

Maximum Likelihood Outperforms Maximum Parsimony Even When Evolutionary Rates Are Heterotachous

Sudhindra R Gadagkar* and Sudhir Kumar†‡

*Department of Biology, University of Dayton; †Center for Evolutionary Functional Genomics, The Biodesign Institute, Arizona State University; and ‡The School of Life Sciences, Arizona State University

Heterotachy occurs when the relative evolutionary rates among sites are not the same across lineages. Sequence alignments are likely to exhibit heterotachy with varying severity because the intensity of purifying selection and adaptive forces at a given amino acid or DNA sequence position is unlikely to be the same in different species. In a recent study, the influence of heterotachy on the performance of different phylogenetic methods was examined using computer simulation for a four-species phylogeny. Maximum parsimony (MP) was reported to generally outperform maximum likelihood (ML). However, our comparisons of MP and ML methods using the methods and evaluation criteria employed in that study, but considering the possible range of proportions of sites involved in heterotachy, contradict their findings and indicate that, in fact, ML is significantly superior to MP even under heterotachy.

Among-site variation in substitution rates, representing differential functional constraints at any given time across the length of a sequence, has been successfully modeled by the use of the gamma distribution (Uzzell and Corbin 1971; Rzhetsky and Nei 1994; Yang 1996). The gamma parameter has since been incorporated into nucleotide-substitution models (see Nei and Kumar 2000; Felsenstein 2003), thus improving the accuracy of phylogenetic inference in such cases. Functional constraints can also change at a given site over time (across lineages), manifesting as a nonuniform substitution rate in different species at that site. Furthermore, these relative rates across species may also be different for different sites in the alignment. This condition where sites in an alignment evolve with different relative rates in different lineages in a phylogeny has been termed heterotachy (Philippe and Lopez 2001). Heterotachy is now increasingly recognized in molecular data sets (Lopez, Casane, and Philippe 2002; Citerne et al. 2003; Gribaldo et al. 2003; Philippe et al. 2003) and has even been specifically implicated in incorrect phylogenetic inference (Inagaki et al. 2004).

Similarity of functional constraints over time is a fundamental assumption in virtually all phylogenetic reconstruction methods, and studies attempting to understand the effect of heterotachy on phylogenetic inference are beginning to be undertaken. Recently, Kolaczkowski and Thornton (2004) (K&T) reported the results of a simulation study where they compared the performance of maximum parsimony (MP) and parametric methods (e.g., maximum likelihood, ML) in inferring the true four-taxa tree when the evolutionary rates were heterotachous and found that MP outperformed ML. Spencer, Susko, and Roger (2005) have since shown that ML can perform at least as well as MP (and sometimes even better) when the sequence evolution is governed by a more sophisticated model (K&T used a one-parameter model) and when many possible relative rate combinations are considered. We have examined the robustness of the conclusions of K&T in another dimension: the influence of the proportion of sites affected by

heterotachy on the performance of MP and ML methods. This is because K&T assumed that 50% of the sites are affected by heterotachy. For simplicity and for the sake of direct comparison with the original study, we have used the exact simulation method and evaluation criteria employed by K&T (see *Methods*).

Using the simulation constructs of K&T, we estimated the minimum interior branch length required for 50% accuracy (BL_{50}) for the entire range of the proportion of heterotachous sites, f_H (see *Methods*). We find that MP performs better (has lower BL_{50}) than ML only in the range $32\% < f_H < 68\%$ (see also Gaucher and Miyamoto 2005). Thus, ML outperforms MP in more than 60% of the cases (fig. 1a), and the BL_{50} required by ML is 20% lower than that required by MP on average. Furthermore, the proportion of trees inferred correctly (P_C) by ML is 50% higher than by MP (fig. 1b). This indicates that ML is superior over a wider range of the proportion of heterotachous sites and by a larger amount when it performs better. This is in contrast to the conclusions of K&T, who implied unqualified superiority of MP over ML under heterotachy.

These simulation studies reveal a peculiar pattern in the performance of MP under heterotachy: BL_{50} does not appear to be significantly affected by a change in the percentage of heterotachous sites, as the interior branch length (BL_{50}) remains at 0.22 substitutions/site. MP is superior to ML at this specific interior branch length (see also K&T). However, it is significantly worse over a majority of the simulation conditions when this optimal branch length is reduced even by only 5% (fig. 1c). Overall, ML is seen to be three times more accurate than MP over the entire range of f_H (fig. 1c), which again contradicts the primary findings reported by K&T.

Because it is typically not easy to determine the extent of heterotachy in most data sets, K&T advocate using both MP and ML when inferring phylogenies, based on their finding that MP is better than ML when there is heterotachy and the fact that ML outperforms MP when there is no heterotachy in the data (Nei and Kumar 2000; Felsenstein 2003). However, our findings and other recent results (Gaucher and Miyamoto 2005; Spencer, Susko, and Roger 2005) invalidate the reasoning of K&T because ML is found to be generally superior to MP even under heterotachy. Therefore, if MP and ML methods produce different phylogenetic trees for the same data set, there appears to be

Key words: heterotachy, phylogenetic inference, maximum likelihood, maximum parsimony.

