Chapter 22

DETECTING DIVERSIFICATION RATE VARIATION IN SUPERTREES

Brian R. Moore, Kai M. A. Chan, and Michael J. Donoghue

Abstract: Although they typically do not provide reliable information on divergence times, supertrees are nevertheless attractive candidates for the study of diversification rates: by combining a collection of less inclusive source trees, they promise to increase both the number and density of taxa included in the composite phylogeny. The relatively large size and possibly more dense taxonomic sampling of supertrees have the potential to increase the statistical power and decrease the bias, respectively, of methods for studying diversification rates that are robust to uncertainty regarding the timing of diversification events. These considerations motivate the development of atemporal methods that can take advantage of recent and anticipated advances in supertree estimation. Herein, we describe a set of whole-tree, topology-based methods intended to address two questions pertaining to the study of diversification rates. First, has a given (super)tree experienced significant variation in diversification rates among its branches? Second, if so, where have significant shifts in diversification rate occurred? We present results of simulation studies that characterize the statistical behavior of these methods, illustrating their increased power and decreased bias. We also applied the methods to a published supertree of primates, demonstrating their ability to contend with relatively large, incompletely resolved (super)trees. All the methods described in this chapter have been implemented in the freely available program, SYMMETREE.

Keywords: cladogenesis; diversification rate shifts; diversification rate variation; equal-rates Markov random branching model; extinction; Primates; speciation; supertrees; tree shape; Yule branching process

Bininda-Emonds, O. R. P. (ed.) Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life, pp. 487–533. Computational Biology, volume 3 (Dress, A., series ed.).

1. Introduction

Supertrees represent somewhat of a mixed bag for the study of diversification rates, providing some kinds of information in unprecedented profusion but inherently limited in their ability to provide other types of pertinent data. Ideally, (super)trees can provide two sources of information relevant to the study of diversification rates: the temporal distribution of branching events through time and the topological distribution of species diversity across its branches[1].

It is generally accepted that, by virtue of directly incorporating information on the timing of diversification, temporal methods enjoy an advantage in power relative to their topological counterparts (e.g., Sanderson and Donoghue, 1996; Paradis, 1998a, b). This power advantage has, in turn, motivated the elaboration of temporal methods to effectively address a relatively wide range of evolutionary questions related to diversification rates. Unfortunately, existing supertree methods typically do not provide reliable branch-length estimates (but note recent progress by Lapointe and Cucumel, 1997; Bryant *et al*., 2004; Lapointe and Levasseur, 2004; Vos and Mooers, 2004), essentially precluding the use of more powerful temporal methods for the inference of diversification rates.

On the other hand, any decrease in power associated with the necessary reliance on topological methods might be offset to some extent by the typically larger size of supertrees because the power of these methods is known to scale with tree size (e.g., Kirkpatrick and Slatkin, 1993; Kubo and Iwasa, 1995; Paradis, 1997, 1998a, b; Agapow and Purvis, 2002). More than just their potentially larger size, however, is the promise of supertrees to greatly increase the density of sampled taxa. Both temporal and topological methods are sensitive to incomplete and/or nonrandom taxon sampling (e.g., Kubo and Iwasa, 1995; Nee *et al*., 1996; Pybus and Harvey, 2000; Barraclough and Nee, 2001) for the simple reason that these methods do not discriminate between species that have been omitted from a phylogenetic analysis and those that have been eliminated by extinction. The relatively broad and dense taxonomic sampling of supertrees should therefore confer

[1] Two corresponding classes of methods have been developed to exploit these different sources of information (Sanderson and Donoghue, 1996). The first class relies exclusively on topological information, comparing the observed difference in species diversity between two (or more) groups descended from a common node to the expectation generated under a stochastic model of diversification (e.g., Slowinski and Guyer, 1989a, b, 1993; Slowinski, 1990). The second class utilizes estimates of branch length or duration to infer the (absolute or relative) timing of speciation events and similarly compares the observed distribution of speciation events through time with that expected under a null model of random diversification (e.g., Harvey *et al*., 1991, 1994a, b; Hey, 1992; Nee *et al*., 1992, 1994a, b, 1995, 1996; Harvey and Nee, 1993, 1994; Sanderson and Bharathan, 1993; Kubo and Iwasa, 1995; Paradis, 1997, 1998a, b; Pybus and Harvey, 2000; Nee, 2001; Pybus *et al*., 2002). We refer to these two approaches as *topological* and *temporal* methods, respectively (Chan and Moore, 2002).

increased statistical power and decreased bias to studies of diversification rates, which motivates the development of methods that do not rely on temporal information.

Furthermore, even when reliable branch-length estimates are available, there might be situations in which it is preferable to omit these data from studies of diversification rates. Several types of evolutionary study entail hypothesized associations (whether correlational or causal in nature) between diversification rates and some other variable that is conditioned on branch lengths/durations. For example, there is considerable interest in exploring the putative correlation between rates of diversification and rates of molecular evolution (e.g., Mindell *et al.*, 1989; Barraclough *et al.*, 1996; Savolainen and Goudet, 1998; Barraclough and Savolainen, 2001; Jobson and Albert, 2002). Similarly, many evolutionary questions pertain to the relationship between rates of diversification and rates (and/or ancestral states) of morphological evolution. Often, rate estimates for such variables are either directly or indirectly conditioned on branch-length estimates (e.g., model-based inference of rates of nucleotide substitution, and model-based inference of rates and/or ancestral states of morphological character evolution, respectively). Consequently, attempts to understand the correlation of such variables to variation in rates of diversification will be confounded if both are conditioned on the same set of branch-length estimates. For such inference problems, it would therefore be desirable to possess methods that do not rely on branch-length data.

Accordingly, the nature of the data at hand and/or the hypotheses of interest will often preclude the inference of diversification rates based on temporal information. Clearly, topological methods warrant further consideration. In this chapter, we extend existing topological methods in new ways to exploit new opportunities. Because different people have different interests in the study of differential diversification rates, we describe a suite of methods intended to address two different questions: 1) has a given tree experienced significant variation in diversification rates among its branches; and, 2) if so, on which branches have significant shifts in diversification rate occurred? We explore the statistical behavior of the various methods by means of simulation and illustrate their application to empirical data using a published supertree of primates (Purvis, 1995). Choice of this data set was motivated by two considerations: the primate supertree is in many respects representative of those published for other groups (e.g., in its size, degree of resolution, and methods of estimation), and this tree has been used previously to explore various aspects of diversification rates in primates (e.g., Purvis *et al.*, 1995), thereby affording comparison of our results to those derived with other methods. All the methods described in this chapter have been implemented in the freely

available software program, SYMMETREE (http://www.kchan.org or (http://www.phylodiversity.net/brian/).

2. The equal-rates Markov random branching model

The ability of phylogenies to inform studies of differential diversification rates has been appreciated for some time. Hennig (1966) reasoned that any difference in species diversity between two sister groups, which are by definition of equal age, must necessarily reflect different rates of diversification (i.e., speciation minus extinction) in those groups. However, other researchers were quick to caution against overly deterministic interpretations of such differences: even if the underlying probability of diversification were identical in all lineages, some degree of variation in their realized diversification rates would be expected to arise because of the inherently stochastic nature of the branching process (e.g., Raup *et al.*, 1973; Gould *et al.*, 1977).

In recognition of the nature of the process under study, stochastic branching process models are frequently employed to generate an expected distribution of differences in diversity against which observed differences can be compared. One of the most elemental and frequently invoked models is the so-called equal-rates Markov (ERM) random branching process (Yule, 1924; Kendall, 1948; Harding, 1971). This is a continuous-time, discrete-state, pure-birth Markov process in which the probability of a branching event, λ, is constant for each tip in a growing tree at any moment in time[2]. Under the ERM model, the allocation of diversity among two sister groups follows a uniform distribution, such that all possible partitions of N species, 1:(N – 1), 2:(N – 2), 3:(N – 3) … (N – 1):1, are equiprobable. Accordingly, given an observed diversity partition of N into ℓ and r species among two sister groups, we can calculate the cumulative probability of realizing a diversity partition as or more extreme under the ERM model as

$$P = \frac{2\ell}{\left(N-1\right)} \tag{1}$$

[2] Note that the ERM model allows λ to vary through time, so long as it is equal across all tips at any instant (e.g., Harding, 1971). This property of the ERM model technically distinguishes it from the more restricted constant-rate, pure-birth Yule branching process model because the latter constrains λ to be constant both across tips and through time (e.g., Yule, 1924). Nevertheless, the two models are operationally identical when branching times are unknown, as is the case for topology-based inferences of diversification rate.

(unless $\ell = N / 2$, in which case $P = 1$), where ℓ is the number of species in the less diverse of the two sister groups (Slowinski and Guyer, 1989a). A significant difference in sister-group diversity constitutes rejection of the ERM null model, and therefore, suggests that the two lineages have diversified under significantly different rates (Slowinski and Guyer, 1989a, b; Slowinski, 1990). For convenience, we refer to these *P*-values as *ERM nodal probabilities* because they pertain to the cumulative ERM probability of realizing a diversity partition between lineages descended from a shared node.

Derivation of an ERM nodal probability incorporates minimal information on the topological distribution of species diversity (only two observations are made). Because the statistical power of a test is a function of sample size, the sensitivity of these single-node tests to differential diversification rates is quite low (e.g., Kirkpatrick and Slatkin, 1993; Fusco and Cronk, 1995; Sanderson and Donoghue, 1996; Sanderson and Wojciechowski, 1996). As we will demonstrate in the following sections, however, these nodal probabilities can serve as building blocks that can be variously generalized to construct methods that harness their collective power.

3. Detecting among-lineage diversification rate variation

In this section we consider the question, "Has a given tree experienced significant diversification rate variation among its branches?" This is the diversification rate analogue to the problem of detecting among-lineage substitution rate variation in studies of molecular evolution. The ability to detect among-lineage diversification rate variation has parallel applications to tests of the molecular clock: tests of rate homogeneity are a prerequisite for the application of several temporal methods that assume negligible levels of among-lineage diversification rate variation (e.g., Hey, 1992; Harvey *et al.*, 1994a, b; Nee *et al.*, 1994a, b; Kubo and Iwasa, 1995; Paradis, 1997, 1998a, b; Pybus and Harvey, 2000). Additionally, and like its molecular counterpart, the study of diversification rate variation has important evolutionary implications that might be of interest in their own right (Chan and Moore, 2002).

Previous work on this problem has largely involved the development of "tree-balance indices", metrics that variously summarize the topological distribution of species diversity as a single number. Approximately 20 such indices have been proposed (e.g., Colless, 1982; Shao and Sokal, 1990; Heard, 1992; Kirkpatrick and Slatkin, 1993; Page, 1993; Fusco and Cronk,

1995; McKenzie and Steel, 2000; Agapow and Purvis, 2002; Purvis *et al.*, 2002). Several authors have noted that these indices appear to capture different but poorly characterized aspects of tree shape (Shao and Sokal, 1990; Kirkpatrick and Slatkin, 1993; Fusco and Cronk, 1995). Consequently, all attempts to test for significant diversification rate variation with these tree-balance indices must grapple with the "agony of choice" between myriad alternatives or opt to use all (or some subset of) the indices and endure issues of multiple-test correction. In any case, interpretation of results under the chosen index (or indices) is apt to be less than straightforward: these indices are not derived explicitly from any model of diversification, such that the biological meaning of "significant imbalance" under these tests is unclear.

Our approach to the problem draws on the analogy to the study of among-lineage substitution rate variation: just as single-node tests (as implemented by the relative-rate test; e.g., Sarich and Wilson, 1967; Wu and Li, 1985) have been variously generalized over the whole tree (e.g., Felsenstein 1988, 1989; Takezaki *et al.*, 1995) to realize substantially increased sensitivity to substitution-rate variation, our strategy is to generalize single-node tests (as implemented by ERM nodal probabilities) over the whole tree with the similar objective of increasing the power to detect diversification rate variation. Our presentation of these whole-tree methods necessarily draws upon our previous work (Chan and Moore, 2002) but includes several new results, including the development of two new statistics and a simulation-based exploration of their statistical behavior.

3.1 Whole-tree tests of diversification rate variation

Generalization of the single-node approach to incorporate information on the relative diversity of all internal nodes of a tree would provide a much more powerful and — by virtue of being based on an explicit model of cladogenesis — also biologically meaningful test of among-lineage diversification rate variation. The development of such whole-tree methods might be achieved by combining individual ERM nodal probabilities on a node-by-node basis over all internal nodes of a given phylogeny (J. Slowinski, pers. comm. to Kirkpatrick and Slatkin, 1993). But how should individual nodal probabilities be combined? A subsequent development by Slowinski and Guyer (1993) suggests a possible solution. They proposed a method for combining individual ERM probabilities from single-node comparisons from many *different* trees using Fisher's combined probability test (FCPT; Fisher, 1932). It would seem relatively straightforward to modify the FCPT protocol to combine probabilities from many nodes within the *same* tree (Figure 1).

