1

2

3

4

EFFECTS OF TREE SHAPE ON THE ACCURACY OF MAXIMUM LIKELIHOOD-BASED ANCESTOR RECONSTRUCTIONS

7

9

Arne Ø. Mooers

Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada V5A 1S6e-mail: amooers@sfu.ca

13

17   INTRODUCTION

18   Ancestral state reconstruction is a powerful tool for testing theories of the order and

19   timing of character evolution (Maddison, 1994; Maddison, 1995; Yang et al., 1995; Schluter et

20   al., 1997; Cunningham et al., 1998). Reconstruction methods that include explicit information

21   both on the relative probabilities of change and the opportunity for change now exist for

22   molecular sequences and amino acids (Yang et al., 1995; Koshi and Goldstein, 1996) and for

23   both multi-state discrete and continuous data (Martins, 1994; 1999; Schluter, 1995; Schluter et

24   al., 1997, Pagel, 1997, 1999, pers. comm.; Mooers and Schluter, 1999). These latter model-

25   based methods have been utilized successfully in ecological and evolutionary studies (see, e.g.,

26   Ryan and Rand, 1999; Hoekstra and Edwards, 2000; Morales, 2000; Matyjasiak et al., 2000;

27   Oakley and Cunningham, 2002; Berendonk et al., 2003), but their behaviors are still largely

28   unknown: Preliminary work highlighted that reconstructions for characters for which there is

29   only a single observation per tip have large confidence limits, and this uncertainty increases with

30   the inferred rate of evolution (Schluter, 1995; Schluter et al., 1997; note one recent study using

31   fossils suggests such confidence limits may be too broad for continuous characters; Polly, 2001);

32   as expected, larger trees offer better (in the sense of more strongly supported) ancestral estimates

33   (Yang et al., 1995; Mooers and Schluter, 1999).

34   The shape of a phylogenetic tree affects both the ease with which it is properly

35   reconstructed (see, e.g., Rohlf et al., 1990; Heijerman, 1993; Mooers et al., 1995; Salisbury,

36   1999) and the ease with which simple Markov traits can be reconstructed under Maximum

37   Parsimony (Salisbury and Kim, 2001). Tree shape also affects the performance of various

38   comparative tests (see, e.g., Martins and Garland, 1991; Harvey and Rambaut, 1998; Lorch and

39   Eadie, 1999), and offers important clues to the process of macroevolution (see, e.g., Purvis,

40   1996): for instance, adaptive radiations might be associated with bursts of speciation early in a

41   clade's history, leading to star-like phylogenies (Schluter, 2000). Tree shape also has

42   implications for conservation biology (Nee and May, 1997; Heard and Mooers, 2000). What

43   aspects of a tree's shape affect our ability to reconstruct history in the likelihood framework?

44   Here I use simple computer simulation and a small set of test trees to explore the effects

45   of aspects of tree shape (balance and stemminess) and the average rate of change on the accuracy

46   of ancestor reconstructions for a simple two-state character.

47

48

METHODS

50   *Reconstructions*: Ancestral states at nodes were estimated with recourse to a two-state

51   Markov process approach as implemented in DISCRETE (Pagel 1997); I refer the reader to

52   Schluter et al. (1997) and Pagel (1999) for clear expositions of the method and justification.

53   Briefly, consider a phylogenetic tree and a set of observations for a discrete character at its tips.

54   We seek some measure of the relative likelihood that a focal internal node (ancestor) was in a

55   particular state. Schluter (1995) introduced two distinct approaches. The first (coined the 'local'

56   approach by Pagel, 1999) compares the fit of the data to the hypothesis that the node is in one

57   state with the fit of the data to the hypothesis that the node is in the alternate state. This is done

58   by fixing the node in one state and maximizing the likelihood of observing the data (at all the

59   tips) by optimizing all other internal nodes and the overall rate of change. The node is then fixed

60   in the alternate state and the maximum likelihood is recalculated, again by optimizing the other

61   parameters. The maximum likelihood state is the one with the higher likelihood, and the support

62   for this state is the ratio of the two likelihoods.