E-mail: gadagkar@notes.udayton.edu.

Mol. Biol. Evol. 22(11):2139–2141. 2005

doi:10.1093/molbev/msi212

Advance Access publication July 13, 2005

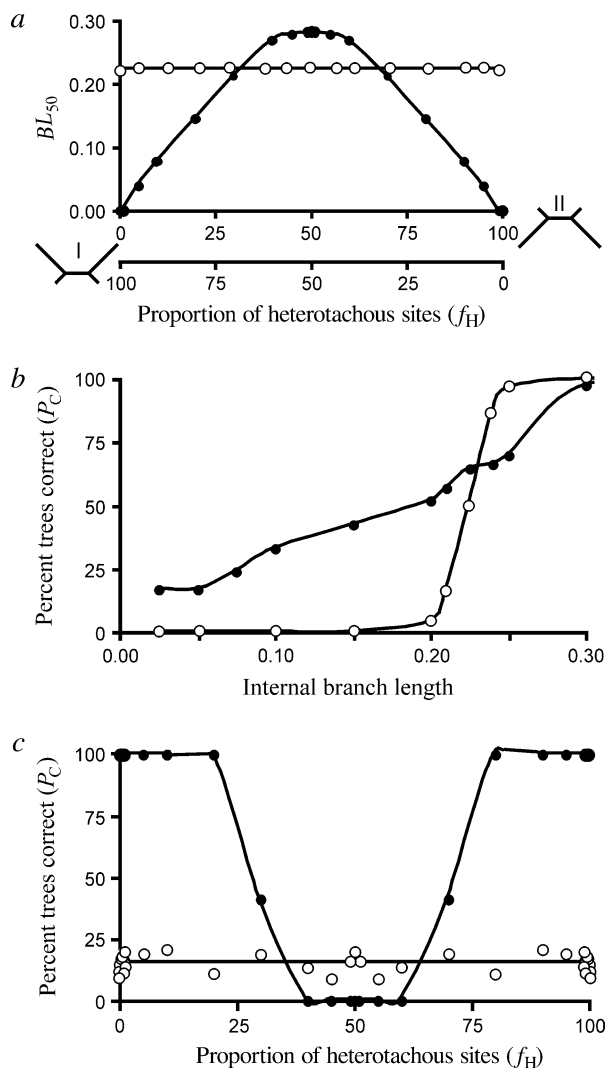


FIG. 1.—Performance of ML (closed circles) and MP (open circles) methods in computer simulations with unequal (heterotachous) site-specific evolutionary rates among species. Heterotachy was introduced in the data following the methods described by K&T. (a) The minimum internal branch length needed for 50% success (BL_{50}) in inferring a four-taxa tree correctly is plotted against the fraction of heterotachous sites, f_H . (b) The percentage of four-species trees inferred correctly in ML and MP analyses for different internal branch lengths. Each data point is the average percent accuracy over 0%, 10%, 20%, 30%, 40%, and 50% heterotachous sites. (c) The percentage of four-taxa trees inferred correctly in ML and MP analyses when the interior branch length is 0.21 (see text), plotted against the proportion of heterotachous sites, f_H .

little justification at present to believe that the MP tree is more accurate, even if evolutionary rates are heterotachous.

Methods

The simulation protocols and evaluation criteria were implemented as in K&T. We describe the basic design here in brief. The simulation was along a four-taxa tree ((A, B), (C, D)). Heterotachy was introduced into the data by conducting simulation along the tree separately for fractions f_H ($0 \leq f_H \leq 1$) and $(1 - f_H)$ of the total sequence length. Two substitution rates, p and q , were then alternated for the f_H fraction, such that A and C received substitutions at rate p

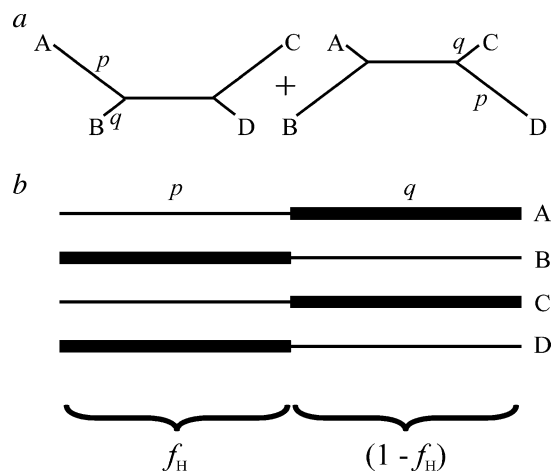


FIG. 2.—The simulation strategy used by K&T to introduce heterotachy in the data. (a) A fraction f_H of the sequence length is evolved along the branches of the first four-taxa tree and the remaining fraction $(1 - f_H)$ along the second. The resulting two sets of sequences are then concatenated to obtain a data set that contains heterotachous sequences. (b) A schematic of the resulting heterotachous data set where each of the four sequences evolves partly (f_H) with rate p (thin line) and partly $(1 - f_H)$ with rate q (thick line). In the illustration, $f_H = 0.50$, the fraction most often used by K&T in their study.

