Syst. Biol. 51(6):835–843, 2002
DOI: 10.1080/10635150290102672

Is Cladogenesis Heritable?

VINCENT SAVOLAINEN,[1,6] STEPHEN B. HEARD,[2] MARTYN P. POWELL,[1,3]
T. JONATHAN DAVIES,[1,4] AND ARNE Ø. MOOERS[5,6]

[1]*Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew TW9 3DS, U.K.;*
E-mail: v.savolainen@rbgkew.org.uk (V.S.), m.powell@rbgkew.org.uk (M.P.P.), j.davies@rbgkew.org.uk (T.J.D.)
[2]*Department of Biology, University of New Brunswick, Fredericton, New Brunswick E3B 6E1, Canada;*
E-mail: sheard@unb.ca
[3]*Centre for Plant Diversity and Systematics, School of Plant Sciences, University of Reading, Reading RG6 6AS, U.K.*
[4]*Department of Biology and NERC Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berkshire SL5 7PY, U.K.*
[5]*Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada;*
E-mail: amooers@sfu.ca

Abstract.—The heritability of speciation rates and extinction risks is a crucial parameter in models of macroevolution, but little direct evidence has been available to assess the occurrence, strength, or generality of this heritability. We tested for heritability using correlations between ancestral and descendent branch lengths in phylogenetic trees, an approach first applied to a bird phylogeny by Harvey et al. (1991, pages 123–137 *in* Genes in ecology [R. J. Berry et al., eds.], Blackwell Scientific, Oxford). We applied Harvey et al.'s test to some of the largest DNA sequence-based phylogenetic analyses published to date for plants, insects, fungi, and bacteria. If one of two parent lineages splits first and if this is the case for any heritable reason, then on average we expect its daughter lineages to also split first. We also used a randomization procedure to assess significance of branch length heritability. Using maximum parsimony and maximum likelihood branch lengths and trees made ultrametric after nonparametric rate smoothing or by enforcing a molecular clock, we found a pattern for most clades consistent with heritable net cladogenesis. Heritability of cladogenesis may be a general phenomenon, detectable across a large number of lineages and a broad range of taxa. [Branch length; cladogenesis; diversification; molecular clock; phylogenetic tree shape; speciation.]

With the wide use of molecular data for reconstructing evolutionary relationships, increasingly broad-scale phylogenetic trees have become available in which patterns of cladogenesis can be profitably explored (Nee et al., 1992, 1994; Harvey et al., 1996; Pagel, 1999). One important question is whether rates of cladogenesis are heritable among lineages, i.e., whether daughter lineages tend to resemble their ancestors in rates of speciation and/or extinction. Fossil evidence (reviewed by Valentine, 1985) and studies of the correlates of speciation (e.g., Mitter et al., 1988; Farrell et al., 1991; Hodges and Arnold, 1995; also see review by Schluter, 2000) suggest that the number of species in a clade can differ because of species-level traits. Such heritable variation among lineages in the net rate of cladogenesis (speciation minus extinction) is essential for the controversial process of species selection (Eldredge, 1989; Stanley, 1998). Several surveys have suggested that topologies of estimated phylogenetic trees are consistent with such heritable variation (reviewed by Purvis, 1996; Mooers and Heard, 1997; see also Heard and Mooers, 2002 [this issue]). However, such surveys have two principal shortcomings: they are based primarily on rather small phylogenies (i.e., those containing few species; but see Purvis and Agapow, 2002 [this issue]), and the inference about heritability is indirect because information about speciation rates per unit time is not used.

A simple, complementary, and more direct phylogenetic test of local heritability in net cladogenesis, which can be applied on a node-by-node basis, is to examine whether the shorter of two branches descended from a single common ancestor tends in turn to lead to shorter descendent branches (Harvey et al., 1991). If branch lengths are proportional to time, such shorter branches indicate higher underlying rates of net cladogenesis. If branch lengths measure the amount of evolution (rate of change $\times$ time), such shorter branches indicate that lineages inherit evolutionary rates. This approach has appeal precisely because it considers nodes one by one, asking whether the phenomenon of autocorrelated branch lengths persists throughout a tree. A decade ago, Harvey et al. (1991) suggested that heritable cladogenesis might hold

[6]Authors for correspondence.

