

Biological Journal of the Linnean Society, 2016, 118, 78-94. With 11 figures.

A multiple variance Brownian motion framework for estimating variable rates and inferring ancestral states

JEROEN B. SMAERS^{1,2,3}*, CARRIE S. MONGLE¹ and ANNE KANDLER⁴

¹Department of Anthropology, Stony Brook University, Circle Rd, Stony Brook, NY 11794-4364, USA ²Department of Anthropology, University College London, 14 Taviton Street, London WC1H 0BW, UK ³Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK

⁴Department of Mathematics, City University London, Northampton Square, London EC1V 0HB, UK

Received 4 July 2015; revised 29 September 2015; accepted for publication 10 December 2015

To understand the nature of the evolutionary process, it is of paramount importance that temporal patterns of change in biological traits are accurately documented. The paleontological record is, however, inherently incomplete, leaving researchers with only a limited set of observed taxonomic units (OTUs) to estimate broader patterns of biological change. In this context, phylogenetic comparative methods have been developed aiming to estimate patterns of phenotypic change through time based on a phylogenetic tree and a limited set of OTUs. Such methods typically employ mathematical models proposing how change is likely to have unfolded over time. The most commonly used model, Brownian motion (BM), assumes that average trait change is proportional to the square root of time and that the rate of evolution is stochastically constant across all branches. This, however, lies in contrast to the commonly agreed notion that many biological traits change at different rates along different branches of the tree of life. We present a method for inferring ancestral states that allows for different evolutionary rates along different branches of the phylogenetic tree. The goal is to include the effects of variation in rates of phenotypic change across phylogenetic space. Based on the available phenotypic and phylogenetic information, we estimate measures of the rate of evolution on each individual branch and, subsequently, these estimates are used to parameterize a multiple variance BM model inferring the phenotypic values at all internal nodes. We demonstrate the validity of our approach with a series of simulations and an empirical example. We show that values for internal nodes inferred using our approach are equivalent to those inferred with a constant variance BM model if phenotypic evolution occurs according to standard BM. When evolution occurs at different rates along different branches of the phylogeny, our approach greatly outperforms constant variance BM. We further demonstrate that our approach accurately detects bursts of change in phylogenetic space. An empirical analysis of the evolution of primate brain and body mass reveals that our approach yields an improved statistical fit relative to both traditional and recent methods, and provides estimates of nodal values that lie within a range expected based on the fossil record. © 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78–94.

KEYWORDS: body size, brain size - multi-rate Brownian motion - phylogenetic comparative method.

INTRODUCTION

Phylogenetic comparative methods comprise a set of statistical approaches used to analyze phylogenetic trees, often in association with trait data. The general aim of many modern phylogenetic comparative methods is to clarify how evolution has shaped biodiversity through time (O'Meara, 2012; Pennell & Harmon, 2013). Common applications include testing for correlated evolution between traits (Felsenstein, 1985; Grafen, 1989), testing hypotheses about the tempo and mode of evolution of morphological traits (Butler & King, 2004; O'Meara *et al.*, 2006), and the study of the dynamics of clade diversification (Rabosky & Crandall, 2006).

In recent years, a wide variety of methods aiming to understand the tempo and mode of trait evolution have been developed. Such methods treat the phylogeny as a historical framework that can be used to model trait change along its branches. These meth-

^{*}Corresponding author. E-mail: jeroen.smaers@stonybrook.edu

ods commonly use models of evolution dictating how to map trait variation in observed taxonomic units onto the branches of the phylogeny. Despite an apparent trend toward model diversification, recent developments can largely be understood as expansions of two standard models: Brownian motion (BM) and Ornstein-Uhlenbeck (OU). The standard BM model assumes that average trait change is proportional to the square root of time and that the rate of evolution is stochastically constant such that it has a single mean and variance across all branches. Expansions of the BM model include the development of multi-rate BM models using reversible-jump Markov chain Monte Carlo (MCMC) procedures to estimate how rates vary across the tree or to test specific hypotheses about where rate shifts occur (O'Meara et al., 2006; Eastman et al., 2011; Venditti, Meade & Pagel, 2011; Revell et al., 2012). Specifically, Venditti et al. (2011) introduced an approach that can be used for the estimation of both variable rates and ancestral states (for implementation, see Pagel & Meade, 2013). OU models allow for a departure from the unconstrained random walk of BM such that traits, in addition to evolving with a random walk component, tend to be pulled towards some optimum value (θ) by the strength of a restraining force (α) . Expansions of the OU framework include 'multi-regime' methods that allow the OU parameters to vary across the tree (Butler & King, 2004; Beaulieu et al., 2012; Ingram & Mahler, 2013; Uyeda & Harmon, 2014). Such 'multi-regime' OU models are particularly useful when selecting among alternative model parameterizations (where each parameterization characterizes a different evolutionary scenario describing the structure of the adaptive landscape through time).

The interpretation of what precisely is explained by the application of these methods is, however, not as straightforward as it may seem. Often, phylogenetic comparative methods are used to test or describe patterns of evolution that align with biological concepts such as 'adaptive radiations' and 'key innovations'. However, the way that BM and OU model parameters are linked with such biological concepts is not always clear (Pennell & Harmon, 2013; Pennell, 2015; Pyron, 2015). Moreover, statistical models may not adequately capture the historical patterns of trait change that we are interested in (Harmon et al., 2010; Pennell et al., 2015). The central assumption of standard BM that trait change is proportional to the square root of time and constant along all branches has long been considered to be incompatible with how most traits evolve (Harvey & Purvis, 1991). The central OU assumption of cladewide stabilizing selection may be equally unrealistic for many traits (Pennell, 2015).

The dominant use of BM and OU based models was proposed to stem from their ability to provide a balanced compromise between three fundamental features (Pennell, 2015): (1) their usefulness for detecting patterns of change through time; (2) their interpretative value in terms of formal processes of evolutionary population genetics; and (3) their being at least loosely tied to biological concepts. It is for these reasons that recent expansions of the two core models undeniably represent a significant advancement in the field of phylogenetic comparative meth-The issue remains, however, that those ods. parameters only loosely tied to the biological processes they purport to model may fail to capture the evolutionary features of greatest interest (Pyron, 2015). Moreover, much more work is needed to increase the link between macroevolutionary parameters and microevolutionary processes (Pennell & Harmon, 2013).

