UNIFYING FOSSILS AND PHYLOGENIES FOR COMPARATIVE ANALYSES OF
DIVERSIFICATION AND TRAIT EVOLUTION

Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary

Graham J. Slater[1]*

[1]*Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, MRC 121, P.O. Box 37012, Washington, DC, 20013-7012, USA*

Summary

1. Phylogenetic comparative methods provide a powerful way of addressing classic questions about tempo and mode of phenotypic evolution in the fossil record, such as whether mammals increased in body size diversity after the Cretaceous-Palaeogene (K-Pg) extinction.

2. Most often, these kinds of questions are addressed in the context of variation in evolutionary rates. Shifts in the mode of phenotypic evolution provide an alternative and, in some cases, more realistic explanation for patterns of trait diversity in the fossil record, but these kinds of processes are rarely tested for.

3. In this study, I use a time-calibrated phylogeny of living and fossil Mammaliaformes as a framework to test novel models of body size evolution derived from palaeontological theory. Specifically, I ask whether the K-Pg extinction resulted in a change in rates of body size evolution or release from a constrained adaptive zone.

4. I found that a model comprising an Ornstein–Uhlenbeck process until the K-Pg event and a Brownian motion process from the Cenozoic onwards was the best supported model for these data. Surprisingly, results indicate a lower absolute rate of body size evolution during the Cenozoic than during the Mesozoic. This is explained by release from a stationary OU process that constrained realized disparity. Despite a lower absolute rate, body size disparity has in fact been increasing since the K-Pg event.

5. The use of time-calibrated phylogenies of living and extinct taxa and realistic, process-based models provides unparalleled power in testing evolutionary hypotheses. However, researchers should take care to ensure that the models they use are appropriate to the question being tested and that the parameters estimated are interpreted in the context of the best fitting model.

Key-words: Brownian motion, body size, comparative methods, evolutionary rates, fossils, macro-evolution, mammals, Ornstein–Uhlenbeck

Introduction

Understanding the pace at which evolution proceeds is one of the most fundamental questions in palaeontology and evolutionary biology (Simpson 1944, 1953). Traditionally, the study of rates of phenotypic evolution has been almost exclusively the domain of palaeobiologists. This tradition is largely practical; only palaeontologists possess a record of morphological variation through time and, until recently, were therefore the only ones able to address questions related to the tempo and mode of phenotypic evolution (Romer 1949). In recent years, however, evolutionary biologists have also become interested in tempo and mode in phenotypic evolution. Although the kinds of data sets used by evolutionary biologists lack temporal information provided by the fossil record, the use of phylogenetic comparative methods goes some way towards compensating for this. Assuming a Brownian motion (BM) model of phenotypic evolution, the shared evolutionary history implied by the branching structure of a time-calibrated molecular phylogeny allows for the estimation of evolutionary rates under criteria such as maximum-likelihood (Pagel 1997, 1999; Mooers, Vamosi & Schluter 1999) or Bayesian inference (Eastman *et al.* 2011; Revell *et al.* 2012; Slater, Harmon & Alfaro 2012a). One particularly attractive aspect of the Brownian motion model is that, by scaling the expected variances and covariances among taxa, one can easily produce alternative models that allow rates to change through time over the history of a clade (Blomberg, Garland & Ives 2003; Harmon *et al.* 2010) or even along specific branches of a phylogeny (O'Meara

*Correspondence author. E-mail: SlaterG@si.edu

et al. 2006; Thomas, Freckleton & Székely 2006; Eastman *et al.* 2011; Venditti, Meade & Pagel 2011; Revell *et al.* 2012; Slater *et al.* 2012b; Thomas & Freckleton 2012). Although power to detect these kinds of processes is limited when restricted to analysis of extant taxa (Slater, Harmon & Alfaro 2012a), the wide availability of these approaches has led to a flood of studies testing whether morphological diversity in particular clades can be explained as the result of shifts in the background rate of phenotypic evolution.

The comparative biologists' emphasis on rate variation may not be altogether well placed. Brownian motion is a diversifying process where morphological disparity increases as a function of time (Fig. 1a). Under BM, rate therefore predicts the expected disparity at any given time point in the history of a clade (O'Meara *et al.* 2006). Although most models of quantitative trait evolution are related to and can be derived from the Brownian motion model, their evolutionary 'rates' are not always directly comparable, at least in terms of their relationship to the accumulation of morphological disparity. This is particularly true for models with additional important parameters, such as the Ornstein–Uhlenbeck (OU) process (Hansen 1997). Under the OU process, a trait evolves towards or about a stationary peak or optimum at some rate. However, the addition of an attraction parameter causes traits to return to their stationary peak should they stray too far from it. Expected disparity through time under the OU process therefore reflects a balance between rate and attraction to the stationary peak (Fig. 1b), and comparisons of OU rates to estimates generated using Brownian motion may be misleading. Hunt (2012) used simulated data sets to show that evolutionary rates in palaeontological lineages could only be accurately estimated under the generating model. He suggested that rates are best thought of as model parameters that describe how trait variance changes through time in evolving lineages, rather than as purely descriptive phenomena about the pace of evolutionary change. This in fact is not a trivial distinction; Hunt's (2012) conclusion implies that simply allowing Brownian rates of phenotypic

evolution to vary across branches of a phylogeny may be insufficient to describe trait variation that arises under more complex processes. Instead, the most accurate rates and, by extension, the best model fits are likely to be obtained under models that closely match the true evolutionary process.

The evolution of mammalian body size provides a illustrative example of the focus on rate to explain patterns of phenotypic evolution. Palaeontologists have long recognized that the Mesozoic mammalian record is dominated by small species [<100g, although for a notable exception, see Hu *et al.* (2005)], but that body size diversity increases greatly after the Cretaceous-Palaeogene (K-Pg) boundary (Lillegraven, Kielan-Jaworowska & Clemens 1979; Alroy 1999; Archibald & Deutschman 2001). These observations have been used to argue that Mesozoic mammals were restricted to small body sizes through competition with or predation by the dominant dinosaur fauna. Freed from this constraint by the mass extinction at the K-Pg event, theory suggests that mammals subsequently underwent ecological release and diversified rapidly. Quantitative evidence for such a pattern is mixed. Palaeontological analyses suggest that rates of body size evolution did indeed increase after the K-Pg boundary (Raia *et al.* 2013), or at least that body size diversity increased substantially in the Palaeogene relative to the Mesozoic or post-Palaeogene Cenozoic (Alroy 1999). Cooper & Purvis (2010) used a phylogenetic comparative data set derived from extant taxa and found support for a general slowdown in rates of mammalian body size evolution, consistent with palaeontological evidence. However, Venditti, Meade & Pagel (2011) fit a model to the same data set that assumed Brownian motion but allowed rates to vary freely along branches of the phylogeny. Their results suggested a different pattern, with great rate heterogeneity among extant mammalian lineages but no evidence for a temporal shift from low to high rates around the K-Pg boundary (Venditti, Meade & Pagel 2011). Despite apparent conflict and differences in scope, one common theme to these studies is that they all assume that the process generating body size diversity,
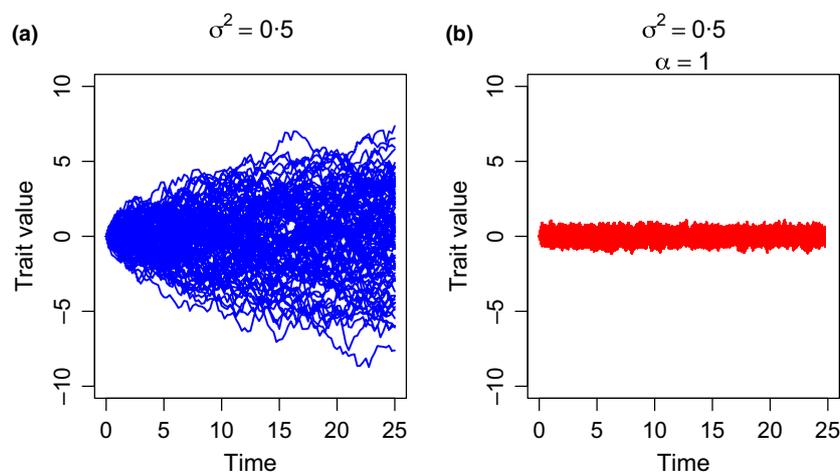
Fig. 1. One hundred realizations of (a) Brownian motion and (b) an Ornstein–Uhlenbeck process. Both processes were generated with a 'rate parameter' (σ^2) of 0·5. Under Brownian motion, the variance of the process clearly increases with time; Brownian motion is an unconstrained, diversifying process. Under the Ornstein–Uhlenbeck process, addition of the rubber band parameter α causes the random walk to return to its central tendency if it strays too far, resulting in a distribution for which time does not predict variance.

