Syst. Biol. 47(1):125–133, 1998

On the Best Evolutionary Rate for Phylogenetic Analysis

ZIHENG YANG[1]

Department of Integrative Biology, University of California, Berkeley, California 94720-3140, USA

Abstract.—The effect of the evolutionary rate of a gene on the accuracy of phylogeny reconstruction was examined by computer simulation. The evolutionary rate is measured by the tree length, that is, the expected total number of nucleotide substitutions per site on the phylogeny. DNA sequence data were simulated using both fixed trees with specified branch lengths and random trees with branch lengths generated from a model of cladogenesis. The parsimony and likelihood methods were used for phylogeny reconstruction, and the proportion of correctly recovered branch partitions by each method was estimated. Phylogenetic methods including parsimony appear quite tolerant of multiple substitutions at the same site. The optimum levels of sequence divergence were even higher than upper limits previously suggested for saturation of substitutions, indicating that the problem of saturation may have been exaggerated. Instead, the lack of information at low levels of divergence should be seriously considered in evaluation of a gene's phylogenetic utility, especially when the gene sequence is short. The performance of parsimony, relative to that of likelihood, does not necessarily decrease with the increase of the evolutionary rate. [Branch lengths; homoplasy; likelihood; parsimony; phylogeny; optimum evolutionary rate; saturation; simulation.]

Neither too similar nor too divergent molecular sequences contain much phylogenetic information. Molecular systematists have long understood that slowly evolving genes such as the small-subunit ribosomal RNA genes should be used to infer relationships among distantly related species (Sogin, 1991), and rapidly changing genes such as those in the mitochondrial genome (Brown et al., 1982) should be sequenced to infer relationships among closely related species. Because time and rate are confounded in molecular phylogenetic analysis, the relevant factor here is the amount of evolution involved in the sequence data. Although very few studies have attempted to quantify the optimum amount of evolution for phylogenetic analysis, many workers have discussed ''saturation'' of substitutions and schemes for downweighting rapidly changing sites and upweighting slowly changing sites in a parsimony analysis. However, without knowing how much evolution is ''too much'' or how little evolution is ''too little'', such weighting schemes can be misleading.

Despite the widespread concerns about saturation, real data analyses suggest that highly divergent sequences may in fact be more informative than sequences of very low divergence. For example, in a phylogenetic analysis of mitochondrial cytochrome *b* genes from strepsirrhine primates, Yoder et al. (1996) noted that the highly variable third codon positions contained the preponderance of phylogenetic signal for the species studied. Similarly, Yang (1996b) estimated the tree lengths to be 0.37, 0.12, and 3.7 for the three codon positions, respectively, in several mitochondrial genes of six primate species. Besides being highly variable with an average of 3.7 substitutions per site, the third positions had a very high transition/transversion rate ratio (with $\hat{\kappa} = 52$). However, the third positions were found to be much more informative than the first or second positions both for the phylogeny and for speciation dates among the species (Yang, 1996b).

It is thus important to gain insight into the optimum evolutionary rate. In this study, I use computer simulations to study the effect of the evolutionary rate on the accuracy of phylogeny reconstruction and to quantify the optimum levels of sequence divergence. Accuracy was measured by the

[1] Present address: Department of Biology, University College London, Galton Lab, 4 Stephenson Way, London NW1 2HE, England.

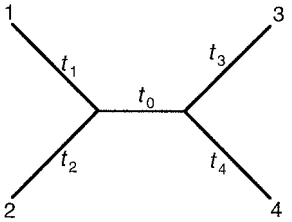

FIGURE 1. Model tree of four species. Branch length (t_i) is defined as the average expected number of nucleotide substitutions per site that have occurred along the branch.

proportion of correctly recovered branch partitions in the model tree. Other measures of accuracy, such as those on estimates of model parameters, may lead to different optimum evolutionary rates. The evolutionary rate or the amount of evolution in the data was measured by the tree length, that is, the sum of branch lengths along the tree. The maximum-parsimony (MP) and maximum-likelihood (ML) methods were used for phylogenetic reconstruction. The effect of the number of species was also examined.