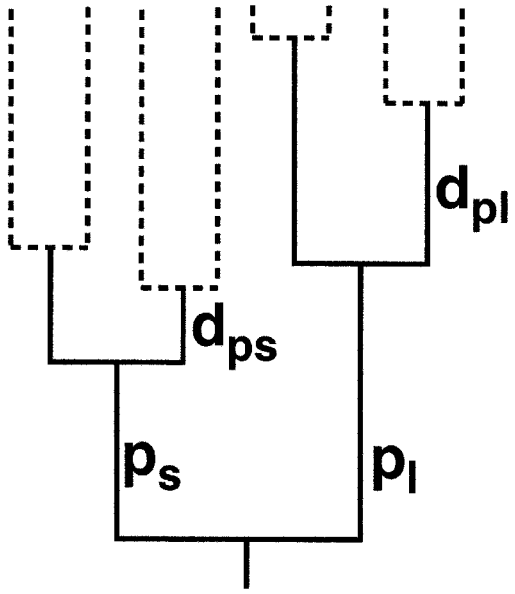

FIGURE 1. Test of whether cladogenesis is locally heritable (Harvey et al., 1991). In this example, branch length is proportional to time, and the parent branch p_s split before p_l and so is the shorter branch. If this is the case for any heritable reason, we expect the daughter d_{ps} to also split before d_{pl}.

for birds: 40 of 68 comparisons from Sibley and Ahlquist's (1990) phylogenetic tree had the shortest of two parent branches leading to the shortest daughter branch (see Fig. 1). We examined whether this pattern can be generalized by analyzing several of the largest available DNA sequence-based phylogenetic trees representing organisms as diverse as insects, plants, bacteria, and fungi.

MATERIALS AND METHODS

Phylogenetic Trees

We examined six phylogenetic trees, chosen because they are to our knowledge the largest published molecular trees in terms of both taxon number and taxonomic diversity encompassed. First, we used the largest published molecular phylogenetic tree of flowering plants (angiosperms), including >75% of all families and based on a combined analysis of the plastid *atpB-rbcL* exons and 18S nuclear ribosomal DNA (rDNA) (Soltis et al., 1999, 2000). Second, we used the phylogenetic tree of monocots based on combined analysis of the same set of genes and including all but one family (Chase et al., 2000). Third, we included a *rbcL*-based phylogenetic tree for the Orchidaceae, the largest angiosperm family, including nearly all tribes and subtribes (Cameron et al., 1999). The second and third trees are nested within the first, but there is little overlap of the orchid tree with the monocot tree (only three species in common) and only moderate overlap of monocot and angiosperm trees (the former with 40% more monocot taxa than the latter). Fourth, we used the largest phylogenetic tree available for fungi, including most families within Agaricales (gilled mushrooms and their allies) and based on analyses of the nucleus-encoded large subunit RNA gene (Moncalvo et al., 2000). Finally, we included broad-scale phylogenetic trees based on 18S rDNA for "bacteria" (Barns et al., 1996) and for holometabolous insects (Whiting et al., 1997).

For each tree, we reanalyzed DNA sequence data provided by the authors (except insects and bacteria, for which both trees and sequences were downloaded from Treebase [http://www.herbaria/harvard.edu/treebase]). Branch lengths were first calculated using maximum parsimony (MP, acctran optimization). Because uncorrected MP branch lengths can underestimate longer branches (Felsenstein, 1981), we also computed maximum likelihood (ML) branch lengths. For ML reconstruction, we used the HKY85+Γ model of DNA evolution (Hasegawa et al., 1985; Yang, 1994) with base frequencies, transition/transversion ratio and the alpha shape parameter estimated from the data (Table 1). Low values for the

TABLE 1. Estimated parameters for the HKY85+Γ model of DNA evolution.

Tree	Transition/transversion ratio	Frequency				Alpha shape parameter
		Adenine	Cytosine	Guanine	Thymine	
Bacteria	1.1538	0.2620	0.2325	0.2534	0.2521	0.6085
Fungi	3.3509	0.2705	0.2380	0.2035	0.2880	0.3057
Insects	1.4731	0.2737	0.2206	0.2096	0.2961	0.3033
Orchids	1.6048	0.2601	0.2262	0.2245	0.2892	0.3031
Monocots	2.1713	0.2536	0.2299	0.2197	0.2968	0.2993
Angiosperms	1.9368	0.2506	0.2396	0.2152	0.2946	0.5823

alpha shape parameters (ca. 0.3; Table 1) indicate that there is rate heterogeneity among sites, the highest figure reaching only about 0.6 for the largest phylogenetic tree (angiosperms) and the broadest taxon coverage ("bacteria"; Table 1). Uncorrected MP is not expected to correctly estimate branch lengths where multiple hits have occurred; however, overall MP and ML branch lengths are highly correlated (at least in the angiosperm tree used here; see Barraclough and Savolainen, 2001). We used HKY85+Γ instead of the more complex general time-reversible (GTR) model because it represents a reasonable compromise between generality of the model, especially for concatenated genes, and computational time required. For *rbcL* and 18S rDNA in angiosperms, HKY85+Γ did not perform significantly worse than GTR+Γ (see Sanderson and Doyle, 2001; not tested for the others).