either along branches of a phylogeny or within time bins, did not vary. Alternatively put, these studies assume that the tempo of body size varies, but mode does not. From a mechanistic perspective, it seems more intuitive to predict that differences in the mode of evolution – the processes generating phenotypic variation – are responsible for the differences in body size diversity between Mesozoic and Cenozoic mammals. Furthermore, if this is the case, then we may not be able to distinguish such variation using variable-rate Brownian models (Hunt 2012).

In this paper, I derive three biologically and ecologically plausible alternative models for mammalian body size evolution over the K-Pg boundary and fit them, along with a set of standard models, to a novel phylogenetic comparative data set comprising Mesozoic and Cenozoic mammalian lineages. One of these models provides a substantially better fit to the comparative data set than the any other model investigated and suggests that the mode of body size evolution did indeed change after the K-Pg extinction. Importantly, results indicate that interpretation of the rate parameter for evolutionary change is strongly dependent on the fitted model (Hunt 2012), suggesting that multirate Brownian models may fail to detect biologically and evolutionarily important variation in some situations.

Materials and methods

PHYLOGENY AND BODY SIZE DATA

I used MrBayes version 3.2 (Ronquist *et al.* 2012b) to simultaneously infer phylogeny and branching times for extant and fossil Mammaliaformes under a relaxed clock model using a combined molecular data set of extant mammalian families (Meredith *et al.* 2011) and a morphological character matrix for living and fossil Mammaliaformes (Luo *et al.* 2011). This approach is attractive in that it allows fossils to be treated as terminal taxa during dating analyses, with their topological positions, divergence times and branch lengths inferred directly from their associated morphological data (Pyron 2011; Ronquist *et al.* 2012a). Full details on the data and phylogenetic methods used are provided in the supplementary materials accompanying this article. Briefly, the molecular matrix comprises 36 049 base pairs from 26 loci for 169 terminal taxa representing extant mammalian families, plus five non-synapsid vertebrate outgroup taxa. The morphological matrix contains 446 discrete morphological characters scored for 103 mammalian and synapsid taxa, 26 of which are extant. Fossil taxa include stem members of Eutheria, Metatheria and Monotremata, as well as entirely extinct mammalian and mammaliaform clades such as multituberculates, triconodonts and docodonts. After removing a few problematic fossil taxa, I retained a matrix of 239 taxa, Fossil taxa were coded as '?' for molecular characters. The same coding was used for all non-overlapping extant taxa in the morphological matrix. The final alignment and MrBayes block are deposited on Dryad.

I computed the mean body mass in natural log (Ln) grams for each extant mammalian family from the PanTHERIA data base (Jones *et al.* 2009), with data for the monotypic family Diatomyidae added from Jenkins *et al.* (2005). To account for body mass variation within families, I also computed the standard error of the mean for each family and incorporated this into model fitting by adding it to the diagonal elements of the variance–covariance matrix (O'Meara *et al.* 2006). For

fossil taxa, I estimated body mass from molar, cranial, mandibular or limb lengths using published regression equations for metatherian or eutherian mammals. I took some estimated body masses from the literature, but only in cases where these had been derived using similar approaches. Dental and bony element lengths, along with estimated masses, are provided in Table S1.

MODELLING BODY SIZE EVOLUTION

I used maximum likelihood to fit five standard, process-based models of trait evolution to the mammalian body size data set using the `fit-Continuous()` function in the `geiger` package (Harmon *et al.* 2008) for R (R Development Core Team 2012). These models are Brownian motion, Brownian motion with a directional trend, a single optimum Ornstein–Uhlenbeck process (also known as a single stationary peak model), accelerating–decelerating evolution (ACDC) and white noise. Their mathematical properties have been described in detail elsewhere (Hansen 1997; Blomberg, Garland & Ives 2003; O'Meara *et al.* 2006; Thomas, Freckleton & Székely 2006; Harmon *et al.* 2010; Hunt 2012; Slater, Harmon & Alfaro 2012a) and will not be expanded upon here. It should be noted, however, that although three of these models allow the evolutionary rate or expected trait value to vary, all assume that the evolutionary mode itself does not change over the history of the clade.

I implemented three alternative models to explicitly test the hypothesis that tempo and/or mode of mammalian body size evolution changed after the K-Pg event. First, I fit a model in which rates of body size evolution were free to vary before and after 66 million years ago (mya). This model was first suggested by O'Meara *et al.* (2006) and described in the supplementary information of Harmon *et al.* (2010) as the EBbreak model. It is here referred to as the K-Pg shift model because the time shift is restricted to the K-Pg boundary. Suppose that a trait evolves under a Brownian motion model with initial rate σ_0^2 but at sometime before present, t_{shift}, shifts to rate σ_1^2. The elements of the model-specific variance covariance matrix, **V**, are then given by

$$V_{ij} = \sigma_0^2 \min(C_{i,j}, t_{shift}) + \sigma_1^2 \max(0, C_{i,j} - t_{shift}), \qquad \text{eqn 1}$$

where C_{ij} is the shared evolutionary history of the ith and jth taxa (O'Meara *et al.* 2006; Harmon *et al.* 2010). The first half of the expression on the right of eqn (1) gives the expected variance accrued due to shared evolutionary history of this pair, if any, prior to t_{shift}, while the second half gives their expected variance, if any, accrued after t_{shift}. This model, as implemented here, has three parameters: the root state of the trait θ, the initial rate of evolution σ_0^2 and the rate after the shift point σ_1^2. The shift point itself, t_{shift}, is not treated as a free parameter as I fixed its value (Burnham & Anderson 2002).

The K-Pg shift model allows the tempo of evolution to vary before and after the K-Pg extinction, but not the evolutionary mode. The second and third alternative models are variations on the first that allow for this alternative. Palaeontological theory (and conventional wisdom) holds that Mesozoic mammals were constrained to small body sizes prior to the K-Pg event through competition with or exclusion by non-avian dinosaurs. If this was the case, then simply assuming a Brownian motion process and allowing rates of phenotypic evolution to vary may not fully capture the dynamics of body size evolution. I therefore first altered the K-Pg shift model to produce an 'ecological release' model. This model assumes that prior to the K-Pg boundary, mammalian body size evolution fluctuated

about a single stationary peak as an OU process. The OU process has a rate of evolution, σ^2, and a so-called rubber band parameter, α, that determines the strength with which the evolving trait is pulled back to its stationary peak as it moves away from it. In the ecological release model, the α parameter is removed after the K-Pg boundary, and trait evolution reverts to an unconstrained Brownian motion process. The elements of the ecological release variance covariance matrix, **V**, are given by

$$V_{ij} = \frac{\sigma^2}{2\alpha} \exp^{-2\alpha(T-t_{ou})}(1 - \exp^{-2\alpha t_{ou}})$$
$$+ \sigma^2 \max(0, C_{i,j} - t_{shift}), \qquad \text{eqn 2}$$

where T is the age of the deepest split in the phylogeny, and t_{ou} is the time spent in the OU process, which is given by

$$t_{ou} = \min(C_{i,j}, (T - t_{shift})). \qquad \text{eqn 3}$$

Here, the second half of the expression on the right of eqn (2) gives the variance accrued under the Brownian motion process occurring after the release point, while the first half of the expression provides the variance accrued under the Ornstein–Uhlenbeck process (Hansen 1997; Butler & King 2004; Harmon *et al.* 2010). This model again has three parameters: the root state of the trait, a single rate of evolution and the rubber band parameter.

