

Bayesian Dating of Shallow Phylogenies with a Relaxed Clock

RICHARD P. BROWN^{1,*} AND ZIHENG YANG²

¹*School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool L3 3AF, UK; and*

²*Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK;*

*Correspondence to be sent to: School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool L3 3AF, UK;
E-mail: r.p.brown@ljmu.ac.uk.

Received 19 February 2009; reviews returned 15 April 2009; accepted 26 October 2009

Associate Editor: Susanne S. Renner

Abstract.—Bayesian methods are increasingly being used to estimate divergence times without the restrictive assumption of a global clock. Little is known about their reliability for shallow phylogenies where DNA sequence divergence is low. We analyzed both simulated and real sequences to evaluate dating methods in phylogenies with mid–late Miocene roots. A large number of data sets (5000) with 10 taxa each were simulated under a rate-drift model for trees with 2 topologies (balanced or unbalanced) and with different sets of divergence times (characterized by long or short external branches). Data were analyzed using Bayesian Markov chain Monte Carlo methods in which the prior on divergence times was specified from a birth–death process with species sampling (BDS) or a Dirichlet distribution using the programs MCMCTREE and MULTIDIVTIME. The programs generally performed well on shallow phylogenies, but posterior mean node ages were biased and 95% posterior intervals included true ages in fewer than 95% of trees in some analyses. This typically occurred when the 95% prior interval did not include the true age and/or sequence lengths were ≤ 1 kbp. Widths of posterior intervals were also very dependent on the position of the calibrated node within the tree, particularly when sequences were short. Different divergence times priors within MCMCTREE, MULTIDIVTIME, and BEAST were used to analyze mitochondrial DNA data sets from a Bovid subfamily (the Caprinae) from Asian *Laudakia* and North African *Chalcides* lizards. Posterior divergence times were quite sensitive to different BDS priors but less sensitive to different Dirichlet priors. Our study demonstrates the impact of the prior on divergence times in shallow phylogenies and shows that 1) prior intervals on nodes should be assessed as a prerequisite to a dating analysis, 2) ≥ 1 kbp of quite rapidly evolving sequence may be required to obtain accurate posterior means and usefully narrow posterior intervals. [Agamidae; Bayesian phylogenetics; Caprinae; phylogeny; posterior; prior; relaxed clock.]

Over the past decade, there has been considerable progress in the development of phylogenetic methods for estimation of lineage divergence times. Unlike earlier molecular clock analyses, Bayesian Markov chain Monte Carlo (MCMC) and maximum likelihood (ML) techniques can incorporate rate heterogeneity between branches and so allow estimation of divergence times even when the clock is violated (Thorne et al. 1998; Sanderson 2002; Yang and Yoder 2003). Bayesian inference has a major advantage over ML because the prior on divergence times can incorporate the uncertainty associated with node age calibrations (Yang 2006, p. 248–251). Also, recent refinements to the technique (e.g., Yang and Rannala 2006), together with a greater availability of software that implement different algorithms (Thorne et al. 1998; Thorne and Kishino 2002; Drummond and Rambaut 2007; Yang 2007; Lepage et al. 2008), have made the Bayesian MCMC approach increasingly popular in phylogenetic studies.

To date, the suitability of Bayesian relaxed clock analyses of divergence times in shallow phylogenies has not been rigorously assessed. Divergence events since the mid–late Miocene generally correspond to lower taxonomic levels, such as the genesis of new intrageneric or intraspecific lineages (Avice 2000). Knowledge of their timing can provide valuable insights into the climatic and physical processes that have shaped genetic diversity within species as well as into speciation events. One obvious difference between these phylogenies and those with older root ages is that taxa are less divergent, and so sequences tend to be less informative. Divergence

time estimation may be less accurate and/or precise as a result. A global clock model may appear sensible in these circumstances because it involves fewer parameters and because of lower rate variation due to shorter time scales and similarity of taxa. Nevertheless, a relaxed clock is often used, possibly because it accommodates both homogenous and heterogenous rates (e.g., Warren et al. 2003; Emerson and Oromí 2005; Wüster et al. 2005; Graziotin et al. 2006; Schmitz et al. 2007; Melville et al. 2009). There is evidence to support this approach. Simulations indicate that global clock analyses perform relatively poorly when the clock is violated but relaxed clock analyses can work well for clock-like trees (Ho et al. 2005; Drummond et al. 2006).

Most previous investigations of relaxed clock Bayesian dating have used longer more divergent sequences in which the likelihood has a greater influence on the posterior. Little is known about the sequence information content required for accurate recovery of divergence times. Posterior estimates of divergence times will reflect the prior when sequences are short (and therefore less informative) but should be increasingly influenced by the likelihood as sequence length is increased. This appears to be the pattern seen in some real data sets (Brown et al. 2008). The prior may therefore have a greater impact in analyses of shallow phylogenies.

In this article, we focus on 2 different methods for specifying the prior on divergence times. The first of these is the recursive procedure implemented in the program MULTIDIVTIME (Thorne et al. 1998; Kishino et al. 2001). The root age is assigned a gamma prior.

Internode times along paths from internal nodes to the tips are then obtained as proportions of the oldest node age on the path (which is the root node on the first iteration) from a Dirichlet density. Generation of internode times is repeated on different paths within the tree until all nodes have been assigned ages. Time constraints are represented as “hard” maximal and/or minimal bounds on node ages (sensu Yang and Rannala 2006).

The second method for specifying this prior on times uses the birth–death process with species sampling (BDS). This is implemented in the program MCMCTREE (Rannala and Yang 2007; Yang and Rannala 2006; Yang 2007). The BDS prior is specified by 3 parameters: per-lineage birth (λ) and death (μ) rates and the sampling proportion (ρ). The joint prior density of divergence times, conditional on root age, and other time constraints are derived from the BDS model. Consider a tree with s tip nodes and $s-1$ internal node ages, one of which is the root, with age t_1 . The nonroot node ages (denoted t_{-1}) comprise calibrated and noncalibrated nodes, for which ages are denoted t_C and t_{-C} , respectively. When time constraints are available for t_1 , the joint distribution of divergence times specified from the BDS is:

$$f(t_{-1} | t_1) = f_{\text{BDS}}(t_{-C} | t_C, t_1) \times f(t_C)$$

(Yang and Rannala 2006). Thus, the prior is specified by first generating the root age (t_1) and then generating ages of noncalibrated nodes (t_{-1}) from the BDS kernel, conditional on the root, and calibration node ages. This is multiplied by the calibration node density, $f(t_C)$ (see Yang and Rannala 2006 for details).

The program BEAST implements special cases of the BDS model to generate a time prior (Drummond and Rambaut 2007). These include the Yule process (lineage birth only, $\mu = 0$, $\rho = 0$) and birth–death models (lineage birth and death only, $\rho = 0$). BEAST assigns hyperpriors on λ and μ , whereas all BDS parameters are specified by the user in MCMCTREE.

We investigated Dirichlet and BDS priors because different specifications of these priors may be suited to particular node age distributions, for example, when most cladogenesis events are ancient and clustered near the root. The probability density function of the Dirichlet is unimodal, symmetrical, and located centrally when the parameter $\alpha > 1$. This leads to “node repulsion,” which should make divergence times evenly spaced (Kishino et al. 2001). A parabolic density function is defined when $\alpha < 1$, causing nodes to be separated by either long or short time durations, and a uniform function is defined when $\alpha = 1$ (the default value in MULTIDIVTIME). The BDS prior allows greater flexibility. The birth (λ) and death (μ) rates and the sampling proportion (ρ) can be adjusted to reflect a variety of tree shapes (Yang and Rannala 1997; Yang 2006, p. 251).

Estimation of divergence times in shallow phylogenies was first investigated by analyzing simulated data. The influence of the BDS and Dirichlet priors on posterior times was assessed in relation to sequence length.

We also assessed the impact of different priors and different programs on divergence time estimation for 3 real data sets. The data were analyzed using different kernel densities for the BDS prior in MCMCTREE and different values of α for the Dirichlet prior in MULTIDIVTIME and also using the Yule process and birth–death priors in the program BEAST.

