Molecular Phylogenetics and Evolution 65 (2012) 974–991

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles

Daniel G. Mulcahy[a,*], Brice P. Noonan[a], Travis Moss[a], Ted M. Townsend[b], Tod W. Reeder[b], Jack W. Sites Jr.[a], John J. Wiens[c]

[a] *Department of Biology, Brigham Young University, Provo, UT 84602, USA*
[b] *Department of Biology, San Diego State University, San Diego, CA 92182-4614, USA*
[c] *Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794-5245, USA*

ARTICLE INFO

Article history:
Received 26 December 2011
Revised 21 August 2012
Accepted 22 August 2012
Available online 7 September 2012

Keywords:
BEAST
Lizards
Penalized likelihood
Phylogeny
r8s
Snakes

ABSTRACT

Recently, phylogenetics has expanded to routinely include estimation of clade ages in addition to their relationships. Various dating methods have been used, but their relative performance remains understudied. Here, we generate and assemble an extensive phylogenomic data set for squamate reptiles (lizards and snakes) and evaluate two widely used dating methods, penalized likelihood in r8s (r8s-PL) and Bayesian estimation with uncorrelated relaxed rates among lineages (BEAST). We obtained sequence data from 25 nuclear loci (~500–1000 bp per gene; 19,020 bp total) for 64 squamate species and nine outgroup taxa, estimated the phylogeny, and estimated divergence dates using 14 fossil calibrations. We then evaluated how well each method approximated these dates using random subsets of the nuclear loci (2, 5, 10, 15, and 20; replicated 10 times each), and using ~1 kb of the mitochondrial *ND2* gene. We find that estimates from r8s-PL based on 2, 5, or 10 loci can differ considerably from those based on 25 loci (mean absolute value of differences between 2-locus and 25-locus estimates were 9.0 Myr). Estimates from BEAST are somewhat more consistent given limited sampling of loci (mean absolute value of differences between 2 and 25-locus estimates were 5.0 Myr). Most strikingly, age estimates using r8s-PL for *ND2* were ~68–82 Myr older (mean = 73.1) than those using 25 nuclear loci with r8s-PL. These results show that dates from r8s-PL with a limited number of loci (and especially mitochondrial data) can differ considerably from estimates derived from a large number of nuclear loci, whereas estimates from BEAST derived from fewer nuclear loci or mitochondrial data alone can be surprisingly similar to those from many nuclear loci. However, estimates from BEAST using relatively few loci and mitochondrial data could still show substantial deviations from the full data set (>50 Myr), suggesting the benefits of sampling many nuclear loci. Finally, we found that confidence intervals on ages from BEAST were not significantly different when sampling 2 vs. 25 loci, suggesting that adding loci decreased errors but did not increase confidence in those estimates.

1. Introduction

In recent years there has been increasing interest in using molecular-based phylogenies to infer the ages of clades (e.g., Sanderson, 2002; Drummond et al., 2006; Rutschmann, 2006; Hedges and Kumar, 2009). Time-calibrated phylogenies have become integral to many evolutionary studies, including analyses of biogeography (e.g., Ree and Smith, 2008), species diversification (e.g., Ricklefs, 2007), and phenotypic evolution (e.g., O'Meara et al., 2006).

Several methods for divergence-time estimation have been developed (e.g., Thorne et al., 1998; Yoder and Yang, 2000; Huelsenbeck et al., 2000; Sanderson, 2003; Thorne and Kishino, 2002; Drummond et al., 2006; Yang and Rannala, 2006; Lepage et al., 2007; Rannala and Yang, 2007; Lartillot et al., 2009). The most widely used methods at present are based on "relaxed" molecular clocks where a general relationship between time and molecular divergence is assumed, and this relationship can vary across the tree.

In the recent literature, two methods in particular have been widely used, penalized likelihood (implemented in r8s; Sanderson, 2002, 2003) and Bayesian estimation with uncorrelated ("relaxed") lognormally distributed rates among branches (implemented in BEAST; Drummond et al., 2006; Drummond and Rambaut, 2007). Penalized likelihood uses an input tree with branch lengths and assumes autocorrelation of rates among lineages, and the uneven-

---

[*] Corresponding author. Address: Laboratory of Analytical Biology, Smithsonian Institution, Museum Support Center, 4210 Silver Hill Road, Suitland, MD 20746, USA. Fax: +1 202 357 3043.

E-mail address: MulcahyD@si.edu (D.G. Mulcahy).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.08.018

ness (roughness) of the change in rates among lineages is penalized. A cross-validation method is used to find the optimal level of rate smoothing, and this "smoothing factor" defines the degree of autocorrelation. Using the Bayesian uncorrelated lognormal approach, rates of change are uncorrelated among branches and the rate on each branch is drawn from a lognormal distribution (Drummond et al., 2006; Drummond and Rambaut, 2007). For brevity, we hereafter use "r8s-PL" to refer to the penalized likelihood approach with r8s and "BEAST" to refer to the Bayesian uncorrelated lognormal method. However, we recognize that these software packages can be used to implement other approaches and that other software packages could potentially be used to implement these approaches. Many other methods for divergence-time estimation have also been frequently used, such as the Bayesian relaxed-clock method using autocorrelated rates among lineages (implemented in MultiDivTime; Thorne and Kishino, 2002).