Looking at this from another perspective, it can be argued that phylogenetic comparative applications measure patterns of trait change over time, irrespective of any microevolutionary processes underlying the change (Pennell, 2015). The primary aim of the present study is to document patterns of change and the putative commonalities and differences among taxa that they reveal, rather than to make inferences about specific evolutionary processes. The adoption of this position opens the possibility of using a wider array of models primarily aimed at providing a better fit to the data. As Pennell (2015) points out, such approaches may be derived from macroevolutionary diffusion processes (Clauset & Erwin, 2008), macroecological theories or statistical learning approaches.

Below, we present a method for inferring phenotypic values for internal nodes that aims to incorporate potential variation in the rate of phenotypic change along different branches of a phylogenetic tree. Our approach differs from standard models such as BM in that it estimates branch-specific rates of evolution in a deterministic manner from the available phenotypic and phylogenetic information. The estimated evolutionary rates are subsequently used to parameterize a multiple variance BM (mvBM) model. Based on this model of evolution, we apply a Bayesian MCMC procedure to stochastically infer the phenotypic values for all internal nodes. The method is partly based on work by Smaers & Vinicius (2009) in that it aims to capture branch-specific change by triangulating between a predicted value of an internal node (comprised of a weighted mean, with weights a function of the patristic distance of the node in question to the OTUs) and the values of its two descendants.

The currently described method can be considered as one possibility within a larger framework of

mvBM estimation. Different estimation procedures of the deterministic rate (see Methods) can be conceived and it is plausible that different procedures may provide a better fit for different datasets. To highlight how the estimation procedure may be modified, we describe one possible alternative modification when evaluating the accuracy of our approach in the context of the evolution of primate brain and body mass.

The accuracy of the proposed mvBM approach is evaluated using a series of simulations and an empirical example analyzing the evolution of primate brain and body mass. In the latter evaluation, the accuracy of mvBM is compared with traditional methods of ancestral estimation (Pagel, 1997, 1999), as well as recent approaches that infer both variable rates and ancestral states (Venditti *et al.*, 2011). Methods based on OU models are not considered because their current formalizations aim to describe the structure of the macroevolutionary landscape, rather than ancestral states and rates for particular nodes/branches in the tree.

METHODS

We assume a phylogenetic tree, φ , characterized by n tips, describing the genetically inferred evolutionary history of the n species at the tips. The sequence $\{i\}_{i=1, \dots, 2(n-1)}$ denotes the node labels of the considered tree and the first n labels describe the nodes at the tips. Further, the matrix $[b_{ij}]_{ij=1, \dots, 2(n-1)}$ describes the patristic distances among all nodes and tips of φ . If node j is a direct descendant of node i then $b_{ij} = (=b_{ji})$ stands for the branch length between the nodes i and j. It holds that $b_{ii} = 0$ for all $i = 1, \dots, 2(n-1)$.

The values of the considered continuous phenotypic trait of the *n* species at the tips are described by the vector $X = [x_1, x_2, ..., x_n]$. The expected phylogenetic covariances for the tree are:

$$V = egin{bmatrix} t_1 & t_{12} & \ldots & t_{1n} \ t_{21} & t_2 & \ldots & t_{2n} \ dots & & \ddots & dots \ t_{n1} & t_{n2} & \ldots & t_n \end{bmatrix}$$

where t_{ij} describes the shared time of evolution of species *i* and *j*, and t_i is the total time of evolution of species *i*.

CONSTANT VARIANCE BM MODEL

Under a constant variance BM model, the joint distribution of X is given by the multivariate normal distribution:

$$\begin{split} \mathbb{P}(X,\mu,\sigma^2) = & \frac{1}{(2\pi\sigma^2)^{\frac{\mu}{2}}} |V|^{-\frac{1}{2}} \text{exp} \\ & (-\frac{1}{2\sigma^2} (X-\mu\mathbf{1})^T V^{-1} (X-\mu\mathbf{1})) \end{split}$$

where the parameters μ and σ^2 describe the value of the root and the variance of the BM, respectively and $\mathbf{1} = [1,1, \ldots 1]$. The corresponding log-likelihood function (Felsenstein, 1973) is given by:

$$\begin{aligned} \ln \mathcal{L}(X,\mu,\sigma^2) &= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} ln |V| - \\ &\frac{1}{2\sigma^2} (X-\mu \mathbf{1})^T V^{-1} (X-\mu \mathbf{1}). \end{aligned} \tag{1}$$

Based on the log-likelihood function (1), we perform a Bayesian MCMC analysis (Revell, 2012). The outcome of this analysis are posterior distributions of the variances σ^2 of the BM model, as well as of the trait values of all internal nodes.

MVBM model

Our aim is to relax the assumption that phenotypic evolution occurs at the same rate on each branch of the given phylogenetic tree φ . Accordingly, we develop a method consisting of a deterministic estimation procedure and a stochastic BM inference framework. Although the deterministic part provides estimations of measures of the rate of evolution on each individual branch based on the given phenotypic and phylogenetic information, the stochastic part uses these estimations to parameterize a mvBM model. Below, we detail the suggested methodology.

Deterministic estimation part

In a first step, we generate estimations of the phenotypic trait values of the n - 1 internal nodes. Two extreme approaches could be conceived to attain this goal. One approach assumes that the incorporation of all phylogenetic and phenotypic information provides an appropriate estimate of the value of an internal node. This approach would combine all available 'global' information from across the phylogeny into an estimate of the value of each internal node. A second approach assumes that nodal values can be calculated without taking tree structure into account but, instead, leveraging information from its closest relatives. Here, we propose to combine both approaches to leverage the 'global' estimate with 'local' information.

