

Appropriate Likelihood Ratio Tests and Marginal Distributions for Evolutionary Tree Models with Constraints on Parameters

Rissa Ota,* Peter J. Waddell,†‡ Masami Hasegawa,† Hidetoshi Shimodaira,† and Hirohisa Kishino§

*The Graduate University for Advanced Studies and †The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, Japan; ‡Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand; and §Department of Agriculture and Life Sciences, University of Tokyo, Tokyo, Japan

We show how to make appropriate likelihood ratio tests for evolutionary tree models when parameters such as edge (internodes or branches) lengths have nonnegativity constraints. In such cases, under the null model of an edge length being zero, the marginal distribution of this parameter is proven to be a “half-normal”, that is, 50% zero values and 50% the positive half of a normal distribution. Other constrained parameters, such as the proportion of invariant sites, give similar results. To make likelihood ratio tests between nested models, e.g., H_0 : homogeneous site rates, and H_1 : site rates follow a gamma distribution with variance $1/k$, then asymptotically as sequence length increases, the distribution under H_0 becomes a mixture of χ^2 distributions, in this case 50% χ_0^2 , and 50% χ_1^2 (where the subscript denotes degrees of freedom, i.e., not the usually assumed 100% χ_1^2 ; which leads to a conservative test). Such mixtures are sometimes called $\bar{\chi}^2$ distributions. Simulations show that even with sequences as short as 125 sites, some parameters, including the proportion of invariant sites, fit asymptotic distributions closely.

Introduction

The maximum-likelihood (ML) method has become increasingly popular over the past decade for evaluating evolutionary trees (e.g., Swofford et al. 1996). So far, two distinct types of algorithms have been developed to calculate the probability of sequence data given a tree and a mechanism of site substitution. The better known of these is Felsenstein's (1981) algorithm, and the second is Hendy's algorithm or the Hadamard conjugation (Hendy 1989; Hendy and Penny 1993), which use Hadamard matrices. Given these algorithms, or their extensions to model unequal substitution rates across sites, the edge (internode or branch) lengths and other parameters of the model are iterated until the likelihood is maximized on a tree of interest. During iteration, the usual constraint is that all edge lengths have nonnegative values. Similar constraints apply to other parameters of interest, such as the proportion of invariant sites and the shape of a gamma distribution of site substitution rates (Yang 1993; Waddell and Penny 1996).

The standard use of likelihood ratio tests is to set up a null hypothesis (H_0) and an alternative hypothesis (H_1) prior to gathering the data. The expected sampling distribution of some statistic is then assessed between H_0 and H_1 , given that H_0 is true. This statistic is often $2 \ln(L(\mathbf{X}|H_1)/L(\mathbf{X}|H_0))$ or $2\Delta \ln L$, where \ln is the natural logarithm and $L(\mathbf{X}|H_i)$ is the likelihood of the data, \mathbf{X} , under hypothesis H_i . If $2\Delta \ln L$ observed from the data is larger than expected in $(1 - \alpha) \times 100\%$ of cases when H_0 is true, then H_0 is rejected at level α (where α equals the probability of rejecting H_0 when it is indeed true, a type I error). Otherwise, the correct terminology

is to say that we have failed to reject H_0 , allowing for the possibility that H_1 may indeed be better than H_0 , but the test-plus-data combination cannot detect this, a so-called type II error. In this paper, we consider the case in which H_0 is a simpler model nested within H_1 (e.g., H_0 can be derived from H_1 by setting certain parameter values to 0). Nestedness is important for deriving the asymptotic distributions of likelihood differences between H_0 and H_1 , as $2\Delta \ln L$ -based tests between non-nested models affected by boundary conditions require a distinct treatment.

For testing H_0 : a star (unresolved) tree, against H_1 : a prespecified resolved tree, it has previously been assumed that the sampling distribution about all edges in a tree is multivariate normal as the sequence length, c , becomes very long (e.g., Felsenstein 1981). Hence, it was assumed that with long sequences, $2\Delta \ln L$ between a resolved tree and the star tree would be distributed as χ_1^2 with four taxa, and, in general, χ_n^2 , for a specific contraction of n edge lengths set to 0 (Felsenstein 1988).