METHODS

Simulations

Sequences were generated for 4 trees of 10 taxa each, representing 2 topologies with 2 sets of divergence times each. The trees were the following: 1) balanced with short external branches (BAL-S), 2) balanced with long external branches (BAL-L), 3) unbalanced with short external branches (UNBAL-S), and 4) unbalanced with long external branches (UNBAL-L) (Fig. 1). Rate heterogeneity was simulated by allowing the logarithm of the substitution rate to drift between branches under a Brownian motion process (Thorne et al. 1998; Kishino et al. 2001). The logarithm of the rate of a branch, $\ln r_i$, conditional on the rate r_A on the ancestral branch, was sampled from a normal distribution: $N(\ln r_A - \frac{\nu t}{2}, s^2)$ (Kishino et al. 2001; Aris-Brosou and Yang 2003). The variance of the rate-drift process is s^2 , where $s^2 = \nu t$. We used $\nu = 0.01$ because it led to a rate density concentrated within the region 0.05–0.015 substitutions/site/unit time, which would appear reasonable. Trees with more clock-like evolution ($\nu = 0.001$) were also generated/analyzed, but results were similar and thus not shown. The rate at the root node was fixed at 0.01 substitutions/site/unit time. When 1 time unit = 1 Ma, our simulated sequences approximated rapidly evolving mitochondrial DNA (mtDNA) genes (e.g., cytochrome b) from a phylogeny with a Miocene root.

Branch lengths were calculated for each tree by multiplying the time duration (t) by the rate, and then, 10 tip sequences were generated using EVOLVER from the PAML package (version 4, Yang 2007). This was done using the F84 model (Hasegawa and Kishino 1989) with base frequencies 0.30 [T], 0.25 [C], 0.30 [A], 0.15 [G]; a transition to transversion ratio $\kappa = 5$; and a gamma shape parameter $\alpha = 0.5$ in order to simulate typical mtDNA sequences. One hundred single locus data sets were generated for each of the following sequence lengths: 250, 500, 1000, 5000, and 10,000 bp for each of the 4 trees.

Simulated data were analyzed using MCMCTREE and MULTIDIVTIME. All 4 trees were analyzed with maximal and minimal constraints on the root (Fig. 1). The hard constraints in MULTIDIVTIME were specified from the uniform distribution $U(8.1, 9.9)$. Soft constraints were specified in MCMCTREE as $'> 8.1 = 9.0 < 9.9'$, which defined the 2.5th percentile, mean, and 97.5th percentile of the gamma prior. The effect of constraint position was assessed by simulating and analyzing an additional 100 data sets for the UNBAL-S tree using the alternative constraints $U(0.45, 0.55)$ in MULTIDIVTIME or $'> 0.45 = 0.5 < 0.55'$ in MCMCTREE on the most

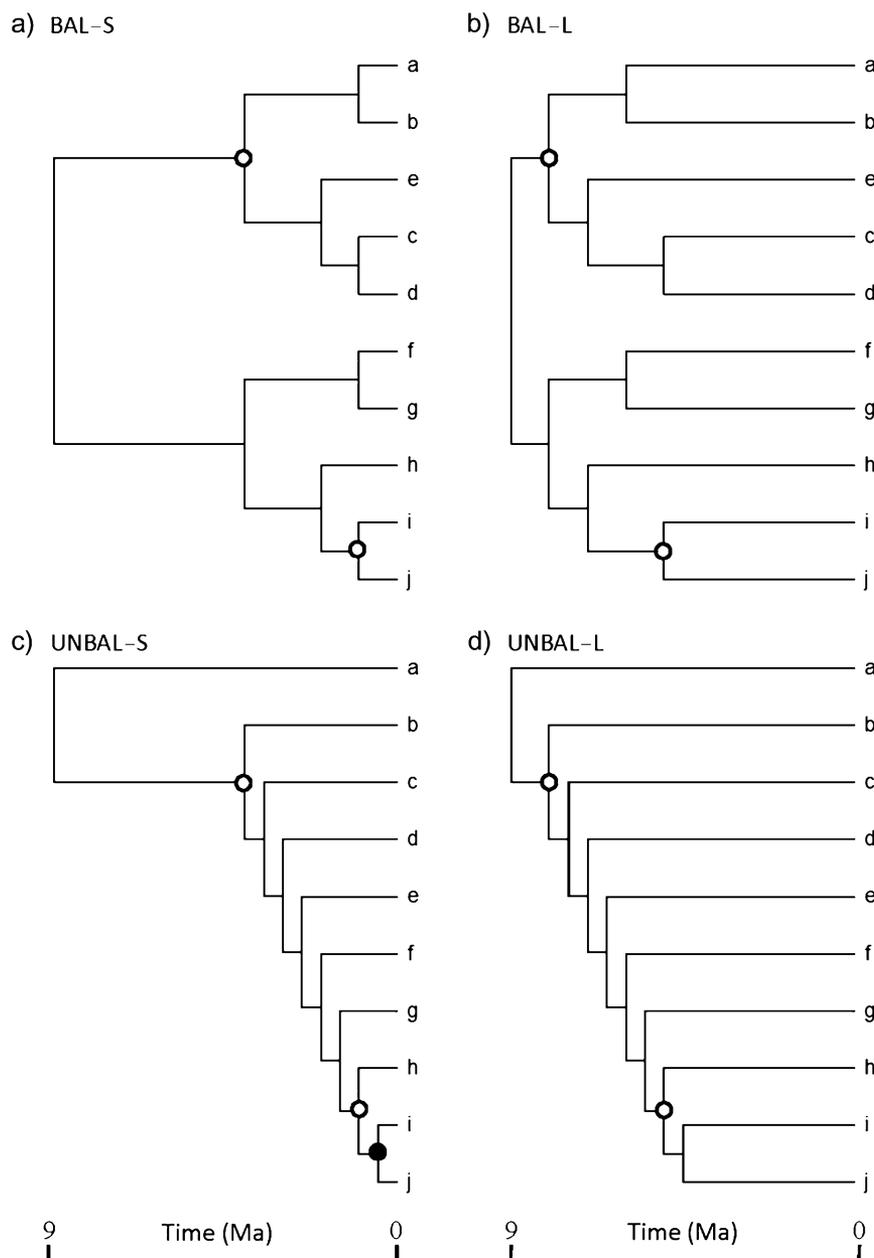


FIGURE 1. The 4 trees that were used to simulate sequence data. Time units were equivalent to Ma before present. The simulated data were analyzed with soft or hard minimal and maximal constraints of 8.1 and 9.9 time units, respectively, specified for the roots of all trees. Additional analyses of sequences simulated for c) were carried out with constraints of 0.45 and 0.55 time units on the most recent node (●). Results were analyzed in detail for 2 nodes on each tree (shown as ○).

recent internal node (Fig. 1). MULTIDIVTIME also requires a gamma prior on the root. We used a gamma distribution with a mean of 9 units and a standard deviation (SD) of 3. The Bayesian MCMC analyses differ between the 2 programs in several other ways, as described below.

A flexible BDS prior was used with $\lambda=2$, $\mu=2$, $\rho=0.1$. This defines a kernel density characterized by an exponential decline in density from 0 to 9 time units. It is likely to be most appropriate for the trees with short external branches (BAL-S and UNBAL-S), which have

high densities of younger nodes. The substitution rate at the root and the hyperparameter, ν , were gamma-distributed with means set equal to the values used to simulate the phylogenies (SDs were set equal to means). The correct substitution model was used to analyze the data. Posteriors from replicated trial runs were analyzed to determine the lengths and sampling intervals for the MCMC chains and consistency of results. Approximately 5–10% of final analyses failed to converge due to poor starting positions. These were diagnosed from the posterior distributions and then repeated.

In the MULTIDIVTIME analyses, the times prior on noncalibrated nodes was specified from a Dirichlet distribution with $\alpha = 1$. The prior on substitution rates was essentially the same as that described for the MCMCTREE analyses, although the Brownian motion process models rate drift over nodes but not over branches (Kishino et al. 2001). Another difference from the MCMCTREE analysis is that the likelihood is calculated by a multivariate normal approximation of the branch lengths. First, the program BASEML was used to obtain ML estimates of branch lengths under the F84+G model for the unrooted topology (Fig. 1) with an additional outgroup (unlike MCMCTREE which does not use an outgroup; Yang 2007). The variance-covariance matrix for branch lengths in the ingroup rooted tree was then calculated using the program ESTBRANCHES. Finally, divergence times were estimated using the MULTIDIVTIME program.

Prior node age distributions were obtained to allow evaluation of their impact on the posterior. Exact evaluation is complex, so approximations were obtained by running the MCMC without data.

