The Robustness of Two Phylogenetic Methods: Four-Taxon Simulations Reveal a Slight Superiority of Maximum Likelihood over Neighbor Joining

John P. Huelsenbeck
Department of Zoology, University of Texas

The robustness (sensitivity to violation of assumptions) of the maximum-likelihood and neighbor-joining methods was examined using simulation. Maximum likelihood and neighbor joining were implemented with Jukes-Cantor, Kimura, and gamma models of DNA substitution. Simulations were performed in which the assumptions of the methods were violated to varying degrees on three model four-taxon trees. The performance of the methods was evaluated with respect to ability to correctly estimate the unrooted four-taxon tree. Maximum likelihood outperformed neighbor joining in 29 of the 36 cases in which the assumptions of both methods were satisfied. In 133 of 180 of the simulations in which the assumptions of the maximum-likelihood and neighbor-joining methods were violated, maximum likelihood outperformed neighbor joining. These results are consistent with a general superiority of maximum likelihood over neighbor joining under comparable conditions. They extend and clarify an earlier study that found an advantage for neighbor joining over maximum likelihood for gamma-distributed mutation rates.

Introduction

The performance of phylogenetic methods is usually evaluated with respect to consistency (the ability to estimate the correct phylogeny with sufficient data), efficiency (the ability to quickly converge on the correct phylogeny), and robustness (the ability to estimate the correct phylogeny even when the assumptions of the phylogenetic method are violated). Of these criteria, robustness may be the most important, because the idealized assumptions underlying phylogenetic methods are most likely violated with real data. Hence, the ability to correctly estimate phylogeny despite model violation is very important. Tateno et al. (1994) have recently performed a simulation study of robustness and found the neighbor-joining method (Saitou and Nei 1987) more robust than the maximum-likelihood method (Felsenstein 1981). The present study is an extension of the Tateno et al. (1994) study comparing the robustness of maximum likelihood and neighbor joining under a large number of evolutionary models. In contrast to their conclusion, I find maximum likelihood to be more robust than neighbor joining. The reason for this discrepancy is that Tateno et al. (1994) based many of their conclusions on inappropriate comparisons of the methods.

Key words: maximum likelihood, neighbor joining, robustness, simulation.

Address for correspondence and reprints: John P. Huelsenbeck, Department of Zoology, University of Texas, Austin, Texas 78712. E-mail: johnh@phylo.zo.utexas.edu.

Mol. Biol. Evol. 12(5):843–849. 1995.

0737-4038/95/1205-0013$02.00

Evaluating Robustness: Some Potential Pitfalls

Evaluating the robustness of phylogenetic methods is not an easy task. The main difficulty is that there are a large number of known (and unknown) ways in which the assumptions of a phylogenetic method can be violated. There are also different degrees to which that assumption can be violated. For example, if a method assumes, in part, a Jukes-Cantor model of DNA substitution, the Jukes-Cantor assumption can be violated by simulating sequences with a transition:transversion bias of 0.51 or 100.0. Invariably, and quite reasonably, researchers focus their attention on the robustness of just a few of the assumptions of a phylogenetic method.

The typical approach taken in evaluating the robustness of a particular method is to generate data in such a way that the assumptions of the method are violated to varying degrees. For example, one could simulate data with and without a transition:transversion bias and examine the performance of maximum likelihood assuming a Jukes-Cantor (no transition:transversion bias) model. The first set of simulations gives an idea of the best-case scenario (all assumptions of the method are met), whereas the second set of simulations gives an idea of the robustness of maximum likelihood when

the assumption of equal rates among transitions and transversions is violated.

When comparing the robustness of competing phylogenetic methods, however, it is important to keep a correspondence between the assumptions of the methods compared. For example, when comparing the performance of maximum likelihood to neighbor joining, one should compare maximum likelihood with a Jukes-Cantor model to neighbor joining with Jukes-Cantor distances, maximum likelihood with a Kimura model to neighbor joining with a Kimura model, and so on. If a correspondence is not kept between the models of the competing methods and one method (say neighbor joining) performs better, it becomes impossible to rule out the possibility that if the same assumptions had been violated in the same manner that the performance of the methods would not be reversed. Such a simulation represents an uncontrolled experiment.

Tateno et al. (1994) found neighbor joining more robust than maximum likelihood. However, they based many of their conclusions on what I feel are inappropriate comparisons between methods. In particular, they compared the performance of neighbor joining assuming gamma distances to maximum likelihood assuming a Jukes-Cantor or Kimura model of DNA substitution. Table 1 summarizes the results from the Tateno et al. (1994) paper (excluding the results from their parsimony analysis). From the information in table 1, Tateno et al. (1994, p. 265) concluded "that the ML [maximum-likelihood] method is slightly more sensitive than the NJG [neighbor-joining-with-gamma-distance] method to violation of the assumptions made in the estimation of topology." In this simulation, the assumptions of both the maximum-likelihood and neighbor-joining method with gamma distances are violated, and inasmuch as this is true, they do address the robustness of the two methods *individually.* However, the *comparison* they make between the two methods is seriously overstated. The assumptions that are violated for the two methods are different in this simulation. In the case of maximum likelihood, an incorrect transition:transversion bias is assumed. In the case of neighbor joining with gamma distances, rate heterogeneity among sites is assumed when in fact no rate heterogeneity exists. It is impossible to determine that if the assumptions of maximum likelihood and neighbor joining were violated in the same manner, that maximum likelihood would not perform better than neighbor joining.

Examination of the simulation results from table 1 shows that maximum likelihood assuming a Jukes-Cantor model is more robust to violations than is neighbor joining assuming Jukes-Cantor distances. Similarly, even though the Kimura model is violated in many of the maximum-likelihood simulations (because an improper transition:transversion ratio was provided), maximum likelihood still performed better than the neighbor-joining method assuming Kimura distances; maximum-likelihood performance varied between 94% and 100% when the assumptions were violated, whereas neighbor-joining performance varied between 90% and 100% even though the assumptions were not violated. Far from casting "some doubts about the ML method" (Tateno et al. 1994, p. 265), these simulations show that maximum likelihood is more robust than the neighbor-joining method when the appropriate comparisons are made, at least when the model generating the simulated data assumed equal rates among sites. A similar conclusion is reached for those simulations that evolved sequences under a gamma model of rate heterogeneity. For example, when comparing the performance of the two methods with the Jukes-Cantor assumptions violated, maximum likelihood performs better than neighbor joining in 24 of the 30 comparisons for all three model trees (i.e., when the rate heterogeneity parameter is 0.5, maximum likelihood with a Jukes-Cantor model correctly estimates the tree 98% of the time, whereas neighbor joining with a Jukes-Cantor distance estimates the correct tree 90%, etc., throughout the table). Similar results are obtained when the comparison is made for the methods when they both assume the Kimura model. I next describe the results of simulations in which neighbor joining and maximum likelihood are compared under comparable gamma-model assumptions.

Material and Methods

The details of the simulation followed the protocol of Tateno et al. (1994), except that here I use maximum likelihood with an explicit gamma model. I used the same three model trees (fig. 1) and simulated data (1,000 sites) with different transition:transversion biases (R = 0.5, 9.0, and 15.0) as well as different degrees of rate heterogeneity among sites (I used a gamma distribution to describe rate variation among sites with shape parameter, $a = \infty$, 2.0, 1.0, and 0.5). When $a = \infty$, no rate heterogeneity among sites exists. A total of 36 simulations were performed (3 trees × 3 transition:transversion biases × 4 levels or rate heterogeneity). A total of 1,000 data sets were simulated for each combination of conditions. Maximum likelihood was implemented assuming a discrete gamma model with five rate categories (Yang 1993, 1994).

Results

In table 2, I summarize the results of simulations in which I replicated the results of the Tateno et al.

Table 1
A Summary of the Results from the Tateno et al. (1994) Study

			NJG			ML		
R	NJD	NJK	$a = 0.5$	$a = 1$	$a = 2$	$R = 0.5$	$R = 9$	$R = 15$
Model tree *A*:								
$a = 0.5$:								
0.5	90	90	100	98	95	98	95	91
9	77	87	95	91	86	88	99	99
15	72	84	95	88	84	84	97	97
$a = 1$:								
0.5	98	98	100	100	100	100	99	99
9	87	96	99	99	97	97	100	100
15	83	93	100	98	95	95	100	100
$a = 2$:								
0.5	100	100	100	100	100	100	100	99
9	96	99	100	100	99	99	99	99
15	93	99	100	99	99	98	100	100
$a = \infty$:								
0.5	100	100	100	100	100	100	100	100
9	99	100	100	100	100	100	100	100
15	99	100	100	100	100	100	100	100
Model tree *B*:								
$a = 0.5$:								
0.5	2	2	89	43	15	48	23	18
9	2	29	75	43	22	20	70	71
15	3	23	70	41	20	20	76	77
$a = 1$:								
0.5	17	17	100	94	67	80	43	44
9	5	60	95	82	56	41	91	92
15	3	48	94	78	52	31	90	94
$a = 2$:								
0.5	51	51	100	100	95	97	80	80
9	9	77	97	94	81	53	94	94
15	5	68	96	94	80	34	94	95
$a = \infty$:								
0.5	98	98	100	100	100	100	96	94
9	21	92	99	99	99	72	100	100
15	16	90	99	99	98	50	98	99
Model tree *C*:								
$a = 0.5$:								
0.5	98	98	98	98	98	97	95	95
9	93	93	95	94	93	96	98	98
15	92	93	94	93	93	93	95	95
$a = 1$:								
0.5	98	98	98	98	98	99	99	99
9	97	97	97	97	97	97	97	97
15	96	96	97	97	97	98	97	97
$a = 2$:								
0.5	98	98	98	98	98	100	98	98
9	97	97	98	98	98	99	99	99
15	97	97	98	97	97	98	98	98
$a = \infty$:								
0.5	98	98	99	99	99	98	98	96
9	98	98	99	99	98	98	99	99
15	98	98	99	98	98	96	98	98

NOTE.—*R*, transition:transversion bias; *a*, shape parameter of gamma distribution; NJD, neighbor joining with Jukes-Cantor distances; NJK, neighbor joining with Kimura distances; NJG, neighbor joining with gamma distances; ML, maximum likelihood.

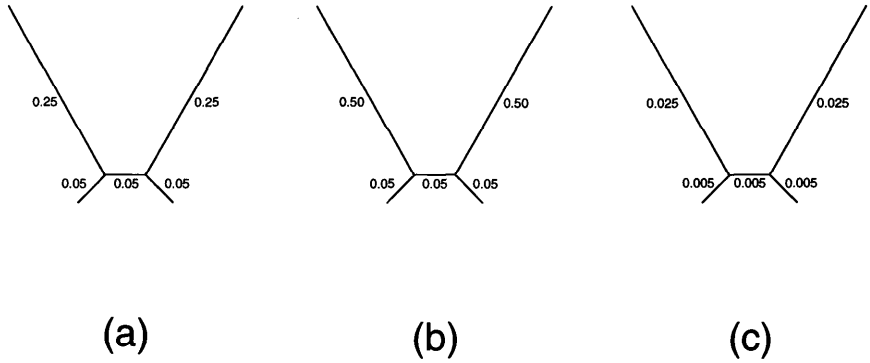

FIG. 1.—The model trees assumed in the Tateno et al. (1994) study and in this study. Branch lengths are expected number of substitutions per site.

(1994) study with the addition of maximum likelihood with gamma-distributed rate heterogeneity among sites. The simulations of table 2 were for model trees *A*, *B*, and *C* of figure 1. For each model tree, 12 different models of evolution were assumed; three different transition: transversion biases and four different levels of rate heterogeneity among sites were simulated (for a total of 12 simulations). Figure 2 summarizes the results of four of the simulations performed in this study in a matrix. The appropriate comparisons between the neighbor-joining and maximum-likelihood methods are along the diagonal, which has been shaded. Inappropriate comparisons, some of which were made by Tateno et al. (1994), are off the diagonal.

Several conclusions can be made about these simulations. The first is that, for any one simulation, one can find cases where neighbor joining performs as well as or better than maximum likelihood if one is willing to accept off-diagonal comparisons. The converse is also true; one can also find off-diagonal comparisons in which maximum likelihood performs better than neighbor joining. In short, if inappropriate comparisons are allowed (i.e., cases in which the model of DNA substitution assumed by the methods are different), virtually any conclusion about the relative performance of phylogenetic methods can be made.

The second result of these simulations is that maximum likelihood generally performs better than neighbor joining when the assumptions of the methods are the same (e.g., in the case where both methods assume a Jukes-Cantor model of evolution). In those cases in which the assumptions of maximum likelihood and neighbor joining were completely satisfied, maximum likelihood performed better in 29 of the 36 simulations (in four of the remaining seven cases, maximum likelihood and neighbor joining tied in performance). Similarly, maximum likelihood appears to be more robust than neighbor joining. In 133 of the 180 cases in which the assumptions of the maximum-likelihood and neighbor-joining methods were violated in the same manner, maximum likelihood outperformed neighbor joining. In 17 of the 180 cases, the performance of the two methods was the same, and in 30 of the 180 cases, neighbor joining performed better than maximum likelihood.

The final conclusion is that both methods had cases in which they failed with high probability. For example, when tree *B* of figure 1 was the model tree and sequences were generated using a Jukes-Cantor model with a high degree of rate heterogeneity among sites ($a = 0.5$), neighbor joining with a Kimura model of evolution (transition:transversion bias, $R = 15$) outperformed maximum likelihood with a Kimura model of evolution (also with a transition:transversion bias, $R = 15$) by 77.4%. Similarly, maximum likelihood outperformed neighbor joining by 97.3% when the same model of DNA substitution is assumed by the methods but the sequences were generated by a Jukes-Cantor model without rate heterogeneity among sites ($a = \infty$). Clearly, neither method is optimal for all the conditions examined in this simulation.

Discussion

Based on the simulations of both the Tateno et al. (1994) study and on this study, maximum likelihood appears to be more efficient and more robust than the neighbor-joining method. This conclusion contradicts the conclusions of Tateno et al. (1994). Although maximum likelihood appears more robust than neighbor joining, one should remember that this conclusion is based on very limited simulations. Only three model trees of four taxa were examined in this study and in the Tateno et al. (1994) study. Given that the branch lengths of the model tree are among the most important determinants of the accuracy of phylogenetic methods (Felsenstein 1978; Jin and Nei 1990; Nei 1991; Huelsenbeck and Hillis 1993; Charleston et al. 1994; Tateno et al. 1994; Huelsenbeck 1995), it is important to simulate many more model trees (as has been done by

Table 2
Probability of Obtaining the Correct Phylogeny (in percent) for the Simulations Performed in This Study

	Neighbor Joining						Maximum Likelihood					
	R			*a*			*R*			*a*		
	0.5	9	15	0.5	1	2	0.5	9	15	0.5	1	2
Model tree *A*:												
$a = 0.5$:												
0.5	90.3	100.0	99.8	100.0	98.2	95.3	98.4	93.1	91.7	99.9	99.6	99.0
9	78.0	88.1	90.5	96.8	91.8	86.7	90.6	98.3	98.7	99.6	99.2	99.5
15	71.0	79.6	82.1	94.6	88.2	82.5	84.3	97.6	97.3	99.3	99.6	98.6
$a = 1$:												
0.5	97.6	99.7	97.9	100.0	100.0	99.3	100.0	98.2	98.3	100.0	100.0	99.9
9	87.0	95.1	96.8	99.1	97.9	96.1	96.5	99.7	99.8	99.9	99.8	99.9
15	84.4	92.8	94.0	99.2	98.2	95.7	94.5	99.7	99.7	99.8	99.7	99.5
$a = 2$:												
0.5	99.9	98.5	98.8	100.0	100.0	100.0	100.0	99.8	99.3	100.0	100.0	100.0
9	94.2	99.0	99.9	99.9	99.9	99.7	99.2	99.9	99.9	100.0	100.0	99.9
15	91.8	97.0	98.3	99.9	99.6	98.8	98.0	99.7	99.8	99.9	100.0	99.8
$a = \infty$:												
0.5	100.0	99.9	99.6	100.0	100.0	100.0	100.0	100.0	100.0	99.9	99.9	100.0
9	99.3	99.9	99.9	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0	100.0
15	98.7	99.9	100.0	100.0	100.0	100.0	99.3	100.0	100.0	99.6	100.0	100.0
Model tree *B*:												
$a = 0.5$:												
0.5	2.7	97.2	94.9	88.8	44.3	16.0	46.8	22.1	17.5	96.7	88.2	72.8
9	3.7	26.9	44.2	71.4	38.7	18.4	23.4	77.0	77.3	97.1	93.5	85.9
15	2.7	13.6	21.5	71.6	38.1	20.1	18.4	71.3	77.5	94.0	91.8	86.1
$a = 1$:												
0.5	15.8	92.7	28.7	99.6	93.5	64.9	79.5	45.8	46.0	99.4	99.0	96.3
9	4.9	56.2	84.0	92.6	78.0	51.2	36.9	92.2	92.0	98.2	98.4	96.4
15	3.5	25.4	48.7	90.7	77.5	52.2	22.7	87.8	91.5	97.4	96.5	95.4
$a = 2$:												
0.5	51.9	51.1	2.9	99.8	99.7	96.1	95.5	77.7	77.3	99.7	99.7	99.7
9	9.5	78.5	97.0	96.5	94.6	81.7	46.5	97.0	97.2	98.9	98.6	99.0
15	5.2	38.4	68.4	96.1	93.8	81.0	35.7	95.0	96.6	98.0	98.6	97.7
$a = \infty$:												
0.5	97.9	56.5	0.2	100.0	100.0	100.0	99.9	97.7	97.5	99.8	100.0	100.0
9	21.7	93.0	99.5	98.3	98.4	97.9	67.4	99.5	99.3	98.5	99.1	99.7
15	11.5	51.3	89.8	98.4	98.4	97.7	51.3	98.8	98.8	98.1	98.7	98.2
Model tree *C*:												
$a = 0.5$:												
0.5	96.7	97.3	97.3	97.4	97.2	96.8	98.1	97.1	96.5	97.8	98.6	97.5
9	98.7	95.7	95.7	96.3	95.8	95.7	94.9	95.2	96.1	96.7	97.1	95.8
15	93.7	94.1	94.1	94.0	93.8	93.7	93.6	95.8	95.3	96.0	96.8	95.6
$a = 1$:												
0.5	98.2	98.6	98.8	98.8	98.6	98.4	99.2	97.7	97.3	98.8	98.4	98.5
9	96.4	96.3	96.3	97.4	96.6	96.6	96.8	98.0	97.5	97.7	97.5	98.2
15	97.0	97.1	97.1	97.9	97.4	97.4	97.1	97.5	96.5	98.0	97.5	96.7
$a = 2$:												
0.5	98.6	98.8	98.9	99.0	98.8	98.7	99.3	98.3	98.1	98.9	99.0	98.5
9	97.6	97.8	97.8	97.8	97.7	97.6	97.6	98.4	98.4	98.3	99.0	98.3
15	96.7	96.9	96.9	97.9	97.6	97.2	97.9	97.6	98.4	98.5	97.9	98.2
$a = \infty$:												
0.5	98.6	99.1	99.2	99.1	99.0	98.7	99.3	99.0	98.8	99.1	99.1	98.8
9	98.1	98.2	98.2	98.7	98.3	98.1	98.3	99.0	98.1	99.1	98.8	98.8
15	98.5	98.6	98.6	98.7	98.5	98.4	97.8	98.8	98.9	99.3	98.7	98.9

NOTE—*R*, transition:transversion bias; *a*, shape parameter of gamma distribution. Results are based on 1,000 simulated data sets for each combination of substitution model.

A

Maximum Likelihood			Neighbor Joining: Kimura ½	Kimura 9	Kimura 15	Gamma ½	Gamma 1	Gamma 2
			91.8	97.0	98.3	99.9	99.6	98.8
Kimura	½	98.0	+7.2	+1.0	−0.3	−1.0	−1.6	−0.8
Kimura	9	99.7	+7.9	+2.7	+1.4	−0.2	+0.1	+0.9
Kimura	15	99.8	+8.0	+2.8	+1.5	−0.1	+0.2	+1.0
Gamma	½	99.9	+8.1	+2.9	+1.6	0.0	+0.3	+1.1
Gamma	1	100.0	+8.2	+3.0	+1.7	+0.1	+0.4	+1.2
Gamma	2	99.8	+8.0	+2.8	+1.5	−0.1	+0.2	+1.0

B

Maximum Likelihood			Neighbor Joining: Kimura ½	Kimura 9	Kimura 15	Gamma ½	Gamma 1	Gamma 2
			97.9	56.5	0.2	100.0	100.0	100.0
Kimura	½	99.9	+2.0	+43.4	+99.7	−0.1	−0.1	−0.1
Kimura	9	97.7	−0.2	+41.2	+97.5	−2.3	−2.3	−2.3
Kimura	15	97.5	−0.4	+41.0	+97.3	−2.5	−2.5	−2.5
Gamma	½	99.8	+1.9	+43.3	+99.6	−0.2	−0.2	−0.2
Gamma	1	100.00	+2.1	+43.5	+99.8	0.0	0.0	0.0
Gamma	2	100.00	+2.1	+43.5	+99.8	0.0	0.0	0.0

C

Maximum Likelihood			Neighbor Joining: Kimura ½	Kimura 9	Kimura 15	Gamma ½	Gamma 1	Gamma 2
			2.7	97.2	94.9	88.8	44.3	16.0
Kimura	½	46.8	+44.1	−50.4	−48.1	−42.0	+2.5	+30.8
Kimura	9	22.1	+19.4	−75.1	−72.8	−66.7	−22.2	+6.1
Kimura	15	17.5	+14.8	−79.7	−77.4	−71.3	−26.8	+1.5
Gamma	½	96.7	+94.0	−0.5	+1.8	+7.9	+52.4	+80.7
Gamma	1	88.2	+85.5	−9.0	−6.7	−0.6	+43.9	+72.2
Gamma	2	72.8	+70.1	−24.4	−22.1	−16.0	+28.5	+56.8

D

Maximum Likelihood			Neighbor Joining: Kimura ½	Kimura 9	Kimura 15	Gamma ½	Gamma 1	Gamma 2
			98.7	95.7	95.7	96.3	95.8	95.7
Kimura	½	94.9	−3.8	−0.8	−0.8	−1.4	−0.9	−0.8
Kimura	9	95.2	−3.5	−0.5	−0.5	−1.3	−0.6	−0.5
Kimura	15	96.1	−2.6	+0.4	+0.4	−0.2	+0.3	+0.4
Gamma	½	96.7	−2.0	+1.0	+1.0	+0.4	+0.9	+1.0
Gamma	1	97.1	−1.6	+1.4	+1.4	+0.8	+1.3	+1.4
Gamma	2	95.8	−2.9	+0.1	+0.1	−0.5	0.0	+0.1

FIG. 2.—Comparing the performance of maximum likelihood and neighbor joining for several of the conditions examined in this study. The conditions simulated for each matrix are (*A*) tree = *A*, $R = 15$, $\alpha = 2$; (*B*) tree = *B*, $R = 1/2$, $\alpha = \infty$; (*C*) tree = *B*, $R = 1/2$, $\alpha = 1/2$; and (*D*) tree = *C*, $R = 9$, $\alpha = 1/2$. The performance of maximum likelihood for different implemented models is shown along the first column, and the performance of neighbor joining is shown along the first row for each matrix. The numbers with + or − represent how much better maximum likelihood performed for each comparison. The diagonal (shown in gray) represents what are called the "appropriate comparisons" in this article (i.e., those cases in which the assumptions of the methods match in terms of the model of DNA substitution assumed).

Huelsenbeck and Hillis [1993]; Huelsenbeck [1995]). Furthermore, the robustness of maximum likelihood and neighbor joining was examined for only a few model violations (transition:transversion bias and rate heterogeneity). Obviously, the assumptions of phylogenetic methods can be violated in many more ways (e.g., nonindependence among sites). It would be interesting to see whether maximum likelihood is also more robust than neighbor joining for other model violations.

The results from this study and the Tateno et al. (1994) study also point out the need in systematics for a way to test the model of evolution assumed by the phylogenetic method (in fact, Tateno et al. [1994] make this point very clearly in their discussion). Because the performance of methods depends so strongly on the match between the evolutionary processes generating the characters and the model assumed by the method, some way of testing one model of DNA substitution against another model seems imperative. Both maximum likelihood and neighbor joining allow many of their assumptions to be tested (Goldman 1993; Rzhetsky and Nei 1995), an important point in favor of both of these methods.

Acknowledgments

I thank Jim Bull, David Hillis, and David Swofford for helpful discussions. This article was improved through the comments of two anonymous reviewers. This work was supported by National Science Foundation grants DEB-9106746 and DEB-9221052 awarded to David Hillis.

LITERATURE CITED

CHARLESTON, M. A., M. D. HENDY, and D. PENNY. 1994. The effects of sequence length, tree topology, and number of taxa on the performance of phylogenetic methods. J. Computational Biol. **1**:133–151.

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

———. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

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.

HUELSENBECK, J. P., and D. M. HILLIS. 1993. Success of phylogenetic methods in the four-taxon case. Syst. Biol. **42**: 247–264.

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

NEI, M. 1991. Relative efficiencies of different tree-making methods for molecular data. Pp. 90–128 *in* M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, Oxford.

RZHETSKY, A., and M. NEI. 1995. Test of applicability of several substitution models for DNA sequence data. Mol. Biol. Evol. **12**:131–151.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

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.

YANG, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol. Biol. Evol. **10**:1396–1401.

———. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. **39**:306–314.

STANLEY SAWYER, reviewing editor

Received December 22, 1994

Accepted March 31, 1995