63   The second ('global') approach can be interpreted as the posterior probability of a

64   particular state at a node in an 'empirical Bayes' framework (Yang et al., 1995; Schluter et al,

65   1997). The empirical estimate is a set of transition rates (for the simplest case, a single transition

66   rate between two states) that maximizes the likelihood of observing the tip data on the tree. Our

67   focal node has contributed to the resulting likelihood in each of the two states (remember that the

68   maximum likelihood will be a sum over all possible paths to the observed data). The support for

69   a particular state at a node is the ratio of the partial likelihood when it was in this state to the

70   partial likelihood when it was in the alternate state. This is the same as computing the 'local'

71   support for a state at a node *conditioned on a single set of maximum likelihood transition rates*

72   (Schluter, 1995). The global estimate has the advantage of allowing all nodes on a tree to be

73   estimated much more efficiently. However, the local estimate might be thought of as a fairer

74   comparison of competing hypotheses at a focal node (Pagel, 1999).

75   *Tree shapes and rates:* The computational intensity of maximum likelihood

76   reconstruction makes simulation across a universe of tree shapes impractical. I therefore created

77   twelve trees of size 32 (Figure 1, Table 1). These trees were not chosen because of their high

78   probability of occurrence under plausible models of diversification. Indeed, one of the unlabelled

79 topologies (the fully unbalanced branching order) has exceedingly low probability under random
80 diversification (Stone and Repka, 1998): it is 10^{14} times less likely than the topology of an
81 average tree grown under the Yule model. This is because it is a constrained shape - very few
82 realized histories of diversification will lead to a perfect comb. (Interestingly, the fully balanced
83 tree, which also looks highly constrained, is only 14 times less likely than any average tree.)
84 Rather, they were chosen in order to capture independent variation in two relevant tree shape
85 parameters (under some plausible assumptions): imbalance, as measured by Colless' index Ic,
86 (Colless, 1982; Heard, 1992) and stemminess, as measured by noncumulative stemminess index
87 Stni (Rohlf et al., 1990). The balance index measures the extent to which lineages fail to break
88 into unequal-sized groups, while the stemminess index measures the average distinctiveness of
89 the taxonomic subsets on a tree: a more unbalanced tree has an Ic value nearer 1.0 and contains
90 sister-groups of dissimilar sizes, while a tree with a stemminess score nearer 1.0 includes taxa
91 that form compact groups with long unbranched stems leading to them. Stemmier trees have
92 more of their nodes nearer the tips. The two measures tend to be highly negatively correlated
93 under simple models of diversification, such that unbalanced trees are not stemmy. To create my
94 trees, I first generated 500 Yule trees (trees created when under a pure birth model with constant
95 rate of splitting) and took four from the outside edges of the envelope describing the distribution
96 of balance vs. stemminess. I then manipulated the branch lengths to either decrease or increase
97 stemminess, keeping topology the same. I then added the two extreme topologies, fully balanced
98 and fully unbalanced. However, the two fully unbalanced trees are perforce also the two least
99 stemmy ones, and so confound my two main variables. I therefore treat these two trees separately
100 in the analysis. The manipulations and tree statistics are summarized in Table 1.

101   Both balance and stemminess have straightforward biological explanations: Heard (1996)
102 has shown by simulation that diversification with varying speciation rates across lineages quickly
103 leads to unbalanced trees. Such variation need not be heritable, though it is generally modelled
104 as being so. Increased rates of extinction, on the other hand, will produce stemmier trees, though
105 here extinction rate need not vary among lineages (see, e.g., figure 1 in Harvey and Rambaut,
106 1998). Of course, various forms of nonrandom extinction of intermediates should also increase
107 stemminess (Simpson, 1953, in Schluter, 2000), while long latencies between speciation events
108 along a lineage can also lead to unbalanced trees (Heard, 1996; Chan and Moore, 1999).

109   I also considered rate of evolution. It is already known that information about ancestral

110 states is quickly lost at higher rates of character evolution as transitions go unrecorded (Schluter,

111 1995; Schluter et al., 1997). More generally, information is lost with more opportunity for

112 change (which is just rate x time; see, e.g. Sober and Steel, 2002). Making trees more

113 unbalanced and less stemmy makes them sample more time overall (since both push nodes

114 nearer the root), and so both provide more opportunity for change (see Figure 1). In order to

115 investigate the effects of rate and shape independently, I scaled each of the twelve tree shapes to

116 be the same total length, with length measured not as time but as total expected number of

117 changes (rate x time). I then arbitrarily made trees with two rates of evolution, such that either 5

118 changes ('low' rate) or 10 changes ('high' rate) per character were expected over the entire tree. It

119 is important to note that this produces trees of different 'depth' or age, but is necessary to

120 consider the effects of topology independently of opportunity for change.