Real Data Sets

Asian Laudakia lizards.—Data originated from a study of the radiation of the *Laudakia caucasia* species group from the Iranian plateau region (Macey et al. 1998) and provided a phylogeny with a 5–10 Ma calibration on the root (Fig. 3a). The original authors used the data to obtain a well-supported phylogeny and estimate divergence times from uncorrected distances. Sequences from protein-coding regions (ND1, ND2, and COI genes: 1155 bp) were partitioned by codon position (3 partitions), whereas tRNA regions (535 bp) were assigned to a fourth partition. The mitochondrial proteins perform similar functions and are under similar selective constraint, but the 3 codon positions have very different substitution rates because of the strong selective constraint on the protein (see, e.g., Kumar 1996 for rate estimates in vertebrate mitochondrial genes). We thus partition sites by codon position rather than by genes. Sites with ambiguous alignment and indels were deleted. The F84+G model was fitted to each partition. This is the most general model implemented in MCMCTREE and accounts for major features of the evolutionary process, such as transition/transversion rate difference, unequal nucleotide frequencies, and rate variation among sites. The sequence alignment and topology used for the analyses are available in TreeBASE (matrix accession M4789).

The sensitivity of the posteriors to different time priors was tested using different programs. Data were divided into 4 partitions according to codon position and tRNA. The F84+G model of sequence evolution was used for each partition, with the following parameters where possible (important differences are specified later): The shape (α) parameter for site heterogeneity was assigned the gamma prior G(0.5, 1), and the priors for the κ parameter (Ts:Tv rate ratio) and the overall

rate were specified as G(5, 1) and G(1, 100), respectively. The rate-change parameter, ν , was also specified as G(1, 100). (Numbers in parentheses refer to the shape, α , and scale, β , parameters of a gamma distribution with mean = α/β and variance = α/β^2 .) Maximal (10 Ma) and minimal (5 Ma) bounds were specified on the root age. The programs/priors were the following: 1) MCMCTREE with the 3 BDS parameters, λ , μ , ρ , specified as follows: i) 5, 5, 1, ii) 5, 3, 0.001, iii) 5, 5, 0.1, and iv) 5, 4.5, 0.001 (kernel densities are shown in Fig. 4); 2) MULTIDIVTIME with the Dirichlet parameter, α , specified as follows: i) 0.3 (parabolic density function), ii) 1.0 (flat density function), and iii) 2.0 (unimodal density function); and 3) BEAST (fixed topology mode) with the prior on times specified from the following: i) the Yule process with a lineage birth rate generated from the uniform distribution U(0, 5) or ii) the birth-death model with the birth minus death rate parameter generated from U(0, 5) and death rate/birth rate parameter specified as U(0, 1). In MULTIDIVTIME, the rttm and rttmsd control variables for root age were specified as 7.5 and 3.0, respectively. In BEAST, the HKY+G model of sequence evolution was used with an uncorrelated lognormal model of rate change (mean 0.022, SD 0.1).

The suitability of the priors on divergence times can only be assessed relative to information on the true divergence times, which are unknown. Hence, we compared priors against a priori estimates of true node ages calculated from simple clock-like divergence times. First, an ML tree was obtained under the molecular clock using BASEML for the 10 taxa and the outgroup *Laudakia lehmanni* (also from Macey et al. 1998). Second, branch durations were calculated by multiplying branch lengths by the rate calculated for an ingroup root age of 7.5 Ma, corresponding to the mean of the 2 constraints on this node. Divergence times for internal nodes were calculated directly from this clock-like tree and compared with corresponding priors.

The Caprinae.—Data were analyzed from a study of mtDNA relationships among the Caprinae, a clade with a mid-late Miocene origin (Lalueza-Fox et al. 2005). The original data comprised the cytochrome b gene (1140 bp) from 26 Caprine species. Some modifications were made to the original data set: 1 taxon was excluded because the entire gene was not available (*Capra pyrenaica*) and 2 other sequences were updated with more recent sequences containing fewer ambiguities, that is, the *Ammotragus lervia* and *Capricornis sumatrensis* sequences were replaced by GenBank accessions NC009510 and DQ459334, respectively. A total of 12 sites were excluded from the analyses due to the presence of ambiguities.

Data were partitioned according to codon position and then divergence times were estimated in the same way as described for the *Laudakia* data, except for node calibrations. Constraints were applied to 2 nodes (Fig. 3b). First, a maximal bound of 14 Ma was applied to the root node, based on the oldest proposed date for the initial divergence of the Caprinae (see Lalueza-Fox et al. 2005). In MULTIDIVTIME, the rttm

and *rttmsd* control variables for the prior on root age were also specified as 6.2 and 3.1. Second, the divergence time of the extinct *Myotragus balearicus* from the Balearic Islands was constrained. Divergence of this species from the *Ovis* (sheep) clade appears to have been initiated by range fragmentation after the refilling of the Mediterranean basin at the end of the Messinian salinity crisis (Lalueza-Fox et al. 2005). This allows a tight calibration because the connection between the Mediterranean and the Atlantic was reestablished 5.33 million years ago (e.g., Krijgsman et al. 1999; Meijer and Krijgsman 2005). We accordingly specified maximal and minimal constraints of 5.38 and 5.28 Ma, respectively. Suitability of priors was assessed in the way described for *Laudakia*: They were compared with simple clock-like estimates of divergence times that assumed a time of 5.33 Ma for the (*Ovis* and *Myotragus*) ancestor. The sequence alignment and topology are available in TreeBASE (matrix accession M4788).

North African Chalcides skinks.—Carranza et al. (2008) established phylogenies of North African skinks to infer the historical biogeography of the *Chalcides* and *Sphenops* genera. Their data comprised sequences from the cytochrome b (396 bp), 12S rRNA (392 bp), and 16S rRNA (537 bp) genes. They used the colonization of the island of El Hierro (from La Gomera) by *Chalcides coeruleopunctatus* soon after its appearance ~ 1.12 Ma (Guillou et al. 1996; Brown et al. 1998) to calibrate the (La Gomera, El Hierro) node. This provides an example data set with external constraints on a very recent node (Fig. 3c).

Individuals from the Western *Chalcides* group that had been sequenced for all 3 genes were selected for analysis, with representatives from all major clades. *Eumeces algeriensis algeriensis* was used as an outgroup, where necessary. Indels were removed leaving 1305 bp of sequence, which was partitioned by gene and codon position (5 partitions). Divergence times were estimated in the same way as described for the *Laudakia* analyses. The (La Gomera, El Hierro) node was calibrated with minimal and maximal bounds of (1.0, 1.2). The root was assigned a maximum age of 20 Ma in MCMCTREE/MULTIDIVTIME (which greatly exceeds the previous estimate of 7.2 Ma; Carranza et al. 2008) and an expected age and SD of 7.2 and 3.6, respectively, in MULTIDIVTIME. The sequence alignment and topology are available in TreeBASE (matrix access M4787).

RESULTS

Simulations

We focus on 2 selected nodes on each simulated tree for conciseness. One of these is the node adjacent to the root of age 8 Ma on the BAL-L and UBAL-L trees and 4 Ma on the BAL-S and UBAL-S trees. The other is a recent node of age 5 Ma and 1 Ma on the -L (long external branch) and -S (short external branch) trees, respectively (Fig. 1).

In the MCMCTREE analyses of the 2 -L trees, the respective 5 Ma nodes were not included within their 95% prior intervals. As a result, posterior means obtained with 250-bp sequence significantly underestimated the true age of these nodes, on average (Fig. 2b,d). Although there is no theory to predict the coverage probabilities of Bayesian credibility intervals in our simulations (which were frequentist simulations), we examined the frequencies at which the Bayesian credibility intervals (CIs) included the true values. In both -L trees, the 95% posterior intervals contained the true node age in fewer than 50 of 100 analyses (Table 1). A greater proportion of analyses captured the true age as sequence lengths were increased. For example, 95 (BAL-L) and 72 (UNBAL-L) intervals contained the true age in the 1 kbp analyses. Nevertheless, the priors continued to influence node age estimates for the UNBAL-L tree for longer sequence lengths: the posterior mean is less than the true age for the 5 Ma node even with 10-kbp sequence. The prior intervals on the 8 Ma nodes in the UNBAL-L (Fig. 2d) and BAL-L (Fig. 2b) trees were wider and included their true age. The prior effect was more moderate as a result. For example, 94 (UNBAL-L) or all 100 (BAL-L) analyses of the 1-kbp data sets yielded 95% posterior intervals that contained the true age (Table 1). Overall, these results show that specification of a BDS prior that favors short external branch lengths will have a significant effect on divergence time estimates for trees with long external branches.