However, as the length is bounded below by 0, part of the tail of the marginal distribution of the length of an internal edge must be truncated, so a χ^2 distribution cannot be appropriate. This point has yet to be appreciated in the very broad field of phylogenetic analysis and the testing of evolutionary trees. For example, Gaut and Lewis (1995) tested the fit of the distribution of the log likelihood ratio between a resolved and an unresolved four-taxon tree to χ_1^2 . While the results did not meet their expectations, they did not diagnose the problem. Herein, we show what the marginal distributions of important parameters with boundary constraints are and how hypothesis tests between nested models need to be constructed to take this into account.

Materials and Methods

For simulations, the data were generated using the SeqGen program of Rambaut and Grassly (1996). It was reformatted using a C program by R.O. and analyzed by

Key words: Phylogenetic log likelihood ratio tests, truncated or half-normal distributions, mixed chi-square, gamma site substitution rate heterogeneity.

Address for correspondence and reprints: Peter J. Waddell, Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand. E-mail: pjwaddell@hotmail.com.

Mol. Biol. Evol. 17(5):798–803, 2000

© 2000 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

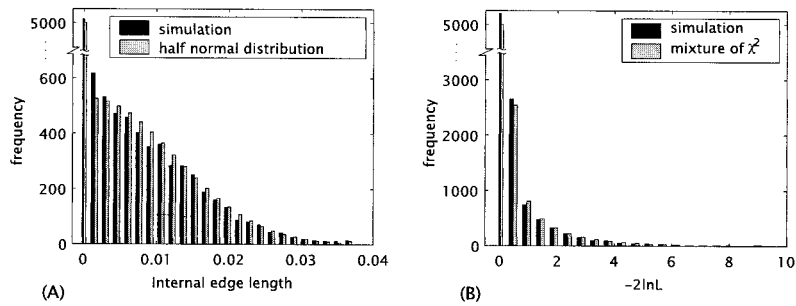


FIG. 1.—The simulated distributions of (A) the estimated length of the internal edge of the tree ((1, 2), 3, 4) and (B) $2\Delta\ln L$ between tree ((1, 2), 3, 4) and the star tree, with $c = 1,000$ sites and 10,000 replications. The theoretically expected distributions are, respectively, a half-normal distribution and a mixture of 50% χ_0^2 and 50% χ_1^2 . Data were simulated under a star tree with external edge lengths equal to 0.25.

PAUP*, version 4.0b2 (Swofford 1998), and Matlab. For all figures, the identical-site-rates Jukes and Cantor, or Poisson, model (see Swofford et al. 1996) was used to generate all the data; to analyze the data, ML estimation of all parameters was employed. Illustrated results were checked using the Hadamard conjugation algorithm and its extensions for unequal site rates (Waddell and Penny 1996), programmed in Matlab.

To evaluate the normality of the tails of distributions of constrained parameters for short sequences (e.g., fig. 3B–D), 50% of the smallest parameter values were subtracted from the distribution before remaining values were reflected about 0, before plotting with a normalized y-axis. This procedure takes into account the distortion that the boundary of the parameter space would otherwise introduce. To confirm that the proposed new distributions were fitting within sampling error, we used a χ^2 goodness-of-fit test, checking that no cells had an expected value of less than 5 (rejection of fit occurred only in fig. 3).

Results

Simulation of Edge Lengths

Figure 1A is an example of the marginal distribution of the reconstructed internal edge length of a resolved tree when the data were generated under a tree for which the true internal edge length was 0. Truncation occurs at the midpoint of a normal distribution, so half of all of the ML edge length estimates should take the value 0, a condition we call a “half-normal” distribution. Consequently, exactly half of the $2\Delta\ln L$ will take the value 0, and half will follow χ_1^2 for a four-taxon tree

(fig. 1B). Here, the mixed χ^2 (or $\bar{\chi}^2$) distribution is 50% χ_0^2 and 50% χ_1^2 , where χ_0^2 is the delta function at 0. The earlier practice of testing using χ_1^2 would result in rejecting H_0 (the star tree) too rarely under the model. Put another way, the earlier $2\Delta\ln L$ test of the positive length of an edge is conservative, with α actually half as small as expected.