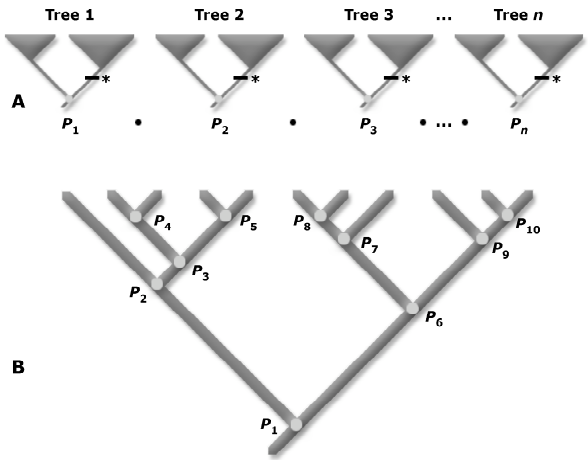

Figure 1. Combining nodal probabilities to develop whole-tree tests of diversification-rate variation. A) Slowinski and Guyer (1993) proposed combining individual ERM nodal probabilities (derived using equation (1)), each from a different tree, using Fisher's combined probability test (FCPT) to test the cumulative effect of a putative key innovation on rates of diversification in the various groups in which it evolved independently (indicated by asterisks). B) Whole-tree tests of diversification-rate variation could seemingly be developed by using FCPT (or ECPT) to combine the individual ERM nodal probabilities from many nodes within the same tree (e.g., $P_1 - P_{10}$). However, the FCPT and ECPT tests assume that the individual probabilities to be combined are independent and can each realize any value between 0 and 1. Nodal probabilities, however, are both non-independent (e.g., P_4 and P_5 are nested phylogenetically within P_3) and valued discretely (they are derived from the comparison of discretely valued species numbers). Nevertheless, approximate solutions can be devised that allow for the combination of nodal probabilities by using Monte Carlo simulation to estimate the appropriate distribution of the test statistics.

Although intuitively appealing, the combination of nodal probabilities under the FCPT is extremely biased. This bias stems from violation of the underlying assumptions of omnibus statistics (i.e., statistics that, like the FCPT, reflect the combined significance of several independent tests of a common hypothesis). The FCPT statistic is calculated by estimating the compound probability that a set of probabilities (in this case, the set of ERM nodal probabilities derived with equation (1)) has a product equal to or smaller than that of the observed set (Fisher, 1932). A less common but equally valid omnibus statistic proposed by Edgington (ECPT: 1972a, b) takes the sum rather than the product of individual probabilities. Both the FCPT and ECPT assume that the individual probabilities to be combined are independent and can realize any value on the interval (0, 1]. However, nodal probabilities are interdependent to the extent that they are derived from phylogenetically nested nodes and these probabilities can realize only a

finite number of discrete values for the simple reason that they are derived (using equation (1)) from the comparison of species diversities, which necessarily occur as whole numbers (i.e., 1, 2, 3, …). This "discreteness" problem is known to cause a discrepancy between the assumed and realizable probability space (Wallis, 1942; Edgington and Haller, 1984), such that the combination of individual nodal probabilities under the FCPT or ECPT will assume a concave function of the true cumulative probabilities.

In view of the complications associated with the use of conventional omnibus statistics for this problem, we pursue a non-analytical solution that avoids the discreteness and interdependence problems while emulating the logic of the FCPT and ECPT statistics. We first review two whole-tree tests of diversification rate variation based on the cumulative ERM probability derived from the product (M_{Π}) and sum (M_{Σ}) of individual nodal probabilities (Chan and Moore, 2002) and then develop two modified versions of these whole-tree statistics, M_{Π}* and M_{Σ}*, that differentially weight the individual ERM nodal probabilities according to their species diversity. Conceptually, these four tests involve mapping the sample space that can be realized by discretely valued, interdependent ERM nodal probabilities. This entails the use of Monte Carlo simulation to estimate the underlying distribution of topologies that can be realized for a tree of a given size.

These tests are implemented with one of two algorithms depending upon the size of the tree in question. For smaller trees ($N < 20$), the appropriate ERM sample space can be mapped exactly by applying the "small-tree" algorithm as follows: 1) Calculate the product (or sum) of all ERM nodal probabilities (derived by equation (1)) in the observed tree. 2) Generate all possible topologies for a tree with the same number of species as the observed tree. For each topology, calculate the product (or sum) of its nodal probabilities and its point probability under the ERM model. 3) Sum the point probabilities of all topologies with nodal probability products (or sums) less than or equal to that of the observed tree. This sum represents the cumulative whole-tree probability based on the nodal probability product, M_{Π} (or on the nodal probability sum, M_{Σ}).

For larger trees ($N > 20$), the appropriate ERM sample space must be approximated using the "large-tree" algorithm owing to the vast number of possible topologies (e.g., only 46 for nine species, but 105 061 603 969 for 35 species; Stone and Repka, 1998). The large-tree algorithm is executed as follows: 1) As in the small-tree algorithm, first calculate the product (or sum) of ERM nodal probabilities in the observed tree. 2) Using the ERM model of cladogenesis, generate a large, random subset of possible topologies for a tree with the same number of species as the observed tree. 3) Count the number of simulated trees with a nodal probability product (or

sum) less than or equal to that of the observed tree and divide by the total number of simulated trees. This quotient is an unbiased estimate of the probability corresponding to M_{Π} (or M_{Σ}).

Note that all nodal probabilities contribute equally to the calculation of the M_{Π} and M_{Σ} whole-tree statistics. However, larger nodes (i.e., those defining more diverse clades) sample a greater number of diversification events and should, therefore, provide more reliable evidence of non-random variation in diversification rates (Figure 2). Accordingly, the power of the whole-tree statistics to detect diversification rate variation should be enhanced by scaling the weight of nodal probabilities according to the size (i.e., species diversity) of their respective nodes. Because diversification is an exponential process, the contribution of each nodal probability is scaled by the natural logarithm of its diversity. The cumulative whole-tree probability based on the product of weighted ERM nodal probabilities, M_{Π}^{*}, involves first calculating the product of weighted ERM nodal probabilities, Π^{*}, for the observed tree and the set of simulated trees using the equation

$$\Pi^{*} = \frac{\sum_{i=1}^{n-1} \ln(n_i)\ln(P_i)}{\sum_{i=1}^{n-1} \ln(n_i)}. \tag{2}$$

(Recall that the sum of the natural logarithms of the ERM nodal probabilities is equivalent to taking their product.) The cumulative whole-tree probability, M_{Π}^{*}, is simply the frequency of simulated trees with Π^{*} values less than that of the observed tree. Similarly, the cumulative whole-tree probability based on the sum of weighted ERM nodal probabilities, M_{Σ}^{*}, involves calculating the sum of weighted ERM nodal probabilities, Σ^{*}, using the equation

$$\Sigma^{*} = \frac{\sum_{i=1}^{n-1} \ln(n_i)P_i}{\sum_{i=1}^{n-1} \ln(n_i)}, \tag{3}$$

where n_i is the diversity of internal node i, and P_i is its corresponding ERM nodal probability derived using equation (1). Given two trees with the same number of tips but different topological shapes, the more asymmetric tree will contain a greater proportion of nodes that are relatively large compared with the more balanced tree. Accordingly, the denominators in equations (2) and (3) normalize the summation of $\ln(n_i)$ over different tree shapes.

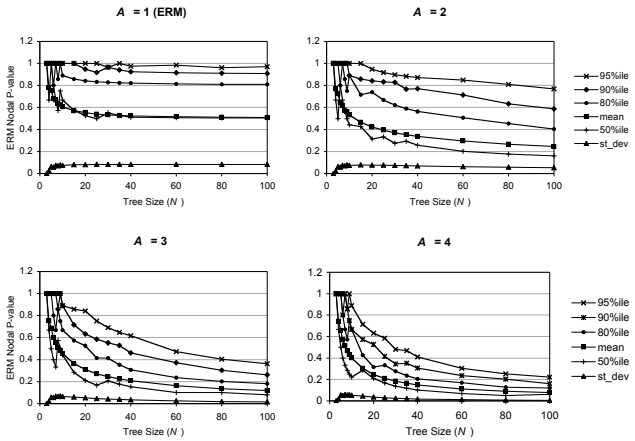

Figure 2. The ability to detect non-ERM diversification increases with tree size. The plots were generated by initiating a stochastic ERM-branching process from a single species with the diversification-rate parameter, λ, set initially to 1. After the first branching event, a diversification-rate shift of magnitude A, where $A \in \{1, 2, 3, 4\}$, was applied deterministically to one of the two lineages descended from the root node. The process was terminated when trees reached size N, where $N \in \{1, 2, 3, \ldots, 10, 15, 20, \ldots, 40, 60, 80, 100\}$. Each combination of parameter settings (magnitude of diversification-rate difference, tree size) was replicated 100 000 times, and, for each tree generated from each such replicate, the ERM nodal probability was calculated for the root node using equation (1). The graphs plot the mean, standard deviation, and various percentiles of the ERM nodal probabilities (where a percentile is the ERM P-value corresponding to the simulated tree for which x% of the set of simulated trees had lower P-values, where $x \in \{50, 80, 90, 95\}$). The plots within each of the four graphs (corresponding to a set of simulations under a given value of A) are concave, with the ERM nodal probabilities for the root node decreasing in value with increasing tree size. For a given value of A, ERM P-values are clustered more tightly around small values for larger trees. Under a diversification-rate difference of three, for example, we are much more likely to obtain a P-value of < 0.1 for $N = 100$ than for $N = 10$, indicating that larger nodes provide more reliable evidence of non-ERM diversification. Note that the apparently stochastic wobbling of the percentile plots near the y-axis is actually a manifestation of the "discreteness" problem. For a tree of a given size, only a finite number of discretely valued diversity partitions can be realized; accordingly, only a finite number of P-values can be realized by their corresponding nodal probabilities. As expected, the discreteness problem is most pronounced for trees of small size.

3.2 The relative sensitivity to diversification rate variation at different phylogenetic scales

Our motivation for developing the whole-tree methods described above is to increase the statistical power of tests to detect diversification rate variation. Power is the ability of a test to reject a null hypothesis when it is false. Nodal

ERM probabilities are the most appropriate measure for tests of differential diversification at individual nodes. Accordingly, we expect the combination of these values — as implemented by the *M* statistics — to provide tests of the ERM model that are exceptionally sensitive to diversification rate variation within whole (super)trees.

Of course, the power of a test is contingent on the nature of the particular alternative hypothesis under consideration. Because there are innumerable possible alternatives to equiprobable diversification rates (frequent rate shifts dispersed throughout the tree or infrequent rate shifts occurring near the base of the tree, among others), it is unrealistic to expect any single statistic to be maximally powerful in all scenarios involving differential diversification. Given the multitude of possible and biologically relevant alternatives to ERM cladogenesis, several different statistics are required. The *M* statistics are intended to provide differential sensitivity to asymmetry arising at different phylogenetic scales (i.e., the relative nodal depth in the tree), permitting their application to a corresponding range of associated evolutionary processes.

The manner in which each statistic summarizes information from individual nodes (i.e., ERM probabilities) will determine the type of diversification rate variation (i.e., the alternative hypothesis) to which it is most sensitive. By considering how the different *M* statistics differentially summarize ERM nodal probabilities, we can theoretically characterize their differential sensitivity to different patterns of diversification rate variation without performing the simulations necessary for a complete characterization of their relative power.

Although M_{Π} and M_{Σ} both consider the relative asymmetry of all internal nodes, these statistics nevertheless exhibit differential sensitivity to large-scale asymmetry. To understand the source of this difference, recall that the potential magnitude of diversity partitions is greater at more inclusive nodes. Consider, for example, that the most extreme diversity partition of an *N*-species tree is a split of $1:(N-1)$, which can only be realized at the root; the next most extreme partition, $2:(N-2)$, can only be realized at the root or at the node just above the root, and so on. Accordingly, the most extreme nodal probabilities (i.e., the smallest) can only be generated by large-scale asymmetry. These extreme probabilities will have a relatively large effect on M_{Π} because calculation of the statistic involves their multiplication. By contrast, M_{Σ} combines nodal probabilities additively, such that the impact of such extreme probabilities is greatly diminished, allowing nodal probabilities associated with small-scale asymmetry to make a more equable contribution to the whole-tree probability under this statistic.