SIMULATION METHODS

Generation of Sequence Data Sets

The Markov process model of Kimura (1980) was used for nucleotide substitution, with the transition/transversion rate ratio fixed at $\kappa = 5$ (α/β in Kimura's notation). Substitution rates were assumed to be variable among nucleotide sites, and the ''discrete-gamma'' model of Yang (1994) was used. The gamma shape parameter α is inversely related to the extent of rate variation, and is fixed at 0.5. The simulation model is referred to as K80+Γ.

Tree topologies and branch lengths were generated using two approaches. The first uses fixed unrooted trees for four species (Fig. 1), and is useful for detecting the effects of different tree topologies. Five sets of branch lengths were used (Table 1), representing five tree shapes identified by Yang (1996a). Branch lengths were scaled such that the total tree length is 1. Because the interior branch length influences greatly the accuracy of phylogeny reconstruction, it is fixed at a small value, 0.05, in all trees. Branch lengths in Table 1 were multiplied by the tree length S before being used to simulate data sets. This mimics the case in which the relative branch lengths are determined by the speciation dates of the species, where different genes with different evolutionary rates give rise to trees with different tree lengths.

TABLE 1. Branch lengths for five trees (tree shapes) of four species (Fig. 1) used in the simulation.

Tree	t_0	t_1	t_2	t_3	t_4
A	0.05	0.2375	0.2375	0.2375	0.2375
B	0.05	0.1	0.1	0.3	0.45
C	0.05	0.1	0.375	0.1	0.375
D	0.05	0.1	0.1	0.1	0.65
E	0.05	0.3	0.3	0.3	0.05

The second approach generates random trees and branch lengths using a model of cladogenesis. This approach is used to obtain a measure of accuracy averaged over different tree topologies and branch lengths. Trees of more than four species were also generated to examine the effect of the number of species. The birth–death process with species sampling was used as a model of cladogenesis (Yang and Rannala, 1997), with birth rate $\lambda = 5$, death rate $\mu = 1$, and sampling fraction $\rho = 0.01$. Under this model, each coalescent tree (rooted tree with the interior nodes ordered according to the ancestral speciation times) has equal probability of occurrence, and the speciation times were generated as random variables. Use of a small sampling fraction (such as the value used here) has the effect of reducing the interior branch lengths in the tree and leads to more difficult trees. To allow for different evolutionary rates among lineages, branch lengths generated under the molecular clock assumption (rate constancy among lineages) were chosen at random, with probability ½ for each case, either to be multiplied or divided by $10^{½} = 3.16$, so that rapidly changing lineages have a rate 10 times that of slowly changing lineages. Branch lengths in the tree were then scaled such that their sum equaled the specified tree length (S). The root in the coalescent

tree was removed and the unrooted tree was considered. With this approach, each replicate was simulated with a different random tree and a different set of random branch lengths.

The number of nucleotide sites in the sequence was fixed at $N = 300$. To simulate a data set with given tree topology and branch lengths, a random sequence of N sites was generated for an arbitrarily chosen interior node in the tree. N random variables were generated from the gamma distribution (Yang, 1994) and were used as rates for the sites. Each site in the sequence was then ''evolved'' along the tree using the branch lengths of the tree and the rate for that site, according to the substitution model (e.g., Goldman, 1993). Sequences at the tips of the tree constitute the data, which were subjected to further analyses. The number of simulated replicates for each parameter combination varies with the tree reconstruction method and the number of species and is given later. Some variations of the parameter values are also examined, as described later.