Our proposed 'global' estimate comprises a weighted mean based on φ and the trait values of the tips. We calculate the estimates of the values for each internal node under the assumption that the

leverage of each tip value on the estimated value of each internal node is a function of their phylogenetic distance. In detail, the weighted mean of node k, denoted by P_k , is determined by

$$P_{k} = \frac{\sum_{i=1}^{n} \frac{x_{i}}{b_{ik}^{2}}}{\sum_{i=1}^{n} \frac{1}{b_{ik}^{2}}}, \quad k = n+1, \dots, 2(n-1).$$
(2)

where b_{ik} describes the patrixtic distance between the tip value *i* and internal node *k*. Utilization of the squared patrixtic distance results in a larger influence of tip values phylogenetically close to node *k*.

Our proposed 'local' estimate of the value of node k, denoted by p_k , is the average of the values $x_{n_1(k)}$ and $x_{n_2(k)}$, representing the values of the two descendants, $n_1(k)$ and $n_2(k)$ of node k:

$$p_k = rac{x_{n_1(k)} + x_{n_2(k)}}{2}, \quad k = n+1, \dots, 2(n-1).$$
 (3)

Based on the global and local estimates given by eqns (2) and (3), the phenotypic trait value of the internal node k is given by:

$$\begin{aligned} x_k &= aP_k + (1-a)p_k = a \frac{\sum_{i=1}^n \frac{x_i}{b_{ik}^2}}{\sum_{i=1}^n \frac{1}{b_{ik}^2}} + (1-a) \frac{x_{n_1(k)} + x_{n_2(k)}}{2} \\ k &= n+1, \dots, 2(n-1). \end{aligned}$$

The coefficient *a* controls the importance of local and global information. Equation (4) thus provides estimates of the values of the internal nodes *k*, $k = n + 1, \ldots, 2(n - 1)$ leveraging both global and local information. In the following, we use $a = \frac{1}{3}$, which corresponds to triangulating P_k with $x_{n_1(k)}$ and $x_{n_2(k)}$ (Fig. 1), and obtain:

$$x_k = rac{P_k + x_{n_1(k)} + x_{n_2(k)}}{3}, \quad k = n+1, \dots, 2(n-1).$$
 (5)

Based on these estimates of the nodal values, we are able to calculate measures of the rate of evolution for each branch of the tree. In particular, we calculate:

$$\sigma_{kn_1(k)}^2 = \frac{(x_{n_1(k)} - x_k)^2}{b_{kn_1(k)}}$$
 and $\sigma_{kn_2(k)}^2 = \frac{(x_{n_1(k)} - x_k)^2}{b_{kn_2(k)}}.$
(6)

As before, x_k represents the value of node k, $x_{n_1(k)}$, $x_{n_2(k)}$ represents the values of its two descendants $n_1(k)$ and $n_2(k)$, and $b_{kn_1(k)}$ and $b_{kn_2(k)}$ represent the corresponding branch lengths.



Figure 1. Illustration of equation (5), assuming $b_{Xn_1(k)} = b_{Xn_2(k)} = b$.

Summarizing, the deterministic part of our framework provides estimates of the values of all internal nodes and of measures of the rate of evolution $\{\sigma_{kn_i(k)}^2\}_{k=n+1,\dots,2(n-1),i=1,2}$ on each individual branch based on the given phenotypic and phylogenetic information. We also note that the functional relationships (2), (3), and (4) used for estimating the nodal values can be modified if this is more appropriate for a specific application. A three-taxon worked example of the deterministic estimations described here is provided in the Supporting information (Data S1).

However, point estimates as given by eqns (5) and (6) are of limited use for inference purposes because the stochastic nature of the evolutionary process is neglected. To include uncertainty, we develop, in the second step, a mVBM framework and use the obtained point estimates as input parameters to a BM Bayesian MCMC procedure.

Stochastic inference part

We assume that phenotypic evolution occurs according to BM but allow the variances of the BM process to vary across branches. In particular, we assume that the respective variances are given by the estimates $\{\sigma_{kn_i(k)}^2\}_{k=n+1,\dots,2(n-1),i=1,2}$ determined in eqn (6). Our aim is to infer the values of the internal nodes under the mvBM model of phenotypic evolution.

To apply the same standard Bayesian MCMC procedure as that for the constant variance BM model but still include different rates of evolution, the branch lengths of the given tree φ need to be transformed appropriately. According to our model assumption, the BM process describing the pheno-

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78-94

typic evolution from node k to its descendants $n_1(k)$ possesses the variance $\sigma_{kn_1(k)}^2$ [given by eqn (6)] and the BM process between node $n_2(k)$ possesses the variance $\sigma_{kn_2(k)}^2$.

Basic BM theory states that the distribution of the BM process (starting at value x) with variance σ^2 after time t is given by a normal distribution with mean x and variance $\sigma^2 t$. Therefore, the same distributions of phenotypic values at the nodes $n_1(k)$ and $n_2(k)$ can be achieved by assuming a constant variance $\bar{\sigma}^2$ but adjusting the time of evolution accordingly. In detail, the BM processes with variances $\sigma_{kn_i(k)}^2$ and $\bar{\sigma}^2$ generate the same distribution at times $b_{kn_i(k)}$ and $\beta_{kn_i(k)}$, respectively, if it holds

$$\sigma^2_{kn_i(k)}b_{kn_i(k)} = ar{\sigma}^2eta_{kn_i(k)}, \quad i = 1, 2$$

Consequently, we define the branch lengths of the transformed tree $\hat{\phi}$ by

$$egin{aligned} η_{kn_i(k)} = rac{\sigma_{kn_i(k)}^2}{ar{\sigma}^2} b_{kn_1(k)} = rac{\left(x_{n_i(k)} - x_k
ight)^2}{ar{\sigma}^2}, \; i=1,2 \end{aligned}$$

where $\bar{\sigma}^2$ is a constant (here proposed as the mean of the posterior distribution of inferred σ^2 values from a constant variance BM MCMC procedure based on φ and X). In this way, we ensure that the distributions of the nodal values obtained by assuming a BM process along the original tree φ with the variances $\{\sigma_{kn_i(k)}^2\}_{k=n+1,\dots,2(n-1),i=1,2}$ at each branch coincide with the distributions obtained by assuming a BM process along the transformed tree $\hat{\phi}$ with constant variance $\bar{\sigma}^2$ at each branch. In other words, $\beta_{kn:(k)}$ describes the expected branch length if a BM process with variance $\bar{\sigma}^2$ is considered to have produced the observed squared trait change $(x_{n_i(k)} - x_k)^2$. Naturally, if trait changes are very small, the branch length is equally small. Although this is biologically plausible, this is statistically undesirable. To avoid this situation, we propose to lengthen branches (alternative branch lengthening procedures can be conceived; one possible such alternative is described and analyzed in the section 'Empirical example') as:

$$\hat{\beta}_{kn_i(k)} = b_{kn_i(k)} + \beta_{kn_i(k)}.$$
(7)