The final alternative model is a 'release and radiate' model. Here, as in the ecological release model, body size evolution is assumed to conform to an OU process prior to 66 mya and revert to a BM process after this time. However, this model allows the Brownian rate σ^2_{bm} to vary, relative to the OU rate σ^2_{ou} after the release point such that

$$V_{ij} = \frac{\sigma^2_{ou}}{2\alpha} \exp^{-2\alpha(T-t_{ou})}(1 - \exp^{-2\alpha t_{ou}})$$
$$+ \sigma^2_{bm} \max(0, C_{i,j} - t_{shift}). \qquad \text{eqn 4}$$

The release and radiate model therefore has four parameters: 1 more than the simple release model is described in eqn (2). If mammals underwent an explosive radiation after the extinction of non-avian dinosaurs at the K-Pg boundary, we might expect that $\sigma^2_{bm} > \sigma^2_{ou}$.

I compared the fit of these eight candidate models to three permutations of the mammalian data set. First, I compared their fit to a data set pruned to only extant mammalian clades. Second, I fit the models to a data set comprising extant mammalian clades and all fossil taxa descended from their most recent common ancestor (i.e. living and fossil Mammalia). Finally, I fit the models to a data set comprising all living and fossil Mammaliaformes represented in my data set, excluding non-mammaliaform mammaliamorphs such as *Sinoconodon*. Model support was compared using Akaike weights computed from small-sample-corrected AIC scores (Burnham & Anderson 2002). Model fits were initially computed for the maximum clade credibility tree, but I also fit models to 100 trees drawn at random from the posterior distribution of time-calibrated trees to investigate the influence of topological and branch length variation on model fit. I have deposited R scripts containing functions and code to perform the analyses on Dryad.

SIMULATION TESTS

I also performed a series of simulation tests to determine power to detect shifts in the mode of phenotypic evolution using the three permutations of the comparative data set. For each mode shift model, I generated 1000 realizations of trait evolution on the complete mammaliaform phylogeny using randomly drawn parameter values. I then assessed the fit of the three mode shift models, as well as the standard set of models, to the extant taxa only, living and fossil Mammalia, and Mammaliaformes

data sets by computing Akaike weights for each model. To compare model selection performance, I computed the median weight for the true model at each simulated parameter value.

For the K-Pg shift model and the release and radiate model, I fixed the Mesozoic rate of evolution to 0·1 and allowed the post-Mesozoic rate to vary such that it was 0·01, 0·1, 1, 10 and 100 times the initial rate. For the release and radiate model, I fixed the α parameter for the Mesozoic OU process to 0·01. For the release model, I again fixed the rate of evolution to 0·1 but allowed the α parameter to vary, such that $\alpha \in \{0·001, 0·005, 0·01, 0·05\}$. Values of α larger than 0·05 resulted in a singular variance–covariance matrix for the fossil data set and so were not explored.

Results

PHYLOGENY AND DIVERGENCE TIME ESTIMATION

The topology of the maximum clade credibility tree of living and fossil Mammaliaformes is shown in Fig. 2. The topology is broadly consistent with previous hypotheses of mammaliaform relationships and will not be discussed further. Simultaneous relaxed clock dating of fossil and living mammalian lineages leads to slightly older divergence times among the higher-level mammalian clades than have been estimated from molecular data only. For example, I recovered a mean divergence time of 205·5 mya for metatherians and eutherians, compared with 186·7 mya in Meredith *et al.* (2011) and 149·4 mya in dos Reis *et al.* (2012), and of 243 mya for the monotreme–therian divergence compared with 215·5 mya in Meredith *et al.* (2011) and 177·7 mya in dos Reis *et al.* (2012). This phenomenon has been previously described when treating fossils as terminals in dating analyses (Pyron 2011; Ronquist *et al.* 2012a). Dates for younger divergences are more consistent with molecular-only hypotheses; for example, my analyses place the origins of most placental orders at or after the K-Pg boundary (see Fig. 2). My mean divergence time estimates conflict strongly with the younger ages reported by O'Leary *et al.* (2013). This is unsurprising given that their dates are minimum divergence times based on ghost-lineage minimization (Yoder 2013) and that the inclusion of the older stem eutherian *Juramaia sinensis* in my analysis pulls some deeper nodes back in time. The majority-rule consensus tree, as well as the entire posterior distribution of trees, has been deposited on Dryad.

I obtained body mass estimates for 57 of the fossil mammaliaformes represented in my phylogeny. Of these, 45 belong to Mammalia. A plot of mammaliaform body mass through time is shown in Fig. 3 and qualitatively suggests that body mass diversity increased greatly after the K-Pg extinction. The complete data set is deposited on Dryad.

MODEL FITTING

For the data set comprising extant taxa only, the release and radiate model was the most favoured model. Differences in relative model support using the maximum clade credibility tree and samples from the posterior distribution were limited (Fig.
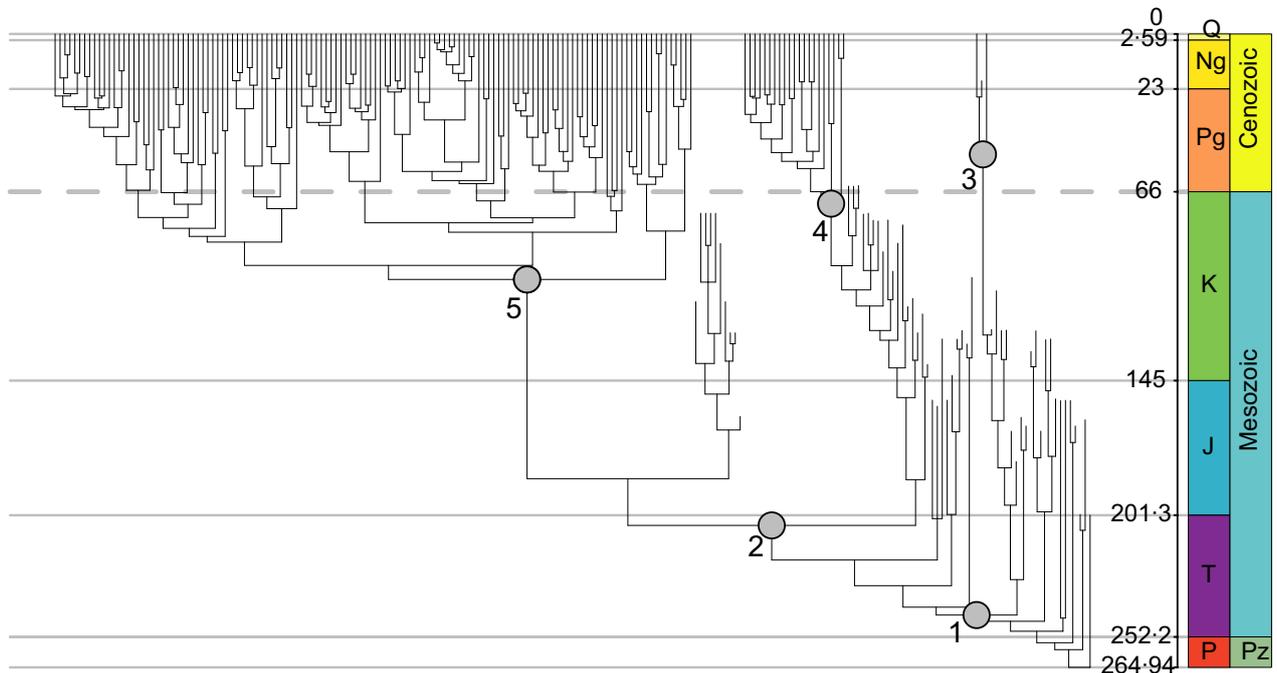
Fig. 2. Time-calibrated maximum clade credibility tree of living and fossil Mammaliaformes inferred for this study. Node labels correspond to the following: 1. Mammalia, 2. Theria, 3. crown Monotremata, 4. crown Metatheria, 5. crown Eutheria. The dashed grey horizontal line corresponds to the K-Pg boundary.