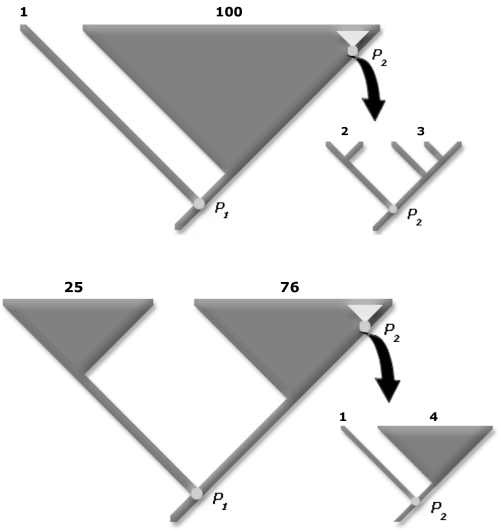

Figure 3. The differential sensitivity of the whole-tree tests to diversification-rate variation manifested at different phylogenetic scales. Trees *A* and *B* exhibit substantial differences in large-scale phylogenetic asymmetry: *A* has a basal split of 1:100 (P_1 = 0.02, $ln\ P_1$ = –3.91) versus a 25:76 split in *B* (P_1 = 0.5, $ln\ P_1$ = –0.69). Now, imagine that the only other difference in asymmetry between the two trees is restricted to a five-species subtree that has a 2:3 split in *A* (P_2 = 1.0, $ln\ P_2$ = 0) and a 1:4 split in *B* (P_2 = 0.5, $ln\ P_2$ = –0.69). The number of such asymmetric five-species subtrees that would be required by each whole-tree statistic to identify *B* as more asymmetric than *A* can be used to characterize their relative sensitivity to small-scale phylogenetic asymmetry. M_Σ identifies *B* as more asymmetric with just a single asymmetric five-species subtree ($P_{1A} + P_{2A}$ = 1.02; $P_{1B} + P_{2B}$ = 1.0); M_Σ* requires three or more equivalent differences; M_Π requires five or more equivalent differences ($lnP_{1A} - lnP_{1B}$ = –3.22, $lnP_{2A} - lnP_{2B}$ = 0.69); and M_Π* requires 14 or more equivalent differences. For comparison, I_C requires 25 or more equivalent differences in small-scale asymmetry, whereas B_1 identifies *B* as *far* more asymmetric than *A* with only a single such difference. Thus, the sensitivity of the whole-tree statistics to diversification-rate variation occurring at large phylogenetic scales is approximately $B_1 < M_\Sigma < M_\Sigma^* < M_\Pi < M_\Pi^* < I_C$.

Predictably, the behavior of the weighted whole-tree statistics, M_Π* and M_Σ*, is similar to that of their equally weighted counterparts. However, because the contribution of each ERM nodal probability to these whole-tree statistics is weighted by the size of its corresponding node, and because larger nodes are realized deeper in the tree, M_Π* and M_Σ* are more sensitive to diversification rate variation at larger phylogenetic scales. Accordingly, the relative sensitivity of the *M* statistics to large-scale diversification rate variation can be approximately characterized as $M_\Sigma < M_\Sigma^* < M_\Pi < M_\Pi^*$ (Figure 3).

3.3 Assessing the statistical behavior of the whole-tree statistics

We performed a simulation study to characterize the relative power of the five whole-tree statistics (M_Σ, M_Σ*, M_Π, M_Π*, *and* M_R) and two previously proposed balance indices: I_C (Colless, 1982; Heard, 1992) and B_I (Shao and Sokal, 1990). Our decision to compare the M statistics with these two balance metrics is based on several considerations. I_C is both the most commonly used index (e.g., Mooers and Heard, 1997) and is also very well characterized mathematically (e.g., Heard, 1992; Rogers 1993, 1994, 1996). By contrast, our inclusion of B_I is motivated by the finding that it is the most powerful of the balance indices (Kirkpatrick and Slatkin, 1993; but see Agapow and Purvis, 2002).

The ability of the seven statistics to detect diversification rate variation was assessed by a simulation design that involved growing trees under a variety of non-ERM conditions intended to simulate plausible and potentially biologically interesting models of cladogenesis. In general, trees were grown under a continuous-time, discrete-state, stochastic branching process in which splitting events were assumed to be both instantaneous and dichotomous. The probability of a branching event was assumed to be independent between tips in a growing tree, with rate shifts being equally likely to involve an increase or a decrease in diversification rate. If no rate shift occurred, a given tip retained the diversification rate of its ancestor.

Diversification rate shifts were applied under three general models of cladogenesis. Under the gradualist model, rate shifts could occur at any instant in time and were inherited by both daughter species. Alternatively, two different punctuated models constrained rate shifts to occur at speciation events, with either one or both daughter species having a chance of experiencing a rate shift. For each evolutionary model, we explored the effects of varying the frequency and magnitude of rate shifts in trees of various sizes. Average diversification rate shift values, λ, included two-, four-, eight-, and 16-fold increases in diversification rate, which were applied under a range of frequencies (0.01, 0.1, 0.2, 0.3, 0.4, 0.5). The branching process was terminated when trees reached the desired size, N, where $N \in \{10, 15, 20, 25, 30, 35, 40, 60, 80, 100\}$. Every permutation of the set of simulation parameters (evolutionary model, rate distribution, tree size, and frequency and magnitude of rate shifts) was replicated 100 000 times, calculating the value for each of seven statistics for each tree generated from each replicate. Power was calculated as the proportion of the replicates in which the null hypothesis of no among-lineage diversification rate variation was correctly rejected at the conventional $\alpha = 0.05$.

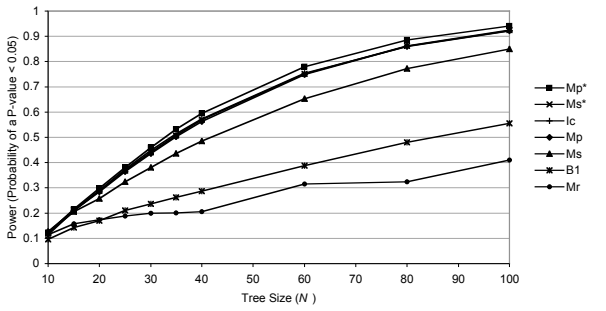

Figure 4. The effect of tree size on the power of several whole-tree methods to detect diversification-rate variation.

Several results of the simulation study were as predicted. First, the use of Monte Carlo simulation to assess significance of the various whole-tree statistics ensured appropriate Type I error rates. The plots for each statistic intersected the *y*-axes at $P \cong 0.05$ (the nominal level of α) when the average diversification rate shift, λ, was 1 (i.e., when the null hypothesis was true). Second, the power of the whole-tree statistics to detect diversification rate variation consistently scaled with tree size (Figure 4). This result is consistent both with theoretical expectations (Figure 2) and findings of previous simulation studies (e.g., Kirkpatrick and Slatkin, 1993; Kubo and Iwasa, 1995; Paradis, 1997, 1998a, b; Agapow and Purvis, 2002) and emphasizes the potential of typically large supertrees to facilitate the study of diversification rate variation. Finally, the observed behavior of the various whole-tree statistics under various rate-shift parameterizations was also unsurprising: power predictably scaled with increases in both the frequency and magnitude of rate shifts applied.

Somewhat more surprising was the response of some whole-tree statistics to various combinations of frequency and magnitude of diversification rate shifts. For instance, we might expect that simulations involving large shifts occurring at low frequencies would enhance the relative power of the M_R statistic given its inherent sensitivity to large-scale diversification rate variation. Similarly, we might predict that the relative performance of the M_Σ or B_1 statistics would be enhanced under conditions involving shifts of small magnitude occurring at relatively high frequencies. Curiously, and despite their rather compelling theoretical basis, no unambiguous patterns supporting these behaviors emerged from the simulation study. A thorough consideration of such intriguing anomalies is beyond the scope of the present analysis but will be treated elsewhere (Moore and Chan, in prep.).

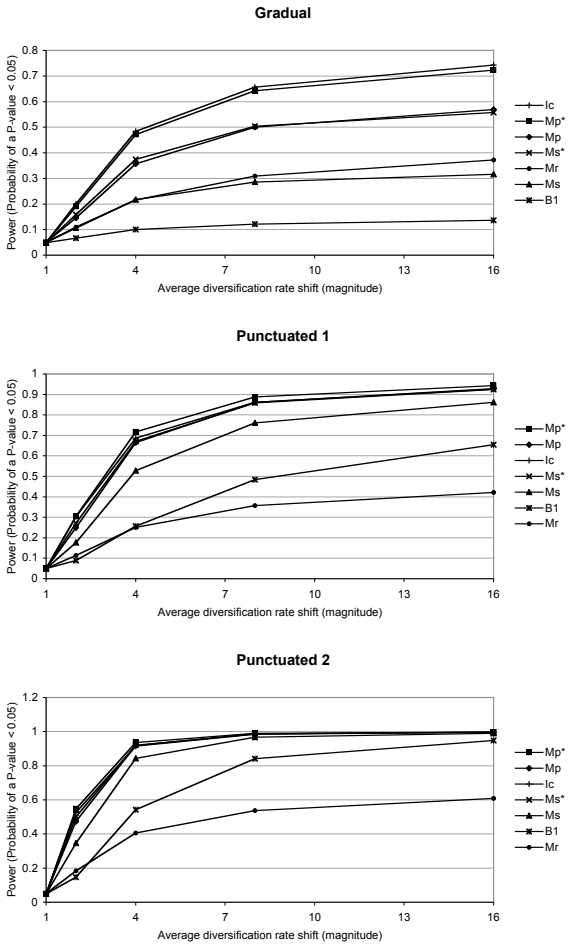

Figure 5. The effects of evolutionary model on the power of several whole-tree methods to detect diversification-rate variation.

Perhaps one of the more interesting findings to emerge from the simulation study was the pronounced effect of the model of diversification on the relative power of the whole-tree methods to detect diversification rate variation. Figure 5 depicts the results for 100-species trees grown under one of three diversification models with rate shifts of various magnitude applied with a constant frequency of 0.1, which were sampled from a uniform rate-shift distribution. Overall, the power of all the statistics tended to be greatest under the Punctuated 2 Model, in which rate shifts were constrained to occur at speciation events with any change in rate inherited by both daughter species (Figure 5, lower graph). By contrast, power was noticeably lower

under the Gradual Model, in which rate shifts were free to occur at any time with any change in rate shared between both daughter species (Figure 5, upper graph). Finally, power was intermediate under the Punctuated 1 Model, in which only one of the descendent species inherited any change in rate (Figure 5, middle graph). Nevertheless, the absolute power of the methods was fairly high even under conditions least favorable to the whole-tree statistics (e.g., two of the statistics, M_{Π}* and I_C, detected a four-fold variation in diversification rates correctly about 50% of the time under the Gradual Model at a very low diversification rate shift frequency).

The conditional nature of conclusions regarding the statistical power of these tests should be emphasized. Despite this caveat, several generalities held over a wide range of the considerable parameter space we explored. Apart from a limited number of extreme conditions, the performance of the M_R and B_1 statistics was uniformly poor. Given its widely accepted status as the most powerful statistic (based on the particular conditions simulated by Kirkpatrick and Slatkin, 1993), the poor performance of B_1 was somewhat surprising. By contrast, the M_{Π}* statistic consistently exhibited maximal (or nearly maximal) power under the vast majority of the simulations.

3.4 Detecting diversification rate variation in primates

The whole-tree *M* statistics described above were used to assess diversification rate variation in a published supertree of primates (Purvis, 1995). Because these data were analyzed for illustrative purposes only, no attempt was made to account for the effect of phylogenetic uncertainty on the results (e.g., Donoghue and Ackerly, 1996; Huelsenbeck *et al.*, 2000b). In addition to analyzing the complete primate tree, we also performed analyses on several clades of primates to facilitate comparison both with the findings of previous temporal studies of diversification rate variation in this group (Purvis *et al*., 1995) and also with results presented in Section 4.3. Note that inference of diversification rate variation in these clades is somewhat confounded: shifts within more nested clades will influence estimates obtained for more inclusive clades. Accordingly, these results should be interpreted cautiously (Purvis *et al*., 1995). Results derived with the whole-tree methods were compared again with those of the tree-shape indices I_C and B_1. All analyses were performed with SYMMETREE, with relevant details and results summarized in Table 1.

