; Díaz-Uriarte & Garland, 1996) that data sets should be explored on a case-by-case basis. This same caution extends to those interested in using the various models when studying the relationship between character evolution and phylogenetic tree reconstruction (e.g. Rohlf et al., 1990; Heijerman, 1992, 1993; Mooers et al., 1995). No single model should be preferred *a priori*. However, the approach presented here can be used to decide on the appropriate model – the model which best fits the data would be that preferred in a subsequent comparative analysis.

Conclusion

We have presented a simple method of comparison of well-defined scenarios of macroevolution within an explicit model framework. For relatively small, species-level vertebrate trees, a model that does not incorporate phylogeny offered the best fit to variation in body size among species. For two large trees of families of birds, a speciation model was preferred. We have shown that the approach can produce new hypotheses (e.g., concerning the relative importance of body size in the diversification of songbirds as compared with the Ciconiiformes), discriminates between alternative modes of evolution

(e.g., within Ursidae) and supports previous ideas of the correlation of certain characters with diversification (e.g., within *Geospiza*). The approach is likely to prove a powerful method for investigating macroevolution.

Acknowledgements

We are grateful to John Avise, Joe Bernardo, Tim Blackburn, Jonathan Losos and Tom Titus for access to unpublished data, and to Joe Felsenstein, Durrell Kapan, Don Ludwig, and Sally Otto for very valuable input. Tim Barraclough and Emmanuel Paradis offered several insightful comments on an earlier version of the manuscript. Fred Schram and Ellinor Michel pointed out many sloppy sentences. A.Ø.M. was supported by the Hamilton Foundation through an E.B. Eastburn Post-Doctoral Fellowship and also by the I.W. Killam Foundation. D.S. acknowledges the support of NSERC Canada.

References

- Arnason U, Gullberg A. 1994. Relationships of baleen whales established by cytochrome b gene sequence comparison. *Nature* 367: 726-728.
- Barraclough TG, Harvey PH, Nee S. 1995. Sexual selection and taxonomic diversity in passerine birds. *Proc. R. Soc. Lond. (B)* 259: 211-215.
- Bell G, Mooers AØ. 1997. Size and complexity among multicellular organisms. *Biol. J. Linn. Soc.* 60: 345-363.
- Benton MJ. 1993. *The fossil record, 2*. London: Chapman & Hall.
- Berrigan D, Purvis A, Charnov EL, Harvey PH. 1993. Phylogenetic contrasts and the evolution of mammalian life histories. *Evol. Ecol.* 7: 270-278.
- Bishop SC. 1943. *Handbook of salamanders*. Ithaca: Comstock Publishers.
- Bjorklund M. 1991. Evolution, phylogeny, sexual dimorphism and mating system in the grackle. *Evolution* 45: 608-621.
- Blackburn TM, Gaston KJ. 1994. The distribution of body sizes of the world's bird species. *Oikos* 70: 127-130.
- Block BA, Finnerty JR, Stewart FR, Kidd J. 1993. Evolution of endothermy in fish: mapping physiological traits on a molecular phylogeny. *Science* 260: 210-213.
- Bowen BW, Nelson WS, Avise JC. 1993. A molecular phylogeny for marine turtles: trait mapping, rate assessment and conservation relevance. *Proc. natl. Acad. Sci. U.S.A.* 90: 5574-5577.
- Brooks DR, McLennan D. 1991. *Phylogeny, ecology and behavior: a research program in comparative biology*. Chicago: Univ. Chicago Press.
- Brown JW. 1995. *Macroecology*. Chicago: Univ. Chicago Press.
- Bustard R. 1973. *Sea turtles: natural history and conservation*. New York: Taplinger Publishing Co.

