Inferring Bounded Evolution in Phenotypic Characters from Phylogenetic Comparative Data

FLORIAN C. BOUCHER[1,*] AND VINCENT DÉMERY[2]

[1]*Institute of Systematic Botany, University of Zurich, Zurich, Switzerland;* [2]*Laboratoire de Physico-Chimie Théorique, UMR Gulliver 7083, CNRS and ESPCI-ParisTech, Paris, France*
Correspondence to be sent to: Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8010 Zurich, Switzerland;
E-mail: florian.boucher@systbot.uzh.ch

Abstract.—Our understanding of phenotypic evolution over macroevolutionary timescales largely relies on the use of stochastic models for the evolution of continuous traits over phylogenies. The two most widely used models, Brownian motion and the Ornstein–Uhlenbeck (OU) process, differ in that the latter includes constraints on the variance that a trait can attain in a clade. The OU model explicitly models adaptive evolution toward a trait optimum and has thus been widely used to demonstrate the existence of stabilizing selection on a trait. Here we introduce a new model for the evolution of continuous characters on phylogenies: Brownian motion between two reflective bounds, or Bounded Brownian Motion (BBM). This process also models evolutionary constraints, but of a very different kind. We provide analytical expressions for the likelihood of BBM and present a method to calculate the likelihood numerically, as well as the associated R code. Numerical simulations show that BBM achieves good performance: parameter estimation is generally accurate but more importantly BBM can be very easily discriminated from both BM and OU. We then analyze climatic niche evolution in diprotodonts and find that BBM best fits this empirical data set, suggesting that the climatic niches of diprotodonts are bounded by the climate available in Australia and the neighboring islands but probably evolved with little additional constraints. We conclude that BBM is a valuable addition to the macroevolutionary toolbox, which should enable researchers to elucidate whether the phenotypic traits they study are evolving under hard constraints between bounds. [BBM; bounds; evolutionary constraints; macroevolution; maximum likelihood estimation; phylogenetic comparative data.]

INTRODUCTION

Models for the evolution of continuous characters on phylogenies are central to comparative analysis and to the study of evolutionary processes over macroevolutionary timescales (Harvey and Pagel 1991; O'Meara 2012; Garamszegi 2014). Among many other applications, these models are frequently used to infer the character states of ancestral species (e.g., Guerrero et al. 2013), identify correlated evolution between traits (e.g. Beaulieu et al. 2007), or measure the effect of a given environment on the evolution of species traits (e.g., Edwards and Smith 2010; Price et al. 2011). The basic model for the evolution of continuous characters on phylogenies is Brownian Motion (hereafter BM; Edwards and Cavalli-Sforza 1964; Felsenstein 1973), in which characters follow a constant-rate random walk with no trend. BM can model drift, drift-mutation balance, or even selection in a rapidly changing environment (Hansen and Martins 1996). In the past 20 years, various extensions of BM have been proposed. The most significant advance was the development of the Ornstein–Uhlenbeck model (hereafter OU, Hansen 1997), which adds a pull toward a central value to BM. OU was designed to model stabilizing selection in an analogy with adaptive landscapes in population genetics (Lande 1976), but is more generally assumed to model adaptive evolution over phylogenies (Hansen 1997; Uyeda and Harmon 2014). Further extensions have allowed the rate of BM to vary in different parts of the tree (O'Meara et al. 2006), parameters of the OU model to vary between clades (Butler and King 2004; Beaulieu et al. 2012), and characters to evolve suddenly

by large jumps either at speciation (Bokma 2008) or during anagenesis (Landis et al. 2013). Recently, complex methods have been developed to detect heterogeneity in the evolutionary processes acting across a clade with no prior information (Ingram and Mahler 2013; Uyeda and Harmon 2014).

Although all of these methods allow for various hypotheses to be tested, the panoply of available models is still limited (Pennell 2015). In particular, the OU model is the only one in which the variance of a character in a clade has limits and cannot grow indefinitely over time (Hansen 1997; Ho and Ané 2014), but no model can accommodate hard bounds on character values. This is a clear lack given that hard bounds on the values of some characters do exist in nature (Garland et al. 1993). Some characters indeed have an obvious lower bound: this is the case for all characters that can only take positive values, such as morphological measurements (including traits that are heavily studied in comparative studies, like body size and body mass) or durations (e.g., generation length, life expectancy). Other continuous characters evolve between two hard boundaries. This is obviously the case for proportions, which evolve between 0 and 1 (e.g., allele frequencies, Edwards and Cavalli-Sforza 1964, proportion of fast glycolytic fibers in muscles, Scales et al. (2009), genomic GC content, Mooers and Holmes (2000), overlap between breeding and wintering climatic niches in migratory animals). Limited genetic variation in a given lineage would also lead the associated phenotypic characters to be confined between two extreme values (Futuyma 2010; Nei 2013). Another example concerns the evolution of species environmental niches: in a given geographic region,
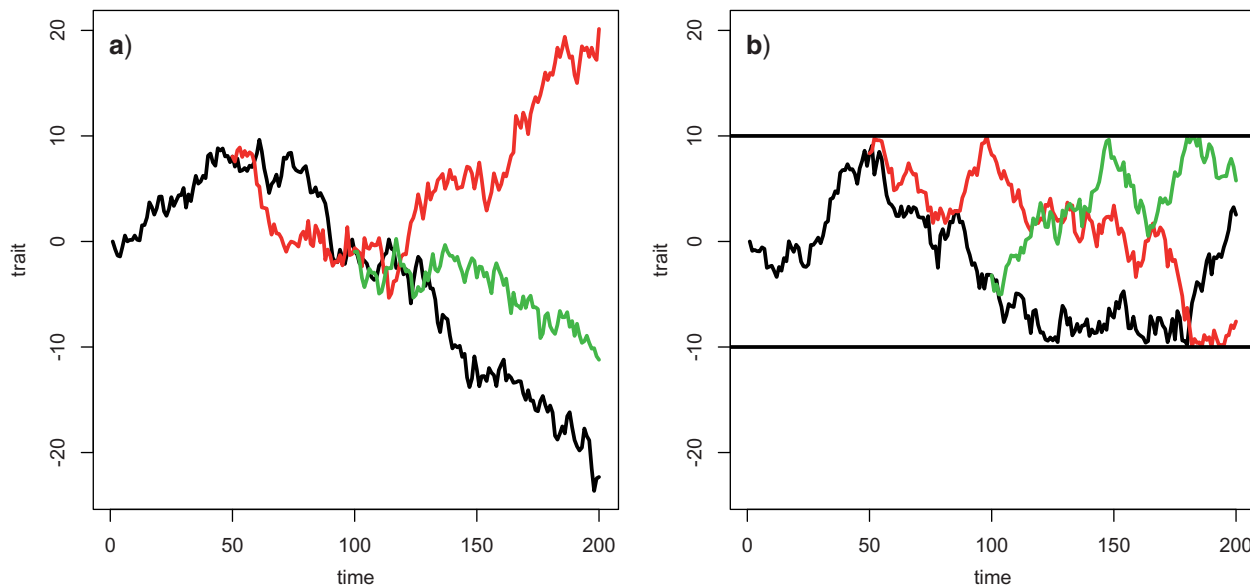
FIGURE 1.   Behaviour of the BBM model. The two panels show the evolution of a trait over time on a tree of three extant species. a) Brownian Motion without bounds (BM). b) Brownian Motion with two bounds (horizontal black lines): the traits of two lineages cross the whole interval, multiple hits on the bounds occur. Both plots were generated using the same value of the evolutionary rate, σ^2: the final variance of the trait is higher in BM.

