doi:10.1111/evo.12628

Placing cryptic, recently extinct, or hypothesized taxa into an ultrametric phylogeny using continuous character data: A case study with the lizard *Anolis roosevelti*

Liam J. Revell,[1,2] D. Luke Mahler,[3] R. Graham Reynolds,[1,4] and Graham J. Slater[5]

[1]*Department of Biology, University of Massachusetts Boston, Boston, Massachusetts 02125*

[2]*E-mail: liam.revell@umb.edu*

[3]*Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045*

[4]*Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138*

[5]*Department of Paleobiology, Smithsonian Institution, Washington, DC 20013*

Received August 17, 2014
Accepted December 16, 2014

In recent years, enormous effort and investment has been put into assembling the tree of life: a phylogenetic history for all species on Earth. Overwhelmingly, this progress toward building an ever increasingly complete phylogeny of living things has been accomplished through sophisticated analysis of molecular data. In the modern genomic age, molecular genetic data have become very easy and inexpensive to obtain for many species. However, some lineages are poorly represented in or absent from tissue collections, or are unavailable for molecular analysis for other reasons such as restrictive biological sample export laws. Other species went extinct recently and are only available in formalin museum preparations or perhaps even as subfossils. In this brief communication we present a new method for placing cryptic, recently extinct, or hypothesized taxa into an ultrametric phylogeny of extant taxa using continuous character data. This method is based on a relatively simple modification of an established maximum likelihood (ML) method for phylogeny inference from continuous traits. We show that the method works well on simulated trees and data. We then apply it to the case of placing the Culebra Island Giant Anole (*Anolis roosevelti*) into a phylogeny of Caribbean anoles. *Anolis roosevelti* is a "crown-giant" ecomorph anole hypothesized to have once been found throughout the Spanish, United States, and British Virgin Islands, but that has not been encountered or collected since the 1930s. Although this species is widely thought to be closely related to the Puerto Rican giant anole, *A. cuvieri,* our ML method actually places *A. roosevelti* in a different part of the tree and closely related to a clade of morphologically similar species. We are unable, however, to reject a phylogenetic position for *A. roosevelti* that places it as sister taxon to *A. cuvieri*; although close relationship with the remainder of Puerto Rican anole species is strongly rejected by our method.

KEY WORDS: Fossils, phylogenetic inference, quantitative characters, time-calibrated tree.

In recent years scientists have made incredible strides toward assembling a comprehensive tree of life. Increasingly, our picture of phylogenetic history is based on molecular data, often from DNA sequences (Hillis et al. 1996; Delsuc et al. 2005). In our modern genomics age, genetic data are more easily and cheaply obtained than ever before. Consequently, we should expect that gaps in the reconstructed tree of life will be increasingly due to species for which molecular data are simply not possible to obtain.

Figure 1. Museum of Comparative Zoology holotype specimen of *Anolis roosevelti*, the Culebra Giant Anole, collected on April 22, 1931, from the Spanish Virgin Island of Culebra, located approximately 17 miles east of the main island of Puerto Rico. Photos courtesy of the Museum of Comparative Zoology, Harvard University, © President and Fellows of Harvard College.

These species could be recently extinct taxa, such as the Columbian Mammoth (*Mammuthus columbi*), for which only minimal ancient DNA data are available (Enk et al. 2011; Gold et al. 2014); or taxa most recently collected in the premolecular era during which the use of DNA damaging fixatives, such as formalin, was commonplace (Srinivasan et al. 2002; Tang 2006). *Anolis roosevelti*, a poorly known Puerto Rican bank island "giant" anole last collected in 1932, is an example of the latter (Fig. 1; Mayer 1989; Ojeda Kessler 2010).

In some cases, such as that of *A. roosevelti*, though DNA characters are unavailable, the preserved type specimen is available and in good condition to study, making morphological measures relatively straightforward to obtain. Before the widespread availability (and near universal adoption) of nucleotide data for phylogeny inference, phylogenetic biologists developed a range of approaches for the inference and analysis of phylogenies using morphological characteristics, typically discrete traits with two or more states. In this brief communication we adapt an approach of Felsenstein (1973, 1981), designed originally to estimate phylogeny from continuous characters, for use in an effort to place recently extinct, cryptic, or hypothesized lineages into an ultrametric molecular phylogenetic base tree using continuously valued phenotypic traits. We assume an ultrametric tree with branch lengths proportional to time because our approach is designed for conditions in which our base tree and the taxon to be added are all contemporaneous and extant, or very recently extinct, taxa. However, our method would also be appropriate for a base tree with edge lengths proportional to time in which some tips were genuinely noncontemporaneous. (For instance, if the base tree contained one or more anciently extinct lineages in whose phylogenetic placement and age we were very confident.) The approach we take is also closely related to one proposed by Felsenstein (2002) for placing fossil species in a molecular phylogeny of extant taxa using a set of quantitative characters and their inferred evolutionary covariances from a base phylogeny obtained from molecular data. We show that the method works quite well in theory, and we apply our approach to the interesting problem of placing *A. roosevelti* on the phylogenetic tree of Greater Antillean *Anolis* lizards.