Branch lengths may not be proportional to divergence times throughout the tree because of DNA rate heterogeneity among lineages. We assessed rate heterogeneity with likelihood ratio tests following Felsenstein (1981). A global clocklike rate of change was rejected for all trees ($P < 0.001$, data not shown). We therefore made the trees ultrametric using the nonparametric rate smoothing method (NPRS; Sanderson, 1997). This method, implemented in the software TreeEdit (version 1.0 alpha 4-61, written by A. Rambaut and M. Charleston, http://evolve.zoo.ox.ac.uk/software/TreeEdit) maximizes rate autocorrelation to correct for rate heterogeneity in the overall tree (Sanderson, 1997). In addition to NPRS, we also made the branch lengths proportional to time by enforcing a molecular clock with the HKY85+Γ model of DNA evolution and the software PAUP 4.0b8 (marginal estimation of ancestral states option; Swofford, 2001). Unfortunately, optimization of the ML+clock model could not be performed with the large angiosperm tree in a reasonable amount of time: the search did not finish after several weeks of calculation using a G3 Macintosh computer (300 MHz, 768 MBytes RAM). For the angiosperm tree, therefore, we only have NPRS branch lengths. Despite some limitations in applying NPRS (Sanderson, 1997, 2002), this method has been efficient in making branch lengths proportional to time in the angiosperm tree used here, with NPRS age estimates in agreement with several external calibration points (Wikström et al., 2001).

Pattern in Branch Lengths

We assembled sets of paired parent (p_l, p_s; Fig. 1) and daughter (d_{pl}, d_{ps}) branches from each phylogeny following Harvey et al. (1991), and we considered only the shorter of the two branches emanating from each parent branch (e.g., d_{pl} and not the unmarked branch from p_l; Fig. 1). This approach allowed us to include nodes where fewer than four daughter branches split, increasing our sample size substantially. Because the trees are not complete, the longer daughter branch emanating from a parent often does not split at all and terminates in the present; these branches were excluded from the analyses. The assumption is that taxon sampling has not been biased with respect to our null hypothesis. Because shorter daughter branches that lead to terminal taxa also cannot be used (they are truncated at the present and so do not measure time to splitting), the number of comparisons is substantially lower than tree size would suggest (Tables 2, 3).

Uneven taxon sampling can obscure or even bias the results (Harvey et al., 1991). Most of our comparisons within trees were made at high taxonomic level (above family level for bacteria, angiosperms, monocots, and insects and above subtribe level for orchids and fungi) before missing taxa become a problem, and such sampling mitigates against bias due to incomplete sampling.

For each tree, every node that satisfied the above conditions (see also Fig. 1) can be considered an independent data point. For each node, we recorded the lengths of the branches denoted p_s, p_l, d_{ps}, and d_{pl} in Figure 1. Besides recording the number of times the shorter parent branch gave rise to the shortest daughter branch ($d_{ps} < d_{pl}$), we performed three different (and complementary) statistical tests of the short-short association: Wilcoxon paired-sample tests, comparisons to short-short fractions for simulated Yule trees, and a randomization test based on reshuffling parent–daughter associations.

Wilcoxon tests.—Our simplest test of branch length heritability was a one-tailed Wilcoxon paired-sample test, with paired d_{pl} and d_{ps} values for each node (Fig. 1) and the alternative hypothesis that $d_{ps} < d_{pl}$. Although the distributions of branch lengths

TABLE 2. Branching patterns in ultrametric phylogenetic trees.

Taxa	No. terminals	Method[a]	Short-short: short-long[b]	*P* values: Wilcoxon paired-sample test	*P* values: Fisher's combined probability[c] 24 tips	*P* values: Fisher's combined probability[c] 100 tips	*P* values: Randomization
Bacteria	64	MP+NPRS	6:1	0.016*	0.223	0.325	<0.001***
		ML+NPRS	3:4	0.711	0.639	0.741	<0.001***
		ML+clock	5:1	0.11	0.09	0.13	<0.001***
Fungi	154	MP+NPRS	14:9	0.044*	0.027*	0.106	<0.001***
		ML+NPRS	15:8	0.074	0.007**	0.038*	0.01**
		ML+clock	6:6	0.73	0.30	0.45	0.085
Insects	85	MP+NPRS	7:5	0.676	0.381	0.520	0.02*
		ML+NPRS	7:5	0.691	0.232	0.398	0.529
		ML+clock	4:2	0.5	0.08	0.11	0.8
Orchids	171	MP+NPRS	14:6	0.010*	0.018*	0.067	0.032*
		ML+NPRS	7:9	0.736	0.604	0.773	0.675
		ML+clock	8:1	0.05*	0.14	0.23	0.282
Monocots	127	MP+NPRS	18:6	0.009**	0.002**	0.013*	<0.001***
		ML+NPRS	15:9	0.063	0.001**	0.009*	0.069
		ML+clock	13:4	0.002**	<0.0001***	0.0005***	0.197
Angiosperms	567	MP+NPRS	70:31	<0.0001***	10^{-12}***	10^{-8}***	<0.001***
		ML+NPRS	63:38	0.001**	10^{-13}***	10^{-9}***	0.028*

[a]MP = maximum parsimony, ML = maximum likelihood, NPRS = nonparametric rate smoothing.