The BDS prior is more suitable for the -S trees (Fig. 2a,c). Prior means are very close to the true age (1 Ma) of the recent nodes on both trees and so generally lead to accurate estimation of age, irrespective of sequence length. The 4 Ma nodes on these trees are overestimated by the prior means, particularly for UNBAL-S, but the effect on the posterior is relatively small, possibly because the true age is included within the prior interval. For example, the 95% posterior intervals include the true age for between 96–98 of the 250 bp and 500 bp analyses on these trees. For 5 and 10 kbp of sequence, an unexpected effect was that the posterior intervals incorporated the true node ages in a slightly lower proportion of analyses (83–95) compared with corresponding analyses on long external branch trees (93–100) (Table 1). Also, the 95% posterior interval contained the true clade age in only 20 of the 500 bp analyses of the BAL-S tree.

The divergence time prior specified by the Dirichlet distribution in MULTIDIVTIME showed a similar pattern of influence on the posterior (Fig. 2a–d), with some differences that we outline here. The Dirichlet distribution with $\alpha = 1$ was slightly more suitable than the BDS (2, 2, 0.1) prior for the -L trees, enabling better estimation of divergence times for UNBAL-L and BAL-L trees for short sequence lengths. Conversely, the Dirichlet was less suitable than the BDS prior for the -S trees. Dirichlet prior intervals were wider than those for the BDS prior on recent nodes but narrower on older nodes. This had a corresponding influence on the widths of the 95% posterior intervals, although the effect was

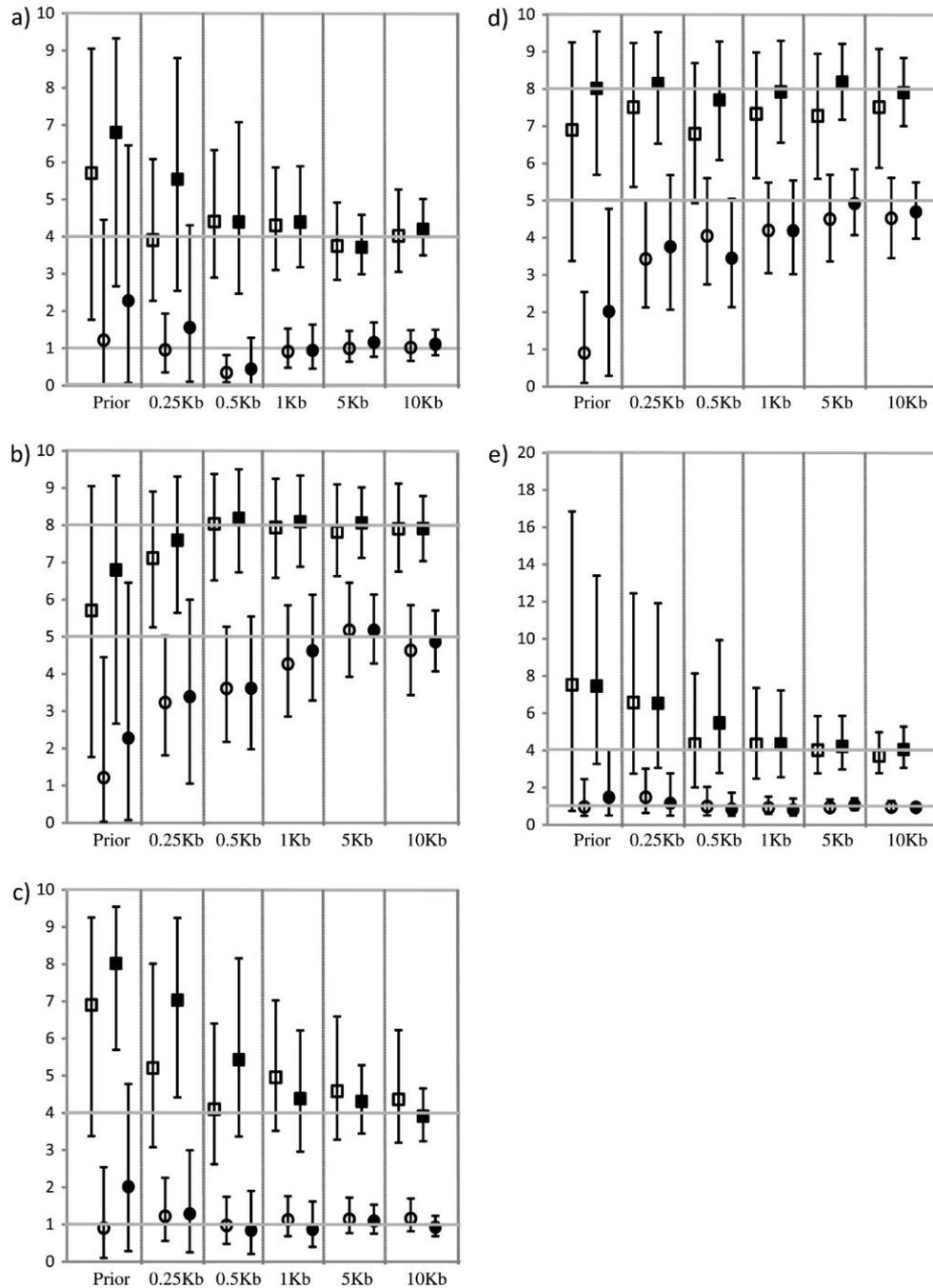


FIGURE 2. Summaries of simulation results for the 4 trees with maximal and minimal bounds on the 9 Ma root: a) BAL-S, b) BAL-L, c) UNBAL-S, d) UNBAL-L, and e) the UNBAL-S tree with maximal and minimal bounds on the 0.5 Ma node. Prior intervals are shown for each node. The means of the 100 posterior means for each sequence length (0.25–10 kbp) sequence are displayed as circles for the younger node and squares for the older node, with open symbols for MCMCTREE analyses and closed symbols for MULTIDIVTIME. Posterior intervals (95%) were also obtained from the means of 100 analyses. Horizontal lines represent the true node ages.

negligible for >500-bp sequence. Also, the posterior intervals were narrower on all nodes (on average) for the 5 and 10 kbp MULTIDIVTIME analyses compared with MCMCTREE. Both these effects were predicted because bounds are hard (and therefore narrower) in MULTIDIVTIME. However, we also found that the 95% posterior intervals included the true age of the nodes in

a slightly lower proportion of MCMCTREE analyses, in general.

Analyses of UNBAL-S trees with constraints on the 0.5 Ma node (Fig. 2e) provided much wider prior intervals on the 4 Ma node and narrower intervals on the 1 Ma node compared with root-constrained trees. This led to corresponding differences in posterior intervals

TABLE 1. Analyses of simulated phylogenies

Tree	Node age	MCMCTREE					MULTIDIVTIME				
		0.25	0.5	1	5	1	0.25	0.5	10	5	10
BAL-S/9	4	98	98	92	83	84	100	100	100	98	99
	1	100	20	100	94	95	100	93	100	97	96
BAL-L/9	8	96	100	100	100	100	100	100	100	100	100
	5	48	65	95	98	99	86	73	100	100	97
UNBAL-S/9	4	96	98	73	88	91	14	88	100	87	96
	1	99	100	100	87	88	100	100	100	98	91
UNBAL-L/9	8	100	85	94	96	97	100	100	100	100	100
	5	47	81	84	93	94	83	54	99	100	96
UNBAL-S/0.5	4	99	100	100	96	99	100	100	100	100	100
	1	100	100	100	96	100	100	100	100	100	100

Notes: MCMCTREE and MULTIDIVTIME analyses of 5 sequence lengths (0.25–10 kbp) for different trees (BAL-S, BAL-L, UNBAL-L, and UNBAL-S) with either an 8.1–9.9 Ma constraint on the root (/9) or a 0.45–0.55 Ma constraint on a recent node (/0.5). One hundred simulations were carried out for each program/tree/sequence length. The table gives the frequencies with which a true node age was included within the 95% posterior intervals. Results are shown for 2 nodes (Fig. 1).

widths, for shorter sequence lengths, and was common to both MULTIDIVTIME and MCMCTREE. Higher proportions of posterior intervals also enclosed the true node ages with a recent constraint in these analyses, with the minimum proportion being 96 of 100 analyses.