Analysis of Sequence Data Sets

Each simulated data set was analyzed using the parsimony and likelihood methods to reconstruct the phylogeny. In the likelihood analysis, two substitution models were assumed: that of Jukes and Cantor (1969), which ignores the transition/transversion rate bias and the rate variation among sites, and the K80+Γ model (the correct model). These analyses will be referred to as MLJC and MLKΓ, respectively. In the MLKΓ analysis, the correct values of κ and α were used and only branch lengths for each tree were optimized to save computation; limited simulations showed that this approach gave results essentially identical to one that estimates those parameters for each simulated data set. A simple stepwise addition algorithm was used for both parsimony and likelihood analyses. This algorithm is not guaranteed to find the best tree for data of more than five species. However, results of this study may not be affected much by use of this algorithm (Russo et al., 1996). Ties encountered in the stepwise addition procedure were arbitrarily resolved; this is equivalent to assigning probability $1/x$ for a polytomy if it can be resolved into x bifurcating trees that include the correct tree. Each interior branch in the model tree defines a bipartition of species. The number of correctly recovered bipartitions in each replicate was counted for each method and the average proportion of correctly recovered bipartitions (p) was calculated.

RESULTS

Fixed Trees of Four Species

Results obtained using the five fixed trees of Table 1 are shown in Figure 2. For four species, the accuracy measure used (the average proportion of correctly recovered branch partitions) is equivalent to the probability that the entire model tree is recovered. The tree shape had a significant effect on the relative performance of the parsimony and likelihood methods. When tree shape B (Fig. 2b) was used, parsimony outperformed the two likelihood methods, and the superiority of parsimony became even greater with the increase of the tree length. For tree shape C (Fig. 2c), parsimony performed much worse than likelihood; this tree shape is known to cause problems for parsimony (Felsenstein, 1978). For the other three trees, the likelihood methods performed better than parsimony, but the differences were small. The performance of MLJC was between parsimony and MLKΓ for all five tree shapes. This is consistent with the proposition (Yang, 1996a) that parsimony is much closer to likelihood with the most stringent assumptions about the evolutionary process (JC69) than to likelihood assuming a more complex model of sequence change. In terms of the relative performance of parsimony and likelihood methods, the results are generally congruent with previous simulation studies (e.g., Kuhner and Felsenstein, 1994; Tateno et al., 1994; Gaut and Lewis, 1995; Huelsenbeck, 1995; Yang, 1996a).

The performance curve as a function of the evolutionary rate has the same shape

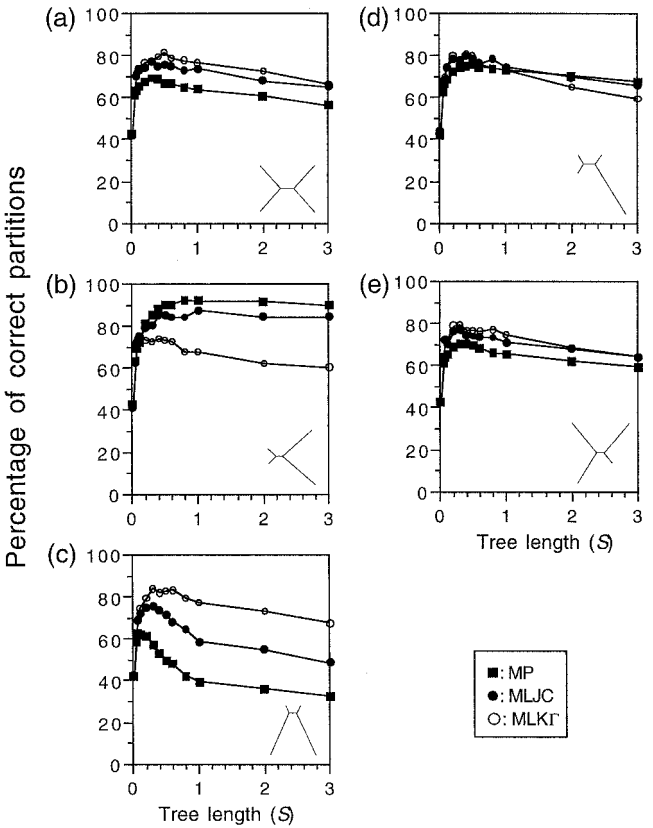

FIGURE 2. Probability of recovering the correct tree as a function of the tree length (S) for different trees of four species. The parsimony (■) and likelihood (● = MLJC; ○ = MLKΓ) methods were used. The number of simulated replicates is 1,000 for the likelihood methods and 5000 for parsimony. The tree shapes are superimposed on the graphs. (a–e) Trees A–E of Table 1.

irrespective of the tree reconstruction method or the correct tree topology. The probability of recovering the correct tree initially increases with the tree length (S), reaches a maximum, and then decreases with the further increase of S. Crude estimates of the optimum tree length (S^*) for tree shape B were 0.8 for parsimony and 0.4–0.5 for the two likelihood methods. For tree shape B, parsimony not only outperforms likelihood but is also more tolerant of multiple substitutions and has a much higher optimum evolutionary rate. For tree shape C, the opposite is true; parsimony performs far more poorly than likelihood and also has a lower optimum tree length ($S^* = 0.1$) than likelihood ($S^* = 0.3$). For the other three tree shapes (A, D, and E), the performance differences among the methods are small, and the optimum tree lengths are also very similar, in the range 0.3–0.5.

An important common characteristic of the performance curves is that the upward slope when $S < S^*$ (and especially when S is very small) is much steeper than the downward slope when $S > S^*$. Increasing the amount of evolution when it is very low improves the accuracy of phylogenetic estimation greatly, but phylogenetic information is not so easily diluted by extra substitutions. The results suggest that lack of information in sequences of low divergence may be a more serious problem than accumulation of noise in highly divergent sequences.

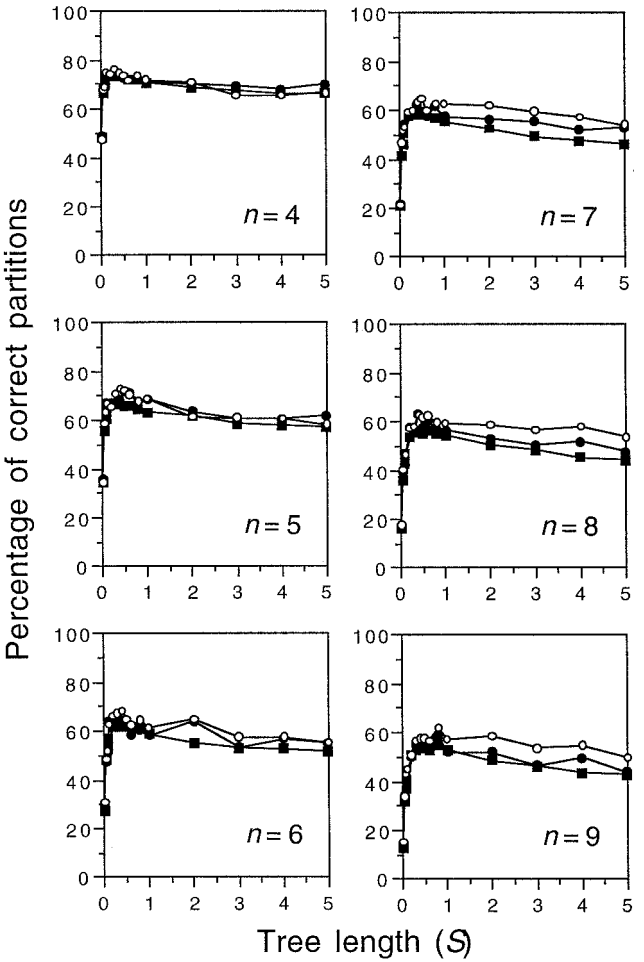

FIGURE 3. Proportion of correctly recovered branch partitions as functions of the tree length (S) for trees of different numbers of species (n). Random trees and branch lengths are generated from a model of cladogenesis, i.e., the birth-death process with species sampling. The parsimony (■) and likelihood (● = MLJC; ○ = MLKΓ) methods were used. The number of simulated replicates is 1,000 for parsimony (MP) and for likelihood with $n = 4$ species, and is 200 for likelihood with $n = 5$, 6, 7, 8, and 9 species.

Random Trees

Simulation results obtained using random trees with random branch lengths are shown in Figure 3 for $n = 4$–9 species. Parsimony and the two likelihood methods performed similarly for small data sets ($n = 4$, 5, 6), but for larger data sets ($n = 7$, 8, 9), the likelihood methods performed slightly better than parsimony. More study is needed to find out whether the performance of likelihood relative to that of parsimony generally increases with the increase of the number of species in the data.

The performance curves as functions of the evolutionary rate have the same shape as that for fixed trees of four species (Fig. 3). The optimum tree lengths were estimated to be 0.2–0.4 for 4 species, 0.4–0.5 for 5 species, 0.3–0.4 for 6 species, 0.4–0.6 for 7 species, 0.4–0.6 for 8 species, and 0.5–0.8 for 9 species (Fig. 3). There were no noticeable differences in the optimum tree lengths among the methods; in particular, the optimum tree length for parsimony is not any smaller than those for the likelihood methods. The optimum tree length

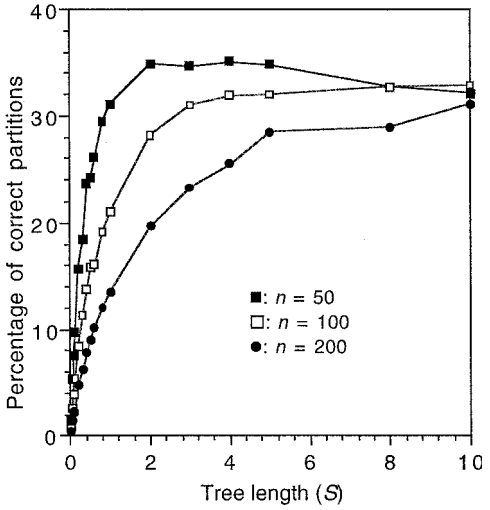

FIGURE 4. Proportion of correctly recovered branch partitions by parsimony as functions of the tree length (S) for trees of $n = 50$, 100, and 200 species. Random trees and branch lengths are generated from the birth–death process with species sampling. The number of replicates is 200, 100, and 20, for $n = 50$, 100, and 200, respectively.

appeared to increase with the increase of the number of species in the data. This pattern is apparent from Figure 4, where the parsimony method was applied to data of $n = 50$, 100, and 200 species. The optimum tree length for parsimony is about 0.2–0.8 for trees of 4–9 species (Fig. 3), and is about 4, 10, and 15 for trees of 50, 100, and 200 species, respectively (Fig. 4). Clearly, a large tree has more branches to disperse multiple substitutions and thus can accommodate more substitutions at the same site.

Effects of Sequence Length and Model Parameters

Additional simulations are performed to examine the effects of the sequence length and model parameters. The parsimony method of tree reconstruction is used and random trees are generated from the birth–death process with species sampling. Figure 5a shows the performance of parsimony when the sequence has $N = 1{,}000$ sites. Apart from the sequence length, all other simulation conditions are the same as in Figures 3 (n = 9) and 4. As expected, the proportions of correctly recovered

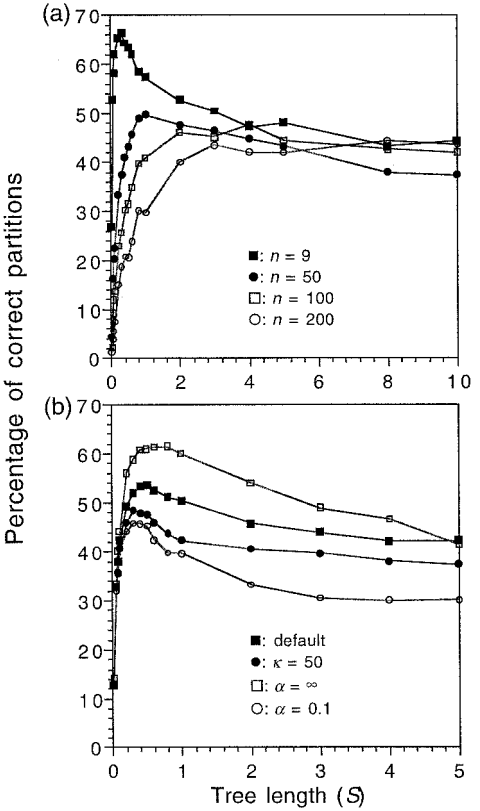

FIGURE 5. The effects of sequence length and model parameters on the performance of parsimony. (a) The sequence had 1,000 sites, and random trees of $n = 9$, 50, 100, and 200 species were generated. The simulation conditions are the same as in Figures 3 ($n = 9$) and 4 except for the number of sites. The number of replicates is 1,000, 100, 50, and 10 for $n = 9$, 50, 100, and 200, respectively. (b) The effect of the transition/transversion rate ratio (κ) and the gamma shape parameter (α). Random trees of $n = 9$ species were simulated with 300 sites in the sequence. The default uses the same values of parameters ($\kappa = 5$, $\alpha = 0.5$) as in Figure 3 ($n = 9$), plotted here for comparison. One parameter is changed at each time in the plots for $\kappa = 50$, $\alpha = \infty$ (constant rate among sites), and $\alpha = 0.1$. The number of replicates is 1,000.

branches at $N = 1{,}000$ sites (Fig. 5a) are much higher than at $N = 300$ sites (Figs. 3 (n = 9) and 4). The performance curve as a function of the tree length has the same shape as identified before. Crude estimates of the optimum tree length (S^*) are 0.3, 1, 4, and 8 for $n = 9$, 50, 100, and 200 species, respectively. These estimates are lower than the corresponding estimates obtained

when there are only 300 sites in the sequence. With $n = 9$ species, the performance appears to deteriorate more quickly with further increase of S when $S > S^*$ for long sequences ($N = 1{,}000$; Fig. 5) than for short sequences ($N = 300$; Fig. 3). Adding more data (sites) improves the performance more significantly at low than high divergence levels. For example, with $S = 0.05$ substitutions per site on the tree, p increases from 32% for $N = 300$ sites to 53% for $N = 1{,}000$ sites, although with $S = 0.5$ substitutions per site on the tree, p increases from 54% for $N = 300$ sites to 63% for $N = 1{,}000$ sites. With very divergent sequences (say, $S = 5$), the improvement is slight (say, from 43% at $N = 300$ to 48% at $N = 1{,}000$). For large trees of $n = 100$ species, the performance improvement with addition of sites is small at extremely low rates (for example, with $S = 0.05$ substitutions per site on the tree, p increases from 3% for 300 sites to 9% for 1,000 sites) or very high rates (with $S = 30$ substitutions per site, p increases from 27% at $N = 300$ to 31% at $N = 1{,}000$). For a wide range of intermediate rates, the performance improvement with addition of sites is great; for example, at $S = 0.5$, p increases from 16% at $N = 300$ to 32% at $N = 1{,}000$, and at $S = 5$, p increases from 32% to 44%. In sum, the sequence length has a significant effect on the performance, especially at low divergence levels. Furthermore, parsimony is quite tolerant of multiple substitutions whether the sequence is long or short.

The effects of the transition/transversion rate ratio (κ) and the gamma shape parameter (α) for variable rates among sites are examined for the case of $n = 9$ species and $N = 300$ sites (Fig. 5b). Estimates of κ from real data range from 1–5 for nuclear genes to over 50 for mitochondrial DNAs (e.g., Wakeley, 1996; Yang, 1996b). The value (5) used in Figures 2–4 may be too high for nuclear genes and too low for mitochondrial genes. Nevertheless, previous simulation studies suggest that the transition/transversion rate bias does not have a great effect on phylogenetic accuracy (e.g., Fukami-Kobayashi and Tateno, 1991; Gaut and Lewis, 1995; Yang, 1996a). A strong transition/transversion bias reduces the ''effective'' number of character states, so that the data will be less informative and less tolerant of multiple substitutions. Comparison of simulation results obtained for $\kappa = 5$ and for $\kappa = 50$ (Fig. 5b) confirms these expectations. The performance of parsimony for the high transition bias ($\kappa = 50$) is lower than for the low transition bias ($\kappa = 5$). The optimum tree length for $\kappa = 50$ is estimated to be about 0.3, lower than the estimate (about 0.5) for $\kappa = 5$. The effect of unequal nucleotide frequencies (not examined here) is expected to be similar to that of the transition/transversion bias (Yang, 1996a).

The gamma shape parameter (α) reflecting the extent of substitution rate variation among sites has been found to influence phylogenetic analysis greatly (see Yang, 1996c). For third codon positions or pseudogenes, estimates of α are usually greater than 1 (indicating little rate variation), although for the first or second positions or functioning genes, the estimates are usually between 0.1 and 1 (Yang, 1996c). With severe rate variation among sites, the data will contain many invariable sites and few informative sites and will not be very tolerant of multiple substitutions. Simulation results (Fig. 5b) obtained for $\alpha = \infty$ (constant rate among sites), 0.5 (default), and 0.1 (severe rate variation) confirm these expectations. The performance deteriorates significantly when α becomes smaller. Crude estimates of the optimum tree lengths are 0.8, 0.5, and 0.3 for $\alpha = \infty$, 0.5, and 0.1, respectively.

Discussion

A major conclusion of this study is that saturation occurs only at a much higher level of sequence divergence than was previously suggested. For example, Meyer (1994) suggested that nucleotide sequence data were saturated with nucleotide substitutions if the overall uncorrected sequence divergence was above 15–20%. Under the K80+Γ model used in this study, the number of substitutions (t) between two sequences and the uncorrected se-

quence divergence (q) is related by the following formula (Jin and Nei, 1990):

$$q = \frac{3}{4} - \frac{1}{2}\left[\frac{\alpha}{\alpha + 2t(\kappa + 1)/(\kappa + 2)}\right]^{\alpha} - \frac{1}{4}\left[\frac{\alpha}{\alpha + 4t/(\kappa + 2)}\right]^{\alpha}. \quad (1)$$

A 15–20% uncorrected sequence divergence corresponds to a pairwise sequence distance of 0.2–0.3 substitutions per site. This is even lower than the optimum evolutionary rates found for small data sets ($n = 4$–9, Figs. 2 and 3) and is much lower than the optimum evolutionary rates for large data sets (e.g., $n = 50$, 100, 200, Fig. 4). By the criterion of 15–20% raw sequence divergence, many data sets would be declared as saturated with substitutions before they had enough substitutions to be most informative. As the accuracy of phylogeny reconstruction deteriorates very slowly with the increase of the evolutionary rate when $S > S^*$, a 30–40% overall uncorrected sequence divergence may be considered a starting point for concerns about saturation (Figs. 2–4). At any rate, pairwise sequence divergence is not a good indicator of the information content in the data, as the accuracy depends on not only the amount of evolution, but also on how many branches the tree has and how the substitutions are distributed among the branches in the tree.

The overconcern about saturation seems largely to be due to the accepted wisdom that phylogeny reconstruction using parsimony requires a small amount of evolution, or absence of convergent changes or homoplasy (e.g., Felsenstein, 1978). However, a small amount of evolution is neither necessary nor sufficient for parsimony to work. Parsimony may perform well in spite of multiple substitutions at the same site along the same branch. For example, for tree shape B of Table 1, parsimony recovered the correct tree with probability 90% when on average $S = 3$ substitutions have occurred at one site. Even with $S = 10$ substitutions per site, the accuracy is as high as 82%. Furthermore, it is known that parsimony may not meet the weak requirement of statistical consistency even with a small amount of evolution (e.g., Takezaki and Nei, 1994).

It is noted that high evolutionary rates are often associated with other problems. One problem is that of alignment. Highly variable segments of the genome are typically more difficult to align than more conserved regions, and yet phylogeny reconstruction may be sensitive to the specific alignment used (Thorne and Kishino, 1992). A second problem that often occurs with highly divergent sequences is the heterogeneity of nucleotide frequencies among different species. Unequal base frequencies among species indicate that the substitution process is not homogeneous among lineages, as is commonly assumed in most phylogeny reconstruction methods. In such cases, tree reconstruction methods tend to group species with similar base content instead of similar genetic background (Steel et al., 1993). A benefit associated with high evolutionary rates is that such genes often have homogeneous substitution rates among sites, and are more tolerant of multiple substitutions than sequence data simulated in this study under the gamma model of severe rate heterogeneity. To evaluate a gene's utility for phylogenetic analysis, all those factors should be considered. However, their effects should not be confused with the effect of saturation of substitutions. For example, the alignment at the third codon positions is no more difficult to obtain than those at the first and second positions, and nucleotide frequencies at the third positions may be quite homogeneous among closely related species such as different primates. In such cases, the highly informative data at the third codon positions should not be thrown away simply because they contain many substitutions.

ACKNOWLEDGMENTS

I thank Anne Yoder, Axel Meyer, David Cannatella, and two anonymous reviewers for helpful comments.

REFERENCES

BROWN, W. M., E. M. PRAGER, A. WANG, AND A. C. WILSON. 1982. Mitochondrial DNA sequences of pri-

mates, tempo and mode of evolution. J. Mol. Evol. 18:225–239.

FELSENSTEIN, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

FUKAMI-KOBAYASHI, K., AND Y. TATENO. 1991. Robustness of maximum likelihood tree estimation against different patterns of base substitution. J. Mol. Evol. 32:79–91.

GAUT, B. S., AND P. O. LEWIS. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. Mol. Biol. Evol. 12:152–162.

GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. J. Mol. Evol. 36:182–198.

HUELSENBECK, J. P. 1995. The performance of phylogenetic methods in simulation. Syst. Biol. 44:17–48

JIN, L., AND M. NEI. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. Mol. Biol. Evol. 7:82–102.

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.

KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.

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

MEYER, A. 1994. Shortcomings of the cytochrome *b* gene as a molecular marker. Trends Ecol. Evol. 9: 278–280.

RUSSO, C. A., N. TAKEZAKI, AND M. NEI. 1996. Effeciencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol. Biol. Evol. 13:525–536.

SOGIN, M. L. 1991. Early evolution and the origin of the eukaryotes. Curr. Opin. Genet. Dev. 1:457–463.

STEEL, M. A., P. J. LOCKHART, AND D. PENNY. 1993. Confidence in evolutionary trees from biological sequence data. Nature 364:440–442.

TAKEZAKI, N. AND M. NEI. 1994. Inconsistency of the maximum parsimony method when the rate of nucleotide substitution is constant. J. Mol. Evol. 39: 210–218.

TATENO, Y., N. TAKEZAKI, AND M. NEI. 1994. Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. Mol. Biol. Evol. 11: 261–277.

THORNE, J. L., AND H. KISHINO. 1992. Freeing phylogenies from artifacts of alignment. Mol. Biol. Evol. 9: 1148–1162.

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. Phylogenetic analysis using parsimony and likelihood methods. J. Mol. Evol. 42:294–307.

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

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

YANG, Z., AND B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: A Markov chain Monte Carlo method. Mol. Biol. Evol. 43:304–311.

YODER, A. D., R. VILGALYS, AND M. RUVOLO. 1996. Molecular evolutionary dynamics of cytochrome *b* in strepsirrhine primates: The phylogenetic significance of third position transversions. Mol. Biol. Evol. 13:1339–1350.

Received: 27 March 1997; accepted: 28 August 1997
Associate Editor: D. Cannatella