[b]Number of occurrence when the shortest of two parent branches gives rise to the shortest of the four daughter branches (short-short) versus occurrences where this is not the case (short-long).

[c]Fisher's combined χ^2, df = 2 × number of comparisons, based on 1,000 Yule trees having 24 tips or 100 tips.

are expected to range from exponential under the simplest Yule process to various forms of the gamma with extinction, the Wilcoxon paired-sample test requires only that the distribution of differences be symmetric about the median (Zar, 1999). Logarithmic transformation of d_{ps} and d_{pl} produced such distributions.

Simulations.—We also converted each nodal comparison to a *P* value based on the expectation from simulated phylogenetic trees in which diversification rates were constant across lineages. This test was useful because there is a bias for $d_{pl} < d_{ps}$ for trees of less than infinite size, contrary to our alternative hypothesis of heritable rates (Harvey et al., 1991). At any given node, both d_{ps} and d_{pl} are constrained to be shorter than total tree length above the node (viz. toward the tips). The constraint is stronger for d_{pl}

TABLE 3. Branching patterns in nonultrametric phylogenetic trees.

Taxa	No. terminals	Method[a]	Short-short: short-long[b]	*P* values: Wilcoxon paired-sample test	*P* values: Fisher's combined probability[c] 24 tips	*P* values: Fisher's combined probability[c] 100 tips	*P* values: Randomization
Bacteria	64	MP	6:1	0.023*	0.278	0.381	<0.001***
		ML	5:2	0.344	0.555	0.779	0.22
Fungi	154	MP	13:8	0.039*	0.041*	0.132	<0.001***
		ML	10:12	0.913	0.731	0.883	<0.001***
Insects	85	MP	6:4	0.244	0.226	0.344	<0.001***
		ML	5:5	0.196	0.196	0.336	0.149
Orchids	171	MP	10:7	0.085	0.027*	0.100	0.062
		ML	8:8	0.430	0.550	0.644	0.129
Monocots	127	MP	14:9	0.283	0.195	0.400	0.009**
		ML	16:8	0.050*	0.004**	0.028*	0.122
Angiosperms	567	MP	66:37	<0.0001***	10^{-9}***	10^{-6}***	<0.001***
		ML	55:46	0.165	10^{-8}***	10^{-5}***	0.015*

[a]MP = maximum parsimony, ML = maximum likelihood.

[b]Number of occurrence when the shortest of two parent branches gives rise to the shortest of the four daughter branches (short-short) versus occurrences where this is not the case (short-long).

[c]Fisher's combined χ^2, df = 2 × number of comparisons, based on 1,000 Yule trees having 24 tips or 100 tips.

because it originates nearer the present. This constraint is attenuated as d_{ps} and d_{pl} become smaller proportions of the total tree length, which is strongly correlated with the number of splits above the node. To account for this bias, we adopted a simple strategy for simulation. For each of 1,000 Yule trees of size 24 or 100 (Yule, 1924; constructed with equal and constant rates of cladogenesis per lineage following Heard, 1996), we recorded $d_{pl} - d_{ps}$ at the root node (expressed as a fraction of $p_s + p_l + d_{ps} + d_{pl}$ because branch lengths from simulated and sample trees are on different scales). For each node, the fraction of the 1,000 simulated trees with ($d_{pl} - d_{ps}$) at the root exceeding that in the real node is a *P* value for a test against the Yule expectation. We combined these individual *P* values across nodes within a tree using Fisher's combined probability test (Sokal and Rohlf, 1995). We used simulated trees of size 24 or 100 because these sizes are typical for the clades arising from our focal nodes; the use of larger simulated trees would not appreciably change our results.