- Caccone A, Milinkovitch MC, Sbordoni V, Powell JR. 1994.** Molecular biogeography: using the Corsica-Sardinia microplate disjunction to calibrate mitochondrial rDNA evolutionary rates in mountain newts (*Euproctus*). *J. evol. Biol.* **7**: 227-245.
- Carr A. 1952.** *Handbook of turtles*. Ithaca: Cornell Univ. Press.
- Cooper A, Penny D. 1997.** Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. *Science* **275**: 1109-1113.
- Corbin KW, Livezey BC, Humphrey PS. 1988.** Genetic differentiation among steamer-ducks (Anatidae: Tachyeres): an electrophoretic analysis. *Condor* **90**: 773-781.
- Díaz-Urriarte R, Garland Th Jr. 1996.** Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Syst. Biol.* **45**: 27-47.
- Dittman DL, Zink RM. 1991.** Mitochondrial DNA variation among phalaropes and allies. *Auk* **108**: 771-779.
- Duncan R, Highton R. 1979.** Genetic relationships of the eastern large *Plethodon* of the Ouachita mountains. *Copeia* **1979**: 95-110.
- Dunn ER, Heinze AA. 1933.** A new salamander from the Ouachita mountains. *Copeia* **1933**: 121-122.
- Dunning JB Jr. 1993.** *CRC handbook of avian body masses*. Boca Raton: CRC Press.
- Edwards AWF. 1992.** *Likelihood*. Baltimore: Johns Hopkins Univ. Press.
- Edwards AWF, Cavalli-Sforza LL. 1964.** Reconstruction of evolutionary trees. In: Heywood VH, McNeill J, eds. Phenetic and phylogenetic classification. *Systematics Ass. Publ.* **6**: 67-76.
- Eisenberg JF. 1989.** *Mammals of the neotropics, 1. The northern neotropics*. Chicago: Univ. Chicago Press.
- Eldredge N, Gould SJ. 1972.** Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, Thomas JM, eds. *Models in paleobiology*. San Francisco: Freeman & Cooper, 305-322.
- Ernst CH, Barbour RW, Lovich JE. 1994.** *Turtles of the U.S. and Canada*. Washington: Smithsonian Inst. Press.
- Felsenstein J. 1981.** Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution* **35**: 1229-1242.
- Felsenstein J. 1983.** Parsimony in systematics: biological and statistical issues. *Ann. Rev. Ecol. Syst.* **14**: 313-333.
- Felsenstein J. 1985.** Phylogenies and the comparative method. *Amer. Nat.* **125**: 1-15.
- Felsenstein J. 1988.** Phylogenies and quantitative characters. *Ann. Rev. Ecol. Syst.* **19**: 445-471.
- Felsenstein J. 1993.** *PHYLIP (Phylogeny inference package)*, 3.0. Distributed by the author, Seattle.
- Felsenstein J. 1995.** *PHYLIP (Phylogeny inference package)*, 3.5. Distributed by the author, Seattle.
- Ferris SD, Portnoy SL, Whitt GS. 1979.** The roles of speciation and divergence time in the loss of duplicate gene function. *Theor. Pop. Biol.* **15**: 114-139.
- Gardner AL. 1982.** Virginia opossum (*Didelphis virginiana*). In: Chapman JA, Feldhamer GA, eds. *Wild mammals of North America*. Baltimore: Johns Hopkins Press, 3-36.
- Garland Th Jr. 1992.** Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Amer. Nat.* **140**: 509-519.
- Garland Th Jr, Dickerman AW, Janis CM, Jones JA. 1993.** Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* **42**: 265-292.
- George MJ, Ryder OA. 1986.** Mitochondrial DNA evolution in the genus *Equus*. *Mol. Biol. Evol.* **3**: 535-546.
- Gerwin JA, Zink RM. 1989.** Phylogenetic patterns in the genus *Heliodoxa* (Aves: Trochilidae): an allozymic perspective. *Wilson Bull.* **101**: 525-543.
- Gittleman JL, Kot M. 1990.** Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* **39**: 227-241.
- Goldman D, Giri PR, O'Brien SJ. 1989.** Molecular genetic-distance estimates among the Ursidae as indicated by one- and two-dimensional protein electrophoresis. *Evolution* **43**: 282-295.
- Goldman N. 1993.** Statistical tests of models of DNA evolution. *J. mol. Evol.* **36**: 182-198.
- Goodman M, Bailey WJ, Hayasaka K, Stanhope MJ, Slightom J. 1994.** Molecular evidence on primate phylogeny from DNA sequences. *Am. J. phys. Anthropol.* **94**: 3-24.
- Gould SJ, Eldredge N. 1993.** Punctuated equilibrium comes of age. *Nature* **366**: 223-227.
- Grafen A. 1989.** The phylogenetic regression. *Phil. Trans. R. Soc. (B)* **326**: 119-157.
- Grafen A, Ridley M. 1997.** A new model for discrete character evolution. *J. theor. Biol.* **184**: 7-14.
- Grant PR, Abbott I, Schluter D, Curry RL, Abbott LK. 1985.** Variation in the size and shape of Darwin's finches. *Biol. J. Linn. Soc.* **25**: 1-39.
- Grzimek B. (ed.) 1990.** *Grzimek's encyclopedia of mammals*, vols. 1-6. New York: McGraw-Hill.
- Hansen TF. 1997.** Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**: 1341-1351.
- Hansen TF, Martins EP. 1996.** Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* **50**: 1404-1417.
- Harvey PH, Mace GM. 1982.** Comparisons between taxa and adaptive trends: problems of methodology. In: Group KCS, ed. *Current problems in sociobiology*. Cambridge, UK: Cambridge Univ. Press, 343-361.
- Harvey PH, Pagel MD. 1991.** *The comparative method in evolutionary biology*. Oxford: Oxford Univ. Press.
- Harvey PH, Purvis A. 1991.** Comparative methods for explaining adaptations. *Nature* **351**: 619-624.
- Hedges SB, Burnell KL. 1990.** The Jamaican radiation of *Anolis* (Sauria: Iguanidae): an analysis of relationships and biogeography using sequential electrophoresis. *Caribb. J. Sci.* **26**: 31-44.
- Heijerman T. 1992.** Adequacy of numerical taxonomic methods: A comparative study based on simulation results. *Z. zool. Syst. Evolut.-forsch.* **30**: 1-20.