The Distribution of Site Rates

We can estimate the proportion of invariant sites (p_{inv}) by maximizing the likelihood of the data on a given tree, conditional on p_{inv} being positive. Figure 2A suggests the distribution of the estimate of the proportion of invariant sites under H_0 : all sites having identical rates (or $p_{\text{inv}} = 0$) is half-normal also.

Suppose the variable sites follow a gamma (Γ) distribution with shape parameter k (Steel et al. 1993; Waddell, Penny, and Moore 1997; others, e.g., Yang [1993], call this parameter α). (We use k to distinguish this parameter from type I error rates and Kimura’s transition rate). The usual test of interest is $H_0: k = \infty$ (or identical site rates), against $H_1: k < \infty$ (e.g., Waddell 1995; Yang, Goldman, and Friday 1995). Parameter k is only constrained to be positive, so it appears there is no boundary problem, since infinity is a long way from zero. However, the real issue is whether the variance of the distribution of site rates (here $1/k$) is >0 . Figure 2B shows that the distribution of $1/k$ fits a half-normal distribution adequately. For likelihood ratio testing this is not a direct concern, since it is not essential to have a parameter in the form that directly yields the half-normal distribution. Thus, the 50% $\chi_0^2 + 50\%$ χ_1^2 test distribution for

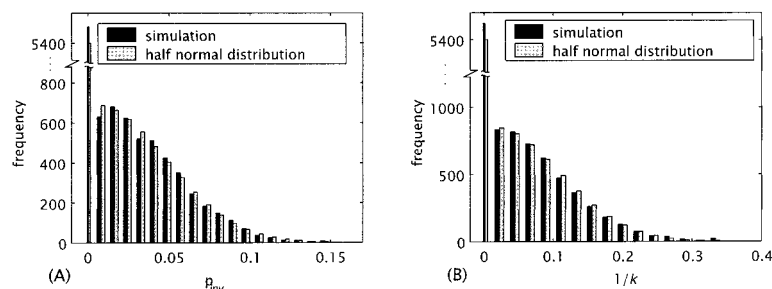


FIG. 2.—The distributions of (A) the estimated proportion of invariant sites (p_{inv}) and (B) parameter $1/k$ of the Γ distribution (10,000 replications each). All data were simulated under the same model used for figure 1, and all free parameters (one shape parameter and four edge lengths in each instance) were estimated via maximum-likelihood.

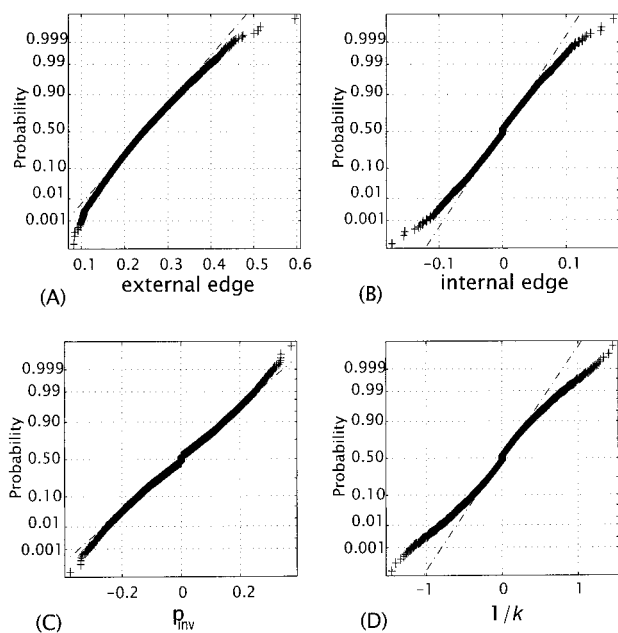


FIG. 3.—Normal plots of the marginal distributions of four types of parameters, respectively, (A) the internal edge of tree ((1, 2), 3, 4), (B) an external edge of the same tree, (C) p_{inv} , and (D) $1/k$, when the generating model is the same as that used in figure 1 except that $c = 125$ (note: there was a separate simulation [i.e., not jointly estimated] for each parameter with value 0). The dotted line is a best fit to the normal distribution not taking into account the discontinuity at the origin due to optimization problems.