Randomizations.—Our final test for branch length heritability was performed on all trees using a randomization technique. The ratio of the difference in branch lengths between parent and daughter branches was calculated at each node using the following equation: $H = 1 - [(l - s)/l]$, where l is the longer branch (whether it is the parent or the daughter) and s is the shorter branch. Every parent branch gives rise to two daughter branches and so was used twice. H was summed for all nodes in the tree to give a measure of branch length heritability. Although double counting of parents in ratios with both daughters means observations are not fully independent, we are comparing our summed result with simulated data measured in the same way. Terminal branches were excluded from the analyses for the same reason given above: they offer no information on times of split. A randomization procedure was then performed to assess whether observed ΣH was larger than expected by chance. A null distribution of ΣH was generated by randomly pairing each parent with any daughter and calculating ΣH from the new set of parent–daughter pairs. The probability that ΣH obtained from the real tree was different from that of the null model was obtained using a one-tailed distribution based on 1,000 such randomizations. If branch length is heritable, the differences in length between parent and daughter branches in the real trees should be small compared with the distribution of differences for the randomized set of parent–daughter pairs, and observed ΣH should be large relative to the randomization.

RESULTS

Results from trees made ultrametric after either NPRS or by enforcing a molecular clock are presented in Table 2; results from nonultrametric trees are presented in Table 3 (MP and ML). For most data sets and tree estimation procedures, shorter branches most often gave rise to shorter daughters. In only 2 of 17 comparisons was the opposite pattern observed (Table 2): ML+NPRS in bacteria and in orchids (but after enforcing a clock the most commonly observed pattern was restored). In one further case (ML+clock in fungi), shorter branches gave rise to shorter daughters exactly half the time (Table 2). If we combined across trees, the pattern for heritable cladogenesis was strong regardless of the means of analysis (see Table 2; over all nodes short-short:short-long, MP+NPRS = 129:58; ML+NPRS = 110:73; ML+clock = 36:14). Significance for some trees was, however, difficult to assess. The pattern was strongest for the angiosperm and monocot trees using Wilcoxon tests and simulations; these taxa are also the largest both in number of terminals and in number of characters (2,928 and 2,221 variable sites vs. 508–1,264 for the other trees). However, randomizations suggested that bacteria and fungi possess at least as strong a signal (Table 2). Patterns were generally stronger in MP trees than in ML trees and when trees were made ultrametric by rate smoothing rather than when enforcing molecular clocks. Tests based on comparisons with simulated Yule trees gave similar results (Table 2). As expected, decreasing the number of terminals for the simulations from 100 to 24 increased the signal, and again combining observations across all trees produced a highly significant result regardless of tree estimation method ($P < 0.001$).

DISCUSSION

Patterns Across Trees and Reconstruction Methods

The simplest pattern (the preponderance of shorter daughters from shorter parental

branches) is also the clearest. As might be expected on statistical grounds, the largest data sets gave the strongest signal. Our largest data sets all consisted of plants, and plants may behave differently from other taxa with regard to net cladogenesis. The impetus for this study was the suggestion of a pattern in a large (albeit potentially biased) DNA–DNA hybridization tree for birds, suggesting that the pattern may have some generality; clearly, more work is needed.

The short-short correlation reported here is also at least superficially consistent with heritability of substitution rates rather than cladogenesis. However, correction of branch lengths using rate smoothing or enforcing molecular clocks generally improved our ability to detect patterns in branch length heritability (compare Tables 2 and 3). The use of NPRS to produce ultrametric trees assumes autocorrelation of substitution rates. If parent lineages differ in heritable rates of substitution, NPRS should stretch short branches (both parent and daughter) and shorten long ones and so should not bias observations in favor of heritability. We do not know if there are other biases in the NPRS (or ML+clock) algorithm with respect to our observations, but our test considers only two consecutive splits at each focal node, so only very local clocks are required in any case. The strongest overall result was with MP branch lengths. Although uncorrected MP branch lengths will underestimate the longest branches most, it is not clear that this underestimation would bias our observations.

The simple randomization test we used may turn out to be a powerful complement to topology-based tests of tree shape. Topology-based tests are hampered by the fact that arbitrary resolutions of polytomies, and poorly supported trees more generally, usually lead to increased imbalance (Guyer and Slowinski, 1991; Mooers et al., 1995; Heard and Mooers, 1996; Huelsenbeck and Kirkpatrick, 1996). However, such resolutions are likely to make branch-length tests conservative if clocks are not enforced, because parent–daughter branch lengths should be assigned arbitrarily. Also, using branch lengths to measure cladogenesis heritability should be independent of the taxonomic rank of the terminal taxa (see Purvis and Agapow, 2002 [this issue]).