Real Data Sets

Divergence times in Laudakia.—For this and other real data sets, we analyze the Bayesian MCMC results in detail for just 4 selected nodes (Table 2) and omit most details of prior intervals for brevity. Different programs gave different estimates of these node ages. Most MCMCTREE and MULTIDIVTIME analyses provided posterior means that were lower and 95% posterior intervals that showed only marginal or no overlap (except for the root) with raw divergence estimates of Macey et al. (1998). In contrast, the original estimates were often lower than the posterior means provided by BEAST, although they were enclosed by the wide posterior intervals generated by this program (Table 2).

A simple analysis of clock-like divergence times indicated a high density of younger nodes in this phylogeny (Fig. 4). The 2 BDS priors with the same birth and death parameter values ($\lambda = 5$, $\mu = 5$, $\rho = 0.1$ and $\lambda = 5$, $\mu = 5$,

$\rho = 0.001$) specified kernel densities that appeared quite suitable for this distribution (Fig. 4). Despite this, the 95% prior intervals specified by $\lambda = 5$, $\mu = 5$, $\rho = 0.001$ only included the simple clock estimate for Node 1 (the root). The intervals specified by the $\lambda = 5$, $\mu = 5$, $\rho = 0.1$ prior were most appropriate and included the clock estimates for Nodes 1–3. The priors specified by $\lambda = 5$, $\mu = 3$, $\rho = 0.001$ and $\lambda = 5$, $\mu = 4.5$, $\rho = 0.001$ provided ostensibly unsuitable kernel densities (Fig. 4) and led to relatively narrow prior intervals that did not enclose any clock-like estimates for nonroot nodes. Posterior means differed among some BDS priors, although the means and 95% posterior intervals were very similar for Nodes 2–4 in the two $\lambda = \mu = 5$ analyses (Table 2). The posterior intervals showed considerable overlap between suitable and unsuitable priors for all nodes. All BDS analyses, except for $\lambda = 5$, $\mu = 4.5$, $\rho = 0.001$, provided 95% posterior intervals that included the raw divergence estimates from the original study (for all nodes).

Prior intervals in MULTIDIVTIME for $\alpha = 2$ or $\alpha = 1$ did not include any of the divergence time estimates obtained under a clock (for nonroot ages). However, prior interval width increased as α decreased (a pattern that was observed in all real data sets), which led to simple clock estimates being included within the prior intervals for Nodes 1, 3, and 4 when $\alpha = 0.3$. Posterior distributions were generally very similar for different values of

TABLE 2. Analyses of divergence times in *Laudakia*

Node	Published estimate	MCMCTREE				MULTIDIVTIME			BEAST	
		$\lambda = 5$, $\mu = 5$, $\rho = 0.001$	$\lambda = 5$, $\mu = 3$, $\rho = 0.001$	$\lambda = 5$, $\mu = 5$, $\rho = 0.1$	$\lambda = 5$, $\mu = 4.5$, $\rho = 0.001$	$\alpha = 0.3$	$\alpha = 1$	$\alpha = 2$	Yule process	Birth–death
1	9	7.98 (5.20–10.05)	9.53 (8.10–10.30)	8.61 (5.91–10.13)	5.59 (4.71–7.73)	7.73 (5.31–9.86)	7.71 (5.30–9.84)	7.71 (5.27–9.86)	7.75 (5.38–10.00)	7.77 (5.37–10.00)
2	3–4	2.68 (1.61–3.99)	3.38 (2.54–4.36)	2.70 (1.73–3.76)	1.94 (1.37–2.89)	1.85 (1.09–2.84)	2.02 (1.21–3.08)	2.29 (1.35–3.54)	5.32 (2.04–8.56)	4.83 (1.68,8.36)
3	2–3	1.58 (0.93–2.40)	2.08 (1.49–2.81)	1.57 (0.97–2.25)	1.15 (0.78–1.75)	1.09 (0.59–1.75)	1.20 (0.67–1.93)	1.38 (0.77–2.22)	3.30 (1.12–5.99)	2.86 (0.83–5.47)
4	1–2	0.77 (0.42–1.23)	1.06 (0.65–1.58)	0.76 (0.46–1.15)	0.58 (0.36–0.91)	0.35 (0.08–0.74)	0.47 (0.19–0.89)	0.60 (0.28–1.08)	1.90 (0.51–3.81)	1.57 (0.38–3.35)

Notes: Bayesian estimates (posterior means, with 95% credibility intervals in parentheses) obtained using different priors in three different programs are compared with published estimates that were obtained from uncorrected distances (Macey et al. 1998). Positions of Nodes 1–4 on the phylogeny are shown in Figure 3a.

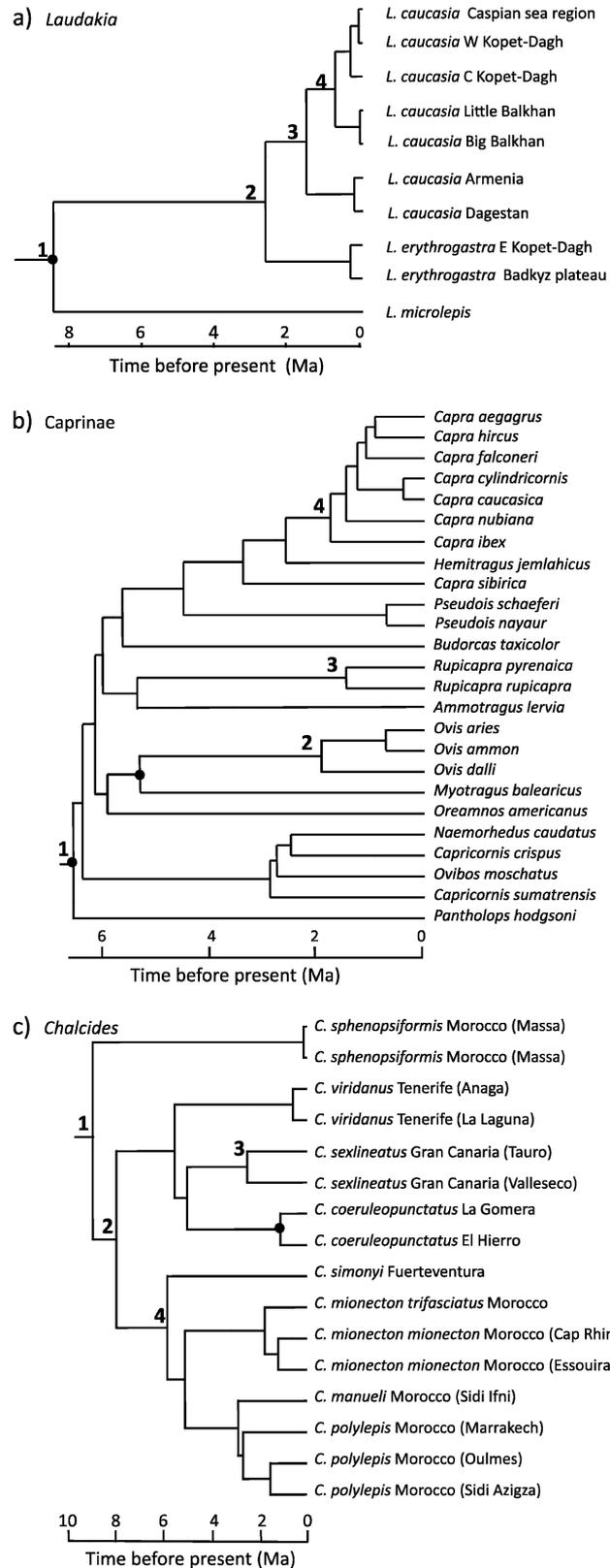


FIGURE 3. a) The *Laudakia* phylogeny, drawn as a chronogram to show the results of the Bayesian MCMCTREE analysis using the most appropriate BDS prior ($\lambda = 5, \mu = 5, \rho = 0.1$). Date estimates are given for Nodes 1–4 in Table 2. b) A Caprinae chronogram for the same analysis and times prior. See Table 3 for date estimates for Nodes 1–4. c) The same analysis of Western *Chalcides*. Date estimates for Nodes 1–4 are given in Table 4. External constraints were placed on nodes marked with filled circles.