- Heijerman T.** 1993. Adequacy of numerical taxonomic methods: further experiments using simulated data. *Z. zool. Syst. Evolut.-forsch.* **31**: 81-97.
- Hey J.** 1992. Using phylogenetic trees to study speciation and extinction. *Evolution* **46**: 627-640.
- Highton R, Larson A.** 1979. The genetic relationships of the salamanders of the genus *Plethodon*. *Syst. Zool.* **28**: 579-599.
- Hirth HF.** 1982. Weight and length relationships of some adult marine turtles. *Bull. mar. Sci.* **32**: 336-341.
- Huey RB, Bennett AF.** 1987. Phylogenetic studies of coadaptation: Preferred temperatures versus optimal performance temperatures of lizards. *Evolution* **41**: 1098-1115.
- Joseph J, Klawe W, Murphy P.** 1980. *Tuna and billfish: fish without a country*. La Jolla, Calif.: Inter-American Tropical Tuna Commission.
- Kirsch JAW, Bleiweiss RE, Dickerman AW, Reig OA.** 1993. DNA/DNA hybridization studies of carnivorous marsupials. III. Relationships among species of *Didelphis* (Didelphidae). *J. mamm. Evol.* **1**: 75-97.
- Krajewski C, Fetzner JW Jr.** 1994. Phylogeny of cranes (Gruiformes: Gruidae) based on cytochrome-B DNA sequences. *Auk* **111**: 351-365.
- Krajewski C, King DG.** 1996. Molecular divergence and phylogeny: rates and patterns of cytochrome b evolution in cranes. *Mol. Biol. Evol.* **13**: 21-30.
- Lande R.** 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**: 314-334.
- Losos JB.** 1990. Ecomorphology, performance capability, and scaling in the West Indian *Anolis* lizards: an evolutionary analysis. *Ecological Monographs* **60**: 369-388.
- Losos JB, Adler FR.** 1995. Stumped by trees? A generalized null model for patterns of organismal diversity. *Amer. Nat.* **145**: 329-342.
- Lynch M.** 1990. The rate of morphological evolution in mammals from the standpoint of neutral expectation. *Amer. Nat.* **136**: 727-741.
- Maddison WP.** 1991. Squared change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. *Syst. Zool.* **40**: 304-314.
- Maddison WP.** 1995. Calculating the probability distributions of ancestral state reconstructed by parsimony on phylogenetic trees. *Syst. Biol.* **44**: 474-481.
- Martins EP.** 1994. Estimating the rate of phenotypic evolution from comparative data. *Amer. Nat.* **144**: 193-209.
- Martins EP.** 1996. Conducting phylogenetic comparative studies when the phylogeny is not known. *Evolution* **50**: 12-22.
- Martins EP, Garland Th Jr.** 1991. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution* **45**: 534-557.
- Martins EP, Hansen TF.** 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information in the analysis of interspecific data. *Amer. Nat.* **149**: 646-667.
- Mayr E.** 1963. *Animal species and evolution*. Cambridge, Mass.: Harvard Univ. Press.
- Moers AØ, 1995.** Tree balance and tree completeness. *Evolution* **49**: 379-384.
- Moers AØ, Cotgreave P.** 1994. Sibley and Ahlquist's tapestry dusted off. *Trends Ecol. Evol.* **9**: 458-459.
- Moers AØ, Nee S, Harvey PH.** 1994. Biological and algorithmic correlates of phenetic tree pattern. In: Eggleton P, Vane-Wright D, eds. *Phylogenetics and ecology*. London: Linn. Soc., 233-251.
- Moers AØ, Page RDM, Purvis A, Harvey PH.** 1995. Phylogenetic noise leads to unbalanced cladistic tree reconstructions. *Syst. Biol.* **44**: 332-342.
- Morrow JE.** 1964. Marlins, sailfish and spearfish of the Indian Ocean. In: *Symposium of scromboid fishes, Part 1*. Mandapam, India: Marine Biol. Ass. of India, 429-440.
- Nee S, Moers AØ, Harvey PH.** 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proc. natl. Acad. Sci. U.S.A.* **89**: 8322-8326.
- Nei M.** 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.
- Nei M, Li WH.** 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. natl. Acad. Sci. U.S.A.* **76**: 5269-5273.
- Nowak RM.** (ed.) 1991. *Walker's mammals of the world*. Baltimore: Johns Hopkins Univ. Press.
- Pagel M.** 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond.* (B) **255**: 37-45.
- Pagel M.** 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scr.* **26**: 331-348.
- Peters RH.** 1983. *The ecological implications of body size*. New York: Cambridge Univ. Press.
- Pritchard PCH.** 1967. *Living turtles of the world*. Jersey City: T.F.H. Publications.
- Pritchard PCH.** 1971. *The leatherback or leathery turtle (Dermochelys coriacea)*. Morges, Switzerland: International Union for Conservation of Nature.
- Purvis A, Rambaut A.** 1995. Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *CABIOS* **11**: 247-251.
- Rice WR.** 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223-225.
- Ricklefs RE, Starck JM.** 1996. Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* **77**: 167-172.
- Ridgway SH, Harrison R.** (eds.) 1985. *Handbook of marine mammals, 3. Sirenians and baleen whales*. London: Academic Press.
- Risch N.** 1992. Genetic linkage: interpreting LOD scores. *Science* **255**: 803-804.
- Rogers JS.** 1972. Measures in genetic similarity and genetic distance. *Studies in Genetics VII* **7213**: 145-153.
- Rohlf FJ, Chang WS, Sokal RR, Kim J.** 1990. Accuracy of estimated phylogenies: effects of tree topology and evolutionary model. *Evolution* **44**: 1671-1684.
- Rosenzweig ML.** 1995. *Species diversity in space and time*. New York: Cambridge Univ. Press.
- Schluter D, Nagel L.** 1995. Parallel speciation by natural selection. *Amer. Nat.* **146**: 292-301.
- Schluter D, Price T, Moers AØ, Ludwig D.** 1997. Likelihood