Our demonstration of local heritablity in net cladogenesis is all the more compelling given that comparisons such as ours have an inherent bias toward the opposite pattern (shorter branches having longer daughters; Harvey et al., 1991) because trees are constrained to terminate at the present. We corrected for this bias by comparing our observed nodes with Yule trees with 24 and 100 tips. Ideally, one would like to make each nodal comparison with simulated trees of exactly equal size; such a procedure would give a more powerful test for nodes involving <24 descendent species. However, incomplete taxonomic sampling in our six molecular phylogenetic trees made constructing tailor-made null tree distributions cumbersome. We chose to use 24-tip and 100-tip simulated trees because doing so is conservative; most nodes used define subclades with 5–70 sampled species. Further, because of extinction, branch lengths from the six trees are not expected to be drawn from a pure Yule distribution (Yule, 1924), but estimating actual rates of speciation and extinction on incomplete trees is difficult (for one avenue, see Pybus and Harvey, 2000). Incomplete sampling in the trees we used might give rise to a further bias because oversampled clades will tend to have shorter branches than will undersampled clades. As a result, if one of a parent's two daughter clades were better sampled than the other, both the shortest parent and shortest daughter branches would likely occur in the oversampled clade. Because we made our comparisons across many individual nodes, however, such sampling problems would have to be extremely widespread to give rise to our results.

Biological Issues

It is well known that taxonomies are strongly skewed in numbers of subtaxa per taxon and that simple null models with constant speciation and extinction rates are rejected by real phylogenies (Dial and Marzluff, 1989; Heard, 1992; Sanderson and Donoghue, 1994, 1996; Harvey et al., 1996; Mooers and Heard, 1997). Heard (1996) showed that patterns in topologies in estimated phylogenies are consistent with the existence of heritable variation among lineages in rates of cladogenesis. Direct evidence for heritability in speciation or extinction rates has, however, been scant, and available for only two major taxa: molluscs (Jablonski, 1987) and birds (Harvey et al.,

1991). Our results suggest that heritability is a common phenomenon, detectable across a broad taxonomic range including bacteria, fungi, animals, and plants.

Why should lineages resemble their ancestors in rates of cladogenesis? We recognize at least two general classes of possible explanation. First, biogeographic processes may favor heritability. For example, because long distance dispersal is relatively rare, habitat is generally heritable and may include habitats that might promote diversification (e.g., multistratified rain forest). If environments are spatially autocorrelated, parapatric speciation could also be compatible with local heritability of cladogenesis (Barraclough and Vogler, 2000), as is the apparent heritability of species' geographic ranges (a trait associated with species duration for molluscs; Jablonski, 1987). Colonization of new areas and subsequent radiations in all available niches may also leave similar signatures in phylogenetic trees, e.g., in the hypothetical case of an archipelago with repetitive waves of speciation on islands.

Second, rates of cladogenesis may depend on traits that are heritable through speciation events. There is a rich literature concerning traits that directly affect speciation rates (e.g., Heard and Hauser, 1995). For instance, clades of monoecious plants have higher rates of net cladogenesis than do their dioecious sister clades (Heilbuth, 2000), and monoecious ancestors will tend to have monoecious descendents. Such traits may even involve molecular details of DNA replication and repair. At least in plants, higher rates of neutral mutations are correlated with an increased rate of morphological change and with more species-rich clades (Barraclough et al., 1996; Savolainen and Goudet, 1998; Barraclough and Savolainen, 2001). This phenomenon would make speciating lineages have longer branches than expected when measured in units of molecular change, making our conclusions conservative with respect to this phenomenon. Heritable factors that might increase the mutation rate, and in turn the rate of cladogenesis, include deficiency in DNA repair systems or higher exposure to ultraviolet radiation (perhaps habitat related).

Our results help establish that persistent differences among lineages in rates of cladogenesis could arise because those rates are heritable at the level of ancestors to descendent taxa. We hope the results reported here spur further and more refined analyses of branch length patterns, complementing what is becoming a rich literature on topological patterns. Illuminating both particular and general causes for nonrandom diversification is a major goal in evolutionary studies (Purvis, 1996; Mooers and Heard, 1997; Schluter, 2000), and indeed progress in this area may have important implications for the preservation of present-day biodiversity (see McKinney, 1997; Heard and Mooers, 2000; Purvis et al., 2000).

ACKNOWLEDGMENTS

We thank Tim Barraclough, Cassio van den Berg, Thomas Buckley, Mark Chase, Andy Purvis, Alan de Queiroz, Chris Simon, Peter Wagner, and several anonymous reviewers for their comments on the manuscript, Andy Purvis and Andrew Rambaut for discussion, and Ken Cameron, Mark Chase, Marc Moncalvo, and Pam and Doug Soltis for sharing their phylogenetic data with us. This work was partially funded by the Swiss, U.S., and Dutch National Science Foundations (grants 31-52378.97 to V.S., DEB 96-28969 and DEB 01-07752 to S.B.H., and PULS 831.48.003 to A.Ø.M.). Further support came from a President's Research Grant to A.Ø.M. from Simon Fraser University.

