Syst. Biol. 52(5):705–716, 2003
Copyright © Society of Systematic Biologists
ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150390235557

Comparison of Likelihood and Bayesian Methods for Estimating Divergence Times Using Multiple Gene Loci and Calibration Points, with Application to a Radiation of Cute-Looking Mouse Lemur Species

ZIHENG YANG[1] AND ANNE D. YODER[2]

[1]*Department of Biology, University College London, Darwin Building, Gower Street, London WC1E 6BT, United Kingdom; E-mail: z.yang@ucl.ac.uk*
[2]*Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, New Haven, Connecticut 06520, USA*

Abstract.—Divergence time and substitution rate are seriously confounded in phylogenetic analysis, making it difficult to estimate divergence times when the molecular clock (rate constancy among lineages) is violated. This problem can be alleviated to some extent by analyzing multiple gene loci simultaneously and by using multiple calibration points. While different genes may have different patterns of evolutionary rate change, they share the same divergence times. Indeed, the fact that each gene may violate the molecular clock differently leads to the advantage of simultaneous analysis of multiple loci. Multiple calibration points provide the means for characterizing the local evolutionary rates on the phylogeny. In this paper, we extend previous likelihood models of local molecular clock for estimating species divergence times to accommodate multiple calibration points and multiple genes. Heterogeneity among different genes in evolutionary rate and in substitution process is accounted for by the models. We apply the likelihood models to analyze two mitochondrial protein-coding genes, cytochrome oxidase II and cytochrome *b*, to estimate divergence times of Malagasy mouse lemurs and related outgroups. The likelihood method is compared with the Bayes method of Thorne et al. (1998, Mol. Biol. Evol. 15:1647–1657), which uses a probabilistic model to describe the change in evolutionary rate over time and uses the Markov chain Monte Carlo procedure to derive the posterior distribution of rates and times. Our likelihood implementation has the drawbacks of failing to accommodate uncertainties in fossil calibrations and of requiring the researcher to classify branches on the tree into different rate groups. Both problems are avoided in the Bayes method. Despite the differences in the two methods, however, data partitions and model assumptions had the greatest impact on date estimation. The three codon positions have very different substitution rates and evolutionary dynamics, and assumptions in the substitution model affect date estimation in both likelihood and Bayes analyses. The results demonstrate that the separate analysis is unreliable, with dates variable among codon positions and between methods, and that the combined analysis is much more reliable. When the three codon positions were analyzed simultaneously under the most realistic models using all available calibration information, the two methods produced similar results. The divergence of the mouse lemurs is dated to be around 7–10 million years ago, indicating a surprisingly early species radiation for such a morphologically uniform group of primates. [Bayes method; divergence times; evolutionary rate; local molecular clock; maximum likelihood; molecular clock.]

The molecular clock hypothesis (Zuckerkandl and Pauling, 1965) has provided a powerful means for estimating species divergence times. However, the clock is often violated, especially when distantly related species are compared. A number of studies have demonstrated that divergence time estimation is highly sensitive to assumptions about the evolutionary rate (e.g., Takezaki et al., 1995; Yoder and Yang, 2000). One approach to the problem is to identify and exclude genes and/or species that appear to have caused violation of the clock so that the clock model can be applied to the rest of the data. This strategy makes an inefficient use of the data if many genes or many species violate the clock. Another approach is to try to accommodate variable rates among lineages when divergence times are estimated. There has been much interest recently in this approach, and both likelihood and Bayes methodologies have been employed.

In a maximum likelihood (ML) method, separate rate parameters can be assumed for lineages that show rate changes, and such rates can be estimated jointly with the divergence times. Kishino and Hasegawa (1990) assigned different transition and transversion rates to individual lineages on a primate tree to estimate both rates and times using a normal approximation to the observed numbers of transitions and transversion in pairwise sequence comparisons. The quartet-dating method of Rambaut and Bromham (1998) implements a specific rate model on a tree of four species. Yoder and Yang (2000) extended the model to the general case, with an arbitrary assignment of rate classes to branches. All of these methods work on one gene and take one calibration point. In a Bayes method, a stochastic model of evolutionary rate change is used to specify the prior distribution of rates, and the Bayes theorem is used to derive the posterior distributions of rates and times. A model of stochastic rate change was proposed by Sanderson (1997, 2002) in his semiparametric rate-smoothing method for divergence time estimation. A formal Bayes approach was implemented by Thorne et al. (1998) and Kishino et al. (2001), using a Markov chain Monte Carlo (MCMC) procedure for efficient computation. A similar algorithm was implemented by Huelsenbeck et al. (2000), who attempted to model rate evolution within a branch and between branches. The Thorne et al. algorithm was recently extended to accommodate multiple gene loci (Thorne and Kishino, 2002), which can have separate evolutionary models. Aris-Brosou and Yang (2002) tested the effects of different prior models of rate change on the posterior estimation of divergence times.

In molecular sequence comparisons, the probability of the data depends on the distances or branch lengths but not on rates and times individually. Thus, rate and time

are intrinsically confounded (Felsenstein, 1981). With the relaxation of the molecular clock assumption, time estimation becomes tricky (e.g., Yoder and Yang, 2000). There is too much flexibility in the model in fitting data of one gene, because a high rate and short time might fit the data as well as a low rate and a long time. It is possible to overcome this problem by analyzing multiple gene loci simultaneously and by using multiple calibration dates. The rates might vary in different ways among multiple loci, but the divergence times are shared, and the internal constraints in the model might lead to reliable estimation of divergence times even when the clock is violated in every gene.

In this article, we extend the local clock models of Yoder and Yang (2000) to accommodate multiple calibration points and to analyze multiple gene loci simultaneously while accounting for their differences in the evolutionary process. We apply the new models to estimation of mouse lemur divergence times and compare results with the Bayes method of Thorne et al. (1998) and Thorne and Kishino (2002).

The mouse lemurs are nocturnal primates endemic to Madagascar and are the world's smallest living primates. Until the late 1970s, only one species, *Microcebus murinus*, had been recognized since its original description in 1795 (Geoffroy Saint-Hilaire, 1795). Further research prompted recognition of two distinct forms: the long-eared gray *M. murinus* from the western regions of Madagascar and the short-eared reddish *M. rufus* from the east (Martin, 1972). Recent morphological (Rasoloarison et al., 2000) and molecular (Yoder et al., 2000) phylogenetic analyses demonstrated that as many as nine species exist and, more strikingly, that they form a north–south rather than west–east divide (Yoder et al., 2000). Divergence times among those species are unknown. By inferring the times of speciation events and correlating them with possible geological and environmental events, we hope to gain a better understanding of the speciation mechanisms of mouse lemurs and other terrestrial vertebrates of Madagascar.

Our main objective in this article is methodological, i.e., to explore the advantages of combined analysis of heterogeneous data sets for date estimation when the molecular clock is violated. We are also interested in exploring the differences between the likelihood and Bayes approaches to the problem, the effects of nucleotide substitution models, and the differences among the three codon positions. Given that the data being analyzed are from a single genetic partition (the mitochondrion), the power of our analyses lies more in the realm of methodological comparison than in confident estimation of geological ages for mouse lemur divergences. This latter issue will be dealt with more fully in subsequent work.

Data and Methods

Sequence Data

We use two protein-coding genes, cytochrome oxidase II (COII) and cytochrome *b*, from the mitochondrial genome from 9 mouse lemur species as well as 26 other mammalian species (Yoder et al., 2000). The sequences are identified in the tree of Figure 1 (GenBank accession numbers given in Table 1). The sequence alignment has been submitted to TreeBase (study accession number S912 and matrix accession number M1510) and is also included in the PAML package (Yang, 1997), which implements the likelihood models described below. Since the two genes are on the same strand of the genome and have similar evolutionary dynamics (Kumar, 1996), we combined them into one data set. We removed the stop codons, leaving 1,812 nucleotides in the alignment. However, the three codon positions have drastically different substitution rates and base compositions. Thus, we analyzed them initially as three separate data sets and later as one combined data set with their heterogeneity accommodated. Some basic statistics for the three codon positions are listed in Table 2. We assume that the tree topology is known and fixed. This is shown in Figure 1, which also highlights the calibration nodes and major nodes of interest for which we present estimated divergence dates.