121   For each of the 24 trees (12 tree shapes x 2 rates) I assigned the ancestral state (0) to the

122 root and simulated 200 characters using a modified version of the simulation program Seq-Gen

123 (v.1.2.3; Rambaut and Grassly, 1997), with equal probability of changes to and from state 0. A

124 UNIX-based version of DISCRETE (Pagel, 1997; 1999, modified by Daniel Barker) was then

125 used to reconstruct the relative likelihoods for each character at each node for both local and

126 global estimates, without information on the state of the root (because this is rarely known in

127 practice), and without constraining forward and backward rates to be equal. In order to make all

128 200 runs comparable, I did not test whether the data generated for each individual run justified

129 fitting different forward and backward rates (see, e.g., Mooers and Schluter, 1999).

130   Each internal node has six pieces of information associated with it for each replicate: (1)

131 the imbalance and (2) stemminess measure for its tree; (3) its relative distance from the tips,

132 scaled from 0 for the present to 1.0 for the root; (4) the simulated ('true') value for the trait (0 or

133 1); and the (5) local and (6) global relative likelihoods of state 0. With these data, we can ask

134 how the parameters Ic, $Stni$, rate, and relative distance from tips affect the accuracy of the

135 ancestral reconstructions. Because the states are coded as 0 or 1, accuracy can be measured as

136 mean[abs(relative likelihood of state 0 - true state)]. This measure does not consider the

137 agreement between a reconstruction at a node and the truth (eg. how many times the favored

138 reconstruction equals the true state). Rather, it scales from 0 to 1. For example, if, for all 200

139 runs, the true state is '1' at a node but DISCRETE returns the relative probability of state '0' to

140   be 1.0 (true states and reconstructions disagree maximally), accuracy is 0.0; if the true state is '1'

141   and DISCRETE returns the relative probability of state '0' as 0.0 (they agree maximally), the

142   accuracy is 1.0. The most straightforward way to analyse these data is to consider the nodes in

143   the tree as independent observations, and to use the tools of General Linear Model analysis. To

144   best meet the assumptions for such a parametric statistical comparison (normal error variance

145   and linear responses), accuracy (the dependent variable) and relative height (an independent

146   variable) were arcsin square root transformed, while the independent variables stemminess and

147   imbalance were each log-transformed. Note that using unscaled independent variables did not

148   change any of the rankings of the strengths of the partial regressions or their nominal

149   significances or the form of the response. Recall also that the fully imbalanced trees were not

150   included in the primary GLM analyses, since they confound stemminess and imbalance.

151

152

153   RESULTS

154   Across 500 simulated Yule trees, imbalance and stemminess are negatively correlated

155   (Pearson's correlation coefficient between ln(Ic) and ln(Stn) = -0.31, p<0.0001) as expected

156   (Rohlf et al., 1990). However, for the 10 trees created for this set of simulations, the relationship

157   was flat enough that GLM approaches are appropriate (r = -0.23, p = 0.52).

158   For these trees, at both high and low rates, average accuracy was high, with global

159   reconstructions being more accurate than local ones (high rate: global mean accuracy = 77.6%

160   vs. local mean accuracy = 74.4%; low rate: global mean accuracy = 87.1% vs. local mean

161   accuracy = 82.6%). For this dataset, these differences are very highly significant, with or

162   without the unlikely balanced and unbalanced topologies (paired t-test on transformed

163   accuracies, all p<0.0001). As expected, global and local accuracy were very strongly correlated

164   (r = 0.975 for transformed accuracies across all topologies). I report results from the global

165   reconstructions, and note differences between correlates of global and local accuracies where

166   appropriate.