REFERENCES

BARNS, S. M., C. F. DELWICHE, J. D. PALMER, AND N. R. PACE. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proc. Natl. Acad. Sci. USA 93:9188–9193.

BARRACLOUGH, T. G., P. H. HARVEY, AND S. NEE. 1996. Rate of *rbcL* gene sequence evolution and species diversification in flowering plants (angiosperms). Proc. R. Soc. Lond. B 263:589–591.

BARRACLOUGH, T. G., AND V. SAVOLAINEN. 2001. Evolution rates and species diversity in flowering plants. Evolution 55:677–683.

BARRACLOUGH, T. G., AND A. P. VOGLER. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. Am. Nat. 155:419–434.

CAMERON, K. M., M. W. CHASE, W. M. WHITTEN, P. J. KORES, D. C. JARRELL, V. A. ALBERT, T. YUKAWA, H. G. HILLS, AND D. H. GOLDMAN. 1999. A phylogenetic analysis of the Orchidaceae: Evidence from *rbcL* nucleotide sequences. Am. J. Bot. 86:208–224.

CHASE, M. W., D. E. SOLTIS, P. S. SOLTIS, P. J. RUDALL, M. F. FAY, W. H. HAHN, S. SULLIVAN, J. JOSEPH, T. GIVNISH, K. J. SYTSMA, AND J. C. PIRES. 2000. Higher-level systematics of the monocotyledons: An assessment of current knowledge and a new classification. Pages 3–16 *in* Monocots systematics and evolution (K. L. Wilson and D. A. Morrison, eds.). CSIRO, Melbourne.

DIAL, K. P., AND J. M. MARZLUFF. 1989. Nonrandom diversification within taxonomic assemblages. Syst. Zool. 38:26–37.

ELDREDGE, N. 1989. Macroevolutionary dynamics: Species, niches, and adaptive peaks. McGraw-Hill, New York.

FARRELL, B. D., D. E. DUSSOURD, AND C. MITTER. 1991. Escalation of plant defence: Do latex and resin canals spur plant diversification? Am. Nat. 138:881–900.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17:368–376.

GUYER, C., AND J. B. SLOWINSKI. 1991. Comparison of observed phylogenetic topologies with null expectations among three monophyletic lineages. Evolution 45:340–350.

HARVEY, P. H., A. J. LEIGH BROWN, J. MAYNARD SMITH, AND S. NEE. 1996. New uses for new phylogenies. Oxford Univ. Press, Oxford, U.K.

HARVEY, P. H., S. NEE, A. Ø. MOOERS, AND L. PARTRIDGE. 1991. These hierarchical views of life: Phylogenies and metapopulations. Pages 123–137 *in* Genes in ecology: The 33rd symposium of the British Ecological Society (R. J. Berry, T. J. Crawford, and G. M. Hewitt, eds.). Blackwell Scientific, Oxford, U.K.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21:160–174.

HEARD, S. B. 1992. Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. Evolution 46:1818–1826.

HEARD, S. B. 1996. Patterns in tree balance with variable and evolving speciation rates. Evolution 50:2141–2148.

HEARD, S. B., AND D. L. HAUSER. 1995. Key evolutionary innovations and their ecological mechanisms. Hist. Biol. 10:151–173.

HEARD, S. B., AND A. Ø. MOOERS. 1996. Imperfect information and the balance of cladograms and phenograms. Syst. Biol. 45:115–118.

HEARD, S. B., AND A. Ø. MOOERS. 2000. Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. Proc. R. Soc. Lond. B 267:613–620.

HEARD, S. B., AND A. Ø. MOOERS. 2002. Signatures of random and selective mass extinctions in phylogenetic tree balance. Syst. Biol. 51:000–000 (this issue).

HEILBUTH, J. C. 2000. Lower species richness in dioecious clades. Am. Nat. 156:221–241.

HODGES, S. A., AND M. L. ARNOLD. 1995. Spurring plant diversification: Are floral nectar spurs a key innovation? Proc. R. Soc. Lond. B 262:343–348.

HUELSENBECK, J. P., AND M. KIRKPATRICK. 1996. Do phylogenetic methods produce trees with biased shapes? Evolution 50:1418–1424.

JABLONSKI, D. J. 1987. Heritability at the species level: Analysis of geographic ranges of Cretaceous mollusks. Science 238:360–363.

MCKINNEY, M. L. 1997. Extinction vulnerability and selectivity: Combining ecological and paleontological views. Annu. Rev. Ecol. Syst. 28:495–516.

MITTER, C., B. FARRELL, AND B. WIEGMANN. 1988. The phylogenetic study of adaptative zone: Has phytophagy promoted insect diversification? Am. Nat. 132:107–128.