In addition to implementing the Bayesian uncorrelated approach, BEAST has other important advantages and is becoming widely used relative to r8s-PL and MultiDivTime. For example, BEAST can incorporate uncertainty in topology and branch lengths in estimating divergence dates and allows for different types of prior distributions (e.g., normal, uniform, and lognormal) on calibration points and other external sources (Drummond et al., 2006; Drummond and Rambaut, 2007). Nevertheless, divergence-time estimates from r8s-PL remain common (e.g., Hugall et al., 2007; Wiens, 2007; Burbrink and Pyron, 2008; Kozak et al., 2009; Spinks and Shaffer, 2009; Schulte and Moreno-Roark, 2010). Furthermore, because BEAST (and MultiDivTime) may not be practical on large data sets, r8s-PL may continue to be commonly used well into the foreseeable future.

We know of no simulation studies that have directly compared divergence-date estimates from r8s-PL and BEAST. For example, a thorough simulation study by Battistuzzi et al. (2010) compared only BEAST and MultiDivTime. Further, although some previous empirical studies have estimated and compared dates from r8s and BEAST (e.g., Ribera et al., 2010; Wielstra et al., 2010), it is difficult to evaluate which method gives "better" or "worse" results without some non-arbitrary criterion. Only a few studies have attempted to systematically address differences in these methods with empirical data (Phillips, 2009; Egan and Doyle, 2010).

In many ways, simulation studies offer the best way to evaluate these methods: one can generate a phylogeny with known dates and then examine how well each method estimates those dates, and under what conditions (e.g., Battistuzzi et al., 2010). However, there are many complications in estimating divergence dates from empirical data sets that may make fully realistic simulations challenging. For example, divergence-date estimation typically depends upon having one or more fossil calibration points (i.e., fossil taxa of "known" clade assignment and age), which tend to be available only sporadically among clades within a given group, but may strongly influence the estimated dates (e.g., Near and Sanderson, 2004; Near et al., 2005; Rutschmann et al., 2007; Marshall, 2008; Inoue et al., 2010). Further, different genes may be used to estimate dates, and these genes may differ not only in their length and rates of change, but also in their underlying histories (e.g., Maddison, 1997). This diverse array of complicated parameters may be difficult to simulate realistically. Thus, as a complement to simulation studies, it would be useful to also evaluate and compare dating methods using empirical data but with a non-arbitrary criterion to evaluate them.

Additionally, divergence times are often estimated using a single locus (e.g., RAG-1; see Hugall et al., 2007; Wiens, 2007; Alfaro et al., 2009). Battistuzzi et al. (2010) demonstrated that Bayesian relaxed clock methods can produce more accurate estimates using multiple loci, but no similar studies have been conducted for r8s-PL. Empirical testing of the robustness of Bayesian dating methods to sampling limited numbers of loci is still needed. It is also unclear how the use of more rapidly-evolving mitochondrial genes (in animals) may influence divergence dating relative to the use of more slowly-evolving nuclear loci. The impact of mitochondrial data may be particularly important in older clades, in which longer branches may be systematically overestimated by rapidly evolving mitochondrial genes (Zheng et al., 2011).

In this paper, we take advantage of our phylogenomic studies of squamate reptiles (lizards and snakes) to evaluate and compare two widely used dating methods: r8s-PL and BEAST. We assemble a data set of 19,020 base-pairs (bp) from 25 protein-encoding nuclear loci for 64 ingroup taxa (representing major squamate clades and most families) and nine outgroup taxa, a data set considerably larger than those used in most dating studies. We then evaluate how well these methods approximate the estimated divergence dates based on all 25 loci, given random subsamples of a limited number of these loci (e.g., 2, 5, 10, 15, or 20 loci). Although we do not know what the true ages are for these clades, a method that gives highly variable estimates from a limited sample of loci may be problematic (i.e., two very different estimates for the same node cannot both be correct), relative to a method for which estimates from 2 or 5 loci are similar to those from 25 loci (even though this pattern does not guarantee that estimates from the latter method are actually correct). Additionally, it is important to understand how subsampling loci (i.e., using fewer loci) influences divergence-time estimates for these methods using empirical data.

We also compare estimated divergence dates from the nuclear loci to those estimated from a single mitochondrial (mtDNA) gene. Although nuclear data are becoming increasingly accessible, many prominent analyses of phylogeny and divergence dates (in animals) continue to be based on mtDNA data alone (e.g., Santos et al., 2009; Schulte and Moreno-Roark, 2010). Few studies have systematically compared divergence dates estimated from mtDNA to those based on multiple nuclear loci (e.g., Wahlberg et al., 2009; Zheng et al., 2011) and our extensive sampling of nuclear loci provides an opportunity to address this issue.

This is not the first study of phylogeny and divergence times in squamates. Recent studies have used molecular data to address higher-level squamate relationships (e.g., Townsend et al., 2004; Vidal and Hedges, 2005; Kumazawa, 2007; Wiens et al., 2010), and divergence dates (e.g., Vidal and Hedges, 2005; Wiens et al., 2006; Hugall et al., 2007). Here, we provide the most extensive analysis of higher-level squamate phylogeny and divergence-time estimates to date, in terms of including many loci, taxa, and fossil calibration points.

2. Materials and methods

2.1. Taxonomic sampling

Our taxon sampling was designed to address higher-level squamate phylogeny. We included at least two representatives of most families, except for some well-established clades (Anguimorpha, Iguania, Serpentes) for which we sampled fewer species. For outgroups, we used the tuatara (*Sphenodon*), the closest living relative to Squamata (e.g., Gauthier et al., 1988; Hugall et al., 2007), two crocodilians (*Alligator* and *Crocodylus*), two birds (*Dromaius* and *Gallus*), two turtles (*Chelydra* and *Podocnemis*), and two mammals (*Homo* and *Mus*). A list of sampled species, vouchers, and GenBank accession numbers is provided in Appendix A.

2.2. Molecular sampling

We obtained DNA sequence data from 25 protein-encoding nuclear loci. Some sequence data were used in our previous studies