4), and so I restrict discussion here to results from the maximum clade credibility tree. Support for release and radiate model was far from unequivocal when applied to extant mammals only (Fig. 4a; Akaike weight = 0·77), but the AICc difference between this and the next most supported model, the K-Pg shift model (Akaike weight = 0·08), was greater than the traditional threshold of 4 (ΔAICc = 4·47) required for strong support (Burnham & Anderson 2002). Among the standard models, BM received the most support (Akaike weight = 0·063). Although I report an Akaike weight for the trend model, note that trends cannot be inferred from ultrametric trees. The AICc score for this model therefore represents the log likelihood of a Brownian motion penalized by an additional, uninferable parameter.

Adding fossil taxa to the comparative data set altered results, but in different ways for different data sets. Using either the data set comprising crown mammals and their fossil relatives (Fig. 4b) or all mammaliaformes (Fig. 4c), the release and radiate model received most support. However, support for release and radiate model increased most markedly in the case of living and fossil Mammalia (Akaike weight = 0·9), while for Mammaliaformes, support increased only slightly relative to extant mammals only (Akaike Weight = 0·79). For both data sets, the second best model received low weight, and AIC differences between best and second best models were >4 (Table 1). Support for the K-Pg shift was reduced in fossil data sets, compared with the extant taxa data set. None of the standard models received much support in the palaeontological data sets, although support for ACDC was elevated, compared with the extant taxa data.
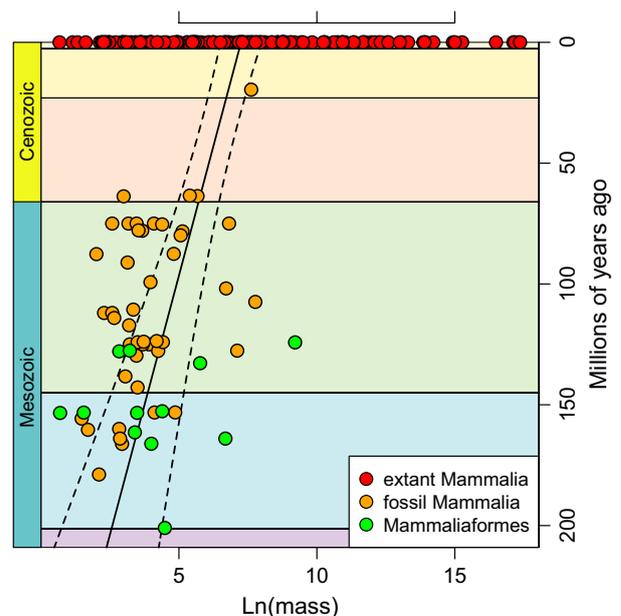
Fig. 3. Natural logarithm of body mass through time for Mammaliaformes used for analyses in this paper. Masses for extant mammals (at 0 millions of years before present) represent family means. The regression line and confidence intervals are derived from a phylogenetic generalized least squares analysis (β = 0·024) and are shown for illustrative purposes only.

Measurement error can reduce phylogenetic signal in comparative data sets, leading to an increase in the potential for erroneously recovering support for low-signal models, such as
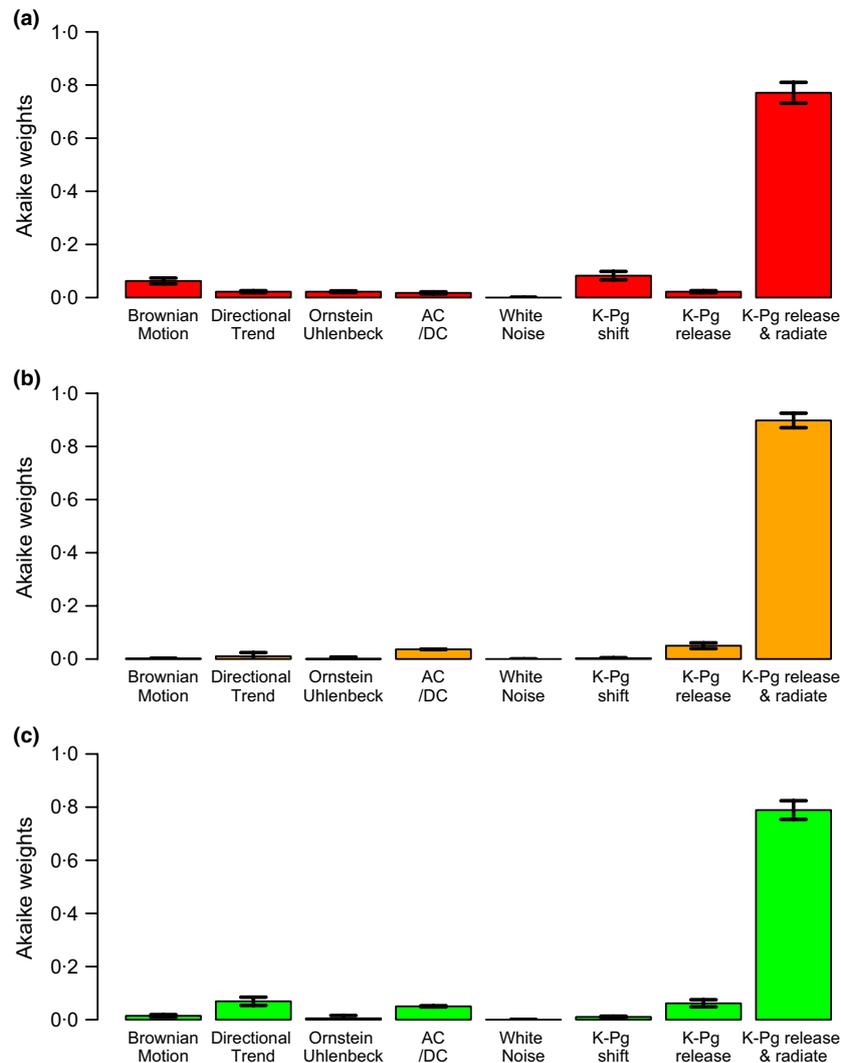
Fig. 4. Akaike weights for analyses performed on the maximum clade credibility tree from Fig. 2 (a) with extant Mammalia only; (b) living and fossil Mammalia; and (c) Mammaliaformes. Colours as in Fig. 3. Whiskers show standard errors for Akaike weights derived from model fits to 100 trees drawn randomly from the posterior distribution of trees.