No fossil data are available for ancestral species closely related to the mouse lemurs because of the paucity of

TABLE 1. GenBank accession numbers for sequences used in this study.

Taxon	COII	Cyt *b*
Didelphis	Z29573	Z29573
Equus	X79547	X79547
Rhinoceros	X97336	X97336
Balaenoptera	X61145	X61145
Physeter	AJ277029	AJ277029
Hippopotamus	AJ010957	AJ010957
Bos	J01394	J01394
Canis	U96639	U96639
Ursus	AF303109	AF303109
Felis	U20753	U20753
Homo	J01415	J01415
Pan	X93335	X93335
Gorilla	D38114	D38114
Pongo	D38115	D38115
Macaca	M74005	U38272
Callithrix	AY321457	AF295586
Galago	M80905	U53579
Loris	AY321458	U53581
Daubentonia	L22776	U53569
Lepilemur	AY321459	AY321456
Propithecus	L22782	U53573
Varecia	L22785	U53578
Hapalemur	L22778	U53574
Lemur	L22780	U53575
Eulemur	AF081045	AF081051
Cheirogaleus	L22775	U53570
Mirza	AY321460	U53571
M. ravelobensis	AF285493	AF285529
M. sambiranensis	AF285518	AF285554
M. rufus 2	AF285516	AF285552
M. tavaratra	AF285497	AF285533
M. rufus 1	AF285515	AF285551
M. berthae	AF285504	AF285540
M. myoxinus	AF285499	AF285535
M. griseorufus	AF321180	AF285567
M. murinus	AF285522	AF285558

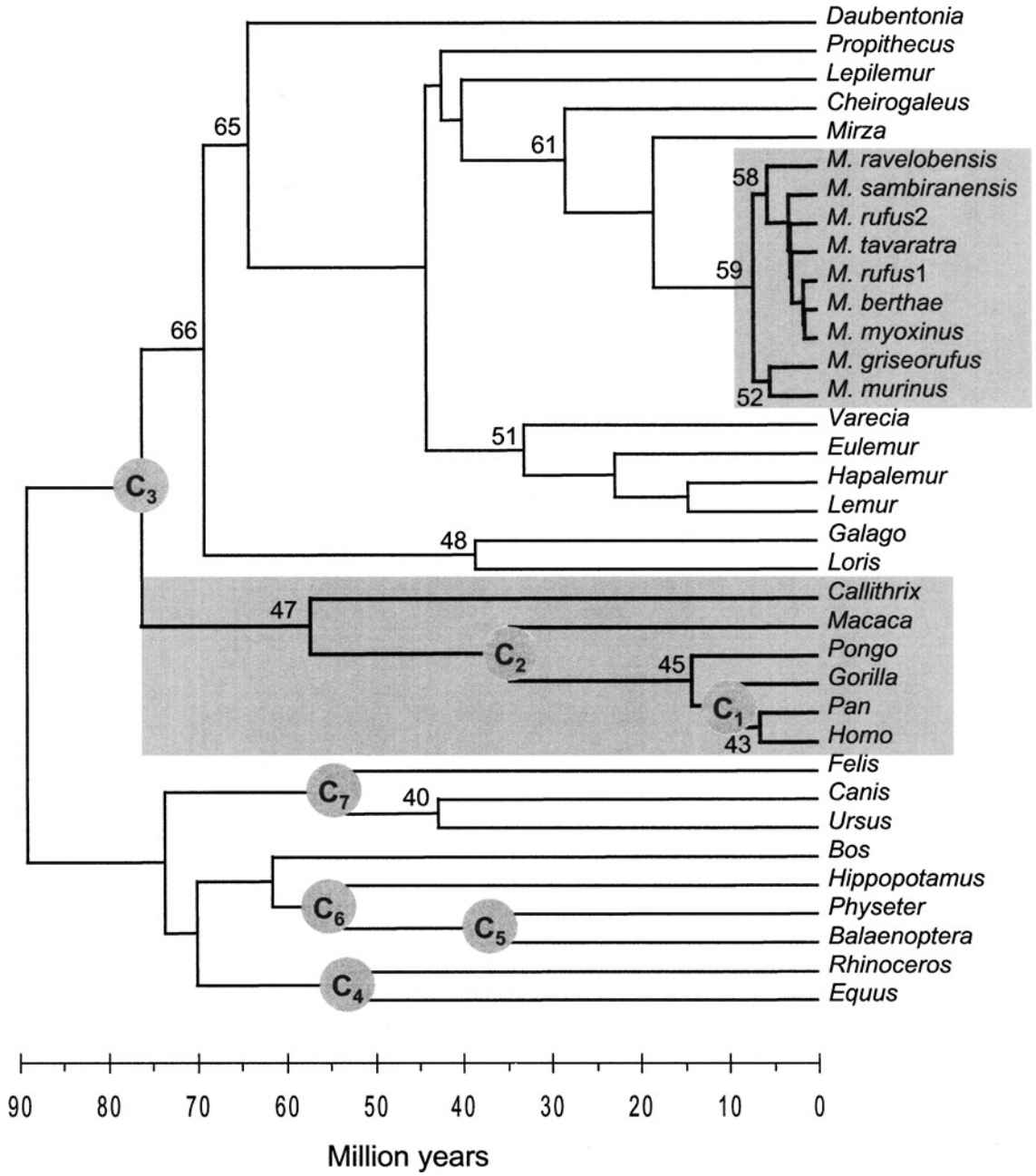

FIGURE 1. Rooted tree topology assumed in estimation of divergence times in this study. Three calibration points within primates (C_1–C_3) and four calibration points outside primates (C_4–C_7) are used (Table 3). Some nodes of interest are labeled using the numbering system of Thorne's programs, for which the divergence date estimates are presented in Tables 4, 5, and 7. In the likelihood local-clock models, branches in the anthropoid clade are assigned rate r_1, mouse lemurs rate r_2, while all remaining branches have the default rate r_0. A few variations of this three-rates model are examined in the text. Branches are drawn to reflect the divergence times estimated in the ML analysis of the combined data set of three codon positions (analysis j, Table 5; F84G), with dates at the seven calibration nodes fixed at the midvalues of Table 3.

the terrestrial fossil record for the Tertiary in Madagascar. Fossil calibration information is available for seven ancestral nodes distributed among the outgroup species, all of which are employed in our analyses. In the Bayes method, lower and upper bounds can be specified for the ancestral node ages, with ranges given in Table 3. Calibrated nodes include the divergence between toothed and baleen whales (Thewissen, 1994), whales and hippos (Thewissen, 1994), horses and rhinoceros (Prothero and Schoch, 1989; Janis et al., 1998), felids and canids (Flynn, 1996), humans and gorillas (Shoshani et al., 1996), monkeys and apes (Shoshani et al., 1996; Yoder and Yang, 2000), and the basal radiation of primates (Martin, 1993; Gingerich and Uhen, 1994; Tavaré et al. 2002). In the ML method, only a fixed date for each calibration node is accepted (see below). We used the midvalue in the table.

TABLE 2. Parameter estimates[a] under the F84G model (Felsenstein, 2002) at the three codon positions.

Position	L	π_T	π_C	π_A	π_G	κ	α	Tree length	$2\Delta\ell$
1	604	0.23	0.26	0.29	0.22	3.49	0.29	3.39	2,479.3
2	604	0.40	0.25	0.22	0.13	2.81	0.17	1.23	1,122.9
3	604	0.21	0.34	0.40	0.04	18.81	1.17	55.11	5,602.7
All	1,812	0.28	0.28	0.31	0.13	4.45	0.28	10.96	12,124.2

[a] L is the number of nucleotide sites. Tree length is the sum of branch lengths, i.e., the number of substitutions per site throughout the tree. $2\Delta\ell$ is twice the log-likelihood difference between the JC69 (Jukes and Cantor, 1969) and F84G models. The unrooted tree includes the 35 ingroup species of Figure 1 as well as the outgroup *Didelphis*. Estimates from the ingroup species only are virtually identical.