Three general findings merit comment. First, the primate tree contains 203 species and is ~80% resolved, illustrating the ability of the whole-tree methods (and their implementation in SYMMETREE) to contend with moderately large and incompletely resolved trees. Second, analysis of the entire primate clade failed to detect significant among-lineage diversification

Table 1. Probability values corresponding to tests of ERM cladogenesis in various primate clades as derived by Monte Carlo simulation of the null distribution for each statistic. All results were obtained using the SYMMETREE program. The null distribution for each statistic was generated with a sample of 100 000 ERM topologies for each tree size. Uncertainty associated with polytomies was assessed by generating 100 000 random resolutions under the size-sensitive ERM taxon-addition algorithm, providing the upper and lower bounds of the confidence interval. These bounds, the "high" and "low" values (for high and low asymmetry), correspond to the tail probabilities for the .025 and .975 frequentiles, respectively. Note that the sensitivity of the whole-tree statistics to large-scale diversification rate variation increases to the right across a given row (i.e., $B_1 < M_\Sigma < M_\Sigma^* < M_\Pi < M_\Pi^* < I_C$). Percent resolution was calculated as $k / (N - 1)$, where k is the number of nodes in a tree of N species; this value assumes implicitly that the underlying phylogeny is strictly dichotomous (i.e., that all polytomies are "soft"; *sensu* Maddison, 1989).

taxon	tree size	resolution	B_1	M_Σ	M_Σ^*	M_Π	M_Π^*	I_C
			high	high	high	high	high	high
			low	low	low	low	low	low
all primates	203	79	0.00020	0.00414	0.04097	0.00468	0.04004	0.21772
			0.09481	0.18733	0.30751	0.12587	0.20208	0.32138
hominoids	14	85	0.00074	0.00507	0.01956	0.01063	0.04627	0.05123
			0.01481	0.03833	0.08997	0.07831	0.17608	0.18630
strepsirhines	39	82	0.14391	0.12380	0.12380	0.24500	0.48850	0.67694
			0.77485	0.67503	0.33586	0.71871	0.82819	0.89569
New World monkeys	65	72	0.01541	0.32875	0.65913	0.44013	0.65815	0.78990
			0.73206	0.97304	0.98205	0.96326	0.96834	0.95474
Old World monkeys	80	81	0.00168	0.00134	0.00473	0.00045	0.00384	0.13364
			0.18815	0.09342	0.08558	0.02488	0.03825	0.30341

rate variation. However, significant diversification rate variation was detected in separate analyses of both hominoids and Old World monkeys. These findings are largely consistent with those reported by Purvis *et al.* (1995), who detected diversification rate variation within both of these clades using temporal methods. Finally, close inspection of the *P*-values for the various whole-tree statistics supports their predicted behavior with respect to diversification rate variation manifest at different phylogenetic scales. The statistics in Table 1 are arranged by their predicted sensitivity to large-scale diversification rate variation (i.e., in the order $B_1 < M_\Sigma < M_\Sigma^* < M_\Pi < M_\Pi^* < I_C$). Looking across a row for any group reveals a trend in the *P*-values; for example, the probabilities for Old World monkeys tend to decrease from B_1 to M_Π and then increase from M_Π^* to I_C (with some shuffling of the order of the statistics resulting from differences in the their absolute power under the particular manner in which the null hypothesis was violated in these data). The most extreme *P*-value (i.e., the smallest) obtained for this clade was returned by M_Π, suggesting that diversification

rate variation in the Old World monkey tree likely occurred at an intermediate phylogenetic scale.

4. Locating shifts in diversification rate

Having provided a means with which to answer the question, "*Has* a given tree experienced significant diversification rate variation among its branches?" in the preceding section, we now address its inevitable sequel: "*Where* have significant shifts in diversification rate occurred in this tree?" Despite its obvious biological significance, this problem has received remarkably little attention (however, Nee *et al*. (1992, 1994b, 1996) developed an approach incorporating temporal information that has been applied to this problem, which we consider in some detail below). By contrast, considerable attention has focused on methods to test hypotheses that specify the location and direction of diversification rate shifts (i.e., "key-innovation" hypotheses[3]). Fortunately, several developments in this hypothesis-testing realm are directly relevant to the issue of localizing shifts in diversification rate. Of particular importance is the iterative maximum likelihood model-fitting approach proposed by Sanderson and Donoghue (1994; see also Sanderson and Bharathan, 1993; Sanderson, 1994; Sanderson and Wojciechowski, 1996).

Following Sanderson and Donoghue (1994), our approach to detecting shifts in diversification rate is developed in a likelihood framework that evaluates the relative fit of models with one or more rate parameters distributed over different parts of a three-taxon tree and assumes an underlying ERM (Yule) branching process. However, our implementation is both significantly simplified (we evaluate only one- and two-rate parameter models and do not integrate their likelihood over all internal branching times) and also substantially generalized (we iterate three-taxon evaluations over all internal branches to survey the whole tree for diversification rate shifts).

In outline, the basic goal is to assess the probability of a shift along the lone internal branch of a given three-taxon tree comprising an outgroup clade and the two basal-most subclades of the ingroup clade. The probability of a diversification rate shift along the internal branch is returned by a shift

[3] Although related, these inference problems are nevertheless distinct. The evaluation of key innovations entails a hypothesis-testing framework in which the location and direction of a diversification rate shift is specified by the hypothesis under consideration (without any knowledge that the tree exhibits significant among-lineage diversification rate variation). By contrast, the search for significant shifts in diversification rate entails a data-exploration framework in which only the existence of significant among-lineage diversification rate variation is specified (without any knowledge of the location or direction of the associated rate shifts).

statistic, which is calculated as a function of two likelihood ratios. One likelihood ratio is calculated at the root of the three-taxon tree (involving the diversity partition between the outgroup and ingroup clades), the other at the root of the ingroup clade (involving the diversity partition between the left and right ingroup clades). Each likelihood ratio compares the likelihood of realizing the observed diversity partition between the two sister clades under a homogeneous (one-rate parameter) model (in which both groups have the same branching rate) versus that under a heterogeneous (two-rate parameter) model (in which the two groups have different branching rates). Different shift statistics can be developed by variously combining information from the resulting inclusive and nested likelihood ratios. Before explicitly deriving these shift statistics, we first review both the details of calculating the likelihoods under one- and two-rate parameter models and also the means of assessing their relative fit to the data using the likelihood ratio.

If the ERM branching process is initiated with a single species and allowed to run for a period of time, t, with a branching probability, λ, the likelihood of realizing N species is (Harris, 1964)

$$P(N \mid \lambda, t) = e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{N-1}. \tag{4}$$

Accordingly, the likelihood of realizing N species partitioned between the left and right descendents of a single node (with ℓ and r species, respectively) under a uniform branching probability after time, t, is

$$P\left(\ell, r \mid H_O\right) = \frac{P\left(\ell \mid \lambda, t\right)P\left(r \mid \lambda, t\right)}{\sum_{i=1}^{N-1} P\left(i \mid \lambda, t\right)P\left(N - i \mid \lambda, t\right)}. \tag{5a}$$

The Markov property of the ERM branching process allows the probabilities for different parts of the tree (such as the two terms in the numerator) to be multiplied. Substituting the expression from equation (4) with $t = 1$ gives the following expansion

$$P\left(\ell, r \mid H_O\right) = \frac{\left(e^{-\lambda}\left(1 - e^{-\lambda}\right)^{\ell-1}\right)\left(e^{-\lambda}\left(1 - e^{-\lambda}\right)^{r-1}\right)}{\sum_{i=1}^{N-1}\left(e^{-\lambda}\left(1 - e^{-\lambda}\right)^{i-1}\right)\left(e^{-\lambda}\left(1 - e^{-\lambda}\right)^{N-i-1}\right)}. \tag{5b}$$

This equation provides the likelihood of observing a partition of ℓ and r species (where $\ell + r = N$) under H_O, the homogeneous, one-rate parameter

model. Similarly, the likelihood of observing a partition of ℓ and r species under the heterogeneous, two-rate parameter model, H_A, is

$$P(\ell, r \mid H_A) = \frac{P(\ell \mid \lambda_\ell, t)P(r \mid \lambda_r, t)}{\sum_{i=1}^{N-1} P(i \mid \lambda_\ell, t)P(N-i \mid \lambda_r, t)}. \tag{6a}$$

Again, substituting the expression from equation (4) with $t = 1$ gives the expansion

$$P(\ell, r \mid H_A) = \frac{\left(e^{-\lambda_\ell}\left(1-e^{-\lambda_\ell}\right)^{\ell-1}\right)\left(e^{-\lambda_r}\left(1-e^{-\lambda_r}\right)^{r-1}\right)}{\sum_{i=1}^{N-1}\left(e^{-\lambda_\ell}\left(1-e^{-\lambda_\ell}\right)^{i-1}\right)\left(e^{-\lambda_r}\left(1-e^{-\lambda_r}\right)^{N-i-1}\right)}, \tag{6b}$$

The denominators in equations (5) and (6) normalize their respective probabilities by defining the relevant probability space. Specifically, this pertains to the sum of the products for all possible partitions of N into ℓ and r species.

The relative fit of the one- and two-rate parameter models to the observed diversity partition is assessed by the difference in the natural logarithm of their respective likelihood values: the log-likelihood ratio (hereafter, simply "likelihood ratio") of the homogeneous and heterogeneous diversification rate models, $LR_{HA:HO}$, is, therefore, calculated as

$$LR_{H_A:H_O} = \ln\left(\frac{P(\ell \mid \lambda_\ell, t)P(r \mid \lambda_r, t)}{\sum_{i=1}^{N-1} P(n_i \mid \lambda_\ell, t)P(N-n_i \mid \lambda_r, t)}\right) - \ln\left(\frac{P(\ell, r \mid \lambda, t)}{\sum_{i=1}^{N-1} P(n_i \mid \lambda, t)P(N-n_i \mid \lambda, t)}\right). \tag{7}$$

As the value of likelihood ratio increases, the evidence increasingly favors acceptance of the heterogeneous model in which the left and right descendants of the node in question diversified under two distinctly different rates, λ_ℓ and λ_r, respectively.

4.1 Presentation of the shift statistics

Having detailed the calculation of likelihoods under the homogeneous and heterogeneous models (and their relative fit with the likelihood ratio), we now have the necessary tools to construct tests to locate significant shifts in

diversification rate. Consider a pair of sister taxa, L and R, with ℓ and r species, respectively (where $\ell < r$). After calculating the likelihood of realizing a partition of ℓ and r species under both the homogeneous and heterogeneous models, we then calculate the difference in their log likelihoods (i.e., the likelihood ratio of H_A:H_O). The discovery of a large likelihood ratio would provide evidence that L and R diversified under two distinctly different rates, λ_L and λ_R, respectively. We might interpret this as evidence of an increase in diversification rate along the internal branch leading to R (i.e., the *stem* branch subtending the R clade; *sensu* Doyle and Donoghue, 1993; Magallón and Sanderson, 2001). However, this interpretation relies on several assumptions, including the key assumption that the diversity of the more diverse group, R, was achieved stochastically under a constant rate, λ_R (e.g., Rakiow, 1986; Sanderson and Donoghue, 1996)[4]. It is possible that an apparent shift in rate along the branch leading to R could be an artifact of a rate shift that occurred within R. This "trickle-down" problem occurs because a bona fide increase in diversification rate along a given internal branch will exert an influence on diversity comparisons made at more inclusive nodes. Accordingly, a local shift in rate is effectively conducted down the tree, creating the illusion of local rate shifts at neighboring internal branches (see Figure 6).

To discriminate between such illusory and real rate shifts, therefore, we must expand the scope of our evaluation to incorporate information not only from the node subtended by L and R but also from the root node of R. Evaluation of these two hierarchically nested nodes thus entails a three-taxon framework comprising an outgroup clade and the two basal subclades that

[4] The other key assumption concerns the inferred direction of the shift in rate: as two-taxon statements, sister-group comparisons are inherently non-directional (e.g., Jensen, 1990; Doyle and Donoghue, 1993; Sanderson and Bharathan, 1993; Sanderson and Donoghue, 1994, 1996; Sanderson and Wojciechowski, 1996). In other words, the observation that clade R contains significantly more species than its sister group, L, can be explained by postulating either a rate increase in R and/or a rate decrease in L. In principle, increases and decreases in diversification rate are likely to have occurred with equal frequency throughout evolutionary history. Nevertheless, our method ignores shifts associated with significant decreases in diversification rate because the detection of such events on the basis of extant diversity is highly problematic given the associated loss of relevant phylogenetic information. That is, while we do not deny the existence of significant decreases in diversification rate, we are unlikely to detect these events because their occurrence effectively ensures the erasure of the evolutionary history necessary for their discovery. The probability that an entire clade will go extinct is governed by the relative extinction rate, ε, which is simply the extinction rate divided by the speciation rate (e.g., Kendal, 1948; Harris, 1964; Nee *et al.*, 1994b; Magallón and Sanderson, 2001). As ε increases, it becomes increasingly likely that a clade will perish before the present; when $\varepsilon \geq 1$, the probability of complete extinction is one. Evidence from the fossil record suggests that ε has historically been quite high for most groups (e.g., Stanley, 1979; Hulbert, 1993). Recall that a significant decrease in the net diversification rate, λ, entails a significant decrease in speciation rate and/or a significant increase in extinction rate. Such a decrease in λ will therefore cause a corresponding increase in ε, which will greatly increase the probability that the clade will go extinct before the present. Accordingly, if a significant decrease in rate actually occurred in a given group, there would likely be no record of such an event in the relationships among extant species.