167   Considering shape traits individually (and ignoring the unbalanced topology for the

168   moment), by far the best predictor of global accuracy was distance from the tips (Rsq = 0.676),

169   followed by rate (Rsq = 0.119) and stemminess (Rsq = 0.070; all nominally significant at

170   p<0.0001). Without controlling for the other variables, balance did not predict accuracy overall

171 (Rsq = 0.002). Interestingly, for local accuracy, while distance from the tips was still by far the
172 most important factor, stemminess was a better predictor than rate: Rsq for distance from tips =
173 0.769; for stemminess = 0.147; for rate = 0.088 (all nominally significant at p<0.0001). If we
174 include the unbalanced topology, distance from the tips and rate hardly change their relative
175 contribution, but stemminess and imbalance now confound each other such that neither explain
176 more than 4% of the variation for either local or global reconstructions.
177   The best-fit General Linear Model for the twenty simulated trees (ignoring the fully
178 unbalanced topology) did identify all four variables as significant predictors of global accuracy
179 (Table 2: for global reconstructions, Rsq for best model = 0.849; for local reconstructions, Rsq
180 for best model = 0.882): increasing rate, distance from the tips and decreasing imbalance all
181 make mean accuracy worse, while increasing stemminess improves it. In addition, several two-
182 way interactions were nominally significant: (i) accuracy drops off more precipitously as you
183 move away from the tips on stemmier trees than it does on less stemmy ones; (ii) more
184 unbalanced trees and stemmier trees enhance each other synergistically to produce more accurate
185 reconstructions on average; and (iii) for the local reconstructions only, accuracy is lost at the low
186 rate more quickly than it is at the high rate as the distance from the tips increases.
187   In order to see how sensitive this GLM was to the inclusion of the fairly unlikely fully
188 balanced topology, the analyses were rerun with these four trees (two stemminess values * two
189 rates) removed. The Rsq values for individual traits were effectively the same. Imbalance (and
190 its interaction) ceased to be a significant predictor of accuracy in the GLM; otherwise the best-fit
191 model was identical.
192
193 DISCUSSION
194
195   Parametric statistical analyses of simulated data is illuminating, but because one is less
196 constrained as to sample-size, significance can be achieved with very small effect-sizes, and so
197 results should be interpreted cautiously. A related issue is that because one set of tip data are
198 used for all reconstructions of internal nodes on a particular tree, these reconstructions are not
199 independent. However, the relative importance of the various parameters and their interactions
200 should still be highlighted in the GLM framework.
201   Overall, the average accuracy of the reconstructions was surprisingly high. Under an

202 asymmetric model (where forward and backward rates are free to vary independently) and where

203 the root state is not fixed a priori, as here, the likelihood surface may be quite flat. These first

204 results suggest that for trees of size 32, the algorithm in DISCRETE does well, but more work is

205 needed.

206   These simulations are the first to compare global and local reconstruction methods. It

207 should be noted that while global reconstructions are true posterior probabilities in an 'empirical

208 Bayes' framework (with an entirely uninformative prior), local reconstructions can only be

209 interpreted in a strict likelihood framework, where alternate hypotheses are compared directly.

210 The common accuracy measure I used, therefore, can not be interpreted as a general posterior

211 probability. That said, why were global reconstructions more accurate by my measure than local

212 ones in this set of simulations? I do not know. However, it may be relevant that local

213 reconstructions allow transition rates to be optimized for alternate states at the focal node and

214 models with very different forward and backward rates could be preferred (Pagel, 1999). When

215 the true rates are equal, as they were here, this would lower accuracy generally. If this

216 interpretation were true, then the amount by which local reconstructions do worse than global

217 ones should increase with decreasing information, measured as both rate and distance from the

218 tips. Interestingly, the average improvement of global over local rates was highest at the 'low'

219 rate of change, so this cannot be the entire story. The prediction that the improvement of global

220 vs. local accuracy increases as one moves towards the root (M. Pagel, pers. comm.), however,

221 does hold. Indeed, the slope is greater at the higher rate of change (under the best GLM model

222 for "global accuracy minus local accuracy", which is the same as that in table 2, the interaction

223 term "distance from tips * rate" has a positive coeffecient and is the most important term,

224 improving the fit of the model by over 10%).

225   Distance from the tips and rate were by far the strongest predictors of accuracy for each

226 reconstruction approach - both reflect the loss of information as multiple changes go unrecorded.

227 The simulations simply offer the first good empirical evidence of a logical necessity suggested

228 previously (Schluter, 1995).