The Method

The problem that this method attempts to address is as follows. We, the investigators, have a time-calibrated, ultrametric (because it consists of contemporaneous extant taxa), molecular phylogeny for $N - 1$ species. For all of these $N - 1$ species, we also have a univariate or multivariate continuous character dataset—probably obtained from external or skeletal measures made on museum or field-collected specimens. These measures could include overall body size, limb proportions, long bone lengths, etc., and may also include count data, such as vertebral number, tooth counts, scale counts, etc., that are meristic but distributed on such a range as to effectively vary on a continuous scale among species. In addition we have the same phenotypic data for an Nth species, not present in our molecular phylogenetic base tree, but whose placement in the tree is of interest. Our leaf to be added must be an evolutionarily distinct lineage, although it technically need not be a nominally distinctive species as long as we are comfortable with the accompanying implicit assumption that the same evolutionary process operates to create phenotypic divergence between lineages within a species and between deeply divergent taxa in our tree.

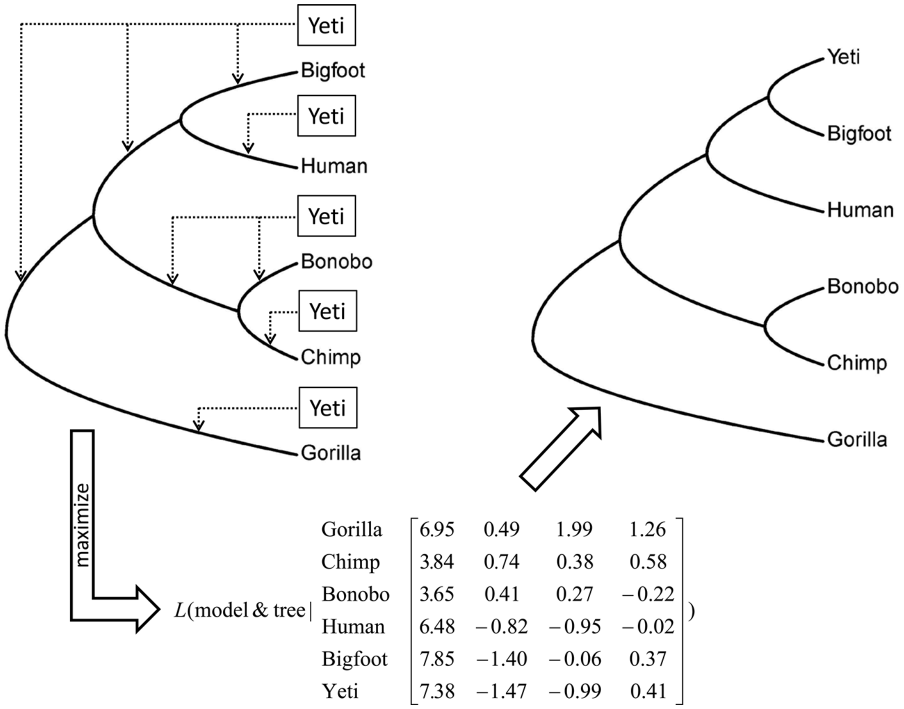

Figure 2. Illustrative figure showing the placement of a cryptic taxon on a hypothetical anthropoid phylogeny by our ML method. Data are simulated and the base tree is (obviously) fictional; however, the result is genuine and virtually identical to the generating tree in this case. The choice of taxa for this simulation is a reference to the phytools function (`locate.yeti`) where this method is implemented—the name of which is a whimsical allusion to the theoretically possible use of this method to place cryptozoological taxa into a base phylogeny.

We assume that the characters in our dataset have evolved on the tree and that most or all of the differences between species means are due to inherited genetic differences between species; however, the traits themselves need not have high heritability within species. For traits with low heritability, the differences between species may still be genetically based (though in these cases we may need more samples from each species to estimate the species mean with minimal error).

The *N*th species that we will add to the tree might be, for example, a cryptic or recently extinct lineage for which we can obtain morphological characteristics but not DNA. This lineage of interest could range from a species that is "recently" extinct only when considered on a geological time scale (say, within one to a few tens of thousands of years before the present), but for which subfossil material is available; to a species that went extinct during the "formalin era" of specimen collection, which may mean that nucleotide data are very difficult or prohibitively expensive to obtain; to a species that is not actually extinct, but that is difficult to sample genetically due to its extreme rarity or restrictive biological sample export laws (Roberts and Solow 2008; Renner et al. 2012; Rocha et al. 2014).

Our objective is to take our continuous character data, consisting of values for one or multiple traits for *N* species, and use them to identify the optimal location to attach a single leaf of unknown position to an $N-1$ species ultrametric molecular phylogenetic base tree (normally this would be a time-calibrated tree of extant species). We do this using the criterion of maximum likelihood (ML).