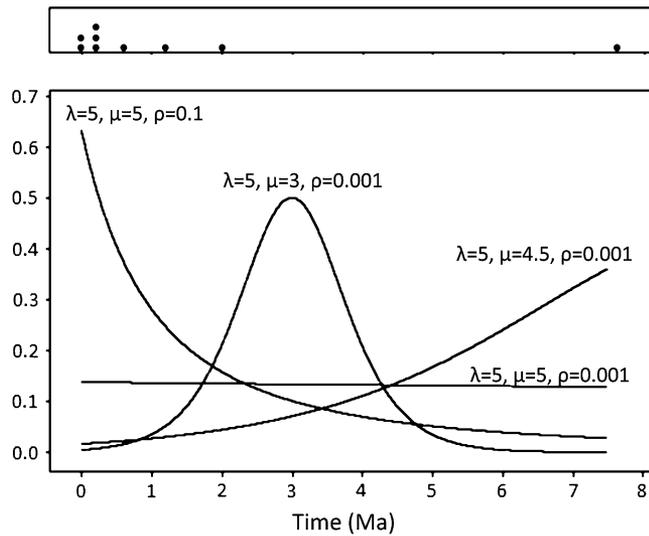


FIGURE 4. Kernel densities for different BDS parameter values (root age of 7.5 Ma) (lower graph). A dot plot of simple clock-like estimates of divergence times in *Laudakia* is provided for comparison (upper graph).

α , although an increase in α was associated with a small increase in divergence time estimates for nonroot ages. Posterior means and intervals were mostly lower than the most suitable BDS prior (and most similar to one of the less suitable MCMCTREE analyses). Posterior intervals on some nonroot nodes included the age estimates published by Macey et al. (1998) when $\alpha = 0.3$ or $\alpha = 1$.

Here and for the other real data sets, the Yule process and birth–death models specified similar prior distributions in BEAST. The 95% prior intervals were wide and enclosed simple clock estimates for all nodes for the birth–death prior but only for Node 2 and the root age for the Yule process. With the exception of the root, 95% posterior intervals were wider for BEAST analyses than for MULTIDIVTIME or MCMCTREE. They enclosed the estimates published by Macey et al. (1998) for all nodes. Posterior means for Nodes 2–4 were notably higher than for other programs and estimates of Macey et al. (1998), particularly when the prior was specified by the Yule process.

Divergence times in the Caprinae.—The simple clock-like analysis showed that node ages are quite evenly distributed between the root and the tips, with a slightly higher density of younger nodes (Fig. 5). Divergence time estimates for Nodes 2–4 were similar among the different analyses, with considerable overlap of the posteriors. The estimated root age showed greater variation among different analyses (Table 3).

As for *Laudakia*, the 2 BDS priors with equal birth and death rates ($\lambda = \mu = 5$) appeared to provide most suitable kernel densities in the MCMCTREE analyses, with the $\lambda = 5, \mu = 5, \rho = 0.1$ prior again being the most appropriate because it specified 95% prior intervals that enclosed the simple clock estimates of age for all nodes (Fig. 5). The $\lambda = 5, \mu = 3, \rho = 0.001$ and $\lambda = 5, \mu = 4.5, \rho = 0.001$ priors were clearly inappropriate because they

specified 95% intervals that did not enclose clock estimates (except for the root). This explains the different results obtained in these analyses (Table 3). Posterior means from the most suitable BDS prior were generally closest to the original estimates, obtained using ML with a global clock.

Prior intervals in the MULTIDIVTIME analyses enclosed the simple clock estimates for all nodes ($\alpha = 0.3$), Nodes 1–3 ($\alpha = 1$), or Nodes 2 and 3 ($\alpha = 2$). This suggests that $\alpha = 0.3$ is also the most suitable Dirichlet prior for the Caprinae data. There were differences in posterior means and 95% intervals for the different priors, which were relatively minor for Nodes 2 and 3 but more substantial for the root (Table 3). Posterior means for the root were considerably lower than those obtained by using MCMCTREE, but the means of the other 3 nodes were similar between the 2 programs.

The Yule process and birth–death prior intervals obtained from BEAST included simple clock estimates for all nodes except for the root. Posterior means for the root were generally higher than those for MULTIDIVTIME and the 2 most suitable MCMCTREE analyses, but posterior means of other nodes were similar to those of the MULTIDIVTIME analyses (Table 3). Posterior intervals were generally wider in the BEAST analyses.

Divergence times in Chalcides.—The *Chalcides* tree had a higher density of younger compared with older nodes, although this tendency was not as pronounced as in *Laudakia*. Most of the divergence time priors led to wide prior and posterior intervals for all programs. Posterior means varied considerably between different programs/priors. As for the Caprinae analyses, the original ML estimates of node ages (Carranza et al. 2008) are useful for comparison.

BDS prior intervals in MCMCTREE analyses were generally wide (extending to 20 Ma) and included simple clock estimates for all nodes, with the exception of the prior specified by $\lambda = 5, \mu = 3, \rho = 0.001$. As for the other 2 data sets, the most appropriate prior was specified by $\lambda = 5, \mu = 5, \rho = 0.1$. Posterior intervals were also very wide (Table 4). The two $\lambda = \mu = 5$ priors provided posterior intervals that enclosed the published ages for all nodes, whereas posterior intervals from the other priors only enclosed Node 3.

The narrowest prior intervals were specified by MULTIDIVTIME analyses. These included simple clock estimates and so appeared suitable for all 4 nodes. Posterior means from these analyses very closely reflected the original ML estimates (Table 4). They showed the opposite trend to the previous 2 data sets inasmuch as posterior means decreased slightly with increasing values of α , except for the root. Posterior distributions of times for Nodes 1, 2, and 4 differed considerably from those in MCMCTREE and BEAST, which is likely due to differences in the method of specifying the prior on the root age (see Discussion section).

Yule process and birth–death prior intervals in BEAST were similar and wider than those specified by other programs. They enclosed the 4 simple clock estimates of

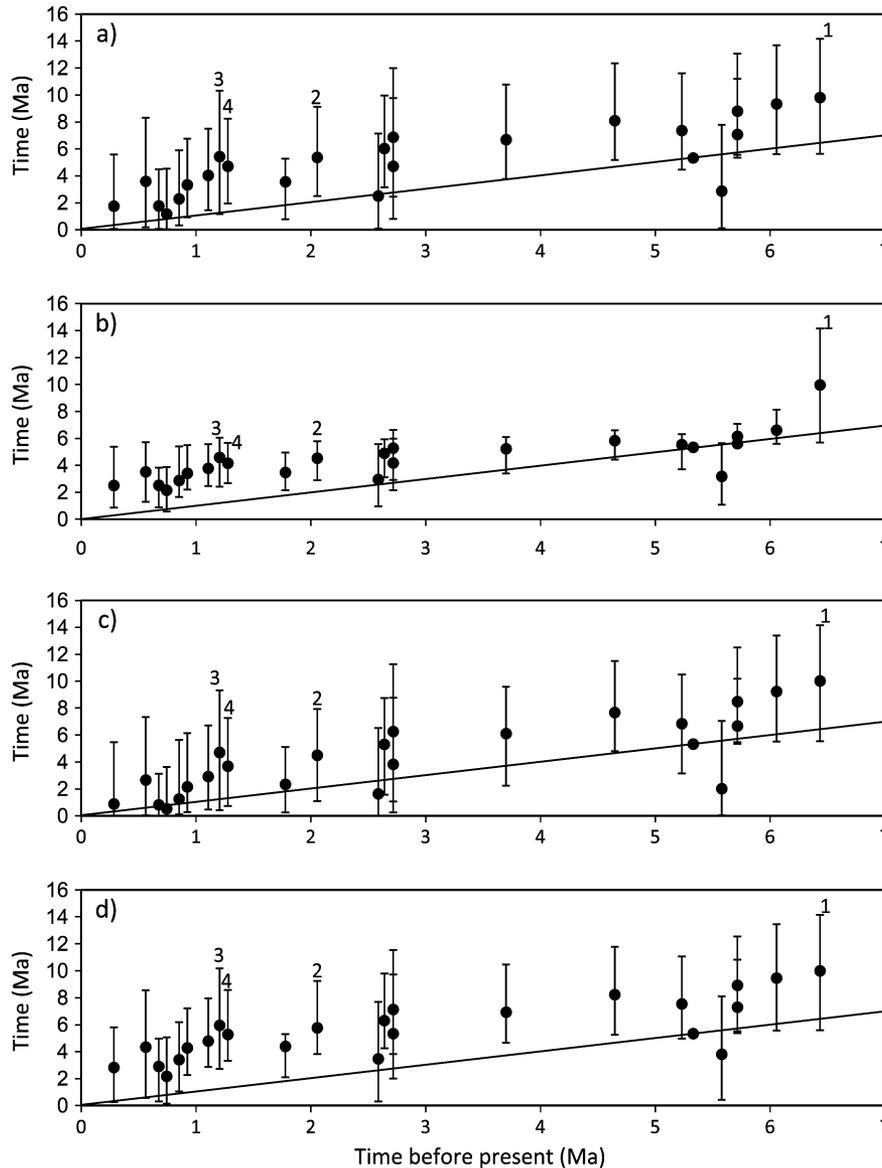


FIGURE 5. Prior means and 95% intervals of times for the Caprinae phylogeny specified by the BDS parameters: a) $\lambda = 5$, $\mu = 5$, $\rho = 0.001$, b) $\lambda = 5$, $\mu = 3$, $\rho = 0.001$, c) $\lambda = 5$, $\mu = 5$, $\rho = 0.1$, and d) $\lambda = 5$, $\mu = 4.5$, $\rho = 0.001$. The x-axis represents simple estimates of divergence times from a clock-like tree. Locations of Nodes 1–4 (Fig. 3b and Table 3) are indicated, as are lines of equality.