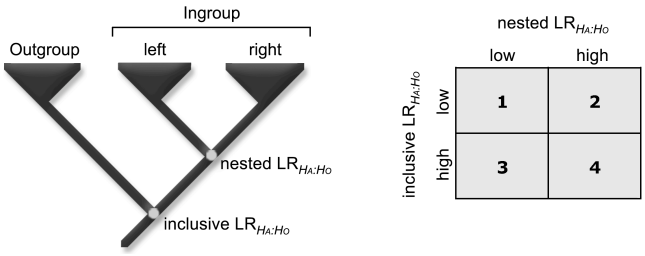

Figure 6. Locating significant shifts in diversification rate in the context of a three-taxon tree. Note that the tree has been rendered in left-light rooting order (Furnas, 1984), such that the more diverse clade is swiveled to the right of every node. Detection of a rate shift along the internal or target branch entails calculation and evaluation of likelihood ratios under the homogeneous and heterogeneous models, $LR_{HA:HO}$, at both the inclusive and nested nodes. The fit of the heterogeneous model to an observed diversity partition at a given node increases with the value of the likelihood ratio. Inspection of the inclusive and nested likelihood ratios entails one of four possible interpretations. Scenarios 1 and 2 indicate that no rate shift occurred along the target branch (although scenario 2 is consistent with a rate shift within the ingroup, which will be assessed as the three-taxon evaluation is iterated up the tree). By contrast, the large likelihood ratios at the inclusive nodes in scenarios 3 and 4 suggest that a rate shift might have occurred along the target branch. In scenario 4, however, the large value of the nested likelihood ratio suggests that rates within the ingroup are significantly heterogeneous. Accordingly, the apparent rate shift along the target branch is likely an artifact of a subsequent rate shift within the ingroup. Thus, scenario 3 represents a bona fide rate shift along the target branch, whereas scenario 4 illustrates the "trickle-down" problem.

together form the ingroup. In outline, the likelihood of a shift along the internal branch of the three-taxon tree (which is based on the likelihood ratio for the observed diversity partition between the outgroup and ingroup clades) must be conditioned by the likelihood of a rate shift within the ingroup (which is based on the likelihood ratio for the observed diversity partition between the left and right ingroup clades). There are many ways one might conceive of conditioning the inclusive likelihood ratio by the nested likelihood ratio, each variant corresponding to a different likelihood ratio–based shift statistic. Indeed, many shift statistics could be imagined that are based on expressions of the data other than their likelihood ratio. In fact, we have developed and experimented with several such alternative shift statistics (see below). Nevertheless, we focus on two shift statistics based on nested likelihood ratios because of their advantageous statistical properties.

The first shift statistic, Δ_1, simply takes the difference in likelihood ratios under the homogeneous and heterogeneous models assessed at the inclusive and nested nodes. It is calculated as

$$\Delta_1 = \left(LR_{H_A:H_O} n_{OG} : n_{IG}\right) - \left(LR_{H_A:H_O} n_{IG_L} : n_{IG_R}\right), \tag{8}$$

where n_i is the number of species in group i, and $LR_{H_A:H_O}n_i:n_j$ is the likelihood ratio of observing a diversity partition $n_i:n_j$ under the homogeneous and heterogeneous models derived using equation (7). The idea is to condition the evidence for a shift at the inclusive node (as reflected by the likelihood ratio of the observed diversity partition between the ingroup and outgroup clades, $n_{OG}:n_{IG}$) by the evidence of a shift at the nested node (as reflected by the likelihood ratio of the observed diversity partition between the left and right ingroup clades, $n_{IG_L}:n_{IG_R}$), thereby reducing the probability of erroneously attributing a local rate shift to the internal branch because of a rate shift within the ingroup clade.

The second shift statistic, Δ_2, is more complicated. Rather than conditioning the inclusive likelihood ratio on the nested likelihood ratio, it attempts to adjust the ingroup diversity used in calculating the inclusive likelihood ratio. The adjusted ingroup diversity excludes the number of ingroup species that can be attributed to a rate increase along the internal branch. This value is calculated as the total ingroup diversity minus the product of the probability of a rate shift at the internal branch, multiplied by the number of species attributable to that shift. The Δ_2 shift statistic is expressed as

(9a) $$\Delta_2 = \left(LR_{H_A:H_O} n_{OG} : n_{IG^*}\right),$$

where

(9b) $$IG^* = n_{IG} - \left(\frac{\left(LR_{H_A:H_O} n_{IG_L} : n_{IG_R}\right)}{\left(LR_{H_A:H_O} n_{IG_L} : n_{IG_R}\right)+1}\right)\left(n_{IG} - \max\left(n_{OG}, 2n_{IG_L}\right)\right).$$

The second term in equation (9b) constrains the adjusted ingroup diversity to assume the larger of two values: the outgroup diversity or two times the diversity of the less diverse (left) ingroup clade. This constraint is imposed to avoid overcorrecting the ingroup diversity in cases for which there is little evidence of a shift along the internal branch. The ERM *P*-values associated with the shift statistics Δ_1 and Δ_2 are assessed by numerical analysis: the cumulative probability of obtaining a shift statistic value as or more extreme than that derived for the observed tree (using equation (8) or (9)) is calculated using the statistic value for the observed topology and the known probabilities of different topologies under the ERM model.

4.2 Assessing the statistical behavior of the shift statistics

We performed a simulation study to explore the behavior of several shift statistics using a simple experimental design in which a rate shift was applied to either the inclusive and/or nested node of a three-taxon tree. The power and bias of the various shift statistics were assessed by their respective abilities to correctly or incorrectly reject the null hypothesis that no rate shift occurred along the internal (target) branch. Specifically, trees were generated under an ERM branching process initiated from a single species with the branching rate parameter, λ, set to 1. During the growth of a simulated tree, a diversification rate shift of a specified magnitude A, where $A \in \{2, 4, 6\}$, occurred deterministically under three different treatments: 1) a shift was applied to the inclusive node (i.e., occurring immediately after the first branching event); 2) a shift was applied to the nested node (i.e., occurring immediately after the first branching event within the ingroup); or 3) a shift was applied both to the inclusive and nested nodes (i.e., occurring immediately after the first and second branching events). The process was terminated when the trees reached the desired size, N, where $N \in \{100, 200, 400\}$. Every permutation of the set of simulation parameters (tree size, magnitude of rate shifts, and location of rate shifts) was replicated 10 000 times, calculating the value for each of seven shift statistics for each tree generated from each replicate. Power and Type I error rates were calculated as the proportion of the replicates in which the null hypothesis of no rate shift along the target branch was correctly or incorrectly rejected, respectively.

We compared the performance of our two likelihood ratio statistics, Δ_1 and Δ_2, to one existing and four other new shift statistics:

1. NP, the ERM nodal probability proposed by Slowinski and Guyer (1989a, b) was calculated for the inclusive node using equation (1).
2. Δ_N, calculated as the difference in "raw" diversity contrasts at the inclusive and nested nodes; that is, $\Delta_N = ((n_{IG} - n_{OG}) - (n_{IGR} - n_{IGL}))$.
3. Δ_R, calculated as the difference in diversification rate contrasts at the inclusive and nested nodes, where the maximum likelihood estimates of diversification rates are calculated as $\hat{\lambda} = (\ln(n))^{-1}$ (Sanderson and Donoghue, 1996); accordingly, $\Delta_R = ((\hat{\lambda}_{IG} - \hat{\lambda}_{OG}) - (\hat{\lambda}_{IGR} - \hat{\lambda}_{IGL}))$.
4. Δ_{NP}, calculated as the difference in the two ERM nodal probabilities calculated at the inclusive and nested nodes, which is somewhat similar to the procedure outlined by Nee and Harvey (1994; see also Nee *et al.*, 1996; Mayhew, 2002).
5. $\Delta_{1*\omega}$, calculated as for Δ_1, but incorporates a scaling parameter, ω, that weights the contribution of the nested likelihood ratio (the second term in

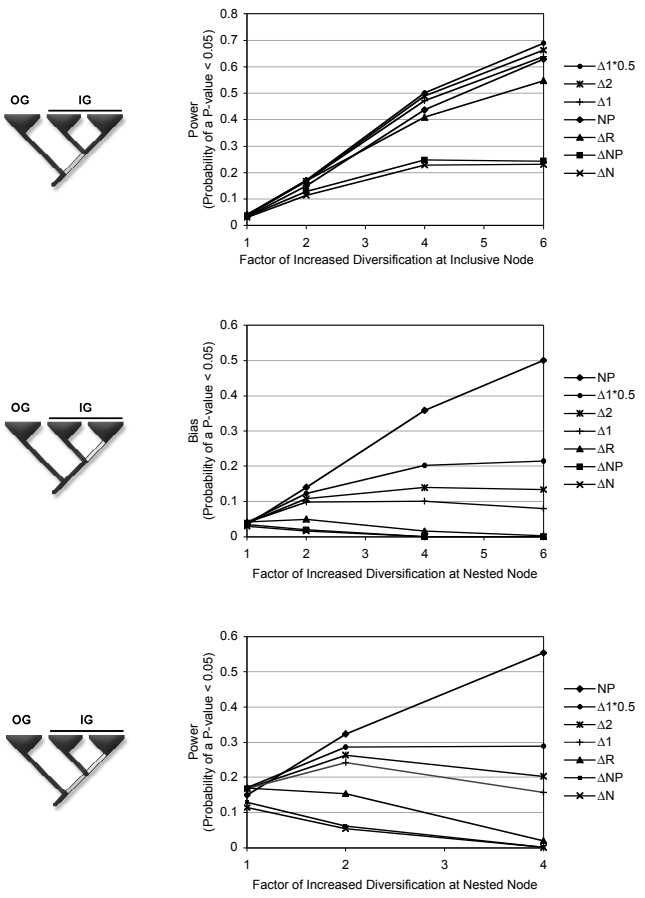

Figure 7. Results of a simulation study of the relative power and bias of several shift statistics in the three-taxon case. The locations of diversification rate shifts in the three-taxon trees are indicated as shaded branches (see text for details).

equation (8)) to the shift statistic; accordingly, ω effectively indexes an infinite array of shift statistics, where $\Delta_{1*\omega}$ is identical to Δ_1 when $\omega = 1$).

The upper graph in Figure 7 illustrates the ability of the shift statistics to detect diversification rate shifts of various magnitudes along the internal branch. These conditions correspond to the idealized case in which potentially confounding diversification rate shifts elsewhere in the tree have not occurred. The plots for each statistic intersect the *y*-axis at $P \cong 0.05$ (the nominal level of α) where the diversification rate is raised by a factor of 1, indicating appropriate Type I error rates when the null hypothesis is true (as

expected under Monte Carlo simulation). For 100-species trees, the likelihood ratio–based shift statistics exhibit the greatest relative power, successfully detecting a four-fold rate increase in ~50% of the replicates, and a six-fold rate increase in ~65% of the replicates. The $\Delta_{1*0.5}$ statistic slightly outperforms Δ_2, which in turn slightly outperforms Δ_1. All three likelihood ratio–based shift statistics enjoy an edge in power over Slowinski and Guyer's (1989a, b) *NP*, which is expected since these Δ statistics possess greater resolution by virtue of incorporating more information. The other shift statistics, Δ_R, Δ_{NP}, and Δ_N, exhibit substantially lower power.

The middle graph in Figure 7 illustrates the bias of the various shift statistics associated with a diversification rate shift of various magnitudes within the ingroup (specifically, along the branch subtending the right ingroup clade). This simulation therefore assesses the relative sensitivity of the various shift statistics to the trickle-down problem. Because no rate increase occurs along the target branch, a completely unbiased statistic should exhibit a flat probability of rejecting the null hypothesis of ~0.05. As expected, *NP* is extremely biased, rejecting the null hypothesis almost as frequently as when a rate increase actually occurred at the target node (compare the plots for *NP* in the upper and middle graphs). The likelihood ratio–based shift statistics fare substantially better, exhibiting Type I error rates ranging between 10–20% under a four-fold rate increase within the ingroup and between 8–21% under a six-fold rate increase within the ingroup. Not surprisingly, the Type I error rates of the three likelihood ratio–based shift statistics mirror their relative power in the upper graph. Accordingly, the slight edge in power exhibited by $\Delta_{1*0.5}$ translates into greater bias. The other two likelihood ratio–based shift statistics exhibited relatively low bias, with Δ_1 consistently outperforming Δ_2. The remaining shift statistics, Δ_R, Δ_{NP}, and Δ_N, are substantially more conservative.