Based on the transformed phylogenetic tree $\hat{\varphi}$ with branch lengths $\{\hat{\beta}_{kn_1(k)}\}_{k=n+1,\dots,2(n-1)}$, we perform, in the last step, the same Bayesian MCMC analysis as for the constant variance BM model.

The outcomes of this analysis are samples of the variance of the constant variance BM model, denoted

by $\{\hat{\sigma}_j^2\}_{j=1,\dots,N}$, where *N* stands for the number of generations sampled by the MCMC algorithm, as well as the trait values of all internal nodes under a multiple variance BM model of phenotypic evolution.

The corresponding samples of evolutionary rates for each branch of the given phylogenetic tree ϕ are determined by:

$$\hat{\sigma}_{kn_{1},j}^{2} = \hat{\sigma}_{j}^{2} \frac{\hat{\beta}_{kn_{1}(k)}}{b_{kn_{1}(k)}} = \hat{\sigma}_{j}^{2} \left(1 + \frac{\beta_{kn_{1}(k)}}{b_{k1(k)}}\right) \quad \text{and } \hat{\sigma}_{kn_{2},j}^{2}$$

$$= \hat{\sigma}_{j}^{2} \frac{\hat{\beta}_{kn_{2}(k)}}{b_{kn_{2}(k)}} = \hat{\sigma}_{j}^{2} \left(1 + \frac{\beta_{k2(k)}}{b_{k2(k)}}\right)$$

for k = n + 1, ..., 2(n - 1) and j = 1, ... N.

The implementation of these algorithms is available in the R package 'evomap' (Smaers, 2014), under the function 'mvBM'.

SIMULATION EXAMPLE

To systematically explore the properties of the proposed mvBM method, we compare its performance with the performance of a constant variance BM method (cvBM) in situations where phenotypic evolution follows and deviates from a standard BM model. The cvBM method used is 'anc.Bayes', as described by Revell (2012), which takes a phylogenetic tree and a vector of terminal states for a continuously valued character and uses Bayesian MCMC to sample from the posterior distribution for the character states at ancestral nodes in the tree. The mvBM method is implemented in the same way but uses the transformed tree $\hat{\varphi}$ with the branch lengths given in eqn (7) as input tree rather than φ .

We start by analyzing the accuracy of mvBM (and cvBM) within a simulation framework but subsequently consider an empirical example as well. The simulation framework provides the advantage of possessing full knowledge of the underlying evolutionary process that has generated the phenotypic data at the tips of the given phylogenetic tree. This allows us to compare the inferred the ancestral states and rates of evolution with the known true values. In detail, for each simulation run, we generate stochastic pure-birth trees (Revell, 2012) and simulate values of a continuous trait at all nodes of the phylogenetic tree (Paradis, Claude & Strimmer, 2004), recursively from the root with value zero, according to the scenarios specified below.

A first class of simulations ('BM simulation') generates 1000 trees with 100 tips along with a continuous trait for each tree. The trait is evolved according to a BM process with variance $\sigma^2 = 0.01$. To evaluate the putative effect of the magnitude of σ^2 , simulations were repeated with $\sigma^2 = 1$, 25, and 100. A second class of simulations ('burst simulations') generates 1000 instances of a continuous trait evolution on a fixed tree with 100 tips. We distinguish three burst scenarios (Fig. 2). The 'one-burst' scenario assumes that evolution occurs according to a BM process with the variance $\sigma^2 = 0.01$ on all branches of the phylogenetic tree but a single branch. The 'burst' branch experiences a higher rate of evolution that is modelled by a BM process with variances $\sigma^2 = 1, 25$, and 100. Similarly, the 'five-burst wide' scenario assumes that evolution occurs according to a BM process with the variance $\sigma^2 = 0.01$ on all branches of the phylogenetic tree except for one branch in each of five major clades in the phylogeny, where σ^2 set to 1, 25, and 100. The 'five-burst close' scenario is equivalent to the 'five-burst wide' scenario, except that all five bursts are allocated to within more closely related clades.

To analyze the accuracy of the cvBM and mvBM method, we determine, for each simulation run, the mean of the posterior distributions (based on 10^6 generations of the MCMC procedure with a 20% burn-in and sampling every 100th generation) of the log-likelihood values and the values of the internal nodes. To quantify the differences, we generate the

distributions of the mean likelihood values and of the sum and SD of the absolute difference between the true and inferred mean node values (Figs 3, 4, 5, 6).

Additionally, for each simulation run, the inferred node values were regressed against the true values. The R^2 value hereby provides an estimate of the overall fit of the inferred and true values $(R^2 = 1)$ indicates that inferred and true values coincide perfectly, whereas $R^2 = 0$ points to situations where the inferred values do not reflect the true values). The slope of the regression line provides information about putative over- or underestimation of smaller or larger true values (a slope of 1 is the benchmark for accuracy). The intercept of the regression line provides information about putative consistent over- or underestimation of true values (an intercept of 0 is the benchmark for accuracy). The distributions of these statistics for the considered evolutionary scenarios are shown in Figures 3, 4, 5, 6.

Figure 3 summarizes the results for the BM simulation. We deduce that cvBM and mvBM possess similar accuracy when inferring values of internal nodes when evolution is simulated according to an underlying BM process (see also Supporting information, Data S2). When comparing the absolute error



Figure 2. The three burst scenarios considered in the simulation example.