Based on Felsenstein (1973, 1981), the ML position using a *single* continuous character is the one in which the resultant tree, τ_N, maximizes the following log-likelihood equation:

$$L(\tau_N, a, \sigma^2 | \tau_{N-1}, \mathbf{x}) = -(\mathbf{x} - \mathbf{1}a)'(\sigma^2 \mathbf{C}_{\tau_N})(\mathbf{x} - \mathbf{1}a)/2 - \log(2\pi)/2 - \log(|\sigma^2 \mathbf{C}_{\tau_N}|)/2 \quad (1)$$

in which $\mathbf{C}_{\tau_N}$ denotes the implied among-species covariance matrix for the *N* species in any specific, hypothesized tree (e.g., see Revell 2008). Here, a indicates the state at the global root of the tree and σ^2 denotes the instantaneous diffusion rate of our Brownian model of evolutionary change. **1** is a conformable vector of ones. Both a and σ^2 have analytic ML solutions given any particular τ_N (e.g., O'Meara et al. 2006). To find the tree, τ_N, that maximizes the likelihood given our base tree (τ_{N-1}) and data (**x**), we can simply attempt to attach the new leaf in many different positions on our base tree and choose the position that maximizes equation (1). This is illustrated schematically in Figure 2. In reality, of course, we use slightly more sophisticated optimization routines, but the truth is that this optimization problem is not particularly difficult nor the calculations especially complicated, so it is fairly straightforward to find the tree τ_N and parameter estimates that maximize equation (1).

Equation (1) gives the likelihood expression for a single character. For multivariable data, our covariance matrix becomes the Kronecker (outer) product of the among-trait and among-species covariance matrices—a matrix of dimension $m \cdot N \times m \cdot N$ for m continuous traits (Felsenstein 2002; Revell and Harmon 2008). For trees of reasonable size and for datasets comprising more than a small numbers of traits, this matrix will become very large and computationally challenging to manage (and, in particular, to decompose or invert). A (barely approximate) solution for the problem of merely attaching a single leaf to an ultrametric base tree is to phylogenetically orthogonalize the m characters in our multivariable dataset before our analysis, compute the log likelihoods separately for each evolutionarily orthogonalized character using equation (1), and then add the results. This can be accomplished by performing phylogenetic principal components analysis (Revell 2009) using our $N-1$ species base tree, and then using the loadings from this PCA to rotate all N samples for the m traits in our dataset—although any other phylogenetic orthogonalization of our original data would produce the same result. The reason this is an approximation is merely because our phylogenetic orthogonalization, such as phylogenetic PCA, depends on the tree, so we have to assume that the rotation required to orthogonalize our data for our $N-1$ species base tree also makes our characters (nearly) evolutionarily orthogonal on the full, true tree (Adams et al. 2011)—an assumption that is probably quite reasonable, especially for relatively large $N-1$. As pointed out by Adams (2014), it is also true that for a number of traits, m, equal to or in excess of the number of taxa in the base tree ($N-1$), phylogenetic PCA will include eigenvalues not different from zero and the likelihood will not be calculable in a meaningful way. It is possible that in this case we could just retain $N-1$ or fewer eigenvectors; however, the effect of so doing has, to our knowledge, not been explored.

We have implemented this method in the phytools (Revell 2012) function `locate.yeti`. This name was suggested by one of the coauthors (G. J. Slater) as a whimsical "tongue-in-cheek" reference to the theoretically possible use of this method to place cryptozoological taxa on a base phylogeny using phenotypic traits; although we fully anticipate that most uses of the method will be much more serious. phytools (Revell 2012) is a contributed R (R Core Team 2014) package that depends internally on the multifunctional base phylogenetics package ape (Paradis et al. 2004), as well as on several other R libraries covering a range of different functionality (Ligges and Mächler 2003; Jackson 2011; Schliep 2011; Gilbert and Varadham 2012; Qiu and Joe 2013; Xie 2013; Adler et al. 2014; Azzalini and Genz 2014; Becker et al. 2014).

In addition to the approximate method in which the character data are orthogonalized using phylogenetic PCA with the base tree, recent versions of phytools ($\geq$0.4-39) also permit the user to run an exact ML optimization in which this preliminary orthogonalization is not performed. In this case, orthogonalization is still used—but is performed separately for each tree and set of branch lengths. The disadvantage, of course, is that optimization of the exact likelihood is more computationally intensive and may be impractical for datasets with many taxa in the base tree, many phenotypic characters, or both.

Simulation Analysis

To test the performance of this method we conducted the following simulations. First, we simulated $m = 10$ correlated characters on pure-birth (Yule) phylogenies containing $N = 21$, 31, 41, 51, 61, 71, 81, 91, and 101 terminal taxa. We simulated correlated character evolution under Brownian motion in which we obtained a random, positive-definite covariance matrix for the Brownian process using the method of Joe (2006) implemented in the R package clusterGeneration (Qiu and Joe 2013). The method of Joe (2006) permitted us to obtain random covariance matrices for simulation in which correlations and variances of the evolutionary process for our continuous traits were drawn from uniform distributions on the intervals $[-1, 1]$ and $[1, 10]$, respectively. To explore the sensitivity of the method to the number of characters simulated, we also simulated $m = 1$, 2, 5, 10, and 20 correlated characters for a total of $N = 51$ terminal taxa in the tree, using the same method for simulating character correlation as described above. For each simulation condition, we performed 100 replicates, and for each replicate we pruned one leaf at random. We assessed the performance of our method in placing the unknown leaf in the (true) base tree by computing both the Robinson–Foulds topological distance (Robinson and Foulds 1981), and the Kuhner and Felsenstein (1994) branch-score distance. The former distance merely counts the minimum number of edges that would need to be dissolved and added to get from one topology to a second (Felsenstein 2004), whereas the branch-score distance computes the sum of squared differences between corresponding edges in two trees, in which edges absent from one tree (but present in the other) count as having a length of 0.0 (Kuhner and Felsenstein 1994; Felsenstein 2004). We computed tree distances using the R packages phangorn (Schliep 2011) and Rphylip (Revell and Chamberlain 2014). The latter is an R interface for Felsenstein's (1989, 2013) PHYLIP phylogeny methods software package.