the realized niche of a species (i.e., the environment it actually experiences, Hutchinson 1957) is bounded between the environmental extremes that exist in the region (Boucher et al. 2014; Wüest et al. 2015). In addition, some characters might have less obvious limits: for example, morphological and physiological processes might impose an upper bound on plant height (Koch et al. 2004), a lower limit on leaf mass per area in plants (Donovan et al. 2011), and an upper limit on insect body size (Kaiser et al. 2007). Taking again the example of species environmental niches, it is likely that biotic interactions frequently limit the amount of niche space that can be used by members of a lineage, for instance when the presence of a competitor or predator prevents a species from occupying some environments (Soberón 2007). This short overview suggests that limits on phenotypic evolution might be rather common in nature, leading some authors to consider the recognition of constraints as one of the main paradigm shifts in evolutionary biology (Futuyma 2010).

   In light of this, being able to detect bounds on the value of a character would be very useful. Unfortunately, the current toolbox does not allow researchers to do so. As a result, traits that have evolved between bounds are currently inferred to have evolved under an OU model, since it is the only model available in which the variance of a trait in a clade does not grow indefinitely over time and reaches stationarity (Revell et al. 2008; Boucher et al. 2014). This is quite worrying since these two models have totally different interpretations: OU would be interpreted as evidence for adaptive evolution towards an optimal value, whereas BM with two bounds would be seen as a rather neutral mode of evolution under hard constraints, which might themselves be set by selection or not.

Here we introduce a new model in which the value of the character of interest evolves under BM with two hard bounds, which we call Bounded Brownian Motion (BBM). In the first section, we detail the general behavior of this model. We then present analytical expressions for the likelihood of BBM given a phylogeny and character values observed at the tips. Finally, we present an algorithm to calculate the likelihood of BBM and assess its performance in model comparison and parameter inference. Our results show that BBM can be easily discriminated from BM, and more importantly from OU, using likelihood. This should make BBM an essential tool for identifying the presence of hard bounds on the evolution of phenotypic traits.

ANALYTICAL TREATMENT OF THE BBM MODEL

We introduce a new model for the evolution of continuous characters that is a simple extension of the classic constant-rate Brownian Motion (BM). Under BM, the value of the character, x, is governed by the following stochastic differential equation: $dx(t) \sim N(0, \sigma^2 dt)$, where σ^2 is the evolutionary rate (Fig. 1a). It follows that the variance of the character grows linearly with time under BM. To model the evolution of a continuous character in bounded space, we assume that the character evolves according to BM, except that there are two reflective bounds (Fig. 1b).

Calculation of the Probability Density

In this section, we consider a continuous trait x starting at $x(t=0)=x_0$ and derive its probability density after a time t has elapsed, $p(x, x_0, t)$. In the remainder of

this article, we call this probability density function the *propagator* to reflect the fact that it propagates the probability density of the trait along one branch of the tree (see below). Without bounds, the probability density of a BM process starting at x_0 is given at time t by a Gaussian with mean x_0 and variance $\sigma^2 t$:

$$p(x, x_0, t) = \frac{1}{\sqrt{2\pi t}\sigma} \exp\left(-\frac{(x-x_0)^2}{2\sigma^2 t}\right). \qquad (1)$$

If only one reflective bound is placed at $x = 0$, and $x_0 > 0$, the probability density for $x > 0$ can be obtained by the method of images (see Jackson 1998, Chap. 2.1). It effectively produces a mirror image Gaussian centered at $-x_0$:

$$p(x, x_0, t) = \frac{1}{\sqrt{2\pi t}\sigma}\left[\exp\left(-\frac{(x-x_0)^2}{2\sigma^2 t}\right) + \exp\left(-\frac{(x+x_0)^2}{2\sigma^2 t}\right)\right], \qquad (2)$$

where we only consider values of $x > 0$. This is equivalent to saying that the part $x < 0$ of the probability density given in Equation (1) is cut out, horizontally reflected, and added to the part $x > 0$ (see Fig. 2). To prove that the probability density is indeed given by Equation (2), it is enough to check that this expression satisfies the diffusion equation with the appropriate boundary conditions.

The case with two bounds at $x = a$ and $x = b$ with $a \le x_0 \le b$ requires the introduction of an infinite set of mirror images, because the set of images (including the original Gaussian) must be symmetric with respect to the two bounds. This leads to the following expression

$$p(x, x_0, t) = \frac{1}{\sqrt{2\pi t}\sigma} \times \sum_{k=-\infty}^{\infty}\left[\exp\left(-\frac{[x-x_0-2k(b-a)]^2}{2\sigma^2 t}\right)\right.$$
$$\left. + \exp\left(-\frac{[x+x_0-2a-2k(b-a)]^2}{2\sigma^2 t}\right)\right], \qquad (3)$$

where we only consider values of x in the interval $[a, b]$. This sum is equivalent to successive cuts and reflections of the Gaussian probability density (1), in a similar fashion as shown for one bound on Figure 2. k represents half the number of reflections and the two terms in the sum are for the even and odd numbers of reflections, respectively. We then check that Equation (3) gives the probability density of BBM (Appendix I, available on Dryad at http://dx.doi.org/10.5061/dryad.k974s).