229   Both imbalance and stemminess predicted reconstruction accuracy in fairly

230 straightforward ways, but did not contribute greatly to model fit. More unbalanced trees lead to

231 higher average accuracy, but the independent contribution was minimal, and the effect

232 disappeared when the fully balanced tree was excluded. At first glance, unbalanced trees might

233 be thought to retain more information about deeper nodes because more tips descend directly

234 from these deeper nodes. However, under the Markov model, opportunity for change should be

235 the prime predictor of information content, and these trees are all ultrametric (tips are all

236 equidistant from the root): imbalanced trees do not contain "ancestral" or "stem-group"

237 observations at the ends of early-splitting lineages (see Salisbury and Kim, 2001 for a similar

238 discussion regarding maximum parsimony reconstructions when Markov traits evolve on

239 ultrametric trees). The explanation for the increase in mean accuracy must be found in the mean

240 and variance in observations contributing to reconstructions at internal nodes: unbalanced trees

241 have a higher mean and variance in the number of descendants from a given node. All

242 descendants from a given node contribute information that increases the accuracy of the estimate

243 for that node, most notably in areas of denser sampling ("bushier" areas of the tree). Consistent

244 with this interpretation, the variance in accuracy across nodes also increases with increasing

245 imbalance, though with only five shapes, this is not significant ($Rsq=0.31$, $p>0.32$). Importantly,

246 this interpretation is also consistent with the observation that stemminess is a fairly strong

247 predictor of mean accuracy, such that compact groups give rise to higher mean accuracies; here,

248 increasing stemminess is a good predictor of increasing coefficient of variation in mean accuracy

249 on a tree (global reconstructions, $Rsq = 0.38$, $p = 0.06$; local reconstructions, $Rsq = 0.48$, $p =$

250 0.03). This interpretation, that increasing the variance in information contributing to

251 reconstructions increases mean accuracy, is also consistent with the two general interaction

252 terms: stemmier trees have denser sampling near the tips, and imbalance and stemminess interact

253 synergistically.

254   The third interaction term further suggests that the local approach is more sensitive to the

255 parameters tested here than is the global one, and that this may be due to the freedom it has to

256 find best fit forward and backward rates. As one moves away from the tips and information is

257 eroded, the local approach might be more likely to prefer rates very different from the generating

258 ones: at the low rate, this problem is more acute because the data near the tips are quite

259 informative while at the high rate, local reconstructions could be off on the wrong track almost

260 immediately. Regardless of whether conjecture is true or not, this third interaction term is over

261 an order of magnitude less important than its constituent main effects.

262   Returning to these main effects, recall that the behaviour of balance and stemminess with

263 respect to accuracy counteract each other for real (rather than contrived) trees: very unbalanced

264 trees cannot also be very stemmy. Unbalanced trees are also expected to sample more time on
265 average than balanced ones of equal depth, which will further erode the accuracy of their
266 constituent reconstructions. A next step in investigating the effects of tree shape on ancestor
267 reconstruction may be to see what the effects of tree shape are for groups of equal age: the
268 prediction from the present analysis is that tree shape will have only marginal effects on mean
269 accuracy, compared with the effects of overall rate and the fit between the process underlying
270 trait evolution and the model used to infer ancestral states. Though this simulation study is too
271 small to make any general statements about the average accuracy of maximum likelihood
272 reconstruction methods, it seems clear that tree shape is less of an issue for average accuracy
273 than these latter variables. Given that we have some control over the models we employ and the
274 traits we choose to study, but little over the shape of the tree of life, this is not unwelcome.
275

285   REFERENCES

286   Berendonk T. U., T. G. Barraclough and J. C. Barraclough 2003 Phylogenetics of pond and lake
287   lifestyles in Chaoborus midge larvae Evolution 57: 2173-2178

288   Chan, K. M. A. and B. R. Moore. 1999. Accounting for mode of speciation increases power and
289   realism of tests of phylogenetic asymmetry. American Naturalist 153:332-346.

290   Colless, D. H. 1982. Phylogenetics: the theory and practise of phylogenetic systematics II (book
291   review). Syst. Zool. 31:100-104.

292   Cunningham, C., K. E. Omland and T. Oakley. 1998. Reconstructing ancestral states: a critical
293   reappraisal. . Trends Ecol. Evol. 13:361-366.