divergence times. Posterior means generally exceeded those obtained from MULTIDIVTIME and MCMCTREE (Table 4). However, as for other data sets, 95% posterior intervals were wider for BEAST and therefore showed considerable overlap with results from other programs and the original analysis.

DISCUSSION

The influence of the divergence time prior on the posterior distribution requires careful consideration when shallow phylogenies are dated. Our MCMCTREE and MULTIDIVTIME analyses of simulated data show that posterior mean ages of individual nodes can be strongly influenced by the corresponding priors. The effect is

most significant for sequence lengths <1 kbp but is also present for longer sequences when the 95% prior interval does not include the true node age. In many practical applications, sequences could be less informative than those simulated here due to lower substitution rates (i.e., <0.01 substitutions/site/Ma) and/or a more recent root (i.e., <9 Ma). In these cases, >1 kbp may be required for reliable dating.

Proportions of 95% posterior intervals that enclosed the true node ages provided an alternative perspective on the performance of the programs. Posterior intervals were very wide for 250-bp sequence and covered the true node age for most simulated data sets, except where the priors were inappropriate. Slightly fewer intervals included the true age for the longest sequence lengths,

TABLE 3. Analyses of divergence times in the Caprinae

Node	Published estimate	MCMCTREE				MULTIDIVTIME			BEAST	
		$\lambda = 5, \mu = 5, \rho = 0.001$	$\lambda = 5, \mu = 3, \rho = 0.001$	$\lambda = 5, \mu = 5, \rho = 0.1$	$\lambda = 5, \mu = 4.5, \rho = 0.001$	$\alpha = 0.3$	$\alpha = 1$	$\alpha = 2$	Yule process	Birth-death
1	6.2 (5.5–7.0)	7.35 (5.84–10.00)	7.63 (6.57–9.21)	6.58 (5.58–8.14)	9.31 (7.06–12.26)	5.80 (5.24–6.84)	6.45 (5.63–7.70)	6.96 (6.02–8.35)	7.60 (5.90–9.58)	7.57 (5.92–9.56)
2	2.2 (1.7–2.8)	2.20 (1.42–3.18)	2.34 (1.58–3.14)	1.94 (1.28–2.78)	2.72 (1.77–3.76)	2.58 (1.93–3.28)	2.64 (1.98–3.37)	2.71 (2.05–3.42)	2.63 (1.73–3.64)	2.60 (1.69–3.63)
3	1.6 (1.1–2.1)	1.72 (0.97–2.82)	1.94 (1.15–2.82)	1.47 (0.88–2.27)	2.40 (1.28–3.96)	1.76 (1.17–2.44)	1.80 (1.19–2.53)	1.84 (1.22–2.59)	1.93 (1.00–3.00)	1.91 (0.99–2.91)
4	1.5 (1.2–1.9)	2.13 (1.47–3.11)	2.40 (1.79–3.04)	1.77 (1.28–2.42)	2.96 (1.99–4.18)	1.83 (1.24–2.59)	2.11 (1.45–2.94)	2.36 (1.64–3.24)	2.10 (1.44–2.84)	2.09 (1.44–2.80)

Notes: Bayesian estimates (posterior means, with 95% credibility intervals in parentheses) obtained using different priors in three different programs are compared with published ML estimates and associated confidence intervals from Lalueza-Fox et al. (2005). Positions of Nodes 1–4 on the phylogeny are shown in Figure 3b.

although the proportions were close to the expected value of 0.95. This can be attributed to the decrease in posterior width associated with increasing sequence length (Rannala and Yang 2007). Coverage probabilities were similarly high for suitable and unsuitable priors for 5- to 10-kbp sequences, demonstrating that the increased influence of the likelihood was sufficient to mitigate the effects of inappropriate divergence time priors.

The calibration on the most recent node of the UBAL-S tree gave different prior intervals to the root calibration on the same tree. These intervals were wider for MCMCTREE than for MULTIDIVTIME because the user-specified rootage control variable specifies a soft maximal limit, but no minimal limit, which causes the root age to be only loosely constrained (in the absence of an external calibration; Inoue et al. 2010). In our simulations, there was a substantial difference between the true age (9 units) and the rootage value (20 units), which explains the considerable increase in the upper 97.5% limit of the prior distribution when only a recent node constraint was used in MCMCTREE. Nodes that are further from the root are less severely affected. In MULTIDIVTIME, the root age is more tightly constrained by a user-specified gamma distribution (rttm and rttmsd control variables) and a maximal limit (bigtime control variable), which explains the more moderate effects of constraint position. The influence of the priors is

reduced for ≥ 1 kbp, leading to smaller differences between the posteriors.

Analyses of real data showed that posterior means were sensitive to the time priors. This emphasizes the importance of assessing the suitability of specified time priors, as opposed to simply using a flexible default prior. Bayes factors can be used to select the best overall model, but this does not help assessment of the suitability of the prior on a particular node. Instead, we show that useful assessment of divergence time priors can be achieved by examining the intervals they specify on nodes of interest relative to simple clock-like estimates of divergence time.

Accurate estimation of divergence times in *Laudakia* is impeded by the wide calibration on the root. Prior intervals were wide in BEAST. This appeared to lead to wide posterior intervals that provided little meaningful information on divergence times. Our simulations show that this pattern can be expected when the influence of the likelihood is weak. Divergence times in MCMCTREE and MULTIDIVTIME are more similar (for the suitable priors). Differences between them are attributable to the methods by which they constrain the root age: Slightly tighter constraints in MULTIDIVTIME were due to hard maximal and minimal bounds for the root calibration (as opposed to soft bounds in MCMCTREE). These two analyses indicated that the estimates of Macey et al. (1998) may be a little high, possibly due to the use of

TABLE 4. Analyses of divergence times in *Chalcides* skinks

Node	Published estimate	MCMCTREE				MULTIDIVTIME			BEAST	
		$\lambda = 5, \mu = 5, \rho = 0.001$	$\lambda = 5, \mu = 3, \rho = 0.001$	$\lambda = 5, \mu = 5, \rho = 0.1$	$\lambda = 5, \mu = 4.5, \rho = 0.001$	$\alpha = 0.3$	$\alpha = 1$	$\alpha = 2$	Yule process	Birth-death
1	7.2	8.67 (5.71–13.28)	5.84 (4.85–7.25)	8.85 (6.08–12.78)	5.68 (4.61–7.02)	7.29 (4.33–11.92)	7.45 (4.54–12.03)	7.29 (4.61–11.57)	9.63 (5.20–15.35)	9.38 (5.01–14.63)
2	6.8	8.00 (5.33–12.27)	5.32 (4.46–6.31)	8.04 (5.55–11.54)	5.10 (4.26–6.14)	6.68 (3.99–10.83)	6.60 (4.02–10.66)	6.38 (4.02–10.01)	8.75 (4.61–13.56)	8.50 (4.65–13.15)
3	2.2	2.66 (1.60–4.19)	1.86 (1.29–2.47)	2.47 (1.58–3.76)	1.82 (1.11–2.42)	2.29 (1.11–4.17)	2.18 (1.07–3.90)	2.06 (1.04–3.56)	2.67 (1.29–4.36)	2.58 (1.17–4.17)
4	5.4	6.03 (3.95–9.22)	3.90 (3.22–4.69)	5.75 (3.86–8.35)	3.75 (3.12–4.48)	5.34 (3.07–9.05)	5.31 (3.11–8.78)	5.15 (3.17–8.31)	6.41 (3.51–10.13)	6.23 (3.35–9.71)

Notes: Bayesian estimates (posterior means, with 95% credibility intervals in parentheses) obtained using different priors in three different programs are compared with ML estimates published by Carranza et al. (2008). Positions of Nodes 1–4 on the phylogeny are shown in Figure 3c.

uncorrected distances. The Bayesian approaches that we discuss provided other advantages over the original analysis. In particular, they incorporated the considerable uncertainty concerning the age of the root node and provided CIs on the divergence times. Nevertheless, the divergence times obtained from the suitable Bayesian analyses were still consistent with the causal geophysical events proposed by Macey et al. (1998).