OU or ACDC (Revell, Harmon & Collar 2008). To account for the possibility that error in estimation of body masses for fossil taxa biased my results towards supporting an OU process during the Mesozoic, I repeated the model-fitting exercise on the maximum clade credibility tree with all fossil taxa assigned the maximum standard error of any extant clade (Dugongidae, SEM = 1·30). This increased support for release and radiate model using the Mammaliaformes data set (Akaike Weight = 0·82) but slightly decreased support for the living and fossil Mammalia data (Akaike Weight = 0·82). Despite these changes in relative model support, the release and radiate model remained the best supported model among the candidate pool.

Inspection of estimated model parameters (Table 1) reveals what at first appears to be a slightly puzzling result. Although the release and radiate model is best supported for each data set, as might be predicted from palaeontological theory, estimated model parameters indicate that the post-K-Pg Brownian rate of evolution was low, between one-tenth and one-fifth

that of the pre K-Pg OU rate. This result seems particularly unexpected given that the K-Pg shift model indicates an increase in rate in the Cenozoic, and among the standard set of models, an accelerating model of evolution is most favoured.

The explanation for this discrepancy lies in the way in which the accumulation of morphological variance under an OU process differs from that under Brownian motion (Fig. 1). The red curve in Fig. 5 shows the expected variance through time for a single particle undergoing an Ornstein–Uhlenbeck process with the maximum-likelihood estimates of the OU parameters for the extant and fossil mammal data (parameter estimates from the other two data sets resulted in visually identical curves). Although the OU process was only fitted to the portion of the tree from the root to 66 million years before present, I have extended the curve to the present day to indicate how variance would have accumulated through the Cenozoic should this process have continued. The curve shows a steady increase in variance from the origin of mammals in the Early Triassic until approximately the Middle Jurassic, at which

Table 1. Maximum-likelihood model fitting results and parameter values for the three permutations of the mammalian body size data set. For the trend model, P1 is mu, the directional change parameter. For OU, P1 is α, the rubber band or attraction parameter. For AC/DC, P1 is *r*, the exponential change parameter. For white noise, P1 is the variance parameter. For K-Pg shift, P1 is the post-Cenozoic rate. For ecological release and release and radiate model, P1 is α. For release and radiate model, P2 is the post-Cenozoic Brownian rate

Taxa	Model	LnL	AICc	ΔAICc	AICcWt	Root	Rate	P1	P2
Extant Mammalia	Brownian motion	−349·11	702·29	5·02	0·06	6·77	0·11	NA	NA
	Trend	−349·11	704·37	7·10	0·02	6·78	0·11	−3·22E-05	NA
	Ornstein Uhlenbeck	−349·11	704·37	7·10	0·02	6·77	0·11	8·54E-09	NA
	ACDC	−349·33	704·82	7·55	0·02	6·75	0·05	0·00	NA
	White noise	−421·23	846·54	149·27	0·00	7·40	14·33	NA	NA
	K-Pg shift	−347·79	701·74	4·47	0·08	6·77	0·27	0·10	NA
	Ecological release	−349·11	704·37	7·10	0·02	6·77	0·11	5·90E-08	NA
	Release and radiate	−344·50	697·27	0·00	0·77	6·71	9·46	0·02	0·01
Extant and fossil Mammalia	Brownian motion	−435·98	876·03	12·81	0·00	3·35	0·10	NA	NA
	Trend	−433·00	872·13	8·92	0·01	0·76	0·10	0·03	NA
	Ornstein Uhlenbeck	−435·98	878·09	14·88	0·00	3·35	0·10	2·06E-09	NA
	ACDC	−431·75	869·62	6·40	0·04	3·30	0·02	0·01	NA
	White noise	−540·38	1084·83	221·61	0·00	6·61	13·66	NA	NA
	K-Pg shift	−434·19	874·50	11·28	0·00	3·35	0·06	0·11	NA
	Ecological release	−431·43	868·99	5·77	0·05	3·30	0·12	3·20E-03	NA
	Release and radiate	−427·50	863·22	0·00	0·90	3·18	0·97	0·01	0·10
Mammaliaformes	Brownian motion	−462·95	929·96	7·99	0·01	4·28	0·10	NA	NA
	Trend	−460·36	926·84	4·87	0·07	2·27	0·10	0·02	NA
	Ornstein Uhlenbeck	−462·95	932·02	10·05	0·01	4·46	0·10	6·70E-04	NA
	ACDC	−460·68	927·48	5·51	0·05	4·26	0·04	4·54E-03	NA
	White noise	−572·07	1148·20	226·24	0·00	6·46	13·41	NA	NA
	K-Pg shift	−462·26	930·63	8·66	0·01	4·28	0·08	0·11	NA
	Ecological release	−460·48	927·07	5·10	0·06	4·26	0·12	1·95E-03	NA
	Release and radiate	−456·89	921·97	0·00	0·79	4·24	0·47	6·50E-03	0·10

point the accumulation begins to level off. By the K-Pg boundary, the expected variance has almost reached its equilibrium point, indicated by the horizontal dashed line and given by $\frac{\sigma^2}{2\alpha}$. The main point to take from this curve is that under an OU process, we would expect mammalian body size diversity to have begun to plateau around the K-Pg boundary. In Fig. 5, I have also projected, in blue, the expected variance accrued by a particle undergoing Brownian motion with the ML-estimated Brownian rate from the K-Pg boundary to the present. Although this rate is only 13% of the OU rate, variance is predicted to increase rapidly after the Palaeogene boundary, consistent with interpretation based on the fossil record.

SIMULATION TESTS

Simulations demonstrate three main points regarding the mode shift models that I used here. First, at least for this data set, it is difficult to strongly prefer the release model when it is the true model of evolution (Fig. 6b). Median weights for this model were highest at strong α values, but did not break 0·7 (Fig. 7). In fact, even the maximum weight for an individual data set under this model did not achieve traditional levels indicating strong support (i.e. ≥ 0·95, results not shown). Among the palaeontologically motivated models, the release and radiate model achieved the next highest level of support when ecological release was the true model, while ACDC provided the best fit among the standard models. However, despite relative low relative support,
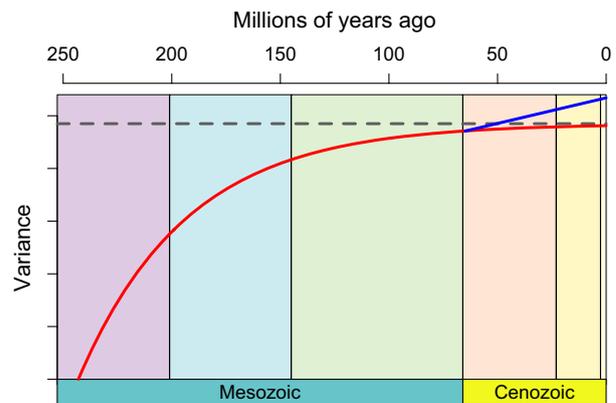

Fig. 5. Accumulated variance through time predicted from maximum-likelihood parameter estimates. The red line gives the variance predicted under an Ornstein–Uhlenbeck process. The blue line gives the predicted variance under a BM process operating from the K-Pg boundary until the present.

ecological release provided a better explanation than these other models at all α values (Fig. 7).