$2\Delta \ln L$ should be used when testing both $p_{inv} = 0$ and $k = \infty$, rather than the earlier assumptions of 100% χ^2_1 (e.g., Waddell 1995; Yang, Goldman, and Friday 1995).

Testing Whether Multiple Parameters Are Zero

It is sometimes useful to test whether combinations of parameters lie on the boundary (e.g., two edges in a tree can be collapsed and thus are both equal to 0, or both p_{inv} and $1/k$ are 0). If parameter estimates are independent (noncorrelated) under H_1 , then the distribution of $2\Delta \ln L$ between H_0 and a prespecified H_1 is easily predicted for long sequences, e.g., for two parameters, it is $1/4 \chi^2_0 + 1/2 \chi^2_1 + 1/4 \chi^2_2$; for three parameters, it is $1/8 \chi^2_0 + 3/8 \chi^2_1 + 3/8 \chi^2_2 + 1/8 \chi^2_3$; and so on. Of course, if you have m additional parameters away from a boundary being simultaneously tested, the distribution will be $\bar{\chi}^2$ with the same coefficients, but the degree of freedom of each component will be raised by m . For example, with two on and one off the boundary, the distribution becomes $1/4 \chi^2_1 + 1/2 \chi^2_2 + 1/4 \chi^2_3$.

If there are n parameters being simultaneously tested, with one or more of them on the boundary, and any of those on the boundary have nonzero correlations with any of the other n parameters, then the distribution may still be $\bar{\chi}^2$, but the weights will no longer be so simple. In order to predict the correct distribution, it is necessary to estimate the information matrix \mathbf{I} (the inverse of the variance-covariance matrix of parameters) under H_1 (for examples of this type of problem, see Self and Liang 1987, pp. 607–609). For example, simultaneously test-

ing $p_{inv} = 0$ and $1/k = 0$ when the site rate distribution is a mixture of invariant sites and the remaining variable sites follow a Γ distribution requires \mathbf{I} (since these parameters are correlated; Waddell 1995; Waddell and Penny 1996). However, if correlations between parameters such as edges are near 0 (e.g., Waddell et al. 1994), then the distributions of the previous paragraph should be reasonable approximations.

It is worth remembering that biologists do not always have the tree (or hypothesis) to test before the data are analyzed, and this will cause a significant deviation from the distributions described here. In this case, H_1 will be better approximated by the maximum of a number of $\bar{\chi}^2$ distributions—an important topic for further study.

Proving the Form of Marginal Distributions

Above were demonstrations of the half-normal character of bounded parameters; proofs are needed for such claims to be adopted. If a parameter is constrained to be nonnegative, it will have a half-normal distribution when its true value is 0 under the following conditions:

1. The model is identifiable.
2. The data points are independent and identically distributed.
3. The central limit theorem applies.

Condition 1 was first proven for Hendy's theorem (Hendy and Penny 1993) and its derivatives for four states (e.g., Steel et al. 1993; Waddell, Penny, and Moore 1997), which includes the case of the Jukes-Cantor model. What this means is that under this model every possible weighted tree produces a different distribution of the data. Condition 2 is entrenched in the Jukes-Cantor model, for which sites are considered to evolve independently.

Condition 3 can be checked by showing that the first three derivatives of the likelihood function are defined and bounded, including at the boundary of the parameter space (e.g., when the parameter takes the value 0; Self and Liang 1987). In appendix 1, we prove that condition 3 holds for edge length parameters under the Jukes-Cantor model and for a proportion of invariant sites. While parameter k itself has a distinctly nonnormal distribution so that the Fisher information matrix (\mathbf{I}) becomes degenerate on the boundary, the likelihood ratio test is unaffected by the form of the parameter used; condition 3 needs to be verified for only one parameterization. For the Γ model, only the parameterization $1/k$ (the variance of the underlying distribution) or its transformation with nonzero first derivative at $k = \infty$ can meet condition 3. In other words, only $p = -1$ is allowed for the form k^p .

Testing with Finite Data

Biologists wanting to use the newly described tests need to know that the asymptotic results described above hold for the sequence lengths they are analyzing. To explore this, take the model used for figure 1 and reduce c until the marginal distributions of parameters

begin to lose their half-normal character, at which point $2\Delta\ln L$ will no longer follow a $\bar{\chi}^2$ distribution. Figure 3 shows one such result. The small kink in the middle of figure 3B–D is due to the iterative optimization of likelihood being affected by the boundary and the act of reflecting the positive values of the half-normal distribution. Here, at sequence length $c = 125$, most parameters fit the expected distribution reasonably (e.g., suitable for testing at the $\alpha = 0.01$ level), but the parameter $1/k$ has begun to lose its half-normal character. (Note: we show an external edge's distribution for those concerned that it might not follow a normal distribution).

The good fit of p_{inv} and the external edge is due to their backbone support being based on a binomial marginal with an expected value of greater than 10, while support of the internal edge is based on a binomial of a smaller expected value, but with a lot of “smoothing” due to the model compensating for multiple substitutions at sites (Waddell et al. 1994; Waddell 1995). What is meant by “backbone” support is the sum of the patterns in the data that correspond most directly to the parameter being considered. For p_{inv} , this is the frequency of constant sites; for an external edge, the frequency of matching singleton patterns (e.g. aggg); and for an internal edge, the frequency of patterns such as (aagg). In the references above, these frequencies are referred to as cs_i , or the sequence length times the probability of a site in the sequence most directly influencing a parameter value.

In finer detail, we can see that in figure 3A, the external edge has a slightly log-normal shape due to the thin tail for smaller values and the fat tail for larger values. Thus, depending on the direction in which a normal-based test is made, it could be either too conservative or too liberal (i.e., under- and overestimate type I errors, respectively) for small values of α (e.g., 0.001). Similarly, figure 3B and D shows that the internal edge and $1/k$ have distributions which will lead to $\bar{\chi}^2$ -based tests that are too liberal, while figure 3C shows a slightly thin tail for p_{inv} and thus slightly conservative tests.

Factors which reduce the quality of the asymptotic approximation to either the normal or the half-normal distribution include unequal edge lengths on the tree, very low or very high rates of substitution (i.e., very small or very large amounts of total evolution), and deviation of substitution type rates away from being equal (e.g., very different transition/transversion rates, highly unequal base compositions). The main thrust of these effects can be gathered from Waddell et al. (1994), which used Hadamard conjugations to examine such distributions. Due to the close relationship of Hadamard conjugations and likelihood methods (e.g., Waddell 1995; Waddell, Penny, and Moore 1997), at very low rates of change, edge lengths reconstructed using ML (and thus $2\Delta\ln L$ also) may even show a multinomial character if their backbone support drops to an expected value near 0. We found that sometimes extreme combinations of the above factors for certain types of data (e.g., mtDNA D-loops) result in a particularly poor convergence to the asymptotic distributions (simulations not shown). For example, if the simulation model used in

figure 1 is modified so that just site rates follow a Γ distribution with $k = 0.2$ (a realistic value with mtDNA sequence), then the convergence for parameters shown in figure 3A–C does not occur until $c > 1,000$.

Discussion

What we show here is based on the assumption that the evolutionary model holds. If the modeling of the phylogenetic process is not adequate, then parameters which should have an expected value tending to 0 often tend to take values which are significantly nonzero (e.g., due to excess homoplasy). This can in turn make statistical tests, including likelihood ratios, reject the null hypothesis too often, i.e., it can make them too liberal (e.g., see Gaut and Lewis 1995). Obviously, biologists still have to be cautious in interpreting test results for any real data, irrespective of a test being correctly prescribed with respect to theory.

Simulations of the distribution of parameters like edge lengths, p_{inv} , and k can run into a number of problems. There is the risk that the iterative likelihood optimization method will terminate prematurely or get stuck at a boundary due to unreliable estimates of gradients in this region (in our figures, there is an excess of values of 0). This can be checked by testing that 50% of parameter estimates with expected values of 0 take the value 0. This is probably not a major practical problem, but it provides a useful check for programmers.

The need to use $\bar{\chi}^2$ distributions can sometimes be avoided if one can diagnose the direction in which the model is deviating with respect to the data and then make a one-tailed test. However, using the appropriate $\bar{\chi}^2$ distribution solves the problem without the need for a specific diagnosis of the model.

It is also worth noting that the appearance of the half-normal and mixed χ^2 or $\bar{\chi}^2$ distributions upset the assumptions under which model selection criteria such as AIC (Akaike 1973), BIC (Schwarz 1978), etc. are derived, so these should be rederived to take this into account when selecting constrained parameters to include in an evolutionary tree model.

An interesting point is that while the expected value of a parameter constrained to be nonnegative is 0 at $c = \infty$, it must converge to this value from above. Thus, constrained ML gives rise to biased phylogenetic parameter estimates, but the small amount of bias introduced is more than offset by reduced variance. Therefore, the mean square error of parameters is always better with constraints in place.

Similarly, likelihood ratio tests of whether entries in rate matrices can be set to 0 will also run into boundary problems. Also note that when testing ratios of parameters (e.g., of transitions to transversions or of two edge lengths), independent errors of the denominator and the numerator slow convergence to a normal distribution, especially if parameters are near a boundary (e.g., Waddell and Penny 1996). Finally, one can also constrain external edges to be 0, e.g., H_0 external edge x is length 0, so that sequence x is directly ancestral to

others, and in this case one must consider the boundary effect under H_0 .

Boundary-like problems can also affect tests of the molecular clock. Basically, researchers should be careful to consider whether or not the marginal distribution of each parameter being constrained is normal under both H_0 and H_1 .

Apparently, the first direct published evidence that constrained parameters in phylogenetic models have half-normal distributions, such that $2\Delta \ln L$ will require $\bar{\chi}^2$ distributions for testing, is that of Ota, Waddell, and Kishino (1999). Independently, Whelan and Goldman (1999) recognized a problem affecting k and followed it up with simulation results (N. Goldman and S. Whelan, personal communication). It is surprising that boundary problems have been overlooked for so long, since Tajima (1992) considered them in the framework of distance methods.

In summary, when testing likelihood ratios between evolutionary tree models, it is important to realize that because of constraints, log ratios do not always tend to a single χ^2 distribution as sequences become long. Accordingly, the distribution on which likelihood ratio tests are based should be considered carefully before tests are made. Fortunately, evolutionary tests made using an a priori H_1 and constrained parameters and the expectation of a χ^2 distribution under H_0 will be conservative unless the degrees of freedom are miscalculated. For example, testing a hypothesis using χ_1^2 at $\alpha = 0.05$ is the same as testing with $1/2 \chi_0^2 + 1/2 \chi_1^2$ at $\alpha = 0.05/2$. This allows one to maintain confidence in the validity of published test results, assuming the model holds and H_1 was specified a priori.

Acknowledgments

R.O. and P.J.W. would like to thank the Marsden Fund of New Zealand. P.J.W. thanks the Japanese Society for the Promotion of Science, Nick Goldman and an anonymous reviewer for comments, and all authors acknowledge the support of Monbusho of Japan for grants in aid of this research. Special thanks go to Avner Bar-Hen for first pointing out to us the relevance of Self and Liang's work to phylogenetic problems.

APPENDIX

Proof of the Three-Times-Differentiable Nature of the Log Likelihood, $\ln L$, as Parameters Go to Zero For Edges

Under the Jukes-Cantor model with identical site rates, the transition probability of base $i \rightarrow j$ may be represented as

$$P_{ij}(\tau) = \begin{cases} \frac{1}{4} + \frac{3}{4} \exp\left(-\frac{4}{3}\tau\right), & i = j \\ \frac{1}{4} \left(1 - \exp\left(-\frac{4}{3}\tau\right)\right), & i \neq j, \end{cases}$$

where τ is the edge length. Obviously, the exp function is three times differentiable and bounded as its argument

goes to 0. Since the likelihoods of site patterns are the sums of products of entries ($P_{ij}(\tau)$) for all edges in the tree (Felsenstein 1981), L is bounded and differentiable. Since $\ln L = \sum f_i \ln(s_i)$, this means that $\ln L$ too is bounded as $\tau \rightarrow 0$ (using L'Hopital's rule), where s_i and f_i are the site pattern probabilities (likelihoods) and frequencies, respectively. Note that the $P_{ij}(\tau)$ of the generalized Kimura 2ST and 3ST model's (e.g., Steel et al. 1993) are similar and the proof extends to such cases.

For the Proportion of Invariant Sites

Under such models, $\ln L$ may be expressed as

$$\ln L = \sum_{l=1}^4 f_l \ln[p_{\text{inv}}/4 + (1 - p_{\text{inv}})s_l] + \sum_{m=5}^{4^t} f_m \ln[(1 - p_{\text{inv}})s_m],$$

where t is the number of taxa. s_1, \dots, s_4 and f_1, \dots, f_4 are the site pattern probabilities (likelihoods) and frequencies, respectively, of the constant sites, and likewise with s_5, \dots and f_5, \dots for the site patterns showing multiple states. Since p_{inv} is involved as a simple product in the likelihood function, L is three times derivable, and as $p_{\text{inv}} \rightarrow 0$, $\ln L$ converges directly to that of the model above.

LITERATURE CITED

- AKAIKE, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267–281 in B. N. PETROV and F. CSAKI, ed. Second International Symposium of Information Theory. Akademiai Kiado, Budapest.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**:368–376.
- . 1988. Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* **22**:521–565.
- GAUT, S. B., and P. O. LEWIS. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. *Mol. Biol. Evol.* **12**:152–162.
- HENDY, M. D. 1989. The relationship between simple evolutionary tree models and observable sequence data. *Syst. Zool.* **38**:310–321.
- HENDY, M. D., and D. PENNY. 1993. Spectral analysis of phylogenetic data. *J. Classif.* **10**:5–24.
- OTA, R., P. J. WADDELL, and H. KISHINO. 1999. Statistical distribution for testing the resolved tree against the star tree. Pp. 15–20 in Proceeding of the annual joint conference of the Japanese biometrics and applied statistics societies. Sinfonica, Minato-ku, Tokyo, Japan.
- RAMBAUT, A., and N. C. GRASSLY. 1996. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* **13**:235–238.
- SCHWARZ, G. 1978. Estimating the dimension of a model. *Ann. Stat.* **6**:461–464.
- SELF, S. G., and K.-Y. LIANG. 1987. Asymptotic properties of maximum likelihood estimator and likelihood test under nonstandard conditions. *J. Am. Stat. Assoc.* **82**:605–609.
- STEEL, M. A., L. SZÉKELY, P. L. ERDŐS, and P. J. WADDELL. 1993. A complete family of phylogenetic invariants for any number of taxa under Kimura's 3ST model. *N. Z. J. Bot.* **31**:289–296.

- SWOFFORD, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0b2. Sinauer, Sunderland, Mass.
- SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. HILLIS, C. MORITZ, and B. K. MABLE, eds. *Molecular systematics*. 2nd edition. Sinauer, Sunderland, Mass.
- TAJIMA, F. 1992. Statistical method for estimating the standard error of branch lengths in a phylogenetic tree reconstructed without assuming equal rates of nucleotide substitution among different lineages. *Mol. Biol. Evol.* **9**:168–181.
- WADDELL, P. J. 1995. Statistical methods of phylogenetic analysis: including Hadamard conjugations, LogDet transforms, and maximum likelihood. Ph.D. thesis, Massey University, New Zealand.
- WADDELL, P. J., and D. PENNY. 1996. Evolutionary trees of apes and humans from DNA sequences. Pp. 53–73 *in* A. J. LOCK and C. R. PETERS, eds. *Handbook of human symbolic evolution*. Oxford University Press, Oxford, England.
- WADDELL, P. J., D. PENNY, M. D. HENDY, and G. ARNOLD. 1994. The sampling distributions and covariance matrix of phylogenetic spectra. *Mol. Biol. Evol.* **11**:630–642.
- WADDELL, P. J., D. PENNY, and T. MOORE. 1997. Hadamard conjugations and modeling sequence evolution with variable rates across sites. *Mol. Phylogenet. Evol.* **8**:33–50.
- WHELAN, S., and N. GOLDMAN. 1999. Distribution of statistics used for the comparison of models of sequence evolution in phylogenetics. *Mol. Biol. Evol.* **16**:1292–1299.
- YANG, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* **10**:1396–1401.
- YANG, Z., N. GOLDMAN, and A. FRIDAY. 1995. Maximum likelihood trees from DNA sequences: a peculiar statistical estimation problem. *Syst. Biol.* **44**:384–399.

MIKE HENDY, reviewing editor

Accepted January 25, 2000