294   Grafen, A. 1989. The phylogenetic regression. Phil. Trans. R. Soc. B 326:119-157.

295   Harvey, P. H. and A. Rambaut. 1998. Phylogenetic extinction rates and comparative
296   methodology. Proceedings of the Royal Society of London Series B-Biological Sciences
297   265:1691-1696.

298   Heard, S. B. 1992. Patterns in tree balance among cladistic, phenetic, and randomly generated
299   phylogenetic trees. Evolution 46:1818-1826.

300   Heard, S. B. 1996. Patterns in tree balance with variable and evolving speciation rates.
301   Evolution 50:2141-2148.

302   Heard, S. B. and A. Ø. Mooers. 2000. Phylogenetically patterned speciation rates and extinction
303   risks change the loss of evolutionary history during extinctions. Proceedings of the Royal
304   Society of London Series B-Biological Sciences 267:613-620.

305   Heijerman, T. 1993. Adequacy of numerical taxonomic methods: further experiments using
306   simulated data. Z. Zool. Syst. Evolut.-forsch. 31:81-97.

307   Hey, J. 1992. Using phylogenetic trees to study speciation and extinction. Evolution 46:627-
308   640.

309   Hoekstra, H. E. and S. V. Edwards. 2000. Multiple origins of XY female mice (genus Akodon):
310   phylogenetic and chromosomal evidence. Proceedings of the Royal Society (London) B
311   Biological Sciences 267:1825-1831.

312   Koshi, J. M. and R. A. Goldstein. 1996. Probabilistic reconstruction of ancestral protein
313   sequences. Journal Of Molecular Evolution 42:313-320.

314   Lorch, P. D. and J. M. Eadie. 1999. Power of the concentrated changes test for correlated
315   evolution. Systematic Biology 48:170-191.

316   Maddison, D. R. 1994. Phylogenetic methods for inferring the evolutionary history and
317   processes of change in discretely valued characters. Annu. Rev. Entomol. 39:267-292.

318   Maddison, W. P. 1995. Calculating the probability distributions of ancestral states reconstructed
319   by parsimony on phylogenetic trees. Syst. Biol. 44:474-481.

320   Martins, E. P. and T. J. Garland. 1991. Phylogenetic analyses of the correlated evolution of
321   continuous characters. Evolution 45:534-557.

Martins, E. P. 1994. Estimating the rate of phenotypic evolution from comparative data. Amer. Nat. 144:193-209.

Martins, E. P. 1999. Estimation of ancestral states of continuous characters: A computer simulation study. Systematic Biology 48:642-650.

Matyjasiak, P., P. G. Jablonski, I. Olejniczak and P. Boniecki. 2000. Imitating the initial evolutionary stage of a tail ornament. Evolution 54:704-711.

Mooers, A. Ø. 1995. Tree balance and tree completeness. Evolution 49:379-384.

Mooers, A. Ø. and D. Schluter. 1999. Reconstructing ancestor states with maximum likelihood: Support for one- and two-rate models. Systematic Biology 48:623-633.

Morales, E. 2000. Estimating phylogenetic inertia in Tithonia (Asteraceae): A comparative approach. Evolution 54:475-484.

Nee, S. and R. M. May. 1997. Extinction and the loss of evolutionary history. Science 278:692-694.

Oakley. T. H. and C. W. Cunningham. 2002. Molecular phylogenetic evidence for the independent evolutionary origin of an arthropod compound eye. Proc. Natl. Acad. Sci. USA 99: 1426-1430

Pagel, M. 1997. Inferring evolutionary process from phylogenies. Zool. Scripta 26:331-348.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Systematic Biology 48:612-622.

Polly, P. D. 2001. Paleontology and the comparative method: Ancestral node reconstructions versus observed node values. American Naturalist 157:596-609.

Purvis, A. 1996. Using interspecies phylogenies to test macroevolutionary hypotheses. 153-168 in New Uses for New Phylogenies (P. H. Harvey, A. J. L. Brown, J. M. Smith and S. Nee, ed.) Oxford University Press, Oxford.

Rambaut, A. and N. C. Grassly. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic frees. Computer Applications in the Biosciences 13:235-238.

Rohlf, F. J., W. S. Chang, R. R. Sokal and J. Kim. 1990. Accuracy of estimated phylogenies: effects of tree topology and evolutionary model. Evolution 44:1671-1684.