Second, our ability to favour the true model generally depends on the strength of the parameter values involved. For example, support for the K-Pg shift model increases as the difference between Mesozoic and Cenozoic rates increases (Fig. 6a). Notably, analyses based on extant taxa only performed slightly differently under the K-Pg shift and release and radiate models; here, ability to detect the true
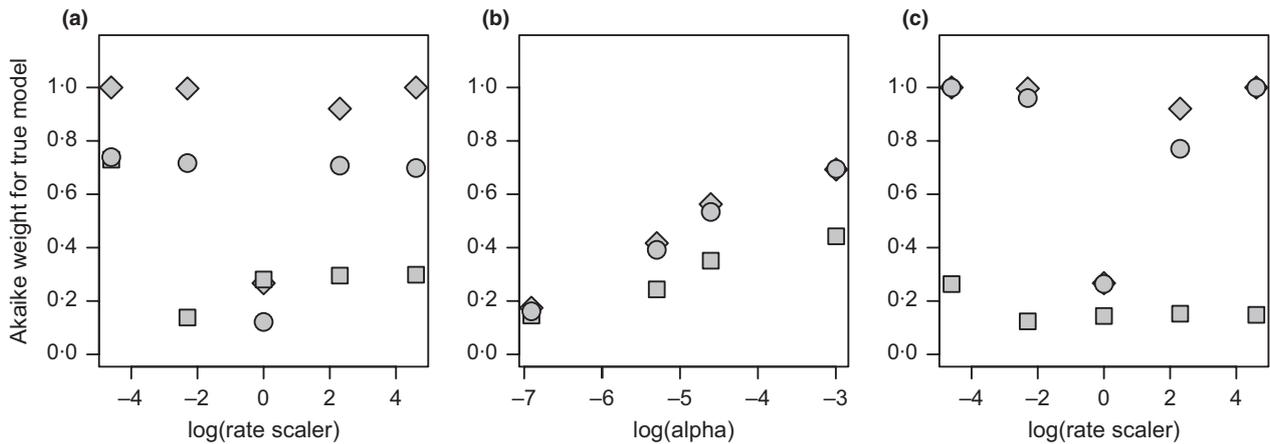
Fig. 6. Results of model fitting to simulated data sets. The plots show mean Akaike weights for the true model when that process is (a) the K-Pg shift model, (b) the release model and (c) the release and radiate model. Symbols indicate mean weights for extant taxa only (squares), living and fossil Mammalia (circles) and Mammaliaformes (diamonds) at each of the sampled parameter values.
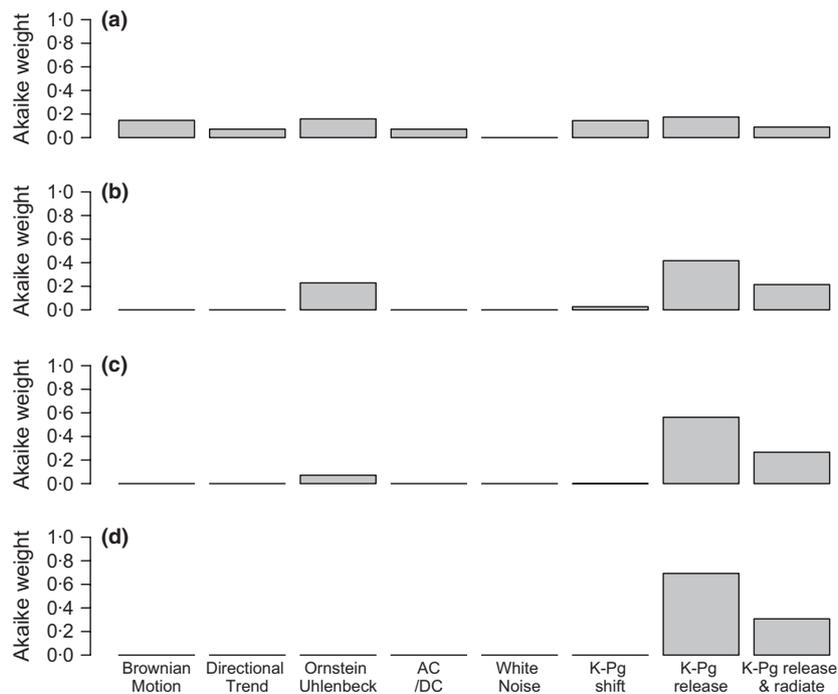
Fig. 7. Median Akaike weights from 1000 simulated data sets for all models when the ecological release model is the generating model. Panels show weights at $\alpha = 0.001$ (a), 0.005 (b), 0.01 (c) and 0.05 (d). Support for the release model increases with the strength of α.

model did not increase appreciably as the post-Mesozoic rate increased relative to the Mesozoic rate (Fig. 6a,c). Finally, support for the true model was always strongest using data sets containing fossil taxa. Use of the Mammaliaformes data set, which contains the greatest number of fossil taxa, always resulted in the highest average weights for the true model, while use of the data set comprising extant taxa only always resulted in the lowest average weights.

Discussion

The results presented in this study provide strong support for a shift in the mode of mammalian body size evolution at the K-Pg boundary. This finding is consistent with interpretations based directly on the fossil record and from ecological and palaeobiological theory (Simpson 1944, 1953; Lillegraven, Kielan-Jaworowska & Clemens 1979; Alroy 1999; Archibald & Deutschman 2001; Cooper & Purvis 2010; Raia *et al.* 2013). Mesozoic mammals were, for the most part, small, but underwent a significant radiation after the extinction of the dinosaurs at the end of the Cretaceous. Two distinctions exist between my study and previous work on this question; first, I incorporated data from Mesozoic mammals in a phylogenetic context, and second, I formulated an ecologically reasonable model for mammalian evolution that does not rely on variation in evolutionary rate.

The importance of fossils for phylogenetic comparative methods, particularly for ancestral state reconstruction, has

long been recognized (Oakley & Cunningham 2000; Polly 2001; Finarelli & Flynn 2006; Albert, Johnson & Knouft 2009; Losos 2011). The impact of fossil data on ancestral body mass estimates for crown mammals is readily seen in Table 1. Using data from only extant mammals, the model-averaged ancestral mass estimate was 957 g. Incorporation of fossil information reduced this estimate to a more palaeontologically reasonable value of 114 g. This result is to be expected; Adding more data from extant taxa does not increase the effective sample size associated with ancestral state estimates in phylogenetic comparative data. Rather, the length of the shortest branch subtended by the root is key in determining the effective sample size (Ané 2008), and including fossil taxa provides one way of accomplishing this. Accommodating variation in rates of phenotypic evolution provides an analogous solution and can result in more realistic ancestral state estimates in some cases (Elliot & Mooers 2013).

The importance of fossil taxa for fitting and comparing models of trait evolution has received less attention. Slater, Harmon & Alfaro (2012a) showed that adding fossil information, either as terminal taxa or as informative node value priors, increased power to detect some models of trait evolution and facilitated selection among models that cannot be distinguished using extant taxa only. Here, the addition of fossil taxa clearly improved model selection performance in simulated data sets (Fig. 6) and in my mammalian data. Using extant mammals, the release and radiate model received most weight, and although the AICc difference between this model and the next best model (the K-Pg shift model) was greater than the nominal difference of four required for 'strong' support, comparison of Akaike weights suggests that this model is not a runaway favourite among the candidate pool (Fig. 4a). Incorporating fossil Mammalia into the comparative data set substantially increased support for the release and radiate model, however (Fig. 4b). Interestingly, the addition of non-mammalian mammaliaformes to the data set decreased support for the release and radiate model relative to the living and fossil mammalia data set (Fig. 4c). This may suggest that body size evolution in these stem groups followed different dynamics to that of the crown mammals and their fossil relatives.