MONCALVO, J. M., F. M. LUTZONI, S. A. REHNER, J. JOHNSON, AND R. VILGALYS. 2000. Phylogenetic relationships of agaric fungi based on large subunit ribosomal DNA sequences. Syst. Biol. 49:278–305.

MOOERS, A. Ø., AND S. B. HEARD. 1997. Evolutionary process from phylogenetic tree shape. Q. Rev. Biol. 72:31–54.

MOOERS, A. Ø., R. D. M. PAGE, A. PURVIS, AND P. H. HARVEY. 1995. Phylogenetic noise leads to unbalanced cladistic tree reconstruction. Syst. Biol. 44:332–342.

NEE, S., R. M. MAY, AND P. H. HARVEY. 1994. The reconstructed evolutionary process. Philos. Trans. R. Soc. Lond. B 344:305–311.

NEE, S., A. Ø. MOOERS, AND P. H. HARVEY. 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proc. Natl. Acad. Sci. USA 89:8322– 8326.

PAGEL, M. 1999. Inferring historical patterns of biological evolution. Nature 401:877–884.

PURVIS, A. 1996. Using interspecies phylogenies to test macroevolutionary hypotheses. Pages 153–168 *in* New uses for new phylogenies (P. H. Harvey, A. J. Leigh Brown, J. Maynard Smith, and S. Nee, eds.). Oxford Univ. Press, Oxford, U.K.

PURVIS, A., AND P.-M. AGAPOW. 2002. Phylogeny imbalance: Taxonomic level matters. Syst. Biol. 51:000–000 (this issue).

PURVIS, A., P.-M. AGAPOW, J. L. GITTLEMAN, AND G. M. MACE. 2000. Nonrandom extinction and the loss of evolutionary history. Science 288:328–330.

PYBUS, O. G., AND P. H. HARVEY. 2000. Testing macroevolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. B 267:2267–2272.

SANDERSON, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218–1231.

SANDERSON, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19:101– 109.

SANDERSON, M. J., AND M. J. DONOGHUE. 1994. Shifts in diversification rate with the origin of angiosperms. Science 264:1590–1593.

SANDERSON, M. J., AND M. J. DONOGHUE. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. 11:15–20.

SANDERSON, M. J., AND J. A. DOYLE. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. Am. J. Bot. 88:1499–1516.

SAVOLAINEN, V., AND J. GOUDET. 1998. Rate of gene sequence evolution and species diversification in flowering plants: A re-evaluation. Proc. R. Soc. Lond. B 265:603– 607.

SCHLUTER, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.

SIBLEY, C. J., AND J. E. AHLQUIST. 1990. Phylogeny and classification of birds: A case study in molecular evolution. Yale Univ. Press, New Haven, Connecticut.

SOKAL, R. R., AND F. J. ROHLF. 1995. Biometry, 3rd edition, W. H. Freeman, New York.

SOLTIS, D. E., P. S. SOLTIS, M. W. CHASE, M. E. MORT, D. C. ALBACH, M. ZANIS, V. SAVOLAINEN, W. H. HAHN, S. B. HOOT, M. F. FAY, M. AXTELL, S. M. SWENSEN, K. C. NIXON, AND J. S. FARRIS. 2000. Angiosperm phylogeny inferred from a combined data set of 18S rDNA, *rbcL*, and *atpB* sequences. Bot. J. Linn. Soc. Lond. 133:381–461.

SOLTIS, P. S., D. E. SOLTIS, AND M. W. CHASE. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402–404.

STANLEY, S. M. 1998. Macroevolution: Pattern and process. Johns Hopkins Univ. Press, Baltimore, Maryland.

SWOFFORD, D. L. 2001. PAUP* 4.0b8: Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Massachusetts.

VALENTINE, J. W. (ed.). 1985. Phanerozoic diversity patterns. Princeton Univ. Press, Princeton, New Jersey.

WHITING, M. F., J. C. CARPENTER, Q. D. WHEELER, AND W. C. WHEELER. 1997. The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. 46:1–68.

WIKSTRÖM, N., V. SAVOLAINEN, AND M. W. CHASE. 2001. Evolution of the angiosperms: Calibrating the family tree. Proc. R. Soc. Lond. B 268:2211–2220.

YANG, Z. H. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39:306–314.

YULE, G. U. 1924. A mathematical theory of evolution based on the conclusions of Dr J. C. Willis. Philos. Trans. R. Soc. Lond. A 213:21–87.

ZAR, J. H. 1999. Biostatistical analysis. Prentice-Hall, Englewood Cliffs, New Jersey.

First submitted 15 January 2002; revised manuscript returned 6 May 2002; final acceptance 14 August 2002
Associate Editor: Alan de Queiroz