ML estimates of divergence times in the Caprinae by Lalueza-Fox et al. (2005) showed generally close agreement with all posterior means from the most suitable MCMCTREE and MULTIDIVTIME analyses. Posterior mean ages were also similar for BEAST, except for the root. As expected, 95% confidence intervals from the ML analyses were narrower than the Bayesian posterior intervals. Posteriors in BEAST are slightly wider than for the other 2 programs. The root was most tightly constrained in MULTIDIVTIME due to a hard maximal bound from a fossil calibration and a user-specified gamma-distributed root age, whereas only a soft maximal bound constrained the root in MCMCTREE. We suggest that this enables the tighter and possibly more accurate posterior for this node in MULTIDIVTIME.

The similarity between all posterior means and the ML estimates of Carranza et al. (2008) indicates that MULTIDIVTIME provided the most reliable Bayesian analysis of the *Chalcides* data. The data differ from the *Laudakia* and the Caprinae data because only a very recent node is calibrated. Our simulations predicted wide posteriors for older nodes under these conditions (unless the data are highly informative), and this was the pattern observed for all programs. BEAST again provided the widest posterior intervals, but MULTIDIVTIME and MCMCTREE intervals were also wide. Hence, a much larger amount of sequence may be required to obtain more informative posteriors.

Accurate estimation of divergence times in relatively shallow phylogenies is an important component of phylogeography. Despite good justification for applying a global clock when species have recently diverged, we show that accurate dating with a relaxed clock can also provide reliable results. However, assessment of the divergence time prior appears to be an important prerequisite to any relaxed clock analysis. Typical sequence lengths of 1–2 kbp mtDNA may provide reasonably accurate posterior means when the priors on nodes are appropriate but may not always provide usefully narrow posterior intervals. We show that restricting the root with a user-specified gamma distribution may help improve the accuracy of the posterior mean for this and adjacent nodes, particularly when calibrations are only available for the younger nodes on the tree.

FUNDING

R.P.B. was supported by a Higher Education Funding Council for England sabbatical awarded by Liverpool John Moores University. Z.Y. was supported by a Biotechnology and Biological Sciences Research Council grant.

ACKNOWLEDGEMENTS

Bruce Rannala kindly provided source code that was modified to simulate phylogenies. Salvador Carranza and José Castresana helpfully supplied their respective *Chalcides* and Caprinae data.

REFERENCES

- Aris-Brosou S., Yang Z. 2003. Bayesian models of episodic evolution support a late precambrian explosive diversification of the Metazoa. *Mol. Biol. Evol.* 20:1947–1954.
- Avise J.C. 2000. *Phylogeography: the history and formation of species*. Cambridge (MA): Harvard University Press.
- Brown R.P., Pestano J. 1998. Phylogeography of Canary Island skinks inferred from mtDNA sequences. *Mol. Ecol.* 7:1183–1191.
- Brown R.P., Terrasa B., Pérez-Mellado V., Castro J.A., Hoskisson P.A., Picornell A., Ramon M.M. 2008. Bayesian estimation of post-Messinian divergence times in Balearic Island lizards. *Mol. Phylogenet. Evol.* 48:350–358.
- Carranza S., Arnold E.N., Geniez P., Roca J., Mateo J.A. 2008. Radiation, multiple dispersal and parallelism in the skinks, *Chalcides* and *Sphenops* (Squamata: Scincidae), with comments on *Scincus* and *Scincopus* and the age of the Sahara Desert. *Mol. Phylogenet. Evol.* 46:1071–1094.
- Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
- Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Emerson B.C., Oromí P. 2005. Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origins of island endemics. *Evolution*. 59:586–598.
- Grazziotin F.G., Monzel M., Echeverrigaray S., Bonatto S.L. 2006. Phylogeography of the Bothrops jararaca complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. *Mol. Ecol.* 15:3969–3982.
- Guillou H., Carracedo J.C., Perez-Torrado F., Rodriguez Badiola E. 1996. K–Ar ages and magnetic stratigraphy of a hotspot-induced, fast-grown oceanic island: El Hierro, Canary Islands. *J. Volcanol. Geotherm. Res.* 73:141–155.
- Hasegawa M., Kishino H. 1989. Confidence limits on the maximum-likelihood estimate of the hominoid tree from mitochondrial-DNA sequences. *Evolution*. 43:672–677.
- Ho S.Y., Phillips M.J., Drummond A.J., Cooper A. 2005. Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. *Mol. Biol. Evol.* 22:1355–1363.
- Inoue J., Donoghue P.C.J., Yang Z. 2010. The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Syst. Biol.* 59:74–89.
- Kishino H., Thorne J.L., Bruno W.J. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol. Biol. Evol.* 18:352–361.
- Krijgsman W., Hilgen F.J., Raffi I., Sierro F.J., Wilson D.S. 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature*. 400:652–655.
- Kumar S. 1996. Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. *Genetics*. 143:537–548.
- Lalueza-Fox C., Castresana J., Sampietro L., Marques-Bonet T., Alcover J.A., Bertranpetit J. 2005. Molecular dating of caprines using ancient DNA sequences of *Myotragus balearicus*, an extinct endemic Balearic mammal. *BMC Evol. Biol.* 5:70.
- Lepagec T., Bryant D., Philippe H., Lartillot N. 2007. A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* 24:2669–2680.
- Macey J.R., Schulte J.A., Ananjeva N.B., Larson A., Rastegar-Pouyani N., Shammakov S.M., Papenfuss T.J. 1998. Phylogenetic relationships among agamid lizards of the *Laudakia caucasia* species group: testing hypotheses of biogeographic fragmentation and an area cladogram for the Iranian Plateau. *Mol. Phylogenet. Evol.* 10:118–131.

- Meijer P., Krijgsman W. 2005. A quantitative analysis of the desiccation and re-filling of the Mediterranean during the Messinian Salinity Crisis. *Earth Planet. Sci. Lett.* 240:510–520.
- Melville J., Hale J., Mantziou G., Ananjeva N.B., Milto K., Clemann N. 2009. Historical biogeography, phylogenetic relationships and intraspecific diversity of agamid lizards in the Central Asian deserts of Kazakhstan and Uzbekistan. *Mol. Phylogenet. Evol.* 53:99–112.
- Rannala B., Yang Z. 2007. Bayesian estimation of species divergence times from multiple loci using multiple calibrations. *Syst. Biol.* 56:453–466.
- Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.
- Schmitz P., Cibois A., Landry B. 2007. Molecular phylogeny and dating of an insular endemic moth radiation inferred from mitochondrial and nuclear genes: the genus *Galagete* (Lepidoptera: Autostichidae) of the Galápagos Islands. *Mol. Phylogenet. Evol.* 45:180–192.
- Thorne J.L., Kishino H. 2002. Divergence time estimation and rate evolution with multilocus data sets. *Syst. Biol.* 51:689–702.
- Thorne J.L., Kishino H., Painter I.S. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15:1647–1657.
- Warren B.H., Bermingham E., Bowie R.C.K., Prys-Jones R.P., Thébaud C. 2003. Molecular phylogeography reveals island colonization history and diversification of western Indian Ocean sunbirds (Nectarinia: Nectariniidae). *Mol. Phylogenet. Evol.* 29:67–85.
- Wüster W., Ferguson J.E., Quijada-Mascareñas J.A., Pook C.E., Salomão M.G., Thorpe R.S. 2005. Tracing an invasion: landbridges, refugia and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Mol. Ecol.* 14:1095–1108.
- Yang Z. 2006. *Computational molecular evolution*. Oxford: Oxford University Press.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586–1591.
- Yang Z., Rannala B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol. Biol. Evol.* 14:717–724.
- Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23:212–226.
- Yang Z., Yoder A.D. 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cutelooking mouse lemur species. *Syst. Biol.* 52:705–716.