Figure 3 shows the results from these analyses. Compared to attaching the unknown leaf randomly in the tree (gray bars in each panel), our method (white bars) results in much higher topological and branch-length accuracy except in the case of very small base trees (Fig. 3A, B), and when only one simulated character trait is used (Fig. 3C, D). In fact, for trees containing 30 or more tips, the

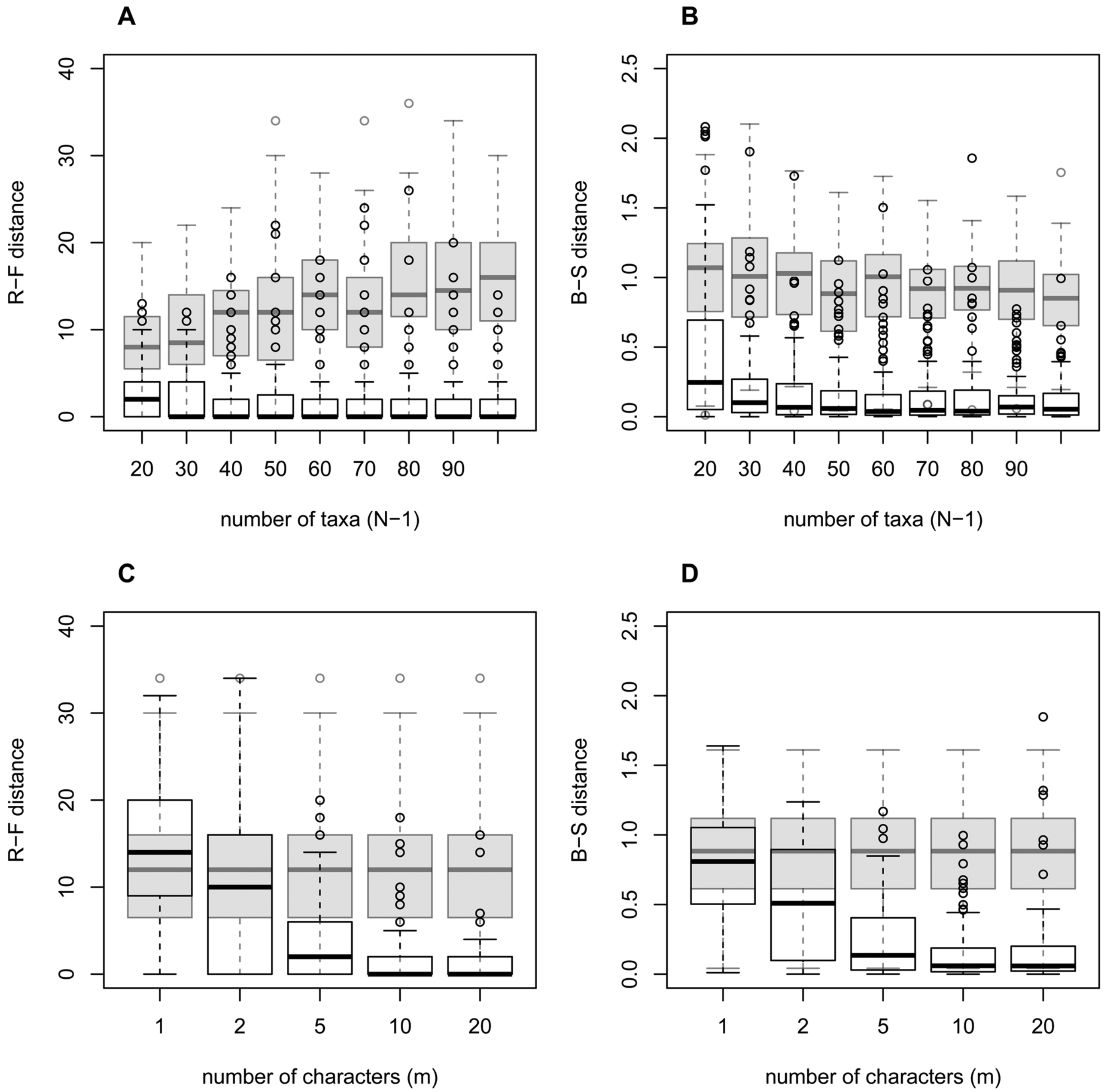

Figure 3. White bars show the results from simulation analysis of the method. In each case, the semitransparent gray bars show the result obtained by randomly placing the missing taxon in the base tree. (A) Robinson–Foulds topological distance (Robinson and Foulds 1981) between the true and estimated phylogeny for various numbers of taxa. (B) Branch-score distances (Kuhner and Felsenstein 1994) for the same simulations as (A). (C) Robinson–Foulds distance for 51 taxon trees for various numbers of simulated correlated characters. (D) Branch-score distance for the same simulations as in (C). In (C) and (D), the vertical axis has been scaled to [0, 2.5] for clarity, but this resulted in the exclusion of six and three outliers from the two plots, respectively.

estimated position of the unknown leaf is very often topologically identical to its true position (Fig. 3A, C).