Ryan, M. J. and A. S. Rand. 1999. Phylogenetic influence on mating call preferences in female tungara frogs, Physalaemus pustulosus. Animal Behaviour 57:945-956.

Salisbury, B. A. 1999. Misinformative characters and phylogeny shape. Systematic Biology 48:153-169.

Salisbury, B.A., and J. Kim. 2001. Ancestral state estimation and taxon sampling density. Systematic Biology 50:557-564.

Schluter, D. 1995. Uncertainty in ancient phylogenies. Nature (Lond.) 377:108-109.

Schluter, D., T. Price, A. Ø. Mooers and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699-1711.

360   Schluter, D. 2000. Ecology of Adaptive Radiation. Oxford University Press, Oxford.

361   Sober, E. and M. Steel. 2002. Testing the hypothesis of common ancestry. Journal of
362   theoretical Biology 218:395-408.

363   Stone, J. and J. Repka. 1998. Using a nonrecursive formula to determine cladogram
364   probabilities. Systematic Biology 47:617-624.

365   Yang, Z., S. Kumar and M. Nei. 1995. A new method of inference of ancestral nucleotide and
366   amino acid sequences. Genetics 141:1641-1650.

367

368   Table 1. Tree shape measures for the twelve test trees. See Figure 1 for more detail.

369

Tree	Ic[1]	Branch lengths[2]	Stni[3]
LL	0.101	Yule	0.322
LL[+]	0.101	Proportional	0.509
LH	0.0925	Yule	0.537
LH[-]	0.0925	ln(i)	0.299
HL	0.247	Yule	0.297
HL[+]	0.247	Proportional	0.474
HH	0.237	Yule	0.523
HH[-]	0.237	ln(i)	0.300
B[+]	0	Proportional	0.613
B[-]	0	ln(i)	0.351
U[+]	1	Constant size	0.195
U[-]	1	Proportional	0.101

370   1. Colless' measure of imbalance (Heard, 1992)
371   2. Model used to generate branch lengths: Yule: sampled from distribution of Yule trees;
372   Proportional: lengths proportional to clade size; ln(i); lengths proportional to ln(number of extant
373   lineages i); Constant size: lengths proportional to 1/i(i-1).
374   3. Noncumulative stemminess index (Rohlf et al., 1990)
375
376
377   Table 2. Results from best-fit General Linear Model of tree shape effects on reconstruction
378   accuracy for both global and local type reconstructions

379

Source[1]	global SS[2]	F	local SS[2]	F
rate	2.614	485.77	0.853	201.2
distance from tips	3.597	668.30	2.222	524.08
stemminess	0.0364	6.76	0.0900	17.22
balance	0.0769	14.30	0.073	21.22
distance * stemminess	0.763	141.77	0.457	112.08
balance * stemminess	0.0600	11.15	0.0563	13.27
rate * distance	-	-	0.0393	9.26
Error[2]	3.299		2.595	

380   2. Adjusted Sums of Squares
381   1. Model terms each have 1 df: Error term has 613 nominal df for global model, and 612 for the
382   local model.

383
384   Figure 1. The six topologies used for the 12 test trees. Because stemminess and balance are

385   negatively correlated (Rohlf et al., 1990), four Yule trees were chosen from a sample of 500 to

386   represent extremes on the two shape axes: LL (low Ic, low Stn), LH (low Ic, high Stn), HL (high

387   Ic, low Stn), and HH (high Ic, high Stn). Two extreme tree topologies were also used: B (fully

388   balanced, Ic=0,) and U (fully unbalanced, Ic=1). Making sequential splitting times = ln(i), i =

389   number of extant branches produced low stemminess for the balanced tree (B$^-$). A model of

390   constant tree size (expected splitting times = 1/i(i-1); Hey 1992, Nee and May, 1997) produced

391   relatively high stemminess for the unbalanced tree (U$^+$). The six additional trees (Table 1) use

392   these topologies but have contrasting stemminess: Making sequential splitting times = ln(i)

393   produced trees LH$^-$ and HH$^-$ with lowered stemminess. Making branch lengths proportional to

394   clade size (Grafen, 1989) produced trees LL$^+$, HL$^+$ and B+ with increased stemminess, and U$^-$

395   with lowered stemminess.

LL

LH

HL

HH

B-

U+