Only one of the alternative models that I implemented here was an explicitly rate-based model. The K-Pg shift model allowed rates of body size evolution to vary between the Mesozoic and Cenozoic, as previous analyses have supposed may have happened (Venditti, Meade & Pagel 2011; Raia *et al.* 2013). Indeed, data sets comprising Mesozoic taxa did result in faster estimated rates under this model for Cenozoic mammals than Mesozoic lineages (Table 1). Estimates for Cenozoic rates derived from only extant lineages were lower than those estimated for the Mesozoic, although this may result from a lack of power as Cenozoic rates remained fairly constant across data sets. However, for no data set was this model preferred (Fig. 4; Table 1). It is also worth noting that I did not recover evidence in favour of a trend towards increasing body size in mammals (Cope's rule) using any of the data sets. This result makes sense in the context of the best fitting models. Body size diversity increased substantially after the K-Pg extinction, and

mammals began this radiation from a small body mass. Because physiological and metabolic demands place a lower limit on body mass in homoeothermic (West, Woodruff & Brown 2002), the evolutionary increase in mean body mass after the K-Pg boundary is likely better explained as passive diffusion away from a limiting lower bound. Alroy (1999) has previously commented on this phenomenon, noting that the increase in mean mass of Cenozoic North American mammals tracks the increase in body mass disparity.

Consideration of variation in tempo without consideration of variation in mode is misleading for mammalian body size evolution. Only the model assuming an Ornstein–Uhlenbeck process during the Mesozoic and an unconstrained Brownian motion process during the Cenozoic was clearly favoured among the candidate models that I considered here. Furthermore, this model indicates lower absolute rates of evolution in the Cenozoic than in the Mesozoic, contrary to prevailing thought. A cold reading of the fossil record implies that mammalian body size diversity increased rapidly after the K-Pg extinction (Alroy 1999), and my results indeed support this, despite the lower absolute rates estimated for the Cenozoic. Based on parameters estimated here, the OU process within which mammalian body size evolved was approaching its equilibrium distribution (Hansen 1997) by the Jurassic-Cretaceous boundary around 145 mya. Recent discoveries have indicated great ecomorphological diversity within Mesozoic Mammaliaformes (Hu *et al.* 2005; Luo & Wible 2005; Ji *et al.* 2006; Meng *et al.* 2006; Luo 2007), and the curve shown in Fig. 6 indicates that this ecological diversification was accompanied by a general increase in body size diversity, centred around a stationary peak of between 25 and 66 g (Table 1). Although rates of body size evolution, indicated by the σ^2 parameter, are estimated to be higher during the Mesozoic than during the Cenozoic, evolution under an OU process with the associated rubber band parameter α prevents the continuous increase in morphological disparity through time that we intuitively expect under an unbounded, diversifying process such as Brownian motion (Fig. 1). In microevolutionary terms, the equilibrium point of an OU process represents a balance between the constraining force of selection and the diversifying force of drift (Hansen 1997; Hunt 2012). In macroevolutionary terms, we can think of a large value of σ^2 with a moderate α as reducing 'phylogenetic signal' within the trait data while constraining the total variance realized (Revell, Harmon & Collar 2008). Shifting from a OU process to a Brownian motion process at the K-Pg boundary, even if that process takes a lower absolute rate of evolution, is therefore sufficient to generate increased phenotypic variance (Fig. 6).

A few caveats to the results presented in this study should be acknowledged. First, my phylogeny is far from complete. Extant clades are represented by families only, meaning that much available information on body size within crown mammalian clades cannot be fully leveraged, despite my attempts to account for sampling error. Similarly, taxon sampling in the morphological data set of Luo *et al.* (2011) is focused on understanding the relationships among the main lineages of Mesozoic Mammaliaformes and is therefore agnostic to much

of the variation found within more diverse Mesozoic mammal radiations, such as multituberculates (Wilson *et al.* 2012). Second, I did not attempt to fit more complex models of body size evolution. Additional sampling of early Palaeogene mammals could generate insights into the mode of body size diversification immediately after the K-Pg extinction, in particular whether body size diversity evolved rapidly in the Palaeogene and subsequently slowed (e.g. Alroy 1999; Cooper & Purvis 2010; Raia *et al.* 2013) or evolved at a relatively constant rate through the Cenozoic. Body size information for Palaeogene mammals could be incorporated into future work as node priors by using the Bayesian approach of Slater, Harmon & Alfaro (2012a). Furthermore, although my results suggest that mode shift models might provide a better explanation for mammalian body sizes than rate shift models, I did not explicitly compare my results with those from models that identify shifts in the rate of Brownian evolution along unspecified branches of a phylogeny. This limitation stems from the fact that these approaches are implemented in a Bayesian framework (e.g. Eastman *et al.* 2011; Venditti, Meade & Pagel 2011; Revell *et al.* 2012) or else use restricted maximum likelihoods (Thomas & Freckleton 2012), rendering comparisons with my results inappropriate. Future work exploring the fit of both rate and mode shift models to more complete comparative data sets for living and fossil mammals will undoubtedly lead to further refinement of ideas about tempo and mode in mammalian body size evolution. Finally, speciation and extinction rates are often correlated with body size (e.g. Van Valkenburgh, Wang & Damuth 2004; Clauset & Erwin 2008), and it is possible that such relationships influence the results recovered here. FitzJohn (2010) described an approach that models speciation and extinction rates as a function of a continuously varying trait undergoing Brownian motion. However, current implementations (FitzJohn 2012) assume that all species survive to the present (Nee, May & Harvey 1994) and are therefore restricted to analyses of extant taxa only (but see Pyron & Burbrink 2012). Current work on estimating diversification rates from non-ultrametric trees (Stadler 2010; Didier, Royer-Carenzi & Laurin 2012) has the potential to lead to greater insights into the importance of trait-dependent diversification in macroevolution.

Hunt (2012) noted that rates are strongly dependent on the model under which they are estimated and that there is no universal rate metric. This distinction is key for understanding tempo and mode in mammalian body size evolution. Despite their restricted diversity, we have no reason to predict that Mesozoic mammals simply evolved slowly compared with Cenozoic lineages. Multirate Brownian motion models test this very hypothesis however, and it is unsurprising that little evidence for rate shifts at the K-Pg boundary has been found in previous phylogenetic comparative studies (Venditti, Meade & Pagel 2011). In this study, I formulated a simple model that tests predictions derived from the fossil record – that Mesozoic mammals were constrained to a range of small body sizes and were only able to radiate into their current diversity of body sizes after the K-Pg extinction. This model is freed from the assumption of a homogeneous process albeit with different

rates and does a superior job of explaining the distribution of body sizes in my data. My results reinforce Hunt's (2012) conclusions that tempo and mode are intimately linked and that rates of phenotypic evolution under one model may be misleading when interpreted in the light of a different model. Notably, similar conclusions have been made regarding inferences of speciation and extinction rates under a birth–death process when ecological limits place constraints on lineage diversification (Rabosky 2009a, b; Rabosky, Slater & Alfaro 2012). These results, taken together, suggest that comparative biologists should exercise care when both fitting evolutionary models and interpreting their associated parameters. Attempts to explain historical disparity patterns in terms of rate variation only are unlikely to be successful.

Acknowledgements

Data accessibility

Phylogenetic comparative data and R scripts: DRYAD entry 10.5061/dryad. 1n14c.

References

Albert, J.S., Johnson, D.M. & Knouft, J.H. (2009) Fossils provide better estimates of ancestral body size than do extant taxa in fishes. *Acta Zoologica*, **90**, 357–384.

Alroy, J. (1999) The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Systematic Biology*, **48**, 107–118.

Ané, C. (2008) Analysis of comparative data with hierarchical autocorrelation. *The Annals of Applied Statistics*, **2**, 1078–1102.

Archibald, J.D. & Deutschman, D.H. (2001) Quantitative analysis of the timing of the origin and diversification of extant placental orders. *Journal of Mammalian Evolution*, **8**, 107–124.

Blomberg, S.P., Garland, T. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, **57**, 717–745.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference*. Springer, New York.

Butler, M.A. & King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, **164**, 683–695.