Nucleotide Substitution Models

Both the likelihood and Bayes analyses can be performed under a variety of nucleotide substitution models (see Whelan et al., 2001, for a review). In this article, we use two models in both methods: the JC model (Jukes and Cantor, 1969), which assumes that all rates of nucleotide substitution are equal, and the F84G model, which accounts for the transition/transversion rate bias and unequal base compositions (from the DNAML program of Felsenstein, 2002) and uses a discrete gamma model with five rate classes to accommodate variable rates among sites (Yang, 1994). JC appears to have been rejected in every DNA sequence data set in which it was tested and is much worse than F84G for the data analyzed here (Table 2). However, we used it as a contrast to F84G for examining the robustness of the results to assumptions about the substitution process. From previous studies, we expect rate variation among sites to have much greater effect on divergence time estimation than the transition/transversion rate bias or unequal base frequencies (e.g., Yang et al., 1994; Sullivan and Swofford, 2001), although we do not attempt to examine the individual effects of those factors here.

Likelihood Models of Global and Local Clocks with Multiple Calibrations

We describe first our implementation of the global clock model when multiple calibration points are used and then describe extensions to local clock models and to data of multiple genes. Our focus is on the rate and time parameters. When calibration information is available for an ancestral node in the tree, we assume that the node age is known without error and thus is fixed. An alternative approach is to use lower and upper bounds to constrain the node age, in which case the node age is still a free parameter but its estimate will often be at the boundary. We did not pursue this approach here. Parameters in our model include the mutation (substitution) rate μ and the ages of nodes that are not calibration points. Suppose there are s species and $s-1$ ancestral nodes in the rooted tree and that calibration dates are available for c nodes. The parameters in the model are then the substitution rate μ and the $(s-1)-c$ ages for the nodes without calibration information, with a total of $s-c$ parameters in the model. For example, the tree of five species in Figure 2 has four ancestral node ages: t_1, t_2, t_3, and t_4. Suppose fossil data are available for node ages t_2 and t_3, so that t_2 and t_3 are known. There are then three parameters to be estimated: the substitution rate μ and the node ages t_1 and t_4. Note that μ is the absolute rate, measured by the expected number of substitutions per site per time unit. The time parameters have to satisfy the constraints that the age of any node is not older than the age of its mother node; i.e., $t_1 > \max(t_2, t_3)$, and $0 < t_4 < t_2$. Numerical optimization for ML estimation of parameters has to be performed under such constraints. Algorithms used in the PAML package deal with simple bounds on parameters but not general inequality constraints. Thus, the following variable transform is used. For each node age t_j to be estimated, all its descendent nodes are examined, and the oldest calibration date among them is located. This is the youngest age for the node, t_L; i.e., $t_j \geq t_L$. Let t_A be the age of the mother (ancestral) node. We define a new

TABLE 3. Calibration dates[a] for nodes in Figure 1 (millions of years).

Node	Range	Midvalue
Within primates		
C_1 human/gorilla	8–12	10
C_2 monkey/ape	32–38	35
C_3 basal primates	63–90	77
Outside primates		
C_4 horse/rhinoceros	50–58	54
C_5 toothed/baleen whales	33–40	37
C_6 whale/hippo	51–60	56
C_7 felid/canid	45–65	55

[a] The range is used in Bayes analysis as lower and upper bounds for the node ages, while the midvalue is used as a fixed constant in the ML analysis.

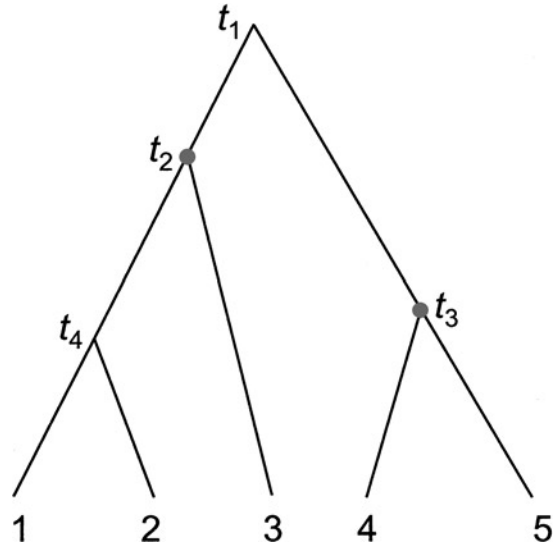

FIGURE 2. Example tree of five species used to explain the implementation of global and local clock models in the ML method. Fossil dates are available for node ages t_2 and t_3, so that these two dates are fixed, while divergence times t_1 and t_4 are estimated by ML through numerical optimization, together with the substitution rate μ.

variable $x_j = (t_j - t_L)/(t_A - t_L)$ or $t_j = t_L + (t_A - t_L)x_j$, so that $0 \le x_j \le 1$. If the age of the root is to be estimated (i.e., if it is not a calibration node), the age is used without transform. Thus, for the example of Figure 2, the parameters used are μ, t_1, and $x_4 = t_4/t_2$, which have to satisfy the following constraints (bounds): $0 < \mu < \infty$, $\max(t_2, t_3) \le t_1 < \infty$, and $0 \le x_4 \le 1$.

Local clock models (Yoder and Yang, 2000) relax the clock assumption by allowing some branches on the tree to have independent rates. Instead of one rate μ, we use k rates, and each branch in the tree can take one of those k possible rates. For example, the trees of Figure 3 show that the mouse lemurs and the anthropoids tend to have very different rates from other lineages in the tree and can be assigned independent rates. We suspect that slight overparametrization, i.e., assigning lineages that share a similar rate to different rate groups, will not cause a serious problem to date estimation, while forcing branches with very different rates into one rate group may lead to biased date estimates.

The length of any branch in the tree is given by the product of the time duration of the branch and the rate for the branch. Given the branch lengths, the likelihood function is calculated using the sequence alignment according to Felsenstein (1981) for models of one rate for all sites and according to Yang (1994) for models of variable rates among sites. Rate and time parameters, as well as parameters in the substitution model are then estimated by maximizing the likelihood. The curvature of the likelihood surface can be used to calculate the approximate variances of ML estimates (Stuart et al., 1999:60–62). The global and local clock models are also implemented for viral sequences with known dates (Rambaut, 2000; Drummond et al., 2001), where the ages of the tips serve as calibration points while the ages of ancestral nodes are estimated by ML.

Accounting for Different Substitution Processes in Multiple Genes

Our description here uses the term *genes* to refer to partitions of sites in the sequence alignment. We stress that site partitioning is for accommodating large-scale heterogeneity among sites in the evolutionary dynamics and does not have to correspond exactly to genes. For the mitochondrial data set analyzed, the differences between the two genes (COII and cytochrome *b*) are small but huge differences exist among the codon positions. Thus, the term *genes* below corresponds to the three codon positions in our data. The general model implemented here assumes that the divergence times and calibration points are shared among genes, while all other parameters involved in the evolutionary process are free to vary. Following Yang (1996b), we assume different genes have different overall substitution rates, different base frequencies and transition/transversion rate ratios, and separate gamma distributions to describe the rate variation among sites within each gene.

Under the local-clock models, we again assume that the branches are classified into k rate classes. We implemented two versions of the local-clock model for multiple genes in the computer program. The first version assumes that the local branch rates are proportional for different genes. Suppose there are g genes. The model then involves $k + g - 1$ rate parameters: k branch rates for the first gene, and $g - 1$ relative rates for the other genes. The model assumes that branches with high rates in one gene have high rates in other genes as well. This assumption is clearly violated in the data analyzed in this study (Fig. 3), and thus this model is not used in the analysis of this paper. The second version allows the branch rates to be estimated separately for different genes, so that there are kg rate parameters.

Bayes Estimation of Divergence Times Using MCMC

The three codon positions were initially analyzed as separate data sets, using a Bayes MCMC package written by Jeff Thorne (Thorne et al., 1998). A new version of the package (Thorne and Kishino, 2002) was used to analyze the three codon positions simultaneously while accounting for their differences in substitution process. The first program in the packages, estbranches, produces ML estimates of branch lengths for the ingroup rooted tree and their approximate variance–covariance matrix. For this purpose, a substitution model should be specified and information about parameters in the substitution model should be provided. We used the baseml program in the PAML package to obtain estimates of the transition/transversion rate ratio κ and the rates for site classes under the discrete-gamma model of rates among sites under the F84G model (Table 2). Those estimates were used as input to the estbranches program. The Bayes program also requires an outgroup clade to locate the root in the ingroup tree, for which we used *Didelphis viginiana*, a marsupial. Note that this outgroup sequence is not used in the ML analysis. The second program in Thorne's packages (divtime5b for one gene and multidivtime for multiple genes) conducts the Bayes MCMC analysis to approximate the posterior distributions of substitution rates and divergence times. The likelihood function is approximated using a multivariate normal distribution of estimated branch lengths and is not calculated from the sequence alignment (Thorne et al., 1998). Thus, the substitution model does not enter the second stage of the analysis.

The prior for substitution rates is specified by a recursive procedure that proceeds from the root of the tree toward the tips. The log rate of the current node or branch follows a normal distribution centered around and conditioned on the log rate of the ancestral node or branch (Thorne et al., 1998; Kishino et al., 2001). The variance of the distribution is given by $\nu\Delta t$, the time duration Δt between the nodes or branches multiplied by parameter ν, which controls how variable the rates are over time (Thorne et al., 1998). The joint prior for all rates is thus given by a gamma density for the rate at the root multiplied by all those conditional densities. The prior for divergence times is specified using another recursive procedure (Kishino et al., 2001), starting from the root and

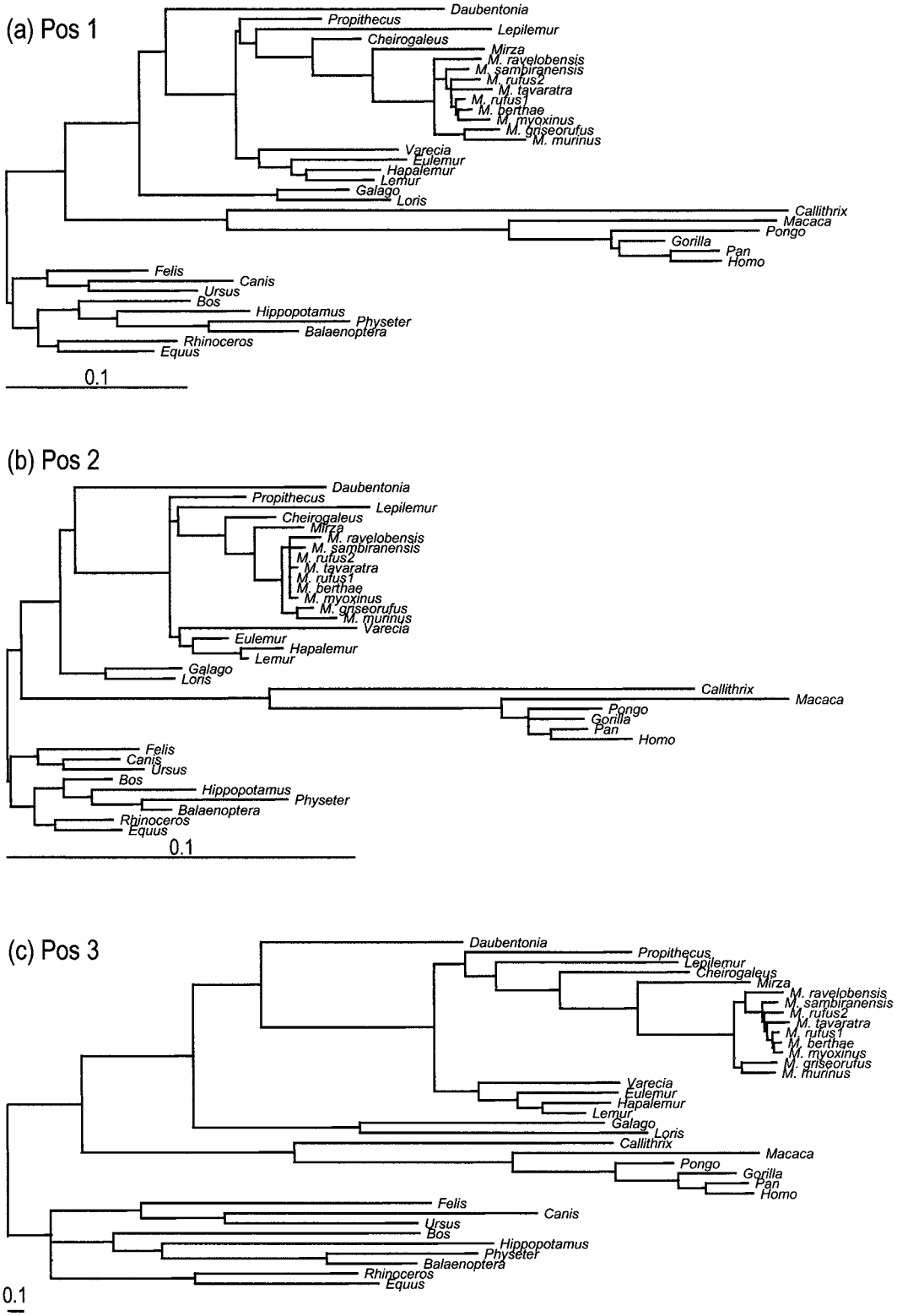

FIGURE 3. Branch lengths estimated under the F84G model for the three codon positions without assuming the molecular clock. The tree is unrooted, and the root is shown for clarity.

moving toward the tips. A gamma density is used for the age of the root. Then each path from an ancestral node to the tip is broken into random segments, corresponding to branches on the path, by using a Dirichlet density with equal probabilities. When fossil calibration information is available for an ancestral node, it is incorporated as constraints on the node age. Several calibration nodes are used (Table 3, Fig. 1).

The MCMC programs require specification of two gamma prior distributions, for the age of the root and for the rate at the root, which is also the overall prior rate on the phylogeny. We found those two priors hard to specify without reference to the sequence data being analyzed; in theory there seems to be no need for specifying those priors since the sequence data and the fossil calibrations should determine the overall rate and the age of the root. We used $m = 70$ million years (MY) and $\sigma = 35$ MY for the mean and SD of the gamma prior for the root date. We used the ML estimate of the substitution rate under the global clock as the prior mean for the rate and set σ to be half of the mean, so that the gamma shape parameter $\alpha = (m/\sigma)^2 = 4$. Since both the rate and the age are strictly positive, we believe it is sensible to require the gamma distribution to have a strictly positive mode, i.e., $\alpha > 1$. However, for the data analyzed in this study, relatively large changes to those prior parameters made little difference to the posterior. As a result, our concerns above did not have much practical relevance. An additional gamma prior is specified for parameter ν, and we used $m = 0.4$ and $\sigma = 0.8$. Since the different codon positions clearly have different patterns of rate variation (Fig. 3), we use different ν values for different codon positions (i.e., commonbrown is set to 0).

Convergence of the MCMC algorithm was monitored mainly through multiple independent runs for each analysis. Stable results, with the difference in posterior means of divergence times <0.1 MY (or 0.5 MY at most) between runs, were obtained using a burn-in of 20,000 iterations and then sampling every 5 iterations, with 20,000 samples taken (a total of 20,000 + 20,000 × 5 iterations). Longer chains were run for a few analyses and were found to produce very similar results.

RESULTS

Likelihood Analysis under the Global and Local Clock Models

ML estimates of divergence times for the major nodes in the tree of Figure 1 obtained under the global molecular clock are listed in Table 4. The JC (Jukes and Cantor, 1969) and F84G (Yang, 1994; Felsenstein 2002) models of nucleotide substitution are used. Divergence times at the seven calibration nodes (Fig. 1) are fixed at the midvalues in Table 3, and times for other nodes are estimated by ML, as are other parameters in the model. When the three codon positions are analyzed separately (analyses b–d for JC and f–h for F84G in Table 4), date estimates are highly variable among codon positions and between the two models. For example, under JC estimated mouse lemur divergence dates range from 11 to 24 MY. Under F84G, those dates are more similar to each other, from 7 to 14 MY. When all three positions are analyzed together without accounting for their differences (analyses a and e [all] in Table 4), the estimates are averages over the three positions. The combined analysis (analyses i and j in Table 4) uses all three codon positions as well, but accounts for their differences. Under JC, the model accounts for differences in substitution rate at the three codon positions, while under F84G the model also accounts for differences in the transition/transversion rate bias, in base compositions, and in the gamma shape parameter (Yang, 1996b). The mouse lemur divergence date is estimated to be 21 MY under JC and 9 MY under F84G. Both of these age estimates are surprisingly old, with the date estimated under JC to be unrealistically ancient.

Likelihood ratio tests of the clock assumption indicates clear rejection of the global clock for all three codon positions under both the JC and F84G models (results not shown). Figure 3 shows the estimated branch lengths for the three codon positions under the F84G model. At the first and second positions, the mouse lemur lineages seem to have relatively high rates, while the anthropoids

TABLE 4. ML estimates of divergence dates under the global clock using multiple calibrations.

	JC				F84G				Combined	
Node	(a) all[a]	(b) pos1	(c) pos2	(d) pos3	(e) all	(f) pos1	(g) pos2	(h) pos3	(i) JC	(j) F84G
40 dog/bear	44.9	38.7	31.5	51.9	43.3	37.2	29.1	42.6	45.9	39.2
43 human/chimp	8.4	6.7	8.3	9.1	8.0	6.8	8.4	6.9	8.5	7.4
45 hominoid	22.6	20.9	16.6	24.1	17.8	20.2	15.8	15.8	22.3	17.4
47 anthropoid	59.1	63.0	59.3	56.8	61.6	65.4	60.5	57.9	60.0	61.8
48 lorisiform	41.4	27.2	31.7	49.2	36.1	23.6	28.8	42.8	40.8	33.9
51 Lemuridae	41.4	38.8	36.7	40.8	31.1	34.3	35.8	23.7	40.6	28.3
52 southern clade	15.3	11.0	7.7	16.7	9.3	9.6	7.0	5.7	14.8	6.9
58 northern clade	18.7	13.7	6.9	20.3	10.8	11.5	5.4	5.8	17.9	7.4
59 mouse lemurs	22.5	16.2	11.2	24.4	13.1	13.8	10.0	7.3	21.7	9.4
61 Cheirogaleidae	39.1	33.4	24.0	42.7	30.7	30.0	22.5	26.5	38.9	27.0
65 lemuriform	62.5	60.3	61.6	62.5	59.1	59.5	64.0	53.8	62.2	58.0
66 Strepsirrhine	66.7	63.8	66.4	67.2	63.0	62.6	64.0	65.8	66.1	63.4

[a]All = concatenated sequences including all three codon positions.

TABLE 5. ML estimates of divergence dates under local clock models with three rates.

	JC				F84G				Combined[a]			
Node	(a) all	(b) pos1	(c) pos2	(d) pos3	(e) all	(f) pos1	(g) pos2	(h) pos3	(i) JC	(j) F84G	(j*) pri	(j**) out
40 dog/bear	46.4	41.8	36.3	52.8	46.4	43.0	36.6	44.4	47.7	43.2	34.9	45.3
43 human/chimp	8.2	6.2	7.8	9.1	7.3	5.6	7.6	6.5	8.3	6.5	6.4	7.6
45 hominoid	20.1	17.7	13.1	21.8	14.5	14.6	10.5	14.7	19.7	14.2	14.1	16.6
47 anthropoid	53.6	58.0	54.3	50.5	58.0	61.7	54.8	54.7	54.4	57.6	58.2	74.8
48 lorisiform	44.8	31.0	38.7	52.1	41.1	29.3	36.9	45.6	44.6	38.9	34.0	48.6
51 Lemuridae	44.7	45.0	44.8	43.3	36.3	43.7	46.2	25.8	44.5	33.3	27.8	39.7
52 southern clade	8.3	7.6	6.4	9.0	6.4	7.9	6.4	3.5	8.5	5.2	4.3	6.1
58 northern clade	10.1	9.3	5.4	11.0	7.4	9.0	4.7	3.6	10.3	5.5	4.6	6.5
59 mouse lemurs	12.7	11.4	9.6	13.6	9.1	11.4	9.3	4.5	12.7	7.1	5.9	8.4
61 Cheirogaleidae	34.9	33.8	26.6	37.3	31.1	33.7	26.9	25.7	35.6	28.6	23.9	34.0
65 lemuriform	65.3	66.4	67.4	63.9	64.8	69.0	71.7	56.7	65.7	64.8	57.4	81.8
66 Strepsirrhine	69.7	69.1	71.7	69.4	68.8	71.8	71.7	68.8	69.7	69.9	63.5	90.8

[a]Analysis j* used only the three calibrations within primates, and j** used only the four calibrations outside primates.

have very high rates. At the third position, the rates are more homogeneous among lineages, although mouse lemur rates are high. We use information about branch lengths to formulate a local clock model, in which the hominoids and the mouse lemur lineages have separate rate parameters r_1 and r_2, while all other "background" branches have the default rate r_0 (see Fig. 1). To examine the sensitivity of date estimation on assumptions about rates, we also considered a few variations of the three-rates model: a two-rates model with one rate for the anthropoids and another rate for all other branches, a four-rates model with an additional rate for branches in the clade (*Hippopotamus*(*Physeter*, *Balaenoptera*)), and a five-rate model with a further rate for the *Daubentonia* branch (see Fig. 3).

ML estimates of divergence times under the local clock model with the three branch rates described above are shown in Table 5. The estimated substitution rates are listed in Table 6. Compared with the time estimates under the global clock (Table 4), estimates under the local clock are considerably less variable among the three codon positions. For example, the estimated mouse lemur divergence times range from 10 to 14 MY over the three codon positions under JC (Table 5), while the range was from 11 to 24 MY under the global clock (Table 4). As the local clock model accounts for the high rates in the anthropoids (particularly at positions 1 and 2) and in mouse lemurs (particularly at position 3), the divergence times for those lineages become younger. In contrast with the global clock models, the local clock model appropriately interpreted the long branches in those clades as high rates rather than ancient divergences. This effect is particularly apparent at the first and third positions. For comparison, we also fitted the two-, four- and five-rates models, as mentioned above, under F84G. The two-rates model produced similar date estimates for the anthropoids to the three-rates model but produced much older dates for the mouse lemur divergence (18.2, 14.0, and 8.2 MY for the three codon positions compared with 11.4, 9.3, and 4.5 MY, respectively, under the three-rates model of Table 5). The four- and five-rates models produced date estimates that are virtually identical to the three-rates model of Table 5. Based on those results, we make two observations: (1) the mouse lemurs have high rates ($r_2 > r_0$ in Table 6) and incorrectly forcing them to have the same rate as the background branches biases date estimates; and (2)

TABLE 6. ML estimates of substitution rates for the three branch classes ($\times 10^{-8}$ substitutions/site/year). One rate is estimated for each analysis under the global clock model. The local clock model assumes three rates: r_1 for the hominoids, r_2 for the mouse lemurs, and r_0 for all other branches. In the combined global clock analysis, three rates ($r^{(1)}, r^{(2)}, r^{(3)}$) are estimated for the three codon positions. In the combined local clock analysis, nine rates are estimated, with $r_j^{(k)}$ ($j = 0, 1, 2$; $k = 1, 2, 3$) the rate in branch group j at codon position k.

	JC				F84G				Combined	
Model	(a) all	(b) pos1	(c) pos2	(d) pos3	(e) all	(f) pos1	(g) pos2	(h) pos3	(i) JC	(j) F84G
Global clock	0.276	0.167	0.058	0.698	0.792	0.218	0.088	4.238	$\hat{r}^{(1)} = 0.158$	0.238
									$\hat{r}^{(2)} = 0.051$	0.084
									$\hat{r}^{(3)} = 0.721$	3.352
Local clock	$\hat{r}_0 = 0.239$	0.133	0.041	0.632	0.595	0.166	0.061	3.598	$\hat{r}_0^{(1)} = 0.131$	0.175
	$\hat{r}_1 = 0.392$	0.281	0.117	0.887	1.077	0.456	0.205	4.705	$\hat{r}_1^{(1)} = 0.281$	0.465
	$\hat{r}_2 = 0.583$	0.273	0.078	1.419	1.147	0.294	0.106	6.801	$\hat{r}_2^{(1)} = 0.259$	0.499
									$\hat{r}_0^{(2)} = 0.040$	0.061
									$\hat{r}_1^{(2)} = 0.113$	0.195
									$\hat{r}_2^{(2)} = 0.044$	0.100
									$\hat{r}_0^{(3)} = 0.634$	3.136
									$\hat{r}_1^{(3)} = 0.884$	4.337
									$\hat{r}_2^{(3)} = 1.514$	4.266

slight overparameterization in the four- and five-rates models does not affect date estimation too much.

The combined analysis of the three positions under the local clock models (analyses i and j in Table 5) is expected to be the most reliable as the model accounts for the local branch rates as well as the different substitution rates at the three codon positions. Given that we have no external means (i.e., no fossil data) for testing the veracity of mouse lemur age estimates, it is interesting to contrast age estimates of the mouse lemurs with those for the human–chimpanzee divergence, for which the fossil record is quite informative. The estimates for mouse lemur divergence are 12.7 MY (with a 95% confidence interval [CI] of 9.4–16.1 MY) under the JC model (analysis i in Table 5) and 7.1 MY (95% CI of 5.2–9.1 MY) under F84G (analysis j in Table 5). The age estimates for the human–chimpanzee divergence are 8.3 MY (95% CI of 7.6–9.0 MY) under JC and 6.5 MY (95% CI of 5.4–7.7 MY) under F84G. (The confidence intervals were calculated by a normal approximation to the ML estimates and are too narrow as the model ignores uncertainties in fossil calibration dates.) We note that the age estimates for the human–chimpanzee divergence are compatible with current interpretations of the fossil record (Brunet et al., 2002), suggesting that the method might be robust. Furthermore, it is notable that mouse lemurs, whose morphological differences are hard to discern, appear to be as old or even older than the human–chimpanzee divergence.

To see how robust the date estimates in the combined analysis are to assumptions about the rates, we also fitted a few variations to the three-rates model, as mentioned above, under F84G. The four- and five-rates models produced date estimates very similar to those of Table 5 under the three-rates model, with date differences within 1 MY. The improvement in the log likelihood, compared with the three-rates model, is 0.1 and 8.2 for the four- and five-rates models, respectively. The two-rates model forcing the mouse lemurs to have the same rate as the background branches produced older estimates for the mouse lemur divergence (11.6 MY compared with 7.1 MY of Table 5, analysis j). The log likelihood is worse that that for the three-rates model by 22.4. Although we examined only a few rate models, those results appear to suggest that three rates are necessary to capture the main features of rate variation in these data.

Besides the nucleotide-based models discussed in this study, an alternative approach to dealing with heterogeneity among the codon positions is to use models of codon substitution (Goldman and Yang, 1994; Muse and Gaut, 1994). Codon models naturally accommodate differences between the three codon positions as they make use of the genetic code and account for synonymous and nonsynonymous rate differences. Application of the codon model (Goldman and Yang, 1994) with the three branch rates produced date estimates similar to those of the combined nucleotide-based analysis. The estimates for the mouse lemur and human–chimpanzee divergence dates were 8.4 MY and 6.9 MY, respectively, compared with 7.1 MY and 6.5 MY in the combined nucleotide-based analysis (Table 5, analysis j, F84G). The estimate of the nonsynonymous/synonymous rate ratio, $\hat{\omega} = 0.0325$, is much less than 1, consistent with the extreme rate differences among codon positions (Table 6).

Bayes Analysis

Bayes posterior means of divergence times are listed in Table 7. The 95% credibility intervals are calculated as well, but are presented for a few models only. The prior

TABLE 7. Bayes estimates of divergence times (including some 95% credibility intervals).

	JC				F84G				Combined[b]			
Node[a]	(a) all	(b) pos1	(c) pos2	(d) pos3	(e) all	(f) pos1	(g) pos2	(h) pos3	(i) JC	(j) F84G	j*	j**
40 dog/bear	50.2 (43.9, 56.0)	42.4	36.7	54.9	49.9 (40.0, 59.4)	43.3	34.7	46.4	51.3	45.2 (34.9, 55.8)	41.4	46.3
43 human/chimp	8.5 (7.2, 9.9)	6.8	8.0	9.9	8.1 (6.0, 10.2)	7.1	5.9	6.9	9.2	7.1 (5.1, 9.3)	7.1	8.3
45 hominoid	18.4 (16.3, 20.6)	16.8	16.0	20.0	16.2 (13.1, 19.5)	15.6	19.2	15.2	18.5	15.2 (12.1, 18.6)	15.1	16.8
47 anthropoid	55.3 (49.6, 61.6)	58.8	56.1	51.1	59.4 (49.8, 69.8)	60.4	56.0	60.5	53.3	61.1 (50.1, 73.0)	59.6	65.9
48 lorisiform	47.4 (39.6, 55.7)	33.6	46.4	48.7	41.7 (31.1, 53.2)	29.9	40.4	45.1	45.0	40.5 (29.3, 53.0)	38.8	45.1
51 Lemuridae	37.6 (30.7, 44.9)	43.8	46.9	35.0	37.0 (28.0, 47.3)	45.0	45.4	30.0	35.8	35.3 (26.2, 46.1)	33.2	39.9
52 southern clade	7.2 (5.0, 9.8)	10.8	13.3	8.9	9.0 (5.4, 13.8)	14.6	14.4	6.7	7.4	7.6 (4.6, 12.0)	7.1	8.6
58 northern clade	8.9 (6.3, 11.9)	15.2	17.5	10.8	10.9 (6.8, 16.2)	19.1	20.5	6.8	9.3	8.0 (5.0, 12.7)	7.5	9.2
59 mouse lemurs	11.0 (7.9, 14.5)	17.6	21.5	13.2	13.1 (8.3, 19.2)	23.1	25.8	08.5	11.3	10.0 (6.4, 15.4)	9.4	11.4
61 Cheirogaleidae	27.1 (21.3, 33.5)	34.1	34.9	29.0	31.1 (22.2, 41.1)	38.0	38.6	29.5	27.4	30.3 (21.9, 40.5)	28.8	34.1
65 lemuriform	61.1 (53.1, 69.5)	66.2	68.8	55.7	65.4 (54.1, 76.4)	66.8	63.7	60.5	57.8	66.9 (55.2, 78.3)	63.5	75.3
66 Strepsirrhine	68.0 (60.3, 76.4)	71.0	74.9	61.7	70.4 (59.2, 81.0)	72.0	69.1	71.5	63.3	73.3 (62.2, 83.6)	69.8	82.2
35 C4	53.7 (50.2, 57.6)	54.6	53.5	52.1	53.1 (50.1, 57.5)	53.6	53.6	52.5	53.3	53.1 (50.1, 57.5)	42.8	53.5
36 C5	35.1 (33.0, 38.8)	35.8	36.3	36.2	35.1 (33.0, 39.1)	36.3	36.4	34.8	35.5	34.8 (33.0, 38.7)	22.2	34.9
37 C6	53.1 (51.0, 57.2)	54.1	54.4	54.1	56.1 (51.4, 59.8)	54.5	54.4	56.6	53.1	56.0 (51.3, 59.7)	46.4	56.3
41 C7	62.1 (56.8, 64.8)	56.2	57.9	60.1	60.6 (52.7, 64.8)	55.5	55.6	59.0	59.9	58.8 (49.7, 64.6)	54.2	60.0
44 C1	11.6 (10.9, 11.9)	11.1	10.7	11.6	11.1 (9.3, 11.9)	10.7	10.3	10.1	11.7	10.8 (8.7, 11.9)	10.8	12.8
46 C2	33.1 (32.0, 35.9)	34.0	34.1	33.8	33.6 (32.0, 37.0)	34.7	35.0	34.6	33.1	33.9 (32.0, 37.4)	33.9	33.2
67 C3	78.1 (70.7, 86.3)	81.6	83.7	69.6	83.6 (73.9, 89.6)	79.1	75.6	84.4	73.1	86.1 (78.0, 89.8)	82.5	96.9

[a] C1–C7 are calibration nodes.
[b] Analysis j* used only the three calibrations within primates, and j** used only the four calibrations outside primates.

mean rate at the root (rtrate) for the separate analysis of the three codon positions are the ML estimates under the global clock (Table 6). The SE of the rate is set to be half of the mean.

Similar to the likelihood analysis, there is considerable variation in date estimates among the three codon positions when they are analyzed separately. For example, the estimates for the mouse lemur divergence date range from 13 to 22 MY under JC and from 8 to 26 MY under F84G over the three codon positions. There are also differences between the Bayes and ML estimates. For the first and second positions, the Bayes estimates of mouse lemur dates (for nodes 59, 52, and 58 in Fig. 1) are all much larger than the ML estimates under the local clock models and are closer to the ML estimates under the global clock, although the two methods are more similar at the third position. The reasons for these differences are not clear. Based on our ML analysis using variations of the three-rates model discussed above, we speculate that the Bayes model of stochastic rate evolution might have difficulty in accommodating rapid rate changes at the first two positions (see Fig. 3). However, this interpretation is not consistent with the robustness, to be discussed below, of the Bayes date estimation to changes to parameter ν, which controls how variable the rates are over time.

When all three codon positions are analyzed without accounting for their differences (analyses a and e in Table 7), the Bayes estimates are similar to the ML estimates. The mouse lemur divergence time is estimated to be 11 MY and 13 MY under JC and F84G, respectively, while the corresponding ML estimates are 11 MY and 9 MY. Date estimates from the combined Bayes analysis that accounts for the heterogeneity of the three codon positions are presented (analyses i and j) in Table 7 for JC and F84G, respectively. In general, those estimates are very similar to corresponding estimates from the ML analysis (Table 5, analyses i and j). Under F84G, the Bayes method estimated the divergence time of the mouse lemurs to be 10 MY with a 95% CI of (6.4 MY, 15.4 MY), slightly older than the ML estimate under the same model (8.5 MY; analysis j in Table 5). The Bayes estimate for the human–chimpanzee divergence is 7.1 MY (5.1 MY, 9.3 MY), identical to the ML estimate. In general, the date estimates from the combined analysis are remarkably similar for the two methods. The likelihood confidence intervals are too narrow but overlap considerably with the Bayes credibility intervals.

DISCUSSIONS

Limitations of the Likelihood Local-Clock Models

The likelihood models implemented in this paper have a number of limitations, some of which are noted here. First, assignment of branches on the phylogeny to the k rate groups has to be done by the researcher. The models are thus suitable for estimating and testing rates for well-defined phylogenetic clades or lineages but are difficult to use if the rate changes affect many branches scattered across the tree. The latter situation is much easier to deal with in the Bayes framework using a prior probabilistic model of rate change (Thorne et al., 1998). In the analysis of the mitochondrial data, we estimated branch lengths without the clock to help specify branch rate classes. Since the hypotheses about rates are derived from the data, it is inappropriate to use the χ^2 approximation to perform the likelihood ratio test to test them, due to the problem of multiple comparisons. We do not consider this a serious problem in this study since our interest is not in testing rate hypotheses but in estimation of divergence times and substitution rates. Second, the likelihood models implemented here consider calibration dates as fixed points and do not account for uncertainties in the fossil record. As a result, the variances of parameter estimates are underestimated and the confidence intervals are too narrow. Explicitly modeling uncertainties in calibration information is possible in theory but introduces high-dimensional integrals in the likelihood calculation, which does not seem feasible computationally. Third, our implementation is restrictive in that the same number of branch rate classes are assumed for all genes. For example, in our combined analysis, we assumed three rate classes for every locus: the mouse lemurs, the anthropoids, and all the other lineages. In real data, a two-rate model might fit one locus while a three-rate model might fit another. We also assume that data are available for every species at each locus. However, when heterogeneous multilocus data are combined, some species might be missing at some loci. These restrictions are implementation details and can be relaxed in theory.

Effects of Priors, Model Assumptions, and Calibrations on Date Estimation

We changed the means and SDs for the gamma priors for the age of the root and the rate at the root, finding that the posterior date estimates are rather stable. For example, doubling the age of the root (so that $m = 140$ MY and $\sigma = 70$ MY) changed the posterior means of divergence times by no more than 0.2 MY. The posterior time estimates are also very insensitive to the parameters in the gamma prior for the Brownian motion parameter ν, except when ν is forced to be extremely small so that a molecular clock is effectively assumed. Similar robustness has been reported in a recent study by Springer et al. (2003) estimating divergence times in mammals. It is unclear whether robustness is generally the case. Furthermore, some prior model assumptions are made in the program and are not evaluated. For example, assumptions about the shape of the tree (how starlike the tree is) in the prior for divergence times might affect time estimation. The specific model of rate evolution might not be very important according to the comparisons of Aris-Brosou and Yang (2002), as long as the rate change is appropriately accounted for by the model.

The nucleotide substitution model had larger effects on the time estimates for both the ML and the Bayes methods, although the patterns are different. For ML, there is little difference between JC and F84G at the first and second positions (Tables 4 and 5). However, big differences exist at the third position, especially concerning

the mouse lemur divergence; for example, the mouse lemur divergence was dated at 24 MY under JC and 7.2 MY under F84G for the clock analysis and at 11 MY and 4.5 MY for the local-clock analysis. For the Bayes method, differences existed between the two models for all three codon positions, although they are not so large as in the ML analysis (Table 7). We do not understand the pattern but suspect that the large differences between JC and F84G seen in the ML analysis of the third codon position might be partly due to the extremely high divergence (Table 2). JC is clearly ineffective in correcting for multiple hits. However, F84G is unlikely to fit the data perfectly and might be sensitive to violations of assumptions at such high divergences. It may be noted that F84G does not always produce older date estimates than JC (through more effective correction for multiple hits), as can be expected if a fixed substitution rate were used to calculate dates under both models. As substitution rates are estimated under each model, the differences in date estimation between the two models are due to the relative distances from the internal nodes to the tips and to the positions of the nodes relative to the calibration points (Yang, 1996a).

Clearly the ages of nodes close to calibration points should be estimated more reliably than the ages of nodes far away. In all our analyses (Tables 4, 5, 7), the estimated dates for the human–chimpanzee divergence, which is close to calibration points within primates, are much more stable than those for the mouse lemur divergences, which are more distant from calibration points. To evaluate this effect further, we performed the combined analysis using only the three calibration points within primates or only the four calibration points outside primates. The results are listed as analyses j* and j** in Table 5 for ML and in Table 7 for the Bayes analysis, which clearly demonstrated the importance of calibration information in both methods. For example, in the likelihood analysis, the dog/bear divergence date changed from 43 MY to 34 MY when the calibration information outside primates was withheld (compare analyses j and j* in Table 5), while withholding primate calibration information caused big changes in the hominoid and anthropoid divergence dates (compare analyses j and j** in Table 5). Similar effects are seen in the Bayes analysis (Table 7). The results of cross-calibration are particularly interesting. For example, when only the three primate calibration nodes, C_1, C_2 and C_3, are used, the posterior means of the ages of the four calibration nodes outside primates (C_4, C_5, C_6, and C_7) are all too young, with C_4, C_5, and C_6 outside their ranges (Table 3). Similarly, when C_4, C_5, C_6, and C_7 outside primates are used, the dates for C_1 and C_3 are outside their ranges. Notably, the mouse lemur divergence dates are all quite stable for the three sets of calibration information (analyses j, j*, and j** in Table 7).

Combined versus Separate Analysis

In the molecular phylogenetics literature, there has been a lingering debate concerning whether heterogeneous data sets should be analyzed separately or in a combined data set when phylogenetic trees are reconstructed. The "combined" analysis in that debate means merging and mixing data sets without accounting for their heterogeneity and thus differs from the combined analysis advocated in this study. Such a debate does not appear necessary as neither the separate analysis nor the "mixed" analysis is a good method; a proper analysis should combine data sets but at the same time account for their differences. The advantages of a combined analysis in complicated estimation problems are well recognized. In the date-estimation problem, the model can "explain" the data of a single gene rather easily; a high rate and a recent divergence might fit the data just as well as a low rate and an early divergence. A single gene can thus be misleading about divergence times. In our analysis, date estimates varied considerably among the three codon positions. Combining multiple genes that have different patterns of rate change will introduce internal constraints on the model, allowing extraction of information about the common parameters, i.e., the divergence times. We suggest that there is no question that multiple gene data sets should be combined, although with the strong proviso that the heterogeneity of the different data sets be accounted for. Indeed combined analysis of multiple genes and simultaneous use of multiple calibration points appear to provide the best approach to divergence date estimation when the molecular clock is violated.

ACKNOWLEDGMENTS

We thank Jeff Thorne, Hirohisa Kishino, Chris Simon, and an anonymous referee for many constructive comments. This study was supported by BBSRC grant 31/G13580 to Z.Y. and NSF grant DEB-9985205 to A.D.Y. This is Duke University Primate Center publication 769.

REFERENCES

ARIS-BROSOU, S., AND Z. YANG. 2002. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Syst. Biol. 51:703–714.

BRUNET, M., F. GUY, D. PILBEAM, H. T. MACKAYE, A. LIKIUS, D. AHOUNTA, A. BEAUVILAIN, C. BLONDEL, H. BOCHERENS, J.-R. BOISSERIE, L. DE BONIS, Y. COPPENS, J. DEJAX, C. DENYS, P. DURINGERQ, V. EISENMANN, G. FANONE, P. FRONTY, D. GERAADS, T. LEHMANN, F. LIHOREAU, A. LOUCHART, A. MAHAMAT, G. MERCERON, G. MOUCHELIN, O. OTERO, P. P. CAMPOMANES, M. P. DE LEON, J.-C. RAGE, M. SAPANET, M. SCHUSTERQ, J. SUDRE, P. TASSY, X. VALENTIN, P. VIGNAUD, L. VIRIOT, A. ZAZZO, AND C. ZOLLIKOFER. 2002. A new hominid from the Upper Miocene of Chad, Central Africa. Nature 418:145–151.

DRUMMOND, A., R. FORSBERG, AND A. G. RODRIGO. 2001. The inference of stepwise changes in substitution rates using serial sequence samples. Mol. Biol. Evol. 18:1365–1371.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17:368–376.

FELSENSTEIN, J. 2002. PHYLIP: Phylogenetic inference program, version 3.6. Univ. Washington, Seattle.

FLYNN, J. J. 1996. Carnivoran phylogeny and rates of evolution: Morphological, taxic, and molecular. Pages 542–581 *in* Carnivore behavior, ecology, and evolution (J. Gittleman, ed.). Cornell Univ. Press, Ithaca, New York.

GEOFFROY SAINT-HILAIRE, E. 1795. Observations sur une petite espèce de Maki (*Lemur* Linn.). Bull. Soc. Philomath. Correspondans 1:89–90.

GINGERICH, P. D., AND M. D. UHEN. 1994. Time of origin of primates. J. Hum. Evol. 27:443–445.

GOLDMAN, N., AND Z. YANG. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11:725–736.

HUELSENBECK, J. P., B. LARGET, AND D. SWOFFORD. 2000. A compound Poisson process for relaxing the molecular clock. Genetics 154:1879–1892.

JANIS, C. M., M. W. COLBERT, M. C. COOMBS, W. D. LAMBERT, B. J. MACFADDEN, B. J. MADER, D. R. PROTHERO, R. M. SCHOCH, J. SHOSHANI, AND W. W. WALL. 1998. Perrissodactyla and Proboscidea. Pages 511–524 *in* Evolution of Tertiary mammals of North America (C. M. Janis, K. M. Scott, and L. L. Jacobs, eds.). Cambridge Univ. Press, Cambridge, U.K.

JUKES, T. H., AND C. R. CANTOR. 1969. Evolution of protein molecules. Pages 21–123 *in* Mammalian protein metabolism (H. N. Munro, ed.). Academic Press, New York.

KISHINO, H., AND M. Hasegawa. 1990. Converting distance to time: Application to human evolution. Methods Enzymol. 183:550–570.

KISHINO, H., J. L. THORNE, AND W. J. BRUNO. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352–361.

KUMAR, S. 1996. Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. Genetics 143:537–548.

MARTIN, R. D. 1972. A preliminary field-study of the lesser mouse lemur (*Microcebus murinus*). Z. Tierpsychol. Suppl. 9:43–89.

MARTIN, R. D. 1993. Primate origins: Plugging the gaps. Nature 363:223–234.

MUSE, S. V., AND B. S. GAUT. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11:715–724.

PROTHERO, D. R., AND R. M. SCHOCH. 1989. The evolution of perrissodactyls, pages 173–181. Clarendon Press/Oxford Univ. Press, New York.

RAMBAUT, A. 2000. Estimating the rate of molecular evolution: Incorporating non-comptemporaneous sequences into maximum likelihood phylogenetics. Bioinformatics 16:395–399.

RAMBAUT, A., AND L. BROMHAM. 1998. Estimating divergence dates from molecular sequences. Mol. Biol. Evol. 15:442–448.

RASOLOARISON, R. M., S. M. GOODMAN, AND J. U. GANZHORN. 2000. Taxonomic revision of mouse lemurs (*Microcebus*) in the western portions of Madagascar. Int. J. Primatol. 21:963–1019.

SANDERSON, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218–1232.

SANDERSON, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

SHOSHANI, J., C. P. GROVES, E. L. SIMONS, AND G. F. GUNNELL. 1996. Primate phylogeny: Morphological vs. molecular results. Mol. Phylogenet. Evol. 5:102–154.

SPRINGER, M. S., W. J. MURPHY, E. EIZIRIK, AND S. J. O'BRIEN. 2003. Placental mammal diversification and the Cretaceous–Tertiary boundary. Proc. Natl. Acad. Sci. USA 100:1056–1061.

STUART, A., K. ORD, AND S. ARNOLD. 1999. Kendall's advanced theory of statistics. Arnold, London.

SULLIVAN, J., AND D. L. SWOFFORD. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? Syst. Biol. 50:723–729.

TAKEZAKI, N., A. RZHETSKY, AND M. NEI. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12:823–833.

TAVARÉ, S., C. R. MARSHALL, O. WILL, C. SOLIGOS, ET AL. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416:726–729.

THEWISSEN, J. G. M. 1994. Phylogenetic aspects of cetacean origins: A morphological perspective. J. Mammal. Evol. 2:157–184.

THORNE, J. L., AND H. KISHINO. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689–702.

THORNE, J. L., H. KISHINO, AND I. S. PAINTER. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15:1647–1657.

WHELAN, S., P. LIÒ, AND N. GOLDMAN. 2001. Molecular phylogenetics: State of the art methods for looking into the past. Trends Genet. 17:262–272.

YANG, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39:306–314.

YANG, Z. 1996a. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11:367–372.

YANG, Z. 1996b. Maximum-likelihood models for combined analyses of multiple sequence data. J. Mol. Evol. 42:587–596.

YANG, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13:555–556 (http://abacus.gene.ucl.ac.uk/software/paml.html).

YANG, Z., N. GOLDMAN, AND A. FRIDAY. 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Mol. Biol. Evol. 11:316–324.

YODER, A. D., R. M. RASOLOARISON, S. M. GOODMAN, J. A. IRWIN, ET AL. 2000. Remarkable species diversity in Malagasy mouse lemurs (Primates, *Microcebus*). Proc. Natl. Acad. Sci. USA 97:11325–11330.

YODER, A. D., AND Z. YANG. 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17:1081–1090.

ZUCKERKANDL, E., AND L. PAULING. 1965. Evolutionary divergence and convergence in proteins. Pages 97–166 *in* Evolving genes and proteins (V. Bryson and H. J. Vogel, eds.). Academic Press, New York.

First submitted 19 February 2003; reviews returned 13 May 2003; final acceptance 10 June 2003
Associate Editor: Jeff Thorne