The lower graph in Figure 7 illustrates the ability of the shift statistics to detect a doubling in diversification rate along the internal (target) branch given a subsequent rate shift of varying magnitude within the ingroup clade. This simulation therefore assesses the power of the shift statistics in cases where the trickle-down problem applies. Because a doubling in rate is uniformly applied to the internal branch, the plot for each shift statistic intersects the *y*-axis at the ordinate value corresponding to its respective power under a two-fold rate increase in the upper graph. Note that, because a two-fold rate increase is consistently applied to the target node, a perfectly unbiased shift statistic would exhibit a flat power curve over the range of rate increases applied within the ingroup clade. Not surprisingly, *NP* exhibits the highest power under this scenario because shifts within the ingroup contribute to rejection of the null hypothesis; that is, it does a "good" job, albeit for the wrong reasons. The power plots for the likelihood ratio–based

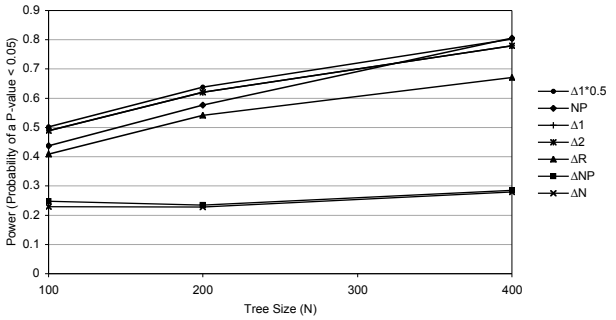

Figure 8. The effect of tree size on the power of several shift statistics to detect significant diversification rate shifts (see text for details).

shift statistics are substantially flatter, their rank order remaining unchanged: $\Delta_1 < \Delta_2 < \Delta_{1*0.5}$. Because none of these Δ shift statistics perfectly condition the inclusive likelihood ratio by the nested likelihood ratio, their power might be inflated slightly by a rate shift within the ingroup clade. Under these conditions, the Δ_1 Δ_2, and $\Delta_{1*0.5}$ statistics appear to slightly undercondition the inclusive by the nested likelihood ratio. This behavior can be seen by comparing the plots of $\Delta_{1*0.5}$ and Δ_1. Because $\Delta_{1*0.5}$ applies a relatively small penalty to the inclusive likelihood ratio when a rate shift occurs at the nested node, its power is consequently more inflated than that of Δ_1. Interestingly, the bias of Δ_1 and Δ_2 appears to decrease as the magnitude of the rate shift at the nested node increases. The remaining shift statistics, Δ_R, Δ_{NP}, and Δ_N, appear to overcompensate for rate shifts within the ingroup, such that their power to detect a rate shift at the target node rapidly diminishes with increasing magnitude of rate shifts at the nested node.

The performance of the various shift statistics under a range of tree sizes is illustrated in Figure 8. These simulations uniformly applied a four-fold diversification rate shift to the internal branch of trees with 100 to 400 tips, a size range reflecting that of supertrees in the literature (e.g., Purvis, 1995; Bininda-Emonds *et al.*, 1999; Wojciechowski *et al.* 2000; Jones *et al.*, 2002; Kennedy and Page, 2002; Salamin *et al.*, 2002; Stoner *et al.*, 2003). Although the power of the shift statistics generally scale with tree size, the increase in power was not realized uniformly by the various tests. The likelihood ratio–based shift statistics, Δ_1, Δ_2, and $\Delta_{1*0.5}$, exhibited the greatest proportional increase in power as tree size increased; by contrast, the power of the Δ_N and Δ_{NP} statistics was essentially flat across the range of tree sizes simulated, whereas Δ_R exhibited an intermediate increase in relative power. Under the range of tree sizes evaluated, the likelihood ratio–based shift statistics consistently exhibited the greatest (and quite similar) absolute

power, accurately identifying a four-fold diversification rate shift ~60% and ~80% of the time in trees with 200 and 400 tips, respectively. Interestingly, although the bias of the $\Delta_{1*0.5}$ shift statistic increased with tree size, the Type I error rates for the Δ_1 and Δ_2 statistics decreased slightly with tree size (not shown). The results discussed above (and illustrated in Figures 7 and 8) were obtained using the three-taxon simulation design; however, we also performed a more elaborate, whole-tree simulation study that allowed rate shifts of various frequency and magnitude to be applied to all the internal branches of simulated trees of various sizes. Results of this more sophisticated study (not shown) were similar qualitatively to those reported for the simpler investigation.

4.3 Locating diversification rate shifts in primates

We used the Δ shift statistics to locate significant diversification rate shifts in the primate supertree published by Purvis (1995). As in the previous analysis of diversification rate variation using the whole-tree statistics, this analysis is intended for illustrative and comparative purposes only. Accordingly, we made no attempt to account for the effects of phylogenetic error. Results for the two likelihood ratio–based shift statistics, Δ_1 and Δ_2, were obtained using the SYMMETREE program. Polytomies were treated by generating 1000 random resolutions using the size-sensitive ERM taxon-addition algorithm, providing an estimate of the confidence intervals for *P*-values associated with each shift statistic. As before, analyses were performed on both the entire primate tree and several of its component clades (e.g., strepsirhines, New World monkeys, colobines, cercopithecines, and hominoids). After summarizing the findings of our analysis, we describe the methods used in a previous investigation of diversification rate shifts in the primate tree (Purvis *et al*., 1995), comparing and contrasting the results obtained by these two studies.

Our analysis using the Δ_1 and Δ_2 statistics detected seven diversification rate shifts in the primate tree. Significant rate shifts (Figure 9; bold black branches) were detected at the base of haplorhines (along branch 1, the root of a clade comprising New World monkeys, Old World monkeys, and hominoids), within New World monkeys (along branch 4, the root of a clade comprising *Callithrix*, *Cebuella*, *Leontopithecus*, and *Saguinus*), and within Old World monkeys (along branch 7, the root of a clade comprising all *Presbytis* species except *P. entellus*). Additionally, several marginally significant rate shifts (Figure 9; bold gray branches) were detected, including two shifts within the Old World monkey clade (the first along branch 5, the root of a clade comprising *Macaca arctoides*, *M. assamensis*, *M. cyclopis*,

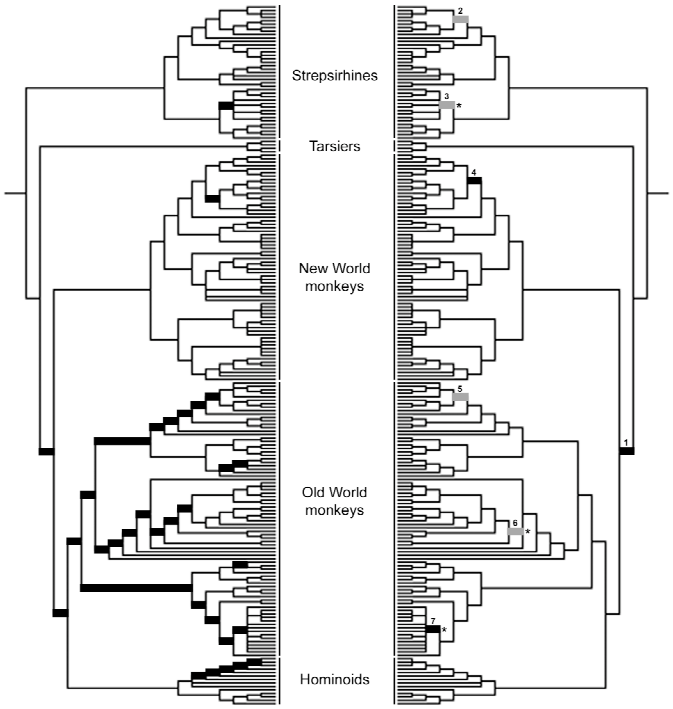

Figure 9. Location of inferred diversification rate shifts in the primate supertree of Purvis (1995). The tree at left depicts results from a previous study by Purvis *et al.* (1995) that identified diversification-rate shifts using the relative-cladogenesis statistic (to identify anomalously diverse lineages) coupled with a parsimony optimization scheme: all 32 diversification-rate shifts are shown, including 23 from the simultaneous analysis of the entire tree and an additional nine non-redundant shifts from the analyses of the five component clades. The tree at right depicts diversification rate shifts identified using the Δ shift statistics. Results obtained under the two approaches are somewhat correspondent: five of the diversification rate shifts identified by the Δ shift statistics are among those identified in the previous study. However, several nested shifts within Old World monkeys (i.e., those involving *Macaca*, *Cercopithecus*, *Presbytis* at branches 5, 6, and 7, respectively) caused a cascade of spurious diversification-rate shifts to be identified at more inclusive nodes throughout the anthropoid clade owing to the trickle-down problem. Bold black branches correspond to significant rate shifts and bold gray branches to marginally significant rate shifts; branches marked with an asterisk identify results involving random resolution of polytomies; numbered branches correspond to clades identified in the text.

M. fascicularis, *M. fuscata*, *M. mulatta*, *M. radiata*, *M. sinica*, and *M. thibetana*; and the second shift along branch 6, the root of a clade comprising all *Cercopithecus* species except *C. aethiops* and *C. solatus*) and two within strepsirhines (the first along branch 2, the root of a clade

comprising *Lemur*, *Hapalemur*, and *Eulemur*; and the second shift along branch 3, the root of a clade comprising *Galago*, *Galagoides*, *Otolemur*, and *Euoticus*). Interestingly, the diversification rate shift along branch 2 was independently identified as a significant radiation in a recent study by Yoder and Yang (in press), which estimated divergence times from several unlinked loci and external fossil calibrations using Bayesian methods.

Several aspects of these findings warrant comment. First, three of the diversification rate shifts were associated with polytomies (i.e., those in *Macaca*, *Presbytis*, and the *Callithrix-Saguinus* clades; Figure 9), demonstrating the applicability of these methods to incompletely resolved (super)trees. Second, in contrast to our previous analysis of diversification rate variation using the whole-tree *M* statistics, results obtained using the Δ shift statistics were insensitive to the specification of taxonomic scope, returning the same *P*-values for the same set of branches regardless of whether the analysis was applied simultaneously to the entire tree or separately to its component clades. Third, the results of the whole-tree *M* statistics and Δ shift statistics are not perfectly correspondent. Specifically, the Δ shift statistics failed to locate significant diversification rate shifts within several clades in which significant among-lineage diversification rate variation had previously been identified by the whole-tree *M* statistics. In these cases, diversification rate variation appears to be rather evenly dispersed across the tree such that, although cumulatively significant under the whole-tree *M* statistics, it is nevertheless insufficiently concentrated along any one branch (or small number of branches) to constitute a significant diversification rate shift under the Δ shift statistics. For example, the topology of the hominoid clade is largely pectinate, indicating significant heterogeneity in diversification rate among its branches. Nevertheless, evaluating the probability of a diversification rate shift along any particular branch is likely to involve a diversity partition of $1:(N-1)$ at the inclusive node and $1:(N-2)$ at the nested node, which is much more consistent with a trickle down in rates than a local shift in rate under the Δ shift statistics. Finally and conversely, significant diversification rate shifts were located within clades for which the whole-tree statistics had previously failed to detect significant among-lineage diversification rate variation. In these cases, diversification rate heterogeneity was largely restricted to a single branch (or small number of branches), constituting a significant local rate shift that was nevertheless below the threshold of detection under the whole-tree *M* statistics. For example, the New World monkey clade is, overall, very balanced: diversity partitions at most nodes in this tree involve splits of approximately $(N/2):(N/2)$. The single prominent exception involves the node at which a rate shift was located (branch 4 in Figure 9), which by itself

was insufficient to cause rejection of the null hypothesis that the whole New World monkey clade diversified under a stochastic ERM branching model.