Clauset, A. & Erwin, D.H. (2008) The evolution and distribution of species body size. *Science*, **321**, 399–401.

Cooper, N. & Purvis, A. (2010) Body size evolution in mammals: complexity in tempo and mode. *The American Naturalist*, **175**, 727–738.

Didier, G., Royer-Carenzi, M. & Laurin, M. (2012) The reconstructed evolutionary process with the fossil record. *Journal of Theoretical Biology*, **315**, 26–37.

dos Reis, M., Inoue, J., Hasegawa, M., Asher, R.J., Donoghue, P.C.J. & Yang, Z. (2012) Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **279**, 3491–3500.

Eastman, J.M., Alfaro, M.E., Joyce, P., Hipp, A.L. & Harmon, L.J. (2011) A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution*, **65**, 3578–3589.

Elliot, M.G. & Mooers, A.Ø. (2013) Inferring ancestral states without assuming neutrality or gradualism using a stable model of continuous character evolution. arXiv:13025104[qbioPE].

Finarelli, J.A. & Flynn, J.J. (2006) Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Systematic Biology*, **55**, 301–313.

FitzJohn, R.G. (2010) Quantitative traits and diversification. *Systematic Biology*, **59**, 619–633.

FitzJohn, R.G. (2012) Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, **3**, 1084–1092.

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

Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan Jennings, W. *et al.* (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, **64**, 2385–2396.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) Geiger: investigating evolutionary radiations. *Bioinformatics*, **24**, 129–131.

Hu, Y., Meng, J., Wang, Y. & Li, C. (2005) Large Mesozoic mammals fed on young dinosaurs. *Nature*, **433**, 149–152.

Hunt, G. (2012) Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology*, **38**, 351–373.

Jenkins, P.D., Kilpatrick, C.W., Robinson, M.F. & Timmins, R.J. (2005) Morphological and molecular investigations of a new family, genus and species of rodent (Mammalia: Rodentia: Hystricognatha) from Lao PDR. *Systematics and Biodiversity*, **2**, 419–454.

Ji, Q., Luo, Z.X., Yuan, C.X. & Tabrum, A.R. (2006) A swimming mammaliaform from the Middle Jurassic and ecomorphological diversification of early mammals. *Science*, **311**, 1123–1127.

Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L. *et al.* (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, **90**, 2648–2648.

Lillegraven, J.A., Kielan-Jaworowska, Z. & Clemens, W.A., eds. (1979) *Mesozoic Mammals: The First Two-thirds of Mammalian History*. University of California Press, Berkeley, CA.

Losos, J.B. (2011) Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *The American Naturalist*, **177**, 709–727.

Luo, Z.X. (2007) Transformation and diversification in early mammal evolution. *Nature*, **450**, 1011–1019.

Luo, Z.X. & Wible, J.R. (2005) A Late Jurassic digging mammal and early mammalian diversification. *Science*, **308**, 103–107.

Luo, Z.X., Yuan, C.X., Meng, Q.J. & Ji, Q. (2011) A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature*, **476**, 442–445.

Meng, J., Hu, Y., Wang, Y., Wang, X. & Li, C. (2006) A Mesozoic gliding mammal from northeastern China. *Nature*, **444**, 889–893.

Meredith, R.W., Janec̆ka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C. *et al.* (2011) Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science*, **334**, 521–524.

Mooers, A.Ø., Vamosi, S.M. & Schluter, D. (1999) Using phylogenies to test macroevolutionary hypotheses of trait evolution in cranes (Gruinae). *The American Naturalist*, **154**, 249–259.

Nee, S., May, R.M. & Harvey, P.H. (1994) The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, **344**, 305–311.

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

O'Leary, M.A., Bloch, J.I., Flynn, J.J., Gaudin, T.J., Giallombardo, A., Giannini, N.P. *et al.* (2013) The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science*, **339**, 662–667.

O'Meara, B.C., Ané, C., Sanderson, M.J. & Wainwright, P.C. (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution*, **60**, 922–933.

Pagel, M. (1997) Inferring evolutionary processes from phylogenies. *Zoologica Scripta*, **26**, 331–348.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

Polly, P.D. (2001) Paleontology and the comparative method: ancestral node reconstructions versus observed node values. *The American Naturalist*, **157**, 596–609.

Pyron, R.A. & Burbrink, F.T. (2012) Trait-dependent diversification and the impact of palaeontological data on evolutionary hypothesis testing in New World ratsnakes (tribe Lampropeltini). *Journal of Evolutionary Biology*, **25**, 497–508.

Pyron, R.A. (2011) Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Systematic Biology*, **60**, 466–481.

R Development Core Team. (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rabosky, D.L. (2009a) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters*, **12**, 735–743.

Rabosky, D.L. (2009b) Ecological limits on clade diversification in higher taxa. *The American Naturalist*, **173**, 662–674.

Rabosky, D.L., Slater, G.J. & Alfaro, M.E. (2012) Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biology*, **10**, e1001381.

Raia, P., Carotenuto, F., Passaro, F., Piras, P., Fulgione, D., Werdelin, L., Saarinen, J. & Fortelius, M. (2013) Rapid action in the Palaeogene, the relationship between phenotypic and taxonomic diversification in Coenozoic mammals. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **280**, 20122244.

Revell, L.J., Harmon, L.J. & Collar, D.C. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, **57**, 591–601.

Revell, L.J., Mahler, D.L., Peres-Neto, P.R. & Redelings, B.D. (2012) A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution*, **66**, 135–146.

Romer, A.S. (1949) Time series and trends in animal evolution. *Genetics, Paleontology and Evolution* (eds G.L. Jepsen, G.G. Simpson& E. Mayr), pp. 103–120. Princeton University Press, New Jersey.

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L. & Rasnitsyn, A.P. (2012a) A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology*, **61**, 973–999.

Ronquist, F., Teslenko, M., der Mark, P., Ayres, D.L., Darling, A., Höhna, S. *et al.* (2012b) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Simpson, G.G. (1944) *Tempo and Mode of Evolution*. Columbia University Press, New York.

Simpson, G.G. (1953) *Major Features of Evolution*. Columbia University Press, New York.

Slater, G.J., Harmon, L.J. & Alfaro, M.E. (2012a) Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution*, **66**, 3931–3944.

Slater, G.J., Harmon, L.J., Wegmann, D., Joyce, P., Revell, L.J. & Alfaro, M.E. (2012b) Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate Bayesian computation. *Evolution*, **66**, 752–762.

Stadler, T. (2010) Sampling-through-time in birth–death trees. *Journal of Theoretical Biology*, **267**, 396–404.

Thomas, G.H. & Freckleton, R.P. (2012) MOTMOT: models of trait macroevolution on trees. *Methods in Ecology and Evolution*, **3**, 145–151.

Thomas, G.H., Freckleton, R.P. & Székely, T. (2006) Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **273**, 1619–1624.

Van Valkenburgh, B., Wang, X. & Damuth, J. (2004) Cope's rule, hypercarnivory, and extinction in North American canids. *Science*, **306**, 101–104.

Venditti, C., Meade, A. & Pagel, M. (2011) Multiple routes to mammalian diversity. *Nature*, **479**, 393–396.

West, G.B., Woodruff, W.H. & Brown, J.H. (2002) Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 2473–2478.

Wilson, G.P., Evans, A.R., Corfe, I.J., Smits, P.D., Fortelius, M. & Jernvall, J. (2012) Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. *Nature*, **483**, 457–460.

Yoder, A.D. (2013) Fossils versus clocks. *Science*, **339**, 656–658.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1. Age ranges (millions of years before present), estimated masses (grams) and sources for masses or raw measurements for fossil Mammaliaformes included in macroevolutionary analyses.