Properties of the Probability Density

The probability density (3), which is written as an infinite sum of Gaussian functions, has simple properties:

- Over short time intervals, that is $t \to 0$, it is close to a Gaussian since the trait most likely did not
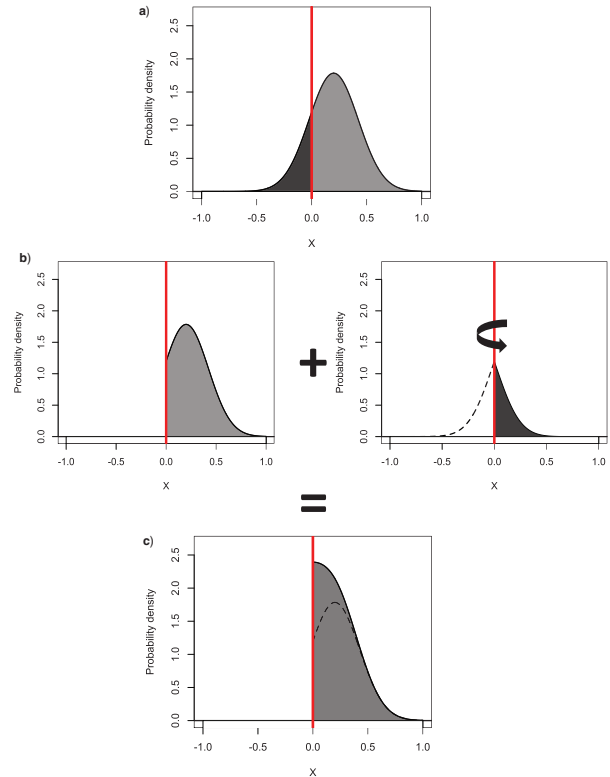

FIGURE 2.   Illustration of the method of images for calculating the probability density of a BM model with one bound: the bound is placed at 0 (vertical line) and the trait x can only take positive values. a) Probability density of a BM model without bounds starting from $x_0 = 0.2$, curves for positive and negative values of x are shown in different shades of grey. b) The probability density of a BM model without bounds is cut to consider only positive values of x (left); the probability density of BM for negative values (dashed line) is 'folded' to positive values (dark grey area). c) The probability density of BM with one bound is obtained by summing the two probability densities obtained in B. The intuitive way to understand the method of images is to imagine that all traits that would have crossed 0 in BM (dashed line in B-right) have bounced back in the interval (dark grey area in B-right). The probability of BBM is obtained in the same manner except that we have to 'fold' the density an infinite number of times at the two bounds since x might cross the interval an infinite number of times.

hit the bounds yet (Fig. 3): the terms coming from the images, which lie outside the interval, become infinitely small. The process still behaves qualitatively as BM without bounds. This is normal since BM is a special case of BBM where the bounds tend to infinity.

- Over long time intervals, that is $t \to \infty$, the density becomes uniform because the bounds have most likely been hit multiple times (Fig. 3). It can be checked that the uniform distribution on the interval $[a, b]$ is the unique stationary distribution for the diffusion equation.

The crossover between behaviors over "short" and "long" time intervals occurs at a characteristic time τ, which is the typical time that it takes for the trait to cross the interval between the two bounds a and b. Because the typical displacement of the trait over a time t is $\sigma\sqrt{t}$, the
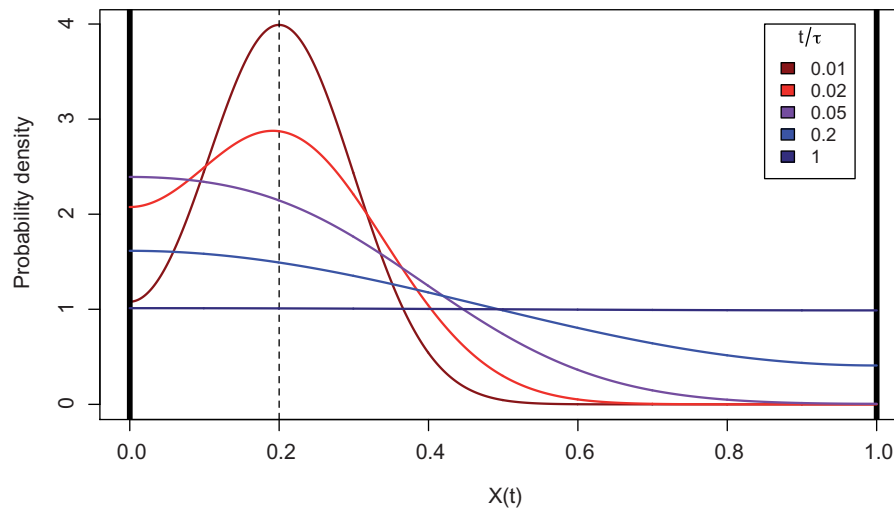
FIGURE 3.   Probability density of the BBM model. The probability density of a trait X evolving under BBM over the interval $[0,1]$ is shown for different times since the initiation of the process. The dashed line shows the initial value of the trait, and σ was set to 0.1 units. As time increases, the probability density flattens and converges to a uniform distribution, reflecting the fact that the value of X cannot be predicted anymore when bounds have been hit multiple times.

characteristic time τ is given by

$$\tau = \frac{(b-a)^2}{\sigma^2}. \qquad (4)$$

Over short time intervals $t \ll \tau$, the trait does not have time to hit the bounds, and over long time intervals $t \gg \tau$, the trait crosses the whole interval many times and the probability density reaches its stationary form.

This behavior of the BBM process on one branch of the tree extends to the whole tree: BBM produces a distribution of trait values at the tips of a phylogeny which flattens as total tree depth increases (Appendix II available on Dryad).

LIKELIHOOD ESTIMATION AND PARAMETER INFERENCE

Computation of the Likelihood

We want to compute the likelihood of observing values for the trait at the tips of a tree for two given bounds a and b, an evolutionary rate σ, and an initial value of the trait at the root x_0. The likelihood is given by

$$\mathcal{L} = \int_a^b \left(\prod_{i \in I \cup T} p(x_i, x_{parent(i)}, t_i - t_{parent(i)}) \right) \prod_{i \in I} dx_i, \qquad (5)$$

where I is the set of internal nodes (excluding the root), T is the set of tips, x_i is the value of the trait at the node i, t_i is the time at node i and $parent(i)$ is the parent of the node i. This expression is the product of the propagators, p, on all the branches of the tree, integrated over the values of the trait at the internal nodes.

Computing the likelihood numerically is challenging. Indeed, this expression involves integrating over all possible values of the trait at internal nodes of the tree, but computing the value of the integrand for all values at the internal nodes taken in a discrete set results in a computational time that is exponential in the number of branches. The problem is the same for other evolutionary models like BM, OU, or their various extensions, but in these cases numerically integrating over all possible ancestral values can be avoided by inverting the variance–covariance matrix produced by the model at the tips of the tree (Generalized Least-Squares method, Grafen 1989) or using even faster algorithms like phylogenetic independent contrasts (PICs, Felsenstein 1973; Freckleton 2012), which saves substantial computing time. However, these techniques can only be applied to traits that have a multivariate normal distribution. As we have seen above, this is obviously not the case for BBM given that trait space is bounded. Below we present an algorithm for computing the likelihood in which computational time is linear in the number of branches.

Algorithm.—We organize the different terms in the integral (5). First, we notice that we can integrate over the value of the trait at all the nodes, including the tips, if we introduce for each tip i the weight function $W^i(x_i) = \delta(x_i - y_i)$, where δ is the Dirac distribution and y_i is the actual value of the process at the tip i. This weight function exactly represents the probability density of the trait at the tip. The Dirac distribution is zero everywhere except when $x_i = y_i$ and has an integral of one: it formalizes the fact that we know with certainty the value of the trait at the tips of the tree.

As an example, we write it explicitly for a subtree consisting of one parent node with trait value x_0 and two descendant nodes with observed trait values x_1 and

x_2, sustained by two branches of lengths t_1 and t_2:

$$\mathcal{L} = p(x_1, x_0, t_1)p(x_2, x_0, t_2) \tag{6}$$

$$= \int_a^b W^1(x_1')W^2(x_2')p(x_1', x_0, t)p(x_2', x_0, t_2)dx_1'dx_2' \tag{7}$$

$$= \left[\int_a^b p(x_0, x_1', t_2)W^1(x_1')dx_1' \right]$$

$$\times \left[\int_a^b p(x_0, x_2', t_2)W^2(x_2')dx_2' \right], \tag{8}$$

where $W^i(x_i') = \delta(x_i' - x_i)$, and we have used the symmetry of the propagator, $p(x, x_0, t) = p(x_0, x, t)$. This expression has a simple interpretation: the weight function of each node is "propagated" to its parent via the BBM process. One can then define the weight function of the parent node by combining the weights propagated along the two descendant branches

$$W^0(x_0) = \left[\int_a^b p(x_0, x_1', t)W^1(x_1')dx_1' \right]$$

$$\times \left[\int_a^b p(x_0, x_2', t)W^2(x_2')dx_2' \right]. \tag{9}$$

We then find the root value that maximizes the weight function and obtain the likelihood

$$\max_{x_0}[\mathcal{L}] = \max_{x_0}\left[W^0(x_0) \right], \tag{10}$$

which should be maximized over σ, a, and b. This procedure generalizes to any tree: starting from the tips, the weight functions are propagated down the branches following the pruning algorithm (Felsenstein 1973) and multiplied at the internal nodes down to the root. The likelihood is finally obtained by maximizing the weight function of the root.

Implementation.—To implement the algorithm, we discretize the value of the trait at the nodes by considering only a set of n points equally spaced between a and b. We introduce a transition matrix specifying the probability for the trait to evolve from one of these n values to another during an infinitesimal time step according to the BBM model. Our algorithm is thus similar to the one commonly used for fitting models of evolution for discrete traits (Lewis 2001), see details in Appendix I available on Dryad). In the rest of the article, we focus on BBM but Appendix III available on Dryad shows how to calculate the likelihood of the most general model possible, allowing any kind of force (i.e., selection of any shape and strength) to be exerted on trait values in addition to drift between bounds.

Likelihood calculations are done for given values of a, b, and σ, but we still need to optimize over these three parameters to obtain the global likelihood of the BBM model. Although there is no rigorous proof that the minimum and maximum of the trait values at the tips are the ML estimators of a and b, there are strong arguments suggesting that it might be the case. First, the minimum and maximum of a sample of uniformly distributed values are the ML estimators of the bounds of an uniform distribution. Second, we verified numerically that the minimum and maximum of the trait values at the tips are indeed the ML estimators of a and b, using simulations over a wide range of parameters (Appendix IV available on Dryad). Moreover, fixing a and b at the extremes of the observed trait distribution enables optimizing over σ only, which leads on average to a 7-fold reduction in computational time but also reduces optimization errors (Appendix IV available on Dryad). In the rest of this article, we thus fit the BBM model by fixing the values of a and b to the minimum and maximum of the trait values observed at the tips of the tree. The most time consuming part of this algorithm is to apply the propagator to weight vectors, which reduces to matrix multiplications. The number of such multiplications is proportional to the number of branches, so that the computational time is linear in the number of branches.

Prior to all analyses done below, we ran preliminary tests to determine which level of discretization is needed to calculate the likelihood of BBM with a reasonable precision. We found that increasing the number of points used for discretizing the interval of trait values (i.e. $[a,b]$) leads to a higher precision in the estimation of the likelihood (Appendix V available on Dryad); on the other hand, the computing time grows quadratically with the number of points used in the discretization (because of the matrix products). Based on these observations, we decided to use a discretization of 100 points in all following analysis since this generally gives a value closer than 0.2 log-likelihood points to the most precise value that we obtained (with 200 points), but we recommend that empiricists working on only one or a few data sets use discretizations of at least 200 points.

Using traits simulated under a BM process with no bounds, we verified that the log-likelihood of the BBM model calculated using our algorithm was generally closer than 0.5 points to the log-likelihood of the BM model calculated by the *fitContinuous* function in the *geiger* package (Pennell et al. 2014, see Appendix IV available on Dryad). Thus, our implementation of BBM is compatible with all models implemented in *geiger*.

R code.—R code to fit the BBM model to empirical data can be freely downloaded from https://github.com/fcboucher/BBM, last accessed February 23, 2016. The code can either estimate the values of a and b along with other parameters or fix them to the minimum and maximum of the trait values observed at the tips of the tree. This later option yields an important reduction in computational time and an increased accuracy of parameter estimation (see Appendix IV available on Dryad). The likelihood function cannot be approximated by a Gaussian around the ML parameter estimates since it is one-sided around the ML estimates of the bounds and often highly

asymmetric around the ML estimate of σ^2, hence extracting standard errors using the Hessian matrix of the system is not possible. Instead, we report confidence intervals that contain the 95% highest probability density around parameter estimates while fixing other parameters to their maximum likelihood estimate. This is technically done by removing the lowest 2.5% density regions on each side of the MLE for sigma and the root value when its MLE does not lie in one of the bounds of the trait interval and removing the lowest 5% density region for the bounds and the root value when its MLE lies in one of the bounds of the trait interval. These confidence intervals can be returned along with likelihood profile plots around parameter estimates. The code only depends on the *ape* package (Paradis et al. 2004).

Accuracy of Parameter Estimation

In order to measure the accuracy with which parameters of BBM are estimated, we used data sets simulated under BBM and inferred parameters using our algorithm. We simulated phylogenetic trees of varying sizes (15, 50, 100, and 200 tips) under a pure birth model using the *geiger* package. All trees were rescaled to a total arbitrary depth of $T_{tot} = 100$ time units. Continuous traits evolving under BBM on the trees were simulated with bounds located at -5 and $+5$ so that total tree depth divided by the characteristic time of the process, τ, was $T_{tot}/\tau = \sigma^2$. The ancestral value was uniformly drawn from the interval and the value of σ was varied in order to explore situations where T_{tot}/τ equalled $1E-6$, $1E-5$, $1E-4$, $1E-3$, $1E-2$, $1E-1$, 1, 10 and 100, thus ranging from BM with no bounds ($T_{tot}/\tau \ll 1$) to BBM with many reflections ($T_{tot}/\tau \gg 1$). For each combination of a value of T_{tot}/τ and a tree size, 100 simulations were performed.

Simulations showed that the root value, x_0, can only be estimated accurately when $T_{tot}/\tau \ll 1$. This corresponds to cases in which the trait did not hit the bounds and thus behaves as BM. However, when T_{tot}/τ increases, information is rapidly lost and x_0 is often estimated to lie at the middle of the interval regardless of its actual value, leading to very low precision (Fig. 4).

In contrast, simulations showed that σ can be estimated with high accuracy for a wide range of values of T_{tot}/τ, even on relatively small trees (Fig. 4). The only cases where σ is rather poorly estimated correspond to simulations with $T_{tot}/\tau \gg 1$. This is not a serious issue since in this case the estimated value of σ essentially inform us that the trait interval has been crossed several times, the exact number of times this has happened being of minor biological importance.

Finally, simulations showed that the bounds of the trait interval are poorly estimated when $T_{tot}/\tau \ll 1$ (Fig. 4c–d). This is again because the trait did not hit the bounds and the process effectively is BM with no bounds: in these cases, both bounds are estimated to be closer than they really are, that is at the minimum and maximum of the trait values observed at the tips of the tree. When $T_{tot}/\tau \geq 0.1$ the bounds have most often been hit and a and b are then estimated with high accuracy. The quality of the estimation dramatically increases with the number of tips in the tree because as the number of extant species increases, it is more and more likely that the traits of some species are very close to the actual bounds.

Comparison to Other Models of Character Evolution

Simulation study.—In order to see whether BBM can be discriminated from other simple models like BM and OU with a single optimum (hereafter OU1), we compared the likelihoods of BBM, BM, and OU1 on traits simulated under a BBM model. To do so, we used the same simulations of BBM as above, that is trees of 15, 50, 100, and 200 tips and values of T_{tot}/τ of $1E-6$, $1E-5$, $1E-4$, $1E-3$, $1E-2$, $1E-1$, 1, 10 and 100 (100 repetitions for each combination).

Simulations showed that in situations where the trait did not hit the bounds of the interval (i.e., $T_{tot}/\tau \ll 1$) BBM has a likelihood close to that of BM and OU1 (Fig. 5a and b). This is expected since BM is a special case of both BBM and OU1: in this case, the three models should have the same likelihood, that is the one of BM, and the Akaike Information Criterion (AIC) would identify BM as the best-fitting model given that it has less parameters, which is satisfactory since the trait evolved under BM in practice. However, as T_{tot}/τ increases BBM gets more and more likely compared to both BM and OU1. As soon as $T_{tot}/\tau > 0.1$, BBM was most often at least twice as likely as OU1 and four times as likely as BM (Fig. 5a and b), which means that BBM would be selected as the best-fit model using AIC (BBM has two more parameters than BM and one more than the implementation of OU1 in *geiger*, in which the root value and the optimum are equated). In simulations where $T_{tot}/\tau > 1$, BBM was often at least 10 times as likely as OU1 and 20 times as likely as BM. The fact that OU1 is a much more likely model than BM on simulations with $T_{tot}/\tau \geq 1$ comes from the fact that it has a stationary variance, like BBM but not BM. Tree size had a strong effect on model discrimination, with larger trees giving more statistical power to discriminate BBM from other models when trait evolution was actually bounded. However, even on trees with only 15 tips, BBM most often had a higher AIC than both BM and OU1 as soon as $T_{tot}/\tau \geq 0.1$ (Fig. 5a and b).

Conversely, we also compared the fit of BBM with traits simulated under an OU1 model. We simulated traits evolving under OU1 on trees of 15, 50, 100, and 200 tips, always using $\sigma = 1$ and an optimum value $\theta = 0$. Values of the attraction strength, α, equalled $1E-4$, $1E-3$, $1E-2$, $1E-1$, and 1, which correspond to phylogenetic half-lives ranging from 0.69 to 6900 time units (the total tree depth was again arbitrarily fixed at 100 time units). The value of the trait at the root of the tree, x_0, was randomly drawn from the stationary distribution of the OU1 process, that is a Gaussian with mean θ and variance $\sigma^2/(2\alpha)$.
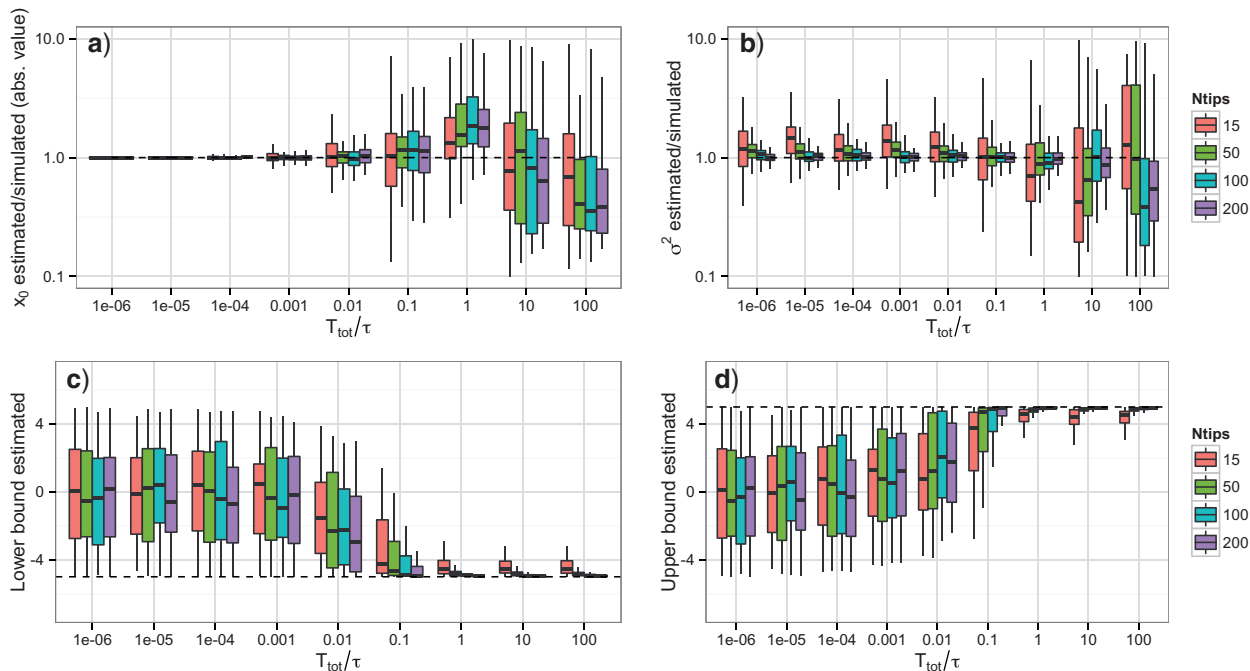
FIGURE 4. Parameter estimation in BBM. a) Estimation of the root value as a function of the expected number of crossings of the trait interval (T_{tot}/τ) for trees of different sizes. Boxplots for the three smallest values of T_{tot}/τ can hardly be seen since in these cases the estimation was very accurate. b) Estimation of the evolutionary rate, σ^2, as a function of the expected number of crossings of the trait interval. c) Estimation of the lower bound of the trait interval, always set to -5 in simulations. d) Estimation of the upper bound of the trait interval, always set to $+5$ in simulations. Dashed lines shows unbiased estimates.
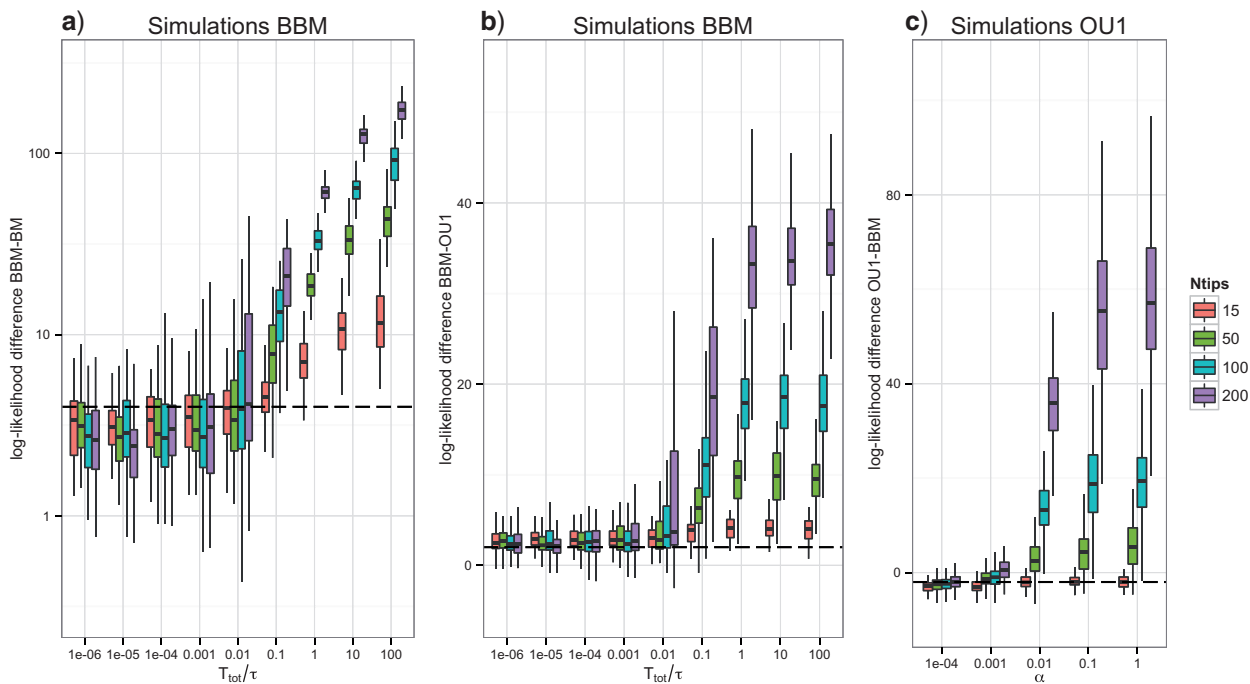

FIGURE 5. Model comparison. The figure shows how BBM can be discriminated from BM and OU1 using likelihood. a) Simulations of the BBM model with increasing expected number of crossings of the interval, T_{tot}/τ. The y-axis shows log-likelihood(BBM)$-$log-likelihood(BM) on a logarithmic scale. The dashed line shows a difference of $+4$: above this line, BBM would be selected using AIC since it has two more parameters than BM. b) Same simulations, but the y-axis shows log-likelihood(BBM)$-$log-likelihood(OU1). The dashed line shows a difference of $+2$: above this line, BBM would be selected using AIC since it has one parameter more than the implementation of OU1 in *geiger*. c) Simulations of OU1 with increasing values of α. The y-axis shows log-likelihood(OU1)$-$log-likelihood(BBM). The dashed line shows a difference of -2: above this line, OU1 would be selected using AIC.

Simulations showed that BBM and OU1 have very similar AICs for values of $\alpha \leq 0.001$ (Fig. 5c). This is normal since in these situations the actual process is very similar to BM. However, for values of $\alpha \geq 0.01$, OU1 most often had a higher AIC than BBM (Fig. 5c). Although tree size again had an effect on model discrimination, OU1 could usually be discriminated from BBM even on trees with 15 tips (Fig. 5c). Put together, these two sets of simulations show that BBM can be effectively discriminated from BM and OU1: whichever one of these three models is simulated, it is most often the one that gets the highest AIC.

Empirical example.—Finally, we investigated an empirical data set. The data set consisted of measurements of the realized climatic niches (mean annual temperature and sum of annual precipitation) for species of the marsupial order Diprotodontia (kangaroos, wallabies, wombats, koalas, and relatives). We chose this clade since diprotodonts occur all over Australia (plus several neighboring islands), hence their realized climatic niches might be bounded by the climate that is available in this region (Boucher et al. 2014; Wüest et al. 2015). This data set should ensure reasonable statistical power since it contains 110 species (out of the 120 described). The phylogeny came from Fritz et al. (2009) and realized climatic niches were measured using range maps from the IUCN realist assessment and climatic data from the Worldclim database (10 arc-minutes resolution, Hijmans et al. 2005). Each species' climatic niche was measured as the mean value of both climatic variables (temperature and precipitation) across the species' range. We then fitted BBM (using 200 points for the discretization), BM, and OU1 to both climatic niche axes.

For both niche axes, BBM had the best fit to the evolution of realized climatic niches. OU1 was always the second model with the best fit ($\Delta_{AIC} = 5.1$ for temperature and $\Delta_{AIC} = 15.5$ for precipitation), and BM came last ($\Delta_{AIC} = 63.3$ for temperature and $\Delta_{AIC} = 20.7$ for precipitation). Estimated rates of evolution were $5.18°C^2/myr$ (confidence interval: 3.67–8.91) for mean annual temperature and 44,053 mm of $rain^2/myr$ (39,647–127,753) for annual precipitation, which correspond to rather high expected numbers of crossings of the climate interval for both niche axes ($T_{tot}/\tau = 0.72$ and 0.15, respectively). Bounds on the trait intervals were estimated to be 7.4°C (3.8–7.4) and 27.2°C (27.2–30.8) for mean annual temperature, and 245 mm (125–245) and 4,195 mm (4195–4314) for annual precipitation. These estimated values correspond to the minimum and maximum climates experienced by species in Diprotodontia.

These results suggest that species' niches in this clade are likely constrained by the temperature and precipitation extremes that are present in Australia and neighboring islands, but are probably not selected toward an "optimal" climatic niche, as already suggested by Boucher et al. (2014).

DISCUSSION

The development of the BBM model was needed since none of the various models available in the toolbox of comparative analysis so far could accommodate bounds on the value of a character (O'Meara 2012), a situation that could be common in nature (Futuyma 2010). We discuss below the opportunities brought by this new model for comparative analysis, as well as its limitations.

Strengths and Weaknesses of the BBM Model

Our simulations have shown that BBM can easily be discriminated from other models using likelihood. Although we expected that BBM and BM should be easy to discriminate, this was less intuitive for BBM and OU1, since these two models have qualitatively similar behaviors. Ultimately, it is most probably the flattening of the probability density under BBM (Fig. 3) that brings the information necessary to discriminate it from OU1, as can be seen from the fact that these two alternative models produce trait distributions at the tips of a phylogeny with very different values of kurtosis (Appendix I available on Dryad).

Parameter inference under BBM has both exciting and disappointing aspects. First of all, since the probability density of BBM rapidly converges to a uniform, ancestral character estimation under BBM is extremely limited: apart from a few very recent nodes for which some information may still exist, values of the trait at internal nodes will often be estimated to lie in the middle of the interval. This is exemplified by the rather poor estimation of x_0 as soon as the bounds of the interval have been reached (Fig. 4a). Intuitively, this makes perfect sense: when the trait starts to hit the bounds repeatedly, most of the information on the trait values of ancestors is rapidly lost. This could seem like bad news but we rather see that as being positive: under BBM researchers will not be tempted to estimate and then interpret the ancestral values of their character of interest in a group, a practice that has already been shown to be associated with high estimation errors for models which retain relatively little phylogenetic information, like OU1 (Martins 2000; Ané 2008).

Although we lack a rigorous proof of it, we have argued that the ML estimators for the bounds of a BBM process should be the minimum and maximum values of the trait in a clade and ran extensive numerical simulations to verify it (Appendix IV available on Dryad). Although this saves a lot of computational time, this is rather disappointing since no situation where a trait is "getting close to a hard bound but not there yet" or where a trait "has just bounced back on the bound" can be recovered. Indeed, simulations have shown that as long as the bounds of the interval have not been reached, their estimation is really inaccurate since they are estimated to lie much closer than in reality (Fig. 4c and d). Bounds estimated using BBM should thus not be interpreted *per se* ; rather, the good fit of BBM relative to BM or OU1 indicates that the trait has most likely

reached the bounds already and that the traits of some species in the clade are probably not very far from the bounds.

Finally, we have shown that the estimation of σ^2 generally achieves good performance. However, variance in the estimation increases for high values of σ^2 and typically spans several orders of magnitude (Fig. 4b). This is of little biological concern since in these situations we have an expected number of crossings of the trait interval, T_{tot}/τ, higher than one: we essentially infer that the bounds of the interval of possible trait values have been hit, even though it is difficult to infer the evolutionary rate with precision. Although T_{tot}/τ suffers exactly from the same estimation errors than σ^2 (it is linearly related to it), we recommend that researchers interpret this parameter rather than σ^2 since we feel it has a more straightforward biological meaning.

Biological Interpretation of Limits on Trait Evolution

Importantly, BBM is not the only way to model limits on phenotypic evolution. BBM indeed assumes that bounds on the phenotypic character can be reached and are reflective. This might be the case for some characters, like realized climatic niches for example (Boucher et al. 2014), and under some circumstances. Indeed, we have shown here that the evolution of realized climatic niches in Diprotodontia is better fit by BBM than by BM or OU1. However, these assumptions might not hold for some characters. First, bounds on the character might not be reachable. This is the case for body mass: this character is actually forced to be positive but a value of zero cannot be reached. In this case, the most straightforward solution is to log-transform body mass: the new variable is not bounded anymore and can be analyzed with standard BM or OU models (e.g., Garland et al. 1992; Harmon et al. 2010). This transformation in addition has the advantage of being more biologically realistic: one would expect that doubling its body mass is as likely for any organism, whether it weights 10 g or 1000 kg, something that would be adequately modeled using BM on log-transformed body mass. The same logic can also apply to traits having both a lower and an upper bound, and some authors have used a logit transformation in this case (Scales et al. 2009). A second situation in which BBM would not be suited is when bounds on the character are absorbing. One good example of such a phenomenon is the evolution of allele frequencies: once one allele gets lost (frequency = 0) or fixed (frequency = 1) in a population its frequency does not evolve anymore, unless mutation or gene flow modify this situation. Thus, it is extremely important that researchers first evaluate the nature of the bounds their character of interest is confronted with before using BBM to model its evolution.

BBM and OU1: Two Very Different Kinds of Constraints

BBM and a single optimum OU model share an important feature: they both impose a limit on the total variance that a trait can accumulate in a clade over time. As a result, they are rather close in essence in that they describe constraints on trait evolution or phylogenetic niche conservatism (PNC, Harvey and Pagel 1991; Wiens et al. 2010). However, these two models have radically different interpretations. The OU model is specifically designed to model adaptation in the form of stabilizing selection around an optimal value (Hansen 1997). In contrast, BBM would describe a scenario in which traits can drift freely over time in phenotypic space, except that they are confined between two bounds. These bounds might be imposed by any kind of constraints, selective or otherwise. Several authors have warned against purely adaptive interpretations of the good fit of an OU1 model to a data set (Revell et al. 2008; Boucher et al. 2014), arguing that many other processes than stabilizing selection, including evolution between bounds, can produce phylogenetic patterns of traits that are very similar to those yielded by an OU1 model. Using simulations, we have shown that BBM and OU1 can indeed be discriminated based on their likelihoods: this opens the way to refining hypotheses about the processes that might have lead to PNC in a clade. Indeed, stabilizing selection might be rejected in some cases in favor of more neutral processes, leading to some kind of "artefactual" PNC, for example resulting from dispersal limitation (Crisp and Cook 2012). In other cases BBM might be rejected, giving more evidence for the role played by stabilizing selection in promoting PNC.

Conclusion and Future Directions

We believe that the development of the BBM model is an important step in comparative analysis. We acknowledge that BBM is a rather simple process, but it could be extended in several different directions. First, we have assumed BBM to be homogeneous across an entire clade. Although our simulation study has shown that parameter estimation and discrimination from alternative models is best achieved on relatively large trees (i.e., of more than 100 tips, Figs. 4 and 5), it is rather unlikely that processes act homogeneously over clades of this size (O'Meara 2012; Rabosky et al. 2013; Uyeda and Harmon 2014). Because we have provided a closed formula for the likelihood of BBM and an efficient algorithm to calculate it, extending it to a version where traits evolve in different bounded intervals for different subclades is straightforward. More importantly, we have provided in the Appendix III available on Dryad a very general formula to include a force acting on trait values in addition to diffusion between bounds. This force can be of any shape, like a directional trend toward one of the bounds, attraction towards one central value (i.e., an Ornstein–Uhlenbeck process with bounds), attraction toward several intermediate values, repulsion from one or several values, or even a combination of these. This opens the way to identifying complex scenarios of evolution between bounds, for example using a MCMC framework.

REFERENCES

Ané C. 2008. Analysis of comparative data with hierarchical autocorrelation. Ann. Appl. Stat. 2:1078–1102.

Beaulieu J.M., Jhwueng D.-c., Boettiger C., O'Meara B.C. 2012. Modeling stabilizing selection : expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution 66:2369–2383.

Beaulieu J.M., Moles A.T., Leitch I.J., Bennett M.D., Dickie J.B., Knight C.A. 2007. Correlated evolution of genome size and seed mass. New Phytol. 173:422–437.

Bokma F. 2008. Detection of "punctuated equilibrium" by Bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. Evolution 62:2718–2726.

Boucher F.C., Thuiller W., Davies T.J., Lavergne S. 2014. Neutral biogeography and the evolution of climatic niches. Am. Nat. 183:573–584.

Butler M.A., King A.A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. Evolution 164: 683–695.

Crisp M.D., Cook L.G. 2012. Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? New Phytologist 196:681–694.

Donovan L.A., Maherali H., Caruso C.M., Huber H., Kroon H.D. 2011. The evolution of the worldwide leaf economics spectrum. Trends Ecol. Evol. 26:88–95.

Edwards A.W.F., Cavalli-Sforza L.L. 1964. Reconstruction of evolutionary trees. In: Phenetic and phylogenetic classification Heywood V. H., McNeill J., editors. London: Systematics Association, p. 67–76.

Edwards E.J., S.A. Smith. 2010. Phylogenetic analyses reveal the shady history of C4 grasses. Proc. Natl Acad. Sci. USA 107:2532–2537.

Felsenstein J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Hum. Genet. 25:471–492.

Freckleton R.P. 2012. Fast likelihood calculations for comparative analyses. Methods Ecol. Evol. 3:940–947.

Fritz S.a., Bininda-Emonds O.R.P., Purvis A. 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. Ecol. Lett. 12:538–549.

Futuyma D.J. 2010. Evolutionary constraint and ecological consequences. Evolution 64:1865–1884.

Garamszegi L.Z.E., ed. 2014. Modern phylogenetic comparative methods and their application in evolutionary biology - concepts and practice. New York, NY: Springer.

Garland T.J., Harvey P.H., Ives A.R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst. Biol. 41:18–32.

Garland T., Dickerman A.W., Janis C.M., Jones J.A. 1993. Phylogenetic analysis of covariance by computer simulation. Syst. Biol. 42: 265–292.

Grafen A. 1989. The phylogenetic regression. Philo. Trans. R. Soc. London. Ser. B Biol. Sci. 326:119–157.

Guerrero P.C., Rosas M., Arroyo M.T.K., Wiens J.J. 2013. Evolutionary lag times and recent origin of the biota of an ancient desert (Atacama-Sechura). Proc. Natl Acad. Sci. USA 110:11469–11474.

Hansen T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Hansen T.F., Martins E.P. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50:1404–1417.

Harmon L.J., Losos J.B., Jonathan Davies T., Gillespie R.G., Gittleman J.L., Bryan Jennings W., Kozak K.H., McPeek M.a., Moreno-Roark F., Near T.J., Purvis A., Ricklefs R.E., Schluter D., Schulte Ii J.a., Seehausen O., Sidlauskas B.L., Torres-Carvajal O., Weir J.T., Mooers A.O. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64:2385–2396.

Harvey P.H., Pagel M. 1991. The comparative method in evolutionary biology. Oxford: Oxford University Press.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25:1965–1978.

Ho L.S.T., Ané C. 2014. A linear-time algorithm for gaussian and non-gaussian trait evolution models. Syst. Biol. 63:397–408.

Hutchinson G.E. 1957. Concluding remarks. Cold Spring Harb. Symp. Quant. Biol. 22:145–159.

Ingram T., Mahler D.L. 2013. SURFACE : detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. Methods Ecol. Evol. 4: 416–425.

Jackson J.D. 1998. Classical electrodynamics. 3rd ed. New York: John Wiley & Sons.

Kaiser A., Klok C.J., Socha J.J., Lee W.-K., Quinlan M.C. and Harrison J.F. 2007. Increase in tracheal investment with beetle size supports hypothesis of oxygen limitation on insect gigantism. Proc. Natl Acad. Sci. USA 104:13198–13203.

Koch G.W., Sillett S.C., Jennings G.M., Davis S.D. 2004. The limits to tree height. Nature 428:852–854.

Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

Landis M.J., Schraiber J.G., Liang M. 2013. Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. Syst. Biol. 62:193–204.

Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.

Martins E.P. 2000. Adaptation and the comparative method. Trends Ecol. Evol. 15:296–299.

Mooers A.O., Holmes E.C. 2000. The evolution of base composition and phylogenetic inference. Trends Ecol. Evol. 15:365–369.

Nei M. 2013. Mutation-driven evolution. Oxford: Oxford University Press.

O'Meara B.C. 2012. Evolutionary inferences from phylogenies: a review of methods. Ann. Rev. Ecol. Evol. Syst. 43:267–285.

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

Paradis E., Claude J., Strimmer K. 2004. Ape: analyses of phylogenetics and evolution in r language. Bioinformatics 20:289–290.

Pennell M.W. 2015. Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice. — edited by László Zsolt Garamszegi. Syst. Biol. 64:161–163.

Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., Harmon L.J. 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics 30:2216–2218.

Price S.A., Holzman R., Near T.J., Wainwright P.C. 2011. Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. Ecol. Lett. 14:462–469.

Rabosky D.L., Santini F., Eastman J., Smith S.A., Sidlauskas B., Chang J., Alfaro M.E. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4:1–8.

Revell L.J., Harmon L.J., Collar D.C. 2008. Phylogenetic signal, evolutionary process, and rate. Syst. Biol. 57:591–601.

Scales J.A., King A.A., Butler M.A. 2009. Running for your life or running for your dinner: what drives fiber-type evolution in lizard locomotor muscles? Am. Nat. 173:543–553.

Soberón J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecol. Lett. 10:1115–23.

Uyeda J.C., Harmon L.J. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. 63:902–918.

Wiens J.J., Ackerly D.D., Allen A.P., Anacker B.L., Buckley L.B., Cornell H.V., Damschen E.I., Jonathan Davies T., Grytnes J.A., Harrison S.P., Hawkins B.a., Holt R.D., McCain C.M., Stephens P.R. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecol. Lett. 13:1310–1324.

Wüest R.O., Antonelli A., Zimmermann N.E., Linder H.P. 2015. Available climate regimes drive niche diversification during range expansion. Am. Nat. 185:640–652.