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78–94



Figure 3. Results of the BM simulation.



Figure 4. Result of the 'one burst' simulation.

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78–94





sd of absolute difference



1.0

Figure 5. Results of the 'five burst wide' simulation.



Figure 6. Results of the 'five burst close' simulation.

(the sum of the absolute difference between estimated and simulated nodal values averaged across all simulations) between methods, the difference between cvBM and mvBM does not exceed 3.3% for any value of σ^2 = 0.01, 1, 25, and 100 (see Supporting information, Data S2). All other accuracy

	One burst		Five burst wide		Five burst close	
Ratio	cvBM	mvBM	cvBM	mvBM	cvBM	mvBM
1 versus BM	1.15	1.03	1.79	1.22	1.72	1.29
5 versus BM	2.63	1.29	6.65	2.58	6.35	3.09
10 versus BM	4.50	1.62	13.53	4.71	11.97	5.34

Table 1. Ratio of the average sum of the absolute difference between simulated and estimated values across 1000 simulations between each burst scenario and BM

BM, Brownian motion; cvBM, constant variance BM; mvBM, multiple variance BM.

measures are equally comparable between the two methods (Fig. 3; see also Supporting information, Data S2). mvBM results further indicate slightly higher log-likelihood values. This result is expected given that mvBM includes more degrees of freedom in its inference.

Results for the burst scenarios (Figs 4, 5, 6 with $\sigma^2 = 25$ on the burst branch) indicate that mvBM outperforms cvBM on all accuracy measures. Considering that, in the BM simulation scenario, the results for cvBM and mvBM are equivalent, Table 1 uses the calculated absolute error of the BM simulation scenario ($\sigma^2 = 0.01$) as a benchmark to evaluate how much the inclusion of burst branches increases the error across scenarios. The results indicate increasingly disparate results as the σ^2 of the burst branches increase.

Table 2 evaluates whether the decreased error in mvBM is indeed a result of the correct inference of the evolutionary burst. For this, the ratio of the average sum of absolute error across simulations of cvBM to mvBM is listed for the node that is directly descendant to the burst branch ('burst') or nodes that are within one generation of either the ancestor or the descendant node of the burst branch ('related'), or neither ('nonrelated'). Table 2 also lists results for the root value. Overall, the results indicate that mvBM is consistently more accurate than cvBM (error ratios between 1.09 and 5.36) and that mvBM is increasingly accurate relative to cvBM for each set of nodes as bursts occur at a higher magnitude (higher σ^2 values produce less error in mvBM relative to cvBM). For the one burst scenario, the nonrelated branches consistently indicate a lower error ratio relative to mvBM, demonstrating that mvBM is increasingly accurate in those nodes that surround the burst branch. The same is true for the five burst

Table 2. Ratio of the average sum of the absolute differ-
ence between simulated and estimated values across 1000
simulations between cvBM and mvBM for different subsets
of nodes (see text for definition of the subsets) for each of
the burst scenarios

σ	Nodes	One burst	Five burst wide	Five burst close
1	All	1.09	1.42	1.29
	Burst	1.51	1.86	1.89
	Related	1.52	1.62	1.46
	Nonrelated	1.02	1.21	1.08
	Root	1.25	1.30	1.24
5	All	1.98	2.51	2.00
	Burst	2.61	2.66	3.04
	Related	3.07	2.46	1.89
	Nonrelated	1.63	2.51	1.84
	Root	3.73	2.78	2.11
10	All	2.70	2.79	2.18
	Burst	2.81	2.72	3.14
	Related	3.51	2.54	1.95
	Nonrelated	2.36	3.31	2.35
	Root	5.55	3.28	2.22

scenarios, although this difference is less pronounced with stronger bursts. Importantly, for all scenarios, the root value is indicated to be increasingly more accurate in mvBM when bursts increase in magnitude.

Figure 7 depicts results from an analysis of the distributions of the evolutionary rates in the 'oneburst' scenario ($\sigma^2 = 25$ for the burst branch; $\sigma^2 = 0.01$ for all other branches) for the burst branch (Fig. 7A), a branch closely related to the burst branch (Fig. 7B), and a branch distant from the burst branch (Fig. 7C). We observe that the cvBM approach explains the single burst of evolution by increasing the rate of evolution across all branches of the phylogenetic tree (Fig. 8); the single instance of a large amount of change is averaged over the tree. By contrast, mvBM identifies the branch where the burst occurs and infers a much higher evolutionary rate for this particular branch (Fig. 7A) and a more accurately low rate for other branches (Fig. 7C), which is more comparable with the true rate of 0.01. This point is further illustrated by Figure 8, which shows the distribution of σ^2 values across the different one burst scenarios obtained by cvBM. It is obvious that, in the cvBM framework, larger burst rates along single branches are compensated for by a larger overall rate of evolution. These results demonstrate that mvBM is able to accurately distinguish between branches with a low vs. high rate of evolution.



Figure 7. Example results of the distribution of rates for the one burst scenario with $\sigma = 5$ for the burst branch and $\sigma = 0.1$ for all other branches. The dotted lines indicate those branches for which the distribution of rates is shown. The distributions shown are representative of those in each category (burst, related, non-related). Note the difference in scale of the X-axis in (C) relative to (A) and (B).

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78-94



Figure 8. The distribution of σ across the different scenarios of the one-burst simulations obtained with cvBM.

We note that, per construction, mvBM exhibits a certain degree of autocorrelation among related branches (Fig. 7A, B). The degree of autocorrelation is, however, not prohibitive for accurately detecting the burst branch in phylogenetic space. To evaluate the accuracy in detecting the burst branch, we calculate the area of overlap between the probability density distribution of the inferred rates of the burst branch and the inferred rates of closely (within two generations) and distantly (outside two generations) related branches. The area of overlap quantifies the certainty with which a higher (or lower) evolutionary rate is inferred among branches. We consider an overlap of < 5% between these distributions (in addition to the burst branch indicating a higher rate) as an indication of the accurate detection of the burst branch. Across both closely and distantly related branches for the example shown in Figure 7, the burst branch is correctly identified in 91-95% of the simulations (i.e. in these situations, the rate of the burst branch is higher and its probability density distribution has an area of overlap with the branch it is compared with of less than 5%). The mean proportion of simulations in which the burst was accurately detected is 93.6% across closely related branches and 94.2% across distantly related branches.

EMPIRICAL EXAMPLE: PRIMATE BRAIN AND BODY SIZE The accuracy of mvBM relative to other methods is evaluated further by analyzing the evolution of primate brain and body mass. Brain and body mass data for 144 extant species were taken from Isler *et al.* (2008) and the phylogeny from Arnold, Matthews & Nunn (2010). Primate brain and body mass are particularly useful when comparing estimated with ancestral values because the primate fossil record provides reasonable estimates for several ancestral nodes. Here, six methods are compared in their estimation of ancestral values of primate brain and body mass. Accuracy is assessed through loglikelihood scores and a comparison of estimated nodal values with the fossil record.

In addition to ancestral estimation through cvBM, two classes of methods are considered. A first class of methods uses maximum likelihood optimization to estimate parameters that rescale all, or a subset of, branch lengths homogenously according to the amount of trait change that is inferred to have occurred along the subset of branches (Pagel, 1997, 1999). A λ model fits the extent to which the phylogeny predicts the covariance among trait values for species. In this model, the lengths of ancestral branches are multiplied by the value of λ . A δ model fits the relative contribution of early vs. late evolution, raising all node depths to the value of δ . A κ model fits a punctuational model where character disparity is related to the number of speciation events among species, raising all branch lengths to the power of κ . A second suite of methods implements heterogonous rescaling of branch lengths considering that the tempo and mode of evolution may be different across phylogenetic space. The first method fits a multi-rate BM model using reversiblejump MCMC procedures to estimate how rates vary across the tree (rjBM) (Venditti *et al.*, 2011). The second method is the method proposed in the present study (i.e. mvBM).

All methods are implemented in a similar manner using MCMC optimization with 5×10^6 iterations with a 20% burn-in (resulting in normally distributed log-likelihood values for all methods). In a first step, branch lengths are rescaled according to the above-described procedures for each method. For the λ , δ , and κ models, parameters are estimated using maximum likelihood (Harmon *et al.*, 2008). For rjBM, branches are rescaled as described by Venditti *et al.* (2011) and implemented in BayesTraits (Pagel & Meade, 2013). For mvBM, branches are rescaled according to the procedure described in the Methods section of the present study. In a second step, the rescaled tree is analyzed using constant variance BM using the MCMC implementation for ancestral estimation as described by Revell (2012). Ancestral values for the rjBM model are also inferred using MCMC optimization, although this procedure is implemented in a different software package (Pagel & Meade, 2013: BayesTraits). All data were log transformed prior to analysis.

For mvBM, we further report results for a possible alternative parameterization of the algorithms described above. The alternative parameterization consists of obtaining $\hat{\beta}_{kn_i(k)}$ by lengthening $\beta_{kn_i(k)}$ by $b_{kn_1(k)} * b_{kn_2(k)}/(b_{kn_1(k)} + b_{kn_2(k)})$ rather than $b_{kn_i(k)}$. This is an approach similar to that proposed by Felsenstein (1985) when computing estimated nodal values with the method of 'Independent Contrasts'. The results of the alternative parameterization are indicated as mvBM₂ in Table 3, whereas those of the parameterizations described in the methods section are indicated as mvBM₁.

Fossil values were selected for those specimens for which reasonable agreement exists with regard to their taxonomic position and the estimation of their brain and body mass. It is clear that fossil values should not be considered as absolute benchmarks of

Method	Log-likelihood	MRCA: Mya Fossil value	Homo-Pan ~6 430–540	Hominoidea ~19 150–170	Cercoptihecoidea ~21 53	Simiiformes ~47 34	Galagidae ~17 NA
Brain mas	s (g)						
$mvBM_2$	168		453	120	72	43	
$mvBM_1$	142		441	126	76	43	
rjBM	127		406	108	69	29	
δ	58		540	192	83	50	
λ	52		530	175	81	39	
κ	52		528	173	80	39	
cvBM	52		530	176	80	38	
		Fossil value:	30–39	10–20	4.8	6.7	0.304
Body mass	s (kg)						
$mvBM_2$	31		48.406	9.252	6.707	2.836	0.355
$mvBM_1$	3		48.842	10.494	7.049	2.873	0.309
rjBM	-18		55.226	8.238	5.320	1.159	0.192
λ	-62		52.123	16.686	7.268	2.621	0.241
δ	-62		53.487	18.457	7.624	3.315	0.224
cvBM	-63		52.557	17.024	7.354	2.563	0.231
κ	-64		51.817	17.017	7.507	2.743	0.240

Table 3. Results of the ancestral estimation analysis. 'MRCA' indicates 'most recent common ancestor'. 'Mya' refers to 'millions of years ago'. ' $vvBM_1$ ' refers to the parameterization of mvBM as described in the methods section. ' $vvBM_2$ ' refers to an alternative parameterization of mvBM described in the text

BM, Brownian motion; cvBM, constant variance BM; mvBM, multiple variance BM; MRCA, most recent common ancestor; NA, not available.

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78–94

accuracy. Traits such as body mass are difficult to infer from fossils. Moreover, in the absence of ancient DNA, the phylogenetic affinity of fossils is mostly inferred based on a gradual assumption of morphological evolution. Recent work, however, indicates that, even within a single anatomical unit, different modules may deviate from a gradual mode in different ways (Kivell, Barros & Smaers, 2013). Despite these uncertainties, the brain and body mass estimates of several primate fossils have been generally agreed to be reasonably representative of specific phylogenetic nodes. These fossil values can thus be used as approximate benchmarks for the accuracy of methods aiming to infer the mode and tempo of trait evolution based on extant information only. For the brain mass analysis, Australopithecus africanus and Australopithecus afarensis are considered for the most recent common ancestor (MRCA) of humans and chimpanzees (Kimbel, Rak & Johanson, 2004; Fleagle, 2013), Proconsul (Radinsky, 1974; Walker et al., 1983) as MRCA of Hominoidea (although recent work places these fossils in between the MRCA of Catarrhini and Hominoidea), Victoriapithecus (Benefit & McCrossin, 1997) as MRCA of the Cercopithecoidea, and Aegyptopithecus (Radinsky, 1973; Simons, 1993) as MRCA of the Simiiformes. For the body mass analysis, *Komba robustus* is also considered as MRCA of the galagidae (Steiper & Seiffert, 2012; Seiffert, Costeur & Boyer, 2015).

Table 3 indicates results for brain mass, indicating that both parameterizations of mvBM yield an improved statistical fit relative to other methods for both brain and body mass. For brain mass, all results lie within the margin of error of the fossil values, although all methods indicate a slight overestimation of the MRCA of Cercopithecoidea. For the body mass analyses, all methods overestimate the MRCA of humans and chimpanzees, although both mvBM parameterizations do so with slightly less error. Furthermore, the MRCA of Cercopithecoidea is slightly overestimated (but less so in rjBM), whereas the MRCA for Simiiformes is underestimated by all methods (particularly by rjBM). The MRCA of Galagidae is accurately estimated by both parameterizations of mvBM but underestimated by all other methods (particularly rjBM). Furthermore, rjBM estimated the root value of both brain and body mass to be markedly lower than other methods.

Figure 9 presents the rescaled branches of each method for the analysis of brain mass. We observe



Figure 9. The rescaled branches of the ancestral estimation of primate brain size according to each method.

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78-94



Figure 10. A visualization of the evolution of brain mass according to mvBM1. Deeper hues of blue represent higher brain mass, deeper hues of red lower brain mass.

© 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 78–94



Figure 11. The distribution of the rates of evolution starting from the ape ancestral branch leading up to the human branch. 'Apes' refers to the distribution of the rate in the ancestral ape branch, referring to the branch ancestral to the indicated extant species and its sister clade (the gibbons, here only indicated as "[8 tips]").

that, for this sample, the rescaled trees of mvBM and rjBM look similar. The difference in the results between these two methods lie in the more optimal statistical fit of mvBM, and the markedly lower estimate for the root value by rjBM. The evolution of brain mass according to mvBM is visualized in Figure 10. Figure 11 highlights the rates for the branches leading up to humans in the ape clade. In line with the fossil record, rates are elevated in the ancestral branch of great apes, relatively lower in subsequent branches towards humans, and again elevated in the human branch.

DISCUSSION

Phylogenetic comparative methods are often used to document patterns of change in biological traits along the branches of a phylogeny. Such methods employ formal mathematical models within a statistical framework. Trait change is described as a continuous stochastic process (e.g. a BM or Ornstein-Uhlenbeck process) and statistical inference techniques determine the parameters of the involved stochastic processes so that theoretical patterns match the observed patterns as closely as possible. Previous work has, however, shown that these models may not adequately capture the historical patterns of trait change that we are interested in (Harmon et al., 2010; Pennell et al., 2015) and also that the parameters are only loosely tied to the biological concepts they purport to explain (Pennell & Harmon, 2013; Pennell, 2015; Pyron, 2015).

Here, we describe a multiple variance BM framework. Our approach differs from standard models like BM in that it combines a deterministic estimation of branch-specific rates of evolution with a stochastic inference of values of all internal nodes. The proposed approach aims to provide a flexible platform for inferring the patterns of trait change along each branch of the phylogeny. This aim lies close to the often-intended use of phylogenetic comparative applications to documenting commonalities and differences among clades through time.

Using a simulation framework, we demonstrate that the accuracy by which mvBM infers values for internal nodes is equivalent to a cvBM model if phenotypic evolution occurs according to standard BM. When evolution occurs at different rates along different branches of the phylogeny, our approach is shown to greatly outperform constant variance BM (Figs 4, 5, 6). The increased accuracy of mvBM is particularly pronounced for values of the root and nodes that are in proximity of a burst of change. Although standard cvBM models deal with evolutionary bursts by increasing the rate of evolution for all branches of the tree (Fig. 8), the mvBM approach accurately infers a higher rate of evolution for a burst branch and much lower rates of evolution for the rest of the tree. Our approach hereby largely overcomes the problem of 'inherited error' that occurs when applying a standard BM framework to modelling a trait that evolves with unequal rates. The mvBM approach accurately infers branches with a higher evolutionary rate relative to those with a lower evolutionary rate. An analysis of the evolution of primate brain and body mass further reveals that our approach yields an improved statistical fit relative to both traditional methods (Pagel, 1997, 1999) and more recent multi-rate BM methods (Venditti *et al.*, 2011) and also provides estimates of nodal values that lie within the range expected based on the fossil record.

We emphasize that what we propose is a framework of multiple variance BM estimation. We describe a method that is a possible formalization of this framework. The results demonstrate that the described set of estimation procedures [specifically the definitions of P_k , p_k , a, and $\hat{\beta}$ in eqns (2–4) and (7)] are useful for capturing patterns of variable change through time. We therefore propose this formalization as a preliminary standard. Future work should explore how modifications to these estimations affect performance. It is plausible that different estimation procedures may provide a better fit for different datasets. In this context, we describe the results of a possible modification to $\hat{\beta}$ (eqn 7) in our analysis of the evolution of primate brain and body size. For the samples considered in this example, this modification considerably improves the statistical fit. Modifications to P_k , p_k , and a may also prove worthwhile. As reported, the current formalization exhibits a degree of autocorrelation among closely related branches, although not prohibitive for accurately detecting bursts of change in phylogenetic space. Although it is clear that modifications to P_k , p_k , and a are expected to affect the degree of inferred autocorrelation, further work is needed to quantify this relationship.

In conclusion, we present a method that comprises a deterministic estimation of branch-specific rates of evolution and a stochastic inference of values of internal nodes through BM Bayesian MCMC procedures. The goal of this method is to include the effects of variation in rates of phenotypic change across phylogenetic space. A series of simulations demonstrate that the proposed method estimates internal node values with equal accuracy to constant variance BM when trait evolution adheres to standard BM but with higher accuracy when evolution occurs at different rates along different branches of the phylogeny. Our approach is suggested to be particularly useful when shifts in the direction and rate of phenotypic change occur a few times on a tree.

ACKNOWLEDGEMENTS

We thank James B. Rossie and Erik R. Seiffert for advice on selecting fossil information, as well as P. David Polly, F. James Rohlf, David J.C. Dennis, and one anonymous reviewer for helpful comments and suggestions on the manuscript. This work was supported by the UK Natural Environment Research Council (grant no. NE/H022937/1).

This paper was a contribution to a Linnean Society symposium on "Radiation and Extinction: Investigating Clade Dynamics in Deep Time" held on November 10–11, 2014 at the Linnean Society of London and Imperial College London and organised by Anjali Goswami, Philip D. Mannion, and Michael J. Benton, the proceedings of which have been collated as a Special Issue of the Journal.

REFERENCES

- Arnold C, Matthews LJ, Nunn CL. 2010. The 10k trees website: a new online resource for primate phylogeny. *Evolutionary Anthropology* 19: 114–118.
- Beaulieu JM, Jhwueng DC, Boettiger C, O'Meara BC. 2012. Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution* 66: 2369–2383.
- Benefit BR, McCrossin ML. 1997. Earliest known Old World monkey skull. *Nature* 388: 368–371.
- Butler MA, King AA. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *American Naturalist* 164: 683–695.
- Clauset A, Erwin DH. 2008. The evolution and distribution of species body size. *Science* **321**: 399–401.
- Eastman JM, Alfaro ME, Joyce P, Hipp AL, Harmon LJ. 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* 65: 3578–3589.
- Felsenstein J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. American Journal of Human Genetics 25: 471.
- Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.
- **Fleagle JG. 2013.** Primate adaptation and evolution. San Diego, CA: Academic Press.
- Grafen A. 1989. The phylogenetic regression. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 326: 119–157.
- Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger
 W. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24: 129-131.
- Harmon LJ, Losos JB, Davies TJ, Gillespie RG, Gittleman JL, Jennings WB, Kozak KH, McPeek MA, Moreno-Roark F, Near TJ, Purvis A, Ricklefs RE, Schluter D, Schulte JA II, Seehausen O, Sidlauskas BL, Torres-Carvajal O, Weir JT, Mooers AO. 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64: 2385–2396.
- Harvey PH, Purvis A. 1991. Comparative methods for explaining adaptations. *Nature* 351: 619–624.
- Ingram T, Mahler DL. 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike information criterion. *Methods in Ecology and Evolution* 4: 416–425.

- Isler K, Kirk EC, Miller JMA, Albrecht GA, Gelvin BR, Martin RD. 2008. Endocranial volumes of primate species: scaling analyses using a comprehensive and reliable data set. Journal of Human Evolution 55: 967–978.
- Kimbel WH, Rak Y, Johanson DC. 2004. The skull of Australopithecus afarensis. Oxford: Oxford University Press.
- Kivell TL, Barros AP, Smaers JB. 2013. Different evolutionary pathways underlie the morphology of wrist bones in hominoids. *BMC Evolutionary Biology* 13: 229.
- O'Meara BC. 2012. Evolutionary inferences from phylogenies: a review of methods. Annual Review of Ecology, Evolution, and Systematics 43: 267–285.
- O'Meara BC, Ané C, Sanderson MJ, Wainwright PC. 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.
- Pagel M, Meade A. 2013. BayesTraits, Version 2.0. Reading: University of Reading.
- Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.
- **Pennell MW. 2015.** Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice. Edited by László Zsolt Garamszegi. *Systematic Biology* **64:** 161–163.
- **Pennell MW, Harmon LJ. 2013.** An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences* **1289**: 90–105.
- Pennell M, FitzJohn RG, Cornwell WK, Harmon LJ. 2015. Model adequacy and the macroevolution of angiosperm functional traits. American Naturalist 186: E33–E50
- **Pyron RA. 2015.** Post-molecular systematics and the future of phylogenetics. *Trends in Ecology & Evolution* **30:** 384–389.
- Rabosky DL, Crandall K. 2006. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60: 1152–1164.

- Radinsky L. 1973. Aegyptopithecus endocasts: oldest record of a pongid brain. *American Journal of Physical Anthropol*ogy **39:** 239–247.
- Radinsky L. 1974. The fossil evidence of anthropoid brain evolution. American Journal of Physical Anthropology 41: 15–27.
- **Revell LJ. 2012.** phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3:** 217–223.
- Revell LJ, Mahler DL, Peres-Neto PR, Redelings BD. 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution* 66: 135– 146.
- Seiffert ER, Costeur L, Boyer DM. 2015. Primate tarsal bones from Egerkingen, Switzerland, attributable to the middle Eocene adapiform Caenopithecus lemuroides. *PeerJ* 3: e1036.
- Simons EL. 1993. New endocasts of Aegyptopithecus: oldest well-preserved record of the brain in Anthropoidea. *Ameri*can Journal of Science 293: 383.
- **Smaers JB. 2014.** evomap: R package for the evolutionary mapping of continuous traits: Github. Available at: https://github.com/JeroenSmaers/evomap
- Smaers JB, Vinicius L. 2009. Inferring macro-evolutionary patterns using an adaptive peak model of evolution. *Evolutionary Ecology Research* 11: 991–1015.
- Steiper ME, Seiffert ER. 2012. Evidence for a convergent slowdown in primate molecular rates and its implications for the timing of early primate evolution. *Proceedings of the National Academy of Sciences of the United States of America* 109: 6006–6011.
- Uyeda JC, Harmon LJ. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology* **63**: 902–918.
- Venditti C, Meade A, Pagel M. 2011. Multiple routes to mammalian diversity. *Nature* 479: 393–396.
- Walker A, Falk D, Smith R, Pickford M. 1983. The skull of *Proconsul africanus*: reconstruction and cranial capacity. *Nature* 305: 525–527.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. 3 taxon worked example.

Data S2. Detailed results of the simulation example.