Case Study

As a case study, we chose to investigate the phylogenetic position of the lizard *A. roosevelti* (the Culebra Island Giant Anole), because, though federally listed as Endangered under the U.S. Endangered Species Act, this species has not been collected since the early 1930s and is thought to be extinct in at least some parts of its original range (Ojeda Kessler 2010), if not globally. *Anolis roosevelti* is an arboreal lizard species that is believed to have once inhabited most of the Spanish, United States, and British Virgin Islands (Grant 1931; Mayer 1989; Ojeda Kessler

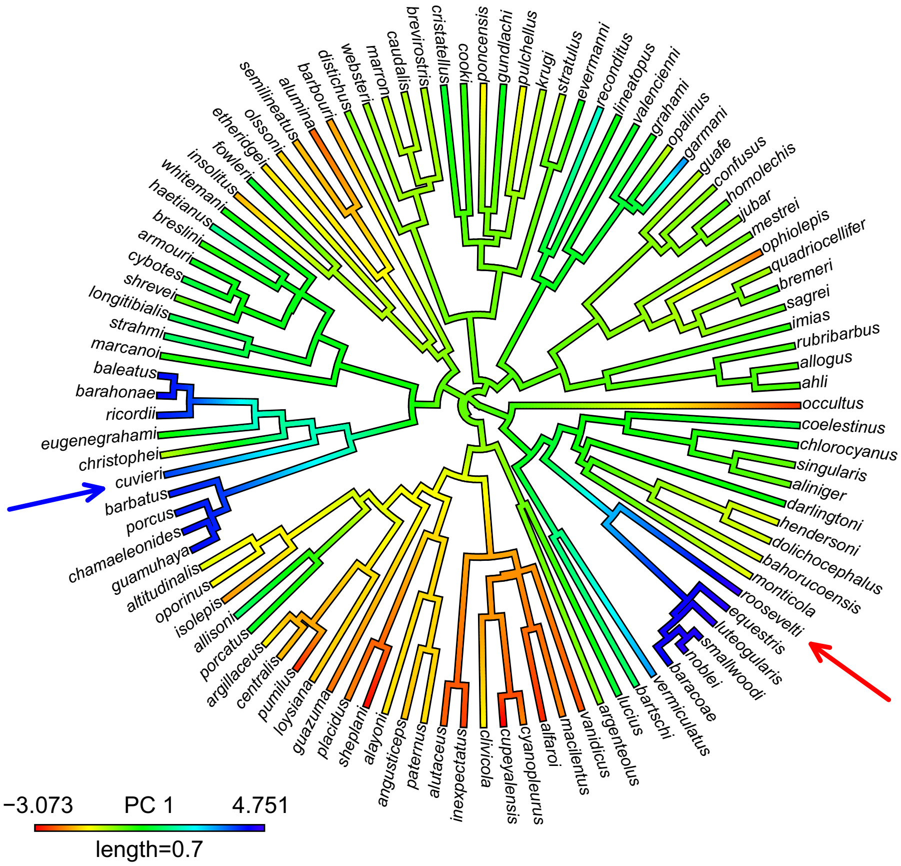

Figure 4. Maximum likelihood tree for the placement of *Anolis roosevelti* based on our 20 character morphological dataset from Mahler et al. (2010). Projected onto the tree is a color gradient mapping of the observed and reconstructed character values for PC 1 from a phylogenetic principal components analysis. Red and blue arrows mark the ML position of *A. roosevelti* and the a priori hypothesized position of *A. roosevelti* (i.e., sister to *A. cuvieri*), respectively.

Table 1. Results from simulation analysis of alternative placements of *Anolis roosevelti* on the Greater Antillean base phylogeny.

Model	log(*L*)	*P*-value (compared to unconstrained model)
Unconstrained	3392.1	–
Sister to *A. cuvieri*	3389.6	0.26
Sister to or nested within other PR species	3385.2	0.02

We cannot reject placement of this species as sister lineage to the Puerto Rican congener, *Anolis cuvieri*; however, a placement sister-to or nested-within the clade containing most other Puerto Rican species (*A. cooki*, *A. cristatellus*, *A. evermanni*, *A. gundlachi*, *A. krugi*, *A. poncensis*, and *A. pulchellus*) is strongly rejected by our method.

2010). For an anole, this species is large—among the biggest of all so-called "crown-giant" ecomorph species (Losos 2009). The (potential) loss of this species represents, in the opinion of at least three of four of the authors (L. J. Revell, R. Graham Reynolds, and D. Luke Mahler), a tragedy of considerable significance. It was long thought to be the case that the first and last collections of the species occurred in the 1930s; however, strong evidence now indicates that a small number of specimens collected in the 1860s or earlier and labeled *A. cuvieri* (or with other binomials) are actually conspecifics (Mayer 1989). It is likely that the 20th century specimens of this species were preserved in formalin—a practice that had become widespread by the 1930s. As such, no DNA sequences have yet been obtained for this species—though hope remains that DNA characters could yet be obtained via improved extraction protocols (e.g., Stuart et al. 2006, Friedman and Desalle 2008), or from the earlier mislabeled specimens, which would likely have been preserved using ethanol (G. Mayer, pers. comm.; Simmons 2014). Because neither molecular nor osteological character data have been obtained for this species, its phylogenetic affinity is unclear (Williams 1976; Poe 2004). As the species was thought to have been a crown-giant ecomorphological habitat specialist ("ecomorph"; Mayer 1989; Losos 2009), it is generally hypothesized to be sister to the Puerto Rican crown-giant *A. cuvieri* (Poe 2004; Losos 2009; Nicholson et al. 2012). However, it is conceivable that the species is more closely related to other Puerto Rican anoles, such as *A. cristatellus*, a trunk-ground ecomorph species that is widespread across the Puerto Rican Bank, or alternatively that it is affiliated with other Greater Antillean giant species (e.g., *A. ricordii* from Hispaniola or *A. equestris* from Cuba).