4.3.1 The relative cladogenesis statistic: potential limitations and comparison to the Δ shift statistics

The location of diversification rate shifts in the primate supertree was previously studied by Purvis *et al.* (1995) using an approach referred to as the "relative cladogenesis statistic," originally described in Nee *et al.* (1992, 1994b, 1994b) and subsequently in Harvey and Nee (1993, 1994), Nee and Harvey (1994), and Nee *et al.* (1994a, 1995, 1996). Like the whole-tree statistics described previously, the relative cladogenesis statistic was originally intended to detect significant diversification rate variation among a set of lineages. In contrast to our strictly topology-based whole-tree statistics, however, the relative cladogenesis statistic relies on temporal information to circumscribe the set of lineages involved in the test. That is, given a phylogeny with estimated divergence times, we can arbitrarily draw a line through the tree at some point in the past, t_k, to identify a set of k contemporary ancestral lineages. Suppose that these k ancestral lineages survive to the present and give rise collectively to N extant descendants, such that the ith ancestral lineage leaves n_i extant species, where $n_i \geq 1$(because all k ancestral lineages have survived) and where the n_i sum to N. If the k lineages all diversified at the same rate, then all vectors of descendant species diversities (n_1, n_2, n_3, …, n_k) are equiprobable[5] (e.g., Nee *et al.*, 1992, 1994b, 1996; Nee and Harvey, 1994; Purvis *et al.*, 1995; Purvis, 1996). This expectation can be used to calculate the probability that one of the ancestral lineages will realize more than r descendants, given a total of N species descended from the set of k ancestral lineages. This probability is given by

$$P = 1 - \frac{\sum_{v=0} (-1)^v \binom{k}{v} \binom{N - rv - 1}{k - 1}}{\binom{N-1}{k-1}}, \tag{10}$$

[5] Curiously, it is often asserted that the relative cladogenesis test "makes no assumptions about how the clades have been growing" (Nee and Harvey, 1994:1550) and that it "does not depend on any particular model of diversification" (Nee *et al.*, 1996:241; see also Nee and Harvey, 1994; Purvis *et al.*, 1995). Clearly, however, the assumptions entailed by this test—that rates of diversification are equal and independent in all lineages at any given point in time—are those specifying the stochastic ERM random branching model. In fact, equation (12) reduces to equation (1) (which provides the ERM nodal probability) when $k = 2$ (i.e., for sister-group comparisons where N descendant species are partitioned among two ancestral sister lineages; e.g., Nee and Harvey, 1994; Nee *et al.*, 1994a, 1995; 1996; Purvis, 1996).

where the summation is for positive $N - rv - 1$ and where $N - rv - 1 \geq k - 1$ (Purvis *et al.*, 1995; Nee *et al.*, 1996). A significant result indicates that the clade in question is anomalously diverse and therefore has diversified under a significantly different rate than its contemporaries.

Although originally intended as a test of significant diversification rate variation, the relative cladogenesis statistic was subsequently extended to infer the location of significant diversification rate shifts by Purvis *et al.* (1995). This extension is based on parsimony optimization: if two sister lineages are inferred to be anomalously diverse under the relative cladogenesis statistic, then a significant shift in rate is inferred to have occurred in their common ancestor. Below we consider several potential limitations associated with the attempt to use the relative cladogenesis statistic to locate diversification rate shifts: some limitations are inherent to the method, others pertain more specifically to divergence-time estimates in supertrees. These limitations are illustrated with reference to the analysis of diversification rate shifts in the primate supertree, and compared with the behavior of the Δ shift statistics where appropriate.

4.3.1.1 Susceptibility of the relative cladogenesis statistic to arbitrary delineation of test window

As described above, the relative cladogenesis statistic requires delineation of a "window" within which the test is to be applied. The dimensions of this window include both its temporal depth and its taxonomic breadth. However, circumscription of this window is arbitrary and therefore potentially problematic given that the results inferred from the test are known to be sensitive to the temporal depth (Purvis, 1996) and taxonomic scope specified. Specification of the temporal dimension can be made less arbitrary by sliding the window over the tree from the root to its tips, recalculating the relative cladogenesis statistic at every point in time, t_k, associated with an increase in k, the number of the ancestral lineages (where $k = 2, 3, 4, \ldots, (N - 1)$). This approach was used by Purvis *et al.* (1995) and has also been implemented in the (now defunct) End-Epi program (Harvey *et al.*, 1996; Rambaut *et al.*, 1997). However, it is considerably more difficult to objectively define (or integrate over) the taxonomic breadth of the comparison, which nevertheless exerts a similarly strong influence on the conclusions obtained. Although the sensitivity of the relative cladogenesis statistic to phylogenetic scope is appropriate when the test is used to detect diversification rate variation, this sample dependency is inappropriate when applied to the problem of locating diversification rate shifts. This problem is manifest in the analysis of diversification rate shifts in the primate supertree: Purvis *et al.* (1995) originally detected 23 significant rate shifts when the

relative cladogenesis statistic was applied to the entire primate tree but subsequently identified an additional nine non-redundant rate shifts when the various component clades were analyzed separately. By contrast, the number of branches identified by the Δ shift statistics (and the *P*-values of the statistics) were unaffected by the phylogenetic scope of the analysis.

4.3.1.2 Susceptibility of the relative cladogenesis statistic to error in divergence times

Application of the relative cladogenesis test requires reliable estimates of divergence times, which is likely to be problematic for the analysis of supertrees. Error in divergence-time estimates can confound the test by causing misspecification of the appropriate set of ancestral lineages present at the specified t_k. Although recent methodological and theoretical advances have greatly improved the accuracy of divergence-time estimates derived from the primary analysis of nucleotide sequence data (e.g., Sanderson, 1997, 2002; Rambaut and Bromham, 1998; Thorne *et al.*, 1998; Huelsenbeck *et al.*, 2000a; Yoder and Yang, 2000; Kishino *et al.*, 2001; Thorne and Kishino, 2002), the extent to which these methods can be extended to the estimation of divergence times in supertrees is presently unknown. Close inspection of the primate phylogeny illustrates some of the challenges of estimating divergence times in supertrees, as well as the undesirable consequences of the associated error for inferences of diversification rates that rely on temporal information. We wish to emphasize, however, that our criticisms are not intended to imply that the dates in this particular supertree are exceptionally unreliable; rather, we believe that the level of uncertainty in these data is similar to that in other published supertrees.

Divergence times were estimated for 90 of the 160 nodes in the primate supertree, all of which were derived directly from or calibrated against the primate fossil record. Under the approach used, the divergence time of a clade was equated with the age of the oldest fossil attributed to that lineage. This approach will tend to systematically underestimate the true divergence times of clades in proportion to their degree of incompleteness in the fossil record. For several reasons, the degree to which a lineage is represented in the fossil record is likely to be phylogenetically biased. For example, preservation potential will be influenced by phylogenetically autocorrelated differences in anatomy and demography, and taphonomic factors will be influenced by phylogenetically autocorrelated differences in habitat preference. Consequently, clades will vary in the degree to which their inferred divergence times will be underestimated. The resulting phylogenetically biased error in divergence-time (under)estimation will

induce a corresponding pathological bias for the study of diversification rates: an underestimate in the age of a clade will cause a corresponding overestimate in its inferred rate of diversification.

As expected of a group with a heterogeneous representation in the fossil record, the number of available fossil calibrations varied markedly across the primate supertree: 15 estimates were used to date one node, whereas the divergence times of many others were based on a single estimate. The uncertainty associated with divergence times based on single estimates was approximately and conservatively estimated to have an average error margin >±50%, prompting Purvis (1995:413) to reasonably conclude that "not too much reliance should be placed on single estimates." Nevertheless, several diversification rate shifts in the primate supertree relied on the single date estimates. For example, Purvis *et al.* (1995:331) were appropriately skeptical of the inferred diversification rate shift in the strepsirhine clade because "the age of the galagid radiation is based on only a single estimate, so it may be inaccurate." However, this caveat applies equally to several other clades in which diversification rate shifts were detected (e.g., *Cercopithecus*, *Colobus*, *Macaca*, *Presbytis*, *Saguinus*) because they were similarly based on a single (or very few) estimates.

Moreover, several of the nodes based on single (or very few) estimates were used to calibrate other nodes in the primate supertree, causing a cascade of error in both estimation of divergence times and the associated inference of diversification rate shifts. For example, the divergence time of Old World monkey-hominoid clade was estimated by Purvis (1995) at 27.5 ± 4.5 million years ago (Mya) based on two fossils. Independent estimates for the age of this node are typically much older. For example, maximum likelihood estimates based on the entire protein-coding region of the mitochondrial genome calibrated with a more reliable external fossil date (the cetacean-artiodactyl divergence at 53–60 Mya) place this divergence in the range of ~38–68 Mya (Arnason *et al.*, 1998; Yoder and Yang, 2000). The discrepancy in the timing of this divergence is somewhat troubling because it was used to calibrate 32 other nodes within the Old World monkey-hominoid clade (A. Purvis, pers. comm.), in which 28 of the 32 total significant diversification rate shifts were detected by Purvis *et al.* (1995).

Uncertainty in divergence times is not restricted to those nodes based on single estimates: dates in the primate supertree based on multiple estimates also had non-trivial error. For example, Purvis (1995:413) reported significant differences in the proportional error in divergence-time estimates among clades in the primate supertree, which would be expected of a group with phylogenetically biased representation in the fossil record. The highest proportional error was found within cercopithecines, in which fully half of the inferred diversification rate shifts occurred. The extent to which the

acknowledged uncertainty in divergence-time estimates influenced this study of diversification rate shifts in the primate supertree is difficult to ascertain; Purvis *et al*. (1995) acknowledged the presence of error in the divergence times and its potential impact on the analysis but did not attempt to quantify the level of uncertainty or assess the sensitivity of the results to this source of error.

Although the divergence times for the 90 dated nodes are likely to be associated with substantial estimation error, the divergence times for the remaining 70 nodes were not estimated at all, but instead generated deterministically under the assumption of what might be called a "branching clock" (*sensu* Sanmartín *et al.*, 2001). Given a deterministic model of exponential diversification, the divergence time of a given node can be calculated as $t_d = (t_a)(\ln N_d / \ln N_a)$, where t_a and t_d are the ages of the ancestral and descendant nodes, with N_a and N_d species, respectively. Given the countless number of hidden parameters influencing diversification rates, the use of a deterministic branching model (particularly one whose fit to the data is not evaluated) is likely to provide an overly simplistic and potentially problematic solution to the problem of specifying the unknown divergence times. The use of a branching model to specify >40% of the divergence times in the primate supertree is likely to bias inferences of diversification rates (Purvis *et al.*, 1995). Many approaches (including the relative cladogenesis test) invoke stochastic branching models to generate the expected distribution of diversification events against which the observed distribution is compared. However, use of a branching clock essentially involves the model-based generation of the "observations" as well. Although it is difficult to ascertain the accuracy of dates generated with this scheme, there is no reason to expect it to be high: these divergence times combine the considerable uncertainty of those estimated from fossil evidence and/or local clocks (from which they are ultimately calibrated) with a branching clock of uncertain justification.

In summary, scrutiny of the primate supertree highlights the challenges of estimating divergence times in supertrees and reveals how the uncertainty in these data can confound attempts to detect diversification rate shifts using the relative cladogenesis statistic or other temporal tests. Although there is reason for optimism that recent efforts will improve the reliability of divergence-time estimates in supertrees (e.g., Lapointe and Cucumel, 1997; Bryant *et al*., 2004; Lapointe and Levasseur, 2004; Vos and Mooers, 2004), the ability of these methods to provide sufficiently accurate temporal information has yet to be demonstrated. By contrast, because they effectively ignore temporal information, the topology-based Δ shift statistics provide a more reliable means with which to infer diversification rate shifts in supertrees.

4.3.1.3 Susceptibility of the relative cladogenesis statistic to the trickle-down problem

In addition to pioneering the development of methods for locating significant diversification rate shifts, Purvis *et al.* (1995:331) were also among the first authors to recognize the potentially confounding influence of what we have termed the trickle-down problem, raising the caveat that any "result must be interpreted cautiously because radiations are not independent: if a given clade is a significant radiation, more inclusive clades will tend to be." In other words, significant diversification rate shifts at more nested nodes will lead to the identification of spurious diversification rate shifts at more inclusive nodes under their proposed parsimony optimization scheme. Results from the primate analysis provide compelling empirical evidence of the susceptibility of the relative cladogenesis statistic to the trickle-down problem. Despite the aforementioned criticisms, the relative cladogenesis statistic identified several diversification rate shifts also indicated by the likelihood ratio–based Δ shift statistics (e.g., shifts located within *Galago*, *Macaca*, *Cercopithecus*, and *Presbytis* at branches 3, 5, 6, and 7, respectively; Figure 9). However, diversification rate shifts detected at relatively nested nodes within the Old World monkey clade (those within *Macaca*, *Cercopithecus*, and *Presbytis* at branches 5, 6, and 7, respectively; Figure 9) caused a trickle-down of diversification rate shifts to be inferred at more inclusive nodes under the relative cladogenesis test. Accordingly, the demonstrable susceptibility of the relative cladogenesis statistic to the trickle-down problem suggests that this test is more appropriately restricted to the inference of diversification rate variation.