and B and D received them at q (fig. 2a). For the $(1 - f_H)$ fraction of the sequence length, the rates were reversed so that p was the rate of substitution for taxa B and D and for A and C it was q . This procedure ensured that, when p and q were different in magnitude, a given site, anywhere in the sequence, evolved with one rate for A and C and another for B and D, thus simulating different site-specific speeds in the phylogeny. The two fractions f_H and $(1 - f_H)$ of the sequences thus generated were then concatenated to produce a four-taxa data set with heterotachous sequences (fig. 2b). We followed the K&T system for assigning site rates deterministically in order to generate results that are directly comparable. This simulation strategy is also biologically realistic because one would expect different functional domains within a gene or different genes in a concatenated sequence alignment to be good candidates for heterotachy. The level of heterotachy in a data set can be manipulated in at least two direct ways: (1) changing the relative values of p and q to bring about different intensities of heterotachy for a fixed percentage of sites ($f_H = 50\%$), and (2) altering the value of f_H to change the extent of heterotachy in terms of the fraction of the sequence length for fixed values of p and q . K&T found that values of $p = 0.75$ and $q = 0.05$ substitutions per site brought about maximal reduction of ML performance when compared to homogeneous controls and MP. We used these values of p and q in our study, thus providing the worst-case scenario for ML. The value of f_H was fixed at 0.50 by K&T when comparing the relative performances of MP and ML. In our simulations, we considered the range $0.1 \leq f_H \leq 0.99$ to allow for heterotachy to manifest over the range of possible proportions. The total concatenated sequence length in our study was 10,000 nt.

For phylogenetic inference with MP and ML we performed an exhaustive search using PAUP* 4.0b10 (Swofford 2001), with default parameter settings. A 50%

majority rule consensus tree was obtained (with the LE50 option turned on in PAUP*) for both methods if there were multiple equally parsimonious or equally likely trees.

Evaluation of the inferred trees was done using the criteria of K&T. Specifically, this involved two quantities: BL_{50} (the minimum internal branch length that was associated with 50% of the simulation replicates being inferred accurately) and P_C (proportion of simulation replicates inferred correctly).

Acknowledgments

This work was supported by start-up funds from the University of Dayton (S.R.G.) and a grant from National Institutes of Health (S.K.). We wish to thank Joe Thornton and Bryan Kolaczkowski for commenting on an earlier draft of the manuscript, Michael Rosenberg for his simulation program, and Mark Nielsen for helpful comments.

Literature Cited

- Citerne, H. L., D. Luo, R. T. Pennington, E. Coen, and Q. C. B. Cronk. 2003. A phylogenomic investigation of CYCLOIDEA-like TCP genes in the Leguminosae. *Plant Physiol.* **131**: 1042–1053.
- Felsenstein, J. 2003. *Inferring phylogenies*. Sinauer Associates, Sunderland, Mass.
- Gaucher, E. A., and M. M. Miyamoto. 2005. A call for likelihood phylogenetics even when the process of sequence evolution is heterogeneous. *Mol. Phylogenet. Evol.* (in press).
- Gribaldo, S., D. Casane, P. Lopez, and H. Philippe. 2003. Functional divergence prediction from evolutionary analysis: a case study of vertebrate hemoglobin. *Mol. Biol. Evol.* **20**: 1754–1759.
- Inagaki, Y., E. Susko, N. M. Fast, and A. J. Roger. 2004. Covarion shifts cause a long-branch attraction artifact that unites microsporidia and archaeobacteria in EF-1 alpha phylogenies. *Mol. Biol. Evol.* **21**:1340–1349.
- Kolaczkowski, B., and J. W. Thornton. 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* **431**:980–984.
- Lopez, P., D. Casane, and H. Philippe. 2002. Heterotachy, an important process of protein evolution. *Mol. Biol. Evol.* **19**:1–7.
- Nei, M., and S. Kumar. 2000. *Molecular evolution and phylogenetics*. Oxford University Press, New York.
- Philippe, H., D. Casane, S. Gribaldo, P. Lopez, and J. Meunier. 2003. Heterotachy and functional shift in protein evolution. *IUBMB Life* **55**:257–265.
- Philippe, H., and P. Lopez. 2001. On the conservation of protein sequences in evolution. *Trends Biochem. Sci.* **26**:414–416.
- Rzhetsky, A., and M. Nei. 1994. Unbiased estimates of the number of nucleotide substitutions when substitution rate varies among different sites. *J. Mol. Evol.* **38**:295–299.
- Spencer, M., E. Susko, and A. J. Roger. 2005. Likelihood, parsimony and heterogeneous evolution. *Mol. Biol. Evol.* **22**: 1161–1164.
- Swofford, D. L. 2001. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Sinauer Associates, Sunderland, Mass.
- Uzzell, T., and K. Corbin. 1971. Fitting discrete probability distributions to evolutionary events. *Science* **172**:1089–1096.
- Yang, Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* **11**:367–372.

William Martin, Associate Editor

Accepted July 4, 2005