While collecting data for the study of morphological diversification among Caribbean anoles (Mahler et al. 2010, 2013), one of the authors (D. Luke Mahler) obtained 20 different morphological measurements from the type specimen available at the Museum of Comparative Zoology, Harvard University, in Cambridge, Massachusetts (Fig. 1). These data are a mixture of external continuous characters that primarily correspond to the dimensions of underlying skeletal elements and lamellar scale counts. (Lamellae are expanded subdigital scales found in anoles and geckos.) More details on the morphological dataset can be seen in Mahler et al. (2010).

We found that the ML position of *A. roosevelti*, based on our method, is as the sister lineage to a clade of morphologically and ecologically similar species in Cuba (*A. equestris* and relatives; Fig. 4), not with *A. cuvieri* as hypothesized in most prior studies. To test the hypothesis that the inferred ML phylogenetic position of *A. roosevelti* is significantly better than alternative positions (specifically, either as sister lineage to *A. cuvieri*, or sister to or nested within the clade of anoles that includes most of the other Puerto Rican species: *A. cooki*, *A. cristatellus*, *A. evermanni*, *A. gundlachi*, *A. krugi*, *A. poncensis*, and *A. pulchellus*), we used simulation on the constraint tree, and then inference without constraint, to generate a null distribution for the likelihood ratio test statistic. (We did this because it is unclear how many degrees of freedom should be consumed by placing topological constraints on the location of the unknown leaf.) The results from this analysis are given in Table 1. Our analysis shows that a sister relationship between *A. cuvieri* and *A. roosevelti* cannot be rejected by our method; however, our data suggest that *A. roosevelti* is most likely not closely related to the clade containing most other Puerto Rican anole species (Table 1).

Discussion

In recent years phylogenetic biologists have developed an increasingly complete picture of the phylogeny of all living things, mostly based on molecular data from DNA sequences. This is due not only to the terrific utility of molecular characters in resolving phylogenetic relationships, but also because molecular genetic data are becoming progressively easier and cheaper to obtain from properly preserved tissue samples. Consequently, we should expect that holes in the tree of life will increasingly represent species for which DNA samples are peculiarly difficult to obtain—for example, recent or historic anthropogenic extinctions, or species found only in countries with limited scientific infrastructure that are either inaccessible to scientists or that have extremely restrictive biological sample export laws. In the present brief communication, we describe a simple ML approach for placing a recently extinct, cryptic, or hypothesized taxon into an ultrametric molecular phylogeny (e.g., a time-calibrated tree of extant species) based on a set of continuous traits. Our ML method is based closely on Felsenstein (1973, 1981, 2002), but also provides a simple and computationally efficient way to deal with evolutionary correlations between characters when most of the phylogeny can be treated as known (e.g., Felsenstein 2002; Adams et al. 2011). We find that the method has good performance in simulation (Fig. 3).

In our case study, we found that although we could not reject a relationship between the putatively extinct *A. roosevelti* and *A. cuvieri* (as has been previously suggested), the ML position of *A. roosevelti* was as sister to the monophyletic clade of Cuban crown-giant anoles that includes *A. equestris* and its relatives (Fig. 4). One might reasonably ask if our placement of *A. roosevelti* could be due to the widespread, multivariable morphological convergence between distantly related clades that is so well known for anoles (Losos 2009; Mahler et al. 2013; but see Poe 2005). Indeed, mapping PC 1 from a phylogenetic PCA of all data onto our estimated ML tree shows that the *A. equestris* clade and *A. cuvieri* are closely convergent on this trait axis (Fig. 4). In a future study it might be interesting to apply the method to characteristics more likely to have arisen under an evolutionary process better approximated by Brownian motion—such as, for example, genetic drift or randomly fluctuating natural selection through time; or, alternatively, to explicitly accommodate different evolutionary processes in different parts of the tree where these have been identified a priori (e.g., Mahler et al. 2013).

We address only the issue of placing one taxon into an otherwise fully resolved base tree. In principle, a similar approach could be used to place two or more leaves onto the tree. Two specific difficulties would arise from this extension. First, the number of places a single leaf can attach is simply the number of edges in the base tree plus the root node: that is, $2N - 1$ for an N taxon fully bifurcating base phylogeny. However, in general, the number of possible attachment points for i leaves will be approximately$(2N - 1)^i$ (ignoring the possibility that new leaves attach to each other or to newly created edges in the tree), an expression that will tend to become very large for even relatively modest N and i. This will naturally create computational challenges that we have not addressed in our relatively simple phytools implementation of this method. Second, and perhaps more seriously, our evolutionary orthogonalization of the continuous traits is an approximation that depends on the tree being almost fully known (Adams et al. 2011). If a large number of leaves are missing from the tree—then this heuristic may risk leading to substantial bias or overconfidence in our inferred topology.

We have also naively behaved as if the base tree is known without error. For circumstances in which the base tree contains substantial uncertainty, we could take this uncertainty into account by repeating the optimization on each tree from a bootstrap sample (Felsenstein 1985a) or from a sample obtained from the posterior distribution of a Bayesian analysis (Huelsenbeck et al. 2001). Because R permits us to read in an arbitrary number of tress from an input file and iterate operations across trees, this is already straightforward to implement using the phytools package.

Although we apply this method specifically to recently extinct missing taxa, a very slight modification would permit the method to be used to place a fossil taxon in the tree (Felsenstein 2002). This simply requires that the terminal edge length leading to the leaf be included as a free parameter that we also optimize using numerical methods. In this case we could impose constraints on the height of this leaf that reflect our knowledge of the age of the fossil or the geological stratum in which it was discovered (e.g., Pyron 2011; Ronquist et al. 2012). Indeed, we have now implemented exactly this method in a new phytools function, `locate.fossil`, that is currently available in *beta* release with recent versions of phytools and, in preliminary testing, seems to work quite well.

We chose to use ML to fit this model, but in some ways that decision was arbitrary. We have also implemented a version of this method (in phytools$\geq$0.4-39) using restricted ML (REML) based on Felsenstein's (1985b) phylogenetically independent contrasts. Using contrasts avoids the estimation of ancestral states for each trait, which helps decrease the rate at which the number of estimated parameters increases with the number of traits (although the Brownian rate, σ^2, must still be estimated for each character). Computation of the restricted likelihood using contrasts is also faster than calculation of the likelihood, and thus REML optimization does not necessitate preliminary phylogenetic orthogonalization of the trait data. We could gain further computational efficiencies by, for instance, rerooting the tree at the attachment point of the unknown lineage, in which case most of the contrasts computed in the previous iteration of the REML optimization will not need to be recalculated (Felsenstein, pers. comm.). Finally, it would also be straightforward to develop an equivalent approach based on Bayesian MCMC. In this case we would obtain a posterior sample of attachment points from which we could compute the posterior probabilities that the leaf of interest arose from each internal or terminal edge in our base tree.

ACKNOWLEDGMENTS

The idea for the method described herein was developed by two of the authors (LJR and GJS) while participating as guest instructors in C. Nunn's AnthroTree workshop at the National Evolutionary Synthesis Center (NESCent) in Durham, NC. This work also benefited from some very helpful discussions with L. Harmon, G. Mayer, and B. Redelings. An earlier version of this article was improved by very helpful comments from D. Adams, J. Felsenstein, D. Polly, and R. Shaw. This work was funded in part by a grant from the National Science Foundation (DEB 1350474).

DATA ARCHIVING

The doi for our data is 10.5061/dryad.1st98.

LITERATURE CITED

Adams, D. C. 2014. Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. Syst. Biol. 63:166–177.

Adams, D. C., A. Cardini, L. R. Monteiro, P. O'Higgins, and F. J. Rholf. 2011. Morphometrics and phylogenetics: principal components of shape from cranial modules are neither appropriate nor effective cladistics characters. J. Hum. Evol. 60:240–243.

Adler, D., D. Murdoch, O. Nenadic, S. Urbanek, M. Chen, A. Gebhardt, B. Bolker, G. Csardi, A. Strzelecki, and A. Senger. 2014. rgl: 3D visualization device system (OpenGL). R package version 0.93.1098.

Azzalini, A., and A. Genz. 2014. The R package 'mnormt': The multivariate normal and 't' distributions. R package version 1.5-1.

Becker, R. A., A. R. Wilks, R. Brownrigg, and T. P. Minka. 2014. maps: draw geographical maps. R package version 2.3-7.

Delsuc, F., H. Brinkmann, and H. Philippe. 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6:361–375.

Enk, J., A. Devault, R. Debruyne, and C. King. 2011. Complete Columbian mammoth mitogenome suggests interbreeding with woolly mammoths. Genome Biol. 12:R51.

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

———. 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution 35:1229–1242.

———. 1985a. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

———. 1985b. Phylogenies and the comparative method. Am. Nat. 125:1–15.

———. 1989. PHYLIP—Phylogeny Inference Package (Version 3.2). Cladistics 5:164–166.

———. 2002. Quantitative characters, phylogenies, and morphometrics. Pp. 27-44 *in* N. MacLeod and P. Forey, eds. Morphology, shape and phylogeny. Taylor and Francis, London.

———. 2004. Inferring phylogenies. Sinauer, Sunderland, MA.

———. 2013. PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Friedman, M., and R. Desalle. 2008. Mitochondrial DNA extraction and sequencing of formalin-fixed archival snake tissue. Mitochondrial DNA 19:433–437.

Gilbert, P., and R. Varadhan. 2012. numDeriv: accurate numerical derivatives. R package version 2012.9-1.

Gold, D. A., J. Robinson, A. B. Farrell, J. M. Harris, O. Thalmann, and D. K. Jacobs. 2014. Attempted DNA extraction from a Rancho La Brea Columbian mammoth (*Mammuthus columbi*): prospects for ancient DNA from asphalt deposits. Ecol. Evol. 4:329–336.