5. Discussion

5.1 Implementation and accommodation of phylogenetic uncertainty

The methods described in this chapter have been implemented in the computer program, SYMMETREE. Executables have been compiled for Macintosh (OS 9 and OS X), Windows, and UNIX operating systems, which are freely available at http://www.phylodiversity.net/brian/ or http://www.kchan.org, or by emailing either of these authors directly.

Methods for detecting diversification rate variation have typically required strictly dichotomous phylogenies; given the empirical reality of polytomies, this limitation has proven to be a serious impediment to their

application. Accordingly, an important feature of SYMMETREE is its facility to deal with incompletely resolved trees. The program recognizes two types of soft polytomies that require different analytical approaches: "collapsed" polytomies, which are caused by internal branches of zero length; and, "consensus" polytomies, which stem from conflict among a set of equally optimal (super)tree estimates. Collapsed polytomies are addressed by randomly (and repeatedly) generating dichotomous solutions using one of several alternative random taxon-addition algorithms. However, this procedure is inappropriate for consensus polytomies. Although such a polytomy might stem from conflict among a small set of source trees, it might nevertheless be consistent with a much larger set of (randomly resolved) binary trees. Accordingly, only those resolutions of a consensus polytomy that belong to the set of conflicting trees should be considered, which can be accomplished by means of a batch-processing option that sequentially analyzes each tree belonging to the set of conflicting trees. For both collapsed and consensus polytomies, the appropriate test can be applied to each tree within the set of (randomly resolved or equally optimal) trees to provide an estimate of the confidence intervals on the inference being made.

More generally, polytomies can be viewed as a manifestation of phylogenetic uncertainty. Although often acknowledged as a crucial assumption, the effect of phylogenetic error on inferences of diversification rate is seldom explicitly taken into account (but see, for example, Sanderson and Wojciechowski, 1996; Baldwin and Sanderson, 1998). In theory, it would be straightforward to assess the confidence interval on an inference by batch processing the bootstrap profile (and/or the posterior probability distribution) of study trees. Although this approach is viable for trees derived from primary analyses (i.e., conventional analysis of the primary character data), supertree estimation methods present a special challenge in this respect because there is currently no comparable means of estimating topological uncertainty in supertrees. Clearly, this area requires further development (Ronquist *et al*., 2004; Moore *et al.*, in prep.).

5.2 Extensions, limitations, and applications

The methods described in this chapter are intended to answer two general questions. Have the branches of this tree experienced differential diversification rates? And, if so, on which branches have those shifts in rate occurred? Accordingly, these methods should find useful application to a range of problems (outlined below) but, of course, will be ill-suited to the investigation of other equally valid and interesting evolutionary questions. For example, we might want to estimate parameters associated with the diversification process (e.g., speciation and extinction rates) or test whether

diversification rates have changed significantly through time. These questions require information on the (relative or absolute) timing of diversification events, and so will necessarily involve the use of temporal methods (e.g., Harvey *et al.*, 1991, 1994a, b; Nee *et al.*, 1992, 1994a, b; Harvey and Nee, 1993, 1994; Kubo and Iwasa, 1995; Paradis, 1997, 1998b; Nee, 2001). In such applications of temporal methods, however, it is first necessary to establish that there has not been significant among-lineage diversification rate variation within the study phylogeny. This requirement can readily be established (or disconfirmed) with our whole-tree tests for diversification rate variation, again emphasizing the inherent complementarity of temporal- and topology-based methods.

Other questions might be profitably addressed by extending the whole-tree methods described herein. For example, we might want to know if shifts in diversification rate are correlated with changes in some other variable (e.g., the origin of morphological or behavioral novelties, ecological associations, or biogeographic events). Topology-based approaches to this problem are available but typically involve replicated sister-group comparisons (e.g., Slowinski and Guyer, 1993; Nee *et al.*, 1996; Barraclough *et al.*, 1998; Goudet, 1999; Simms and McConway, 2003) that incorporate relatively limited phylogenetic information (e.g., Sanderson and Donoghue, 1994, 1996). As has been demonstrated for other diversification rate problems, the power to detect correlates of shifts in diversification rate is likely to be substantially enhanced by incorporating information from more of the tree. We are currently working to extend the methods described in this chapter to provide a whole-tree approach to this problem.

In addition to addressing other types of questions, the whole-tree methods described in this chapter might also be profitably extended to incorporate additional sources of information. Although the whole-tree methods currently utilize exclusively topological information on the distribution of species diversity, they could readily be generalized to incorporate temporal information on the distribution of waiting times between diversification events. It is conceivable that the inclusion of divergence-time estimates, when available and appropriate to the hypothesis of interest, could further enhance the power to detect the presence and locate the position of significant shifts in diversification rate.

Future elaborations notwithstanding, the whole-tree methods presented here have immediate implications for a range of data-exploration and hypothesis-testing scenarios associated with the study of diversification rates. Whole-tree surveys for significant diversification rate variation could provide an effective discovery method for generating causal hypotheses of factors that have caused, are caused by, or are correlated with differential diversification rates. For example, the discovery that diversification rate

variation is often associated with plant clades that are polymorphic for growth form (i.e., woody / herbaceous) might lead us to hypothesize that shifts in growth form are affecting diversification rates (e.g., Eriksson and Bremer, 1991, 1992; Bremer and Eriksson, 1992; Judd *et al.*, 1994; Ricklefs and Renner, 1994; Tiffney and Mazer, 1995; Dodd *et al.*, 1999). The ability to detect clades with significant diversification rate variation and/or diversification rate shifts will also help identify the data relevant to studies of phenomena that are hypothesized to be correlated with differential diversification rates. For example, application of the whole-tree tests could identify the data necessary to evaluate the hypothesized correlation between rates of nucleotide substitution and rates of cladogenesis (e.g., Mindell *et al.*, 1989; Barraclough *et al.*, 1996; Savolainen and Goudet, 1998; Barraclough and Savolainen, 2001; Jobson and Albert, 2002). Additionally, the whole-tree tests could provide more powerful tools for studies that seek to assess the empirical prevalence of diversification rate variation (e.g., Guyer and Slowinski, 1991, 1993; Heard, 1992; Mooers, 1995). Finally, several evolutionary processes could entail hypotheses that predict multiple diversification rate shifts dispersed throughout whole clades, rather than single shifts concentrated at particular nodes. These processes include the effect of various co-evolutionary associations on rates of diversification (e.g., the reciprocal radiations predicted for some insect / plant associations; Farrell, 1998; Farrell and Mitter, 1998; Kelly and Farrell, 1998) and the effect of relative refractory periods associated with "age-biased cladogenesis" (Hey 1992; Harvey and Nee, 1993; Losos and Adler, 1995; Chan and Moore, 1999).

The foregoing discussion suggests that several different evolutionary questions — associated with detecting significant diversification rate variation or locating diversification rate shifts — might be effectively addressed by the separate application of either the whole-tree *M* statistics or the Δ shift statistics, respectively. However, both sets of methods could be applied in concert to address additional evolutionary questions. For instance, the combined application of the *M* and Δ statistics might be used to explore the empirical prevalence of different models of cladogenesis (Figure 10). As demonstrated in the primate analyses, results obtained under the whole-tree *M* statistics and the Δ shift statistics will not always be perfectly correspondent. That is, the whole-tree *M* statistics might occasionally detect significant diversification rate variation within clades for which the Δ shift statistics subsequently fail to locate any significant diversification rate shifts. Conversely, significant diversification rate shifts might sometimes be identified within clades for which the *M* statistics fail to detect significant among-lineage diversification rate variation. The former scenario will arise when diversification rate variation is rather evenly dispersed across the tree

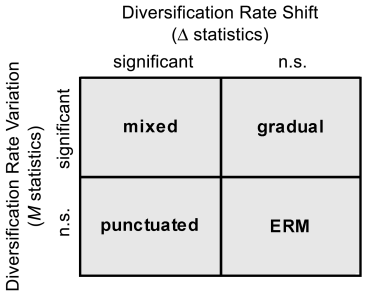

Figure 10. Exploiting discord in results obtained under the whole-tree *M* statistics and the Δ shift statistics to explore modes of diversification. The lower right cell indicates stochastically homogeneous (ERM) diversification rates, whereas the other three scenarios involve diversification-rate heterogeneity consistent with either gradual, punctuated, or mixed evolutionary models of cladogenesis.

(see Section 4.3); such a relatively uniform phylogenetic distribution of diversification rate change is consistent with a gradual evolutionary model of cladogenesis. By contrast, the latter scenario entails a local concentration of diversification rate heterogeneity along a single branch (or small number of branches) that is below the threshold of detection under the whole-tree *M* statistics. This relatively sporadic phylogenetic distribution of diversification rate change is consistent with a punctuated evolutionary model of cladogenesis. Thus, discord in the results obtained under the whole-tree *M* statistics or the Δ shift statistics can be usefully exploited to tease apart modes of diversification rate heterogeneity.

In conclusion, we are optimistic that the methods described in this chapter should enable a range of evolutionary questions to be addressed when reliable temporal information is either unavailable or inappropriate to the problem at hand.

Acknowledgements

We are grateful to Olaf Bininda-Emonds for inviting us to contribute to this volume on supertrees, and for his Buddha-like patience while enduring its elephantine gestation. We are indebted particularly to Mary Moore for assistance with the illustrations, and to Michael Sanderson and Mary Moore for offering perceptive comments on earlier drafts of this chapter. Thanks are also due to Junhyong Kim for insightful discussions on rate estimation, and to Anne Yoder for helpful discussions of divergence-time estimation in

primates and for providing access to preprints. Exceptionally helpful reviews were provided by Olaf Bininda-Emonds, Peter Mayhew, and Andy Purvis. BRM and KC wish to thank Simon Levin for hosting BRM in the Levin Lab during the course of this project and for his perennial guidance and support of their collaborative endeavors. Although this paper has benefited greatly from the suggestions of the aforementioned readers, we accept sole responsibility for any remaining errors that it might contain. Funding for this research was made possible through Natural Science and Engineering Research Council of Canada (NSERC) postgraduate scholarships to BRM and KC and through a Deep Time RCN graduate training award to BRM. This work was completed while KC was at the Department of Ecology and Evolutionary Biology, Princeton University.

References

AGAPOW, P.-M. AND PURVIS, A. 2002. Power of eight tree shape statistics to detect nonrandom diversification: A comparison of two models of cladogenesis. *Systematic Biology* 51:866–872.

ARNASON, U., GULLBERG, A., AND JANKE, A. 1998. Molecular timing of primate divergences as estimated by two non-primate calibration points. *Journal of Molecular Evolution* 47:718–727.

BALDWIN B. G. AND SANDERSON, M. J. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences of the United States of America* 95:9402–9406.

BARRACLOUGH, T. G., HARVEY, P. H., AND NEE, S. 1996. Rate of *rbc*L gene sequence evolution and species diversification in flowering plants (angiosperms). *Proceedings of the Royal Society of London B* 263:589–591.

BARRACLOUGH, T. G., NEE, S., AND HARVEY, P. H. 1998. Sister-group analysis in identifying correlates of diversification: comment. *Evolutionary Ecology* 12:751–754.

BARRACLOUGH, T. G. AND NEE, S. 2001. Phylogenetics and speciation. *Trends in Ecology and Evolution* 16:391–399.

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

BININDA-EMONDS, O. R. P., GITTLEMAN, J. L., AND PURVIS, A. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews* 74:143–175.

BREMER, B. AND ERIKSSON, O. 1992. Evolution of fruit characters and dispersal modes in the tropical family Rubiaceae. *Biological Journal of the Linnean Society* 47:79–95.

BRYANT, D., SEMPLE, C., AND STEEL, M. 2004. Supertree methods for ancestral divergence dates and other applications. In O. R. P. Bininda-Emonds (ed). *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 129–150. Kluwer Academic, Dordrecht, the Netherlands.

CHAN, K. M. A. AND MOORE, B. R. 1999. Accounting for mode of speciation increases power and realism of tests of phylogenetic asymmetry. *American Naturalist* 153:332–346.

CHAN, K. M. A. AND MOORE, B. R. 2002. Whole-tree methods for detecting differential diversification rates. *Systematic Biology* 51:855–865.

COLLESS, D. H. 1982. Review of Phylogenetics: The Theory and Practice of Phylogenetic Systematics, by E. O. Wiley. *Systematic Zoology* 31:100–104.

DODD, M. E., SILVERTOWN, J., AND CHASE, M. