Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds

Kate L. Sanders* and Michael S. Y. Lee
School of Earth and Environmental Sciences, University of Adelaide, Darling Building, Adelaide, Australia 5005
*Author for correspondence (kate.sanders@adelaide.edu.au).

A limiting factor in many molecular dating studies is shortage of reliable calibrations. Current methods for choosing calibrations (e.g. cross-validation) treat them as either correct or incorrect, whereas calibrations probably lie on a continuum from highly accurate to very poor. Bayesian relaxed clock analysis permits inclusion of numerous candidate calibrations as priors: provided most calibrations are reliable, the model appropriate and the data informative, the accuracy of each calibration prior can be evaluated. If a calibration is accurate, then the analysis will support the prior so that the posterior estimate reflects the prior; if a calibration is poor, the posterior will be forced away from the prior. We use this approach to test two fossil dates recently proposed as standard calibrations within vertebrates. The proposed bird–crocodile calibration (approx. 247 Myr ago) appears to be accurate, but the proposed bird–lizard calibration (approx. 255 Myr ago) is substantially too recent.

Keywords: molecular dating; Bayesian analysis; calibration; fossil; reptile

1. INTRODUCTION

Recent advances in molecular clock methods and expanding genetic databases provide an unprecedented opportunity to infer time-scales of organisinal evolution. However, progress is lagging on one vital component of molecular divergence dating—calibration. Methods for choosing calibrations are being developed (e.g. van Tuinen & Dyke 2003; Near et al. 2005) and point calibrations are being replaced by distributions that better represent palaeontological uncertainty (e.g. Yang & Rannala 2005; Drummond et al. 2006). Much of the debate over calibration issues has involved amniotes (mammals, birds and reptiles). The mammal–bird split is probably the most commonly used calibration for these organisms, but its accuracy has been questioned due to a long preceding fossil gap (Reisz & Müller 2004; Müller & Reisz 2005). Instead, Reisz & Müller (2004) proposed using two other splits with tightly bounded divergence times as standard calibrations. They suggested a range of 252–257 Myr ago for the split between Archosauromorpha (birds, crocodiles and relatives) and Lepidosauromorpha (lizards, snakes, tuatara and relatives), and a range of 243–251 Myr ago for the split between Ornithodira (birds and relatives) and Crurotarsi (crocodiles and relatives). These dates, if correct, imply that both divergences occurred in close proximity (less than 14 Myr). More recently, Benton & Donoghue (2007), as part of a review across all metazoans, proposed a much broader interval (259.7–299.8 Myr ago) for the Archosauromorpha–Lepidosauromorpha (bird–lizard) split, but a similar interval (246.5–250.4 Myr ago) for the bird–crocodile split.

If the calibrations proposed by Reisz & Müller (2004) are accurate: (i) the molecular branches between the bird–lizard and the bird–crocodile divergences should be very short and (ii) these calibrations should be broadly consistent with other well-corroborated calibrations. Here, we use Bayesian relaxed clock analyses with soft and hard calibration bounds to demonstrate that the bird–lizard and bird–crocodile divergences were relatively broadly spaced in time; in particular, the bird–lizard divergence occurred much earlier than proposed and should be used with a much deeper maximum bound. The Bayesian approach used here does not treat calibrations as either correct or incorrect, but recognizes that calibrations lie on a continuum between highly accurate (narrow bounds straddling the actual divergence) and very inaccurate (wide bounds that still do not encompass the real date), and places greater weight on the former.

2. MATERIAL AND METHODS

Taxa were selected to span the major amniote clades and a range of well-corroborated calibrations. Two nuclear loci were selected because they provide independent, large single-copy sequences uninterrupted by introns, and evolve at a suitable rate: approximately 3160 bp composing most of the recombination activating gene-1 (RAG-1) and approximately 1100 bp of c-mos (oocyte maturation factor). Sequences were obtained from GenBank, with additional sequences generated to improve sampling for birds and squamates. The tree was rooted with an amphibian (Xenopus). Primers and GenBank accession numbers are provided in electronic supplementary material.

Divergence dating used Beast v. 1.3 (Drummond & Rambaut 2003), which employs a Bayesian Markov chain Monte Carlo (MCMC) to co-estimate topology, substitution rates and node ages. Posterior probability distributions of node ages were obtained for each gene separately and the concatenated 2-gene alignment. The data were partitioned by gene and by codon (first + second versus third). Best-fit substitution models were identified for each data partition using MrModelTest v. 2 (Nylander 2004); the general time-reversible model with rate variation (six gamma categories) was implemented for all partitions, with an invariant site parameter added for one partition (RAG first + second codon). In the combined analysis, model parameters were unlinked across partitions. Each analysis implemented a Yule branching rate prior, the combined analysis, model parameters were unlinked across partitions. Each analysis implemented a Yule branching rate prior, with rate variation across branches assumed to be uncorrelated and lognormally distributed (Drummond et al. 2006). Each MCMC chain was run for 4 000 000 generations (burnin 20%), with parameters sampled every 100 steps. Examination of MCMC samples using Tracer v. 1.2 (Rambaut & Drummond 2003) suggested that the independent chains were each adequately sampling the same probability distribution; effective sample sizes for all parameters of interest were greater than 500.

Four direct fossil calibrations were used in addition to the bird–lizard and bird–crocodile calibrations of Reisz & Müller (2004; table 1). Divergences were estimated using the above MCMC analyses and these calibrations in four ways: (1) First, a relaxed clock was used, with calibrations treated as having a translated lognormal distribution. This yields a skewed distribution consistent with the bias in the fossil record: there is a hard minimum bound meaning zero probability of dates much
Table 1. Prior probability distributions for divergence times enforced for the six calibration nodes based on fossil evidence.

<table>
<thead>
<tr>
<th>Calibration node</th>
<th>prior distribution</th>
<th>bird–crocodile (95% CI)</th>
<th>bird–lizard (95% CI)</th>
<th>scincomorph–anguimorph (95% CI)</th>
<th>pleurodire–cryptodire (95% CI)</th>
<th>Rattus–Mus (95% CI)</th>
<th>bird–mammal (95% CI)</th>
</tr>
</thead>
</table>

Table 2. Posterior probability distributions for divergence times for the bird–crocodile and bird–lizard nodes, obtained from the combined and single gene data analysed using a relaxed clock with normal, lognormal and uniform hard-bounded calibration priors (Table 1), and a globally constant molecular clock with lognormal calibrations.

<table>
<thead>
<tr>
<th>Calibration node</th>
<th>posterior—mean (95% HPD)—millions of years before present</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAG-1 + C-MOS</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>2328 (2308, 2348)</td>
</tr>
<tr>
<td>Lognormal</td>
<td>2473 (2470, 2475)</td>
</tr>
<tr>
<td>Uniform</td>
<td>2448 (2430, 2468)</td>
</tr>
<tr>
<td>Global clock</td>
<td>2358 (2348, 2368)</td>
</tr>
<tr>
<td>RAG-1</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>2319 (2313, 2323)</td>
</tr>
<tr>
<td>Lognormal</td>
<td>2477 (2473, 2481)</td>
</tr>
<tr>
<td>Uniform</td>
<td>2442 (2430, 2464)</td>
</tr>
<tr>
<td>C-MOS</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>2320 (2302, 2332)</td>
</tr>
<tr>
<td>Lognormal</td>
<td>2467 (2475, 250)</td>
</tr>
<tr>
<td>Uniform</td>
<td>2487 (2430, 2510)</td>
</tr>
</tbody>
</table>

combined gene datasets were very similar (table 2) and the discussion below focuses on the combined analyses. Age estimates from the relaxed clock analysis with lognormal calibration priors produced narrow 95% HPD intervals that are mostly consistent with the fossil dates (figure 1, table 2). A notable feature of all these trees is the relatively long branches between the bird–lizard and bird–crocodile divergences, indicating that they are widely spaced in time (contra, Reisz & Müller 2004). Thus, either the bird–lizard divergence is older than proposed or the bird–crocodilian divergence is younger; our analyses indicate the former. The combined gene analysis dates the most recent common ancestor of the bird–crocodile clade at 247.3 Myr ago (95% HPD: 237.7–259.0). The posterior estimate is therefore very similar to the prior (mode = 247; figure 1). However, the bird–lizard split is dated at

3. RESULTS AND DISCUSSION
All MCMC searches converged on a topology highly congruent with recently published molecular phylogenies (e.g. Townsend et al. 2004; Iwabe et al. 2005; van Rheede et al. 2005). Both genes showed high levels of rate heterogeneity (coefficients of branch rate variation greater than 0.8) and low autocorrelation of rates between adjacent branches (rate covariance less than 0.1). Mean and highest posterior density (HPD) node age estimates inferred from individual and

282.8 Myr ago (95% HPD: 263.0–303.9) and greatly exceeds the prior (mode = 254.5). This result suggests that a more liberal lower bound is required for this calibration and is consistent with the older dates proposed by Benton & Donoghue (259.7–299.8 Myr ago).

The relaxed clock analysis with normal calibration priors again produced posterior estimates for the bird–crocodile and bird–lizard divergences that are relatively broadly spaced (table 2). However, as expected (see above), posterior ages for most calibration nodes are younger than the minimum plausible dates implied by fossils, including the bird–crocodile posterior estimate (mean = 232.7 Myr ago; 95% HPD = 213.2–250.9). Only the bird–lizard posterior is older than the prior mode (mean = 269.2 Myr ago; 95% HPD = 250.8–287.8). Similarly, in the global clock analysis, the two divergences of interest were widely spaced: posterior estimates for the bird–crocodile split are slightly shallower than the prior (mean = 235.8 Myr ago; 95% HPD = 232.6–240.0); whereas the bird–lizard posterior exceeds the prior (mean = 261.1 Myr ago; 95% HPD = 249.5–269.1).

The above data suggest that the bird–crocodile and bird–lizard calibrations cannot be as closely spaced in time as proposed: the long intervening molecular branches cause the posterior estimates for the bird–lizard split to move towards older dates. However, we also investigated whether these long branches can be reconciled with the proposed closely spaced divergences by assuming very rapid molecular evolution. Fixing all proposed calibrations (table 1) with hard, narrow bounds implies extremely high branch-specific substitution rates on the branches between the two calibrations of interest. Results for the two-gene analysis are shown (table 2, figure 2). The mean rate of molecular evolution for the branches between the bird–lizard and bird–crocodile nodes is 0.0055 substitutions per site per Myr: this is by far the highest anywhere on the tree, and more than double that of rodents (0.0026), which have the highest rate known in amniotes (e.g. Douzery et al. 2003). There is nothing about the inferred ecology of early archosauromorphs (e.g. generation time, metabolic rate) that would suggest such an exceptionally high rate of molecular evolution.

The present study shows how Bayesian analysis of molecular data with soft and hard bounds can be used to evaluate the accuracy of proposed calibrations. However, for this approach to work, the phylogenetic model must be reasonably accurate, the molecular data informative and most calibrations reliable. All potential calibration information is included in the analysis: the concordant, reliable calibrations contribute most to the final date estimates (priors consistent with posteriors), while less reliable calibrations have less influence (priors inconsistent with posteriors). This is a promising alternative to the existing methods of choosing calibration points (e.g. cross-validation: Near et al. 2005), which often use an arbitrary cut-off to retain a subset of
calibrations, and thus treat calibrations in an all-
or-nothing fashion.


Nylander, J. A. A. 2004 *MrMODELTEST*: program distributed by the author, v2. Evolutionary Biology Centre, Uppsala University.


Townsend, T., Larson, A., Louis, E. & Macey, J. R. 2004 Molecular phylogenetics of Squamata: the position of