Grant, C. 1931. A new species and two new sub-species of the genus *Anolis*. J. Dept. Agric. Puerto Rico 15:219–222.

Hillis, D. M., C. Moritz, and B. K. Mable. 1996. Molecular systematics. 2nd ed. Sinauer, Sunderland, MA.

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310–2314.

Jackson, C. H. 2011. Multi-state models for panel data: the msm package for R. J. Stat. Softw. 38:1–29.

Joe, H. 2006. Generating random correlation matrices based on partial correlations. J. Multivariate Anal. 97:2177–2189.

Kessler A. G. O. 2010. Status of the Culebra Island giant anole (*Anolis roosevelti*). Hepetol. Conserv. Biol. 5:223–232.

Kuhner, M. K., and J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. 11:459–468.

Lemon, J. 2006. Plotrix: a package in the red light district of R. R-News 6:8–12.

Ligges, U., and M. Mächler. 2003. Scatterplot3d—an R package for visualizing multivariate data. J. Stat. Softw. 8:1–20.

Losos, J. B. 2009. Lizards in an evolutionary tree: ecology and adaptive radiation of anoles. University of California Press, Berkeley, CA.

Mahler, D. L., L. J. Revell, R. E. Glor, and J. B. Losos. 2010. Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. Evolution 64:2731–2745.

Mahler, D. L., T. Ingram, L. J. Revell, and J. B. Losos. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. Science 341:292–295.

Mayer, G. C. 1989. Variation and distribution of *Anolis roosevelti* Grant, with comments on Puerto Rican Bank herpetology. Pp. 92–122 *in* Deterministic aspects of community structure in West Indian amphibians and reptiles. PhD dissertation, Harvard University, Cambridge, MA.

Nicholson, K. E., B. I. Crother, C. Guyer, and J. M. Savage. 2012. It is time for a new classification of anoles (Squamata: Dactyloidae). Zootaxa 3477:1–108.

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

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

Poe, S. 2004. Phylogeny of anoles. Herpetol. Monogr. 18:37–89.

———. 2005. A study of the utility of convergent characters for phylogeny reconstruction: do ecomorphological characters track evolutionary history in *Anolis* lizards? Zoology 108:337–343.

Pyron, R. A. 2011. Divergence time estimation using fossils and terminal taxa and the origins of Lissamphibia. Syst. Biol. 60:466–481.

Qiu, W., and H. Joe. 2013. clusterGeneration: random cluster generation (with specified degree of separation). R package version 1.3.1.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Renner, S. C., D. Neumann, M. Burkart, U. Feit, P. Giere, A. Gröger, A. Paulsch, C. Paulsch, M. Sterz, and K. Vohland. 2012. Import and export of biological samples from tropical countries—considerations and guidelines for research teams. Org. Divers. Evol. 12:81–98.

Revell, L. J. 2008. On the analysis of evolutionary change along single branches in a phylogeny. Am. Nat. 172:140–147.

———. 2009. Size-correction and principal components for interspecific comparative studies. Evolution 63:3258–3268.

———. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

———. 2013. Two new graphical methods for mapping trait evolution on phylogenies. Methods Ecol. Evol. 4:754–759.

Revell, L. J., and S. A. Chamberlain. 2014. Rphylip: an R interface for PHYLIP. Methods Ecol. Evol. 5:976–981.

Revell, L. J., and L. J. Harmon. 2008. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evol. Ecol. Res. 10:311–321.

Roberts, D. L., and A. R. Solow. 2008. The effect of the convention on international trade in endangered species on scientific collections. Proc. R. Soc. B 275:987–989.

Robinson, D. R., and L. R. Foulds. 1981. Comparison of phylogenetic trees. Math. Biosci. 53:131–147.

Rocha, L. A., A. Aleixo, G. Allen, F. Almeda, C. C. Baldwin, M. V. L. Barclay, J. M. Bates, A. M. Bauer, F. Benzoni, C. M. Berns, et al. 2014. Specimen collection: an essential tool. Science 344:814–815.

Ronquist, F., S. Klopfstein, L. Vilhelmsen, S. Schulmeister, D. L. Murray, and A. P. Rasnitsyn. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Syst. Biol. 61:973–999.

Schliep K. P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27:592–593.

Simmons, J. E. 2014. Fluid preservation: a comprehensive reference. Rowman and Littlefield, Lanham, MD.

Srinivasan, M., D. Sedmack, and S. Jewell. 2002. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. Am. J. Pathol. 161:1961–1971.

Stuart, B. L., K. A. Dugan, M. W. Allard, and M. Kearney. 2006. Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. Syst. Biodivers. 4:133–136.

Tang, E. P. Y. 2006. Path to effective recovering of DNA from formalin-fixed biological samples in natural history collections, workshop summary. National Academic Press, Washington, DC.

Williams, E. E. 1976. West Indian anoles. A taxonomic and evolutionary summary. 1. Introduction and a species list. Breviora 440:1–21.

Xie, Y. 2013. animation: an R package for creating animations and demonstrating statistical methods. J. Stat. Softw. 53:1–27.

Associate Editor: D. Adams
Handling Editor: R. Shaw