- of ancestor states in adaptive radiation. *Evolution* **51**: 1699-1711.
- Sibley CG, Ahlquist JE. 1990. *Phylogeny and classification of birds: a study in molecular evolution*. New Haven: Yale Univ. Press.
- Slattery JP, Johnson WE, Goldman D, O'Brien SJ. 1994. Phylogenetic reconstruction of South American felids defined by protein electrophoresis. *J. mol. Evol.* **39**: 296-305.
- Tilley SG, Bernardo J. 1993. Life history evolution in plethodontid salamanders. *Herpetologica* **49**: 154-163.
- Titus TA, Larson A. 1996. Molecular phylogenetics of desmognathine salamanders (Caudata: Plethodontidae): A reevaluation of evolution in ecology, life history, and morphology. *Syst. Biol.* **45**: 451-472.
- Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* **26**: 375-400.
- Wilson CA, Dean JM, Prince ED, Lee DW. 1991. An examination of sexual dimorphism in Atlantic and Pacific blue marlin using body weight, sagittae weight, and age estimates. *J. exp. mar. Biol. Ecol.* **151**: 209-226.
- Witzell WN. 1989. Longbill Spearfish *Tetrapturus pfluegeri* incidentally caught by recreational billfishermen in the Western North Atlantic Ocean 1974 to 1986. *Fish. Bull.* **87**: 982-984.
- Yang SY, Patton JL. 1981. Genic variability and differentiation in the Galapagos finches. *Auk* **98**: 230-242.
- Zink RM. 1988. Evolution of Brown Towhees: allozymes, morphometrics and species limits. *Condor* **90**: 72-82.
- Zink RM, Avise JC. 1990. Patterns of mitochondrial DNA and allozyme evolution in the avian genus *Ammodramus*. *Syst. Zool.* **39**: 148-161.
- Zink RM, Dittman DM. 1993. Population structure and gene flow in the Chipping Sparrow and a hypothesis for evolution in the genus *Spizella*. *Wilson Bull.* **105**: 399-413.
- Zink RM, Dittman DL, Rootes WL. 1991a. Mitochondrial DNA variation and the phylogeny of *Zonotrichia*. *Auk* **108**: 578-584.
- Zink RM, Rootes WL, Dittman DL. 1991b. Mitochondrial DNA variation, population structure, and evolution of the Common Grackle (*Quiscalus quiscula*). *Condor* **93**: 318-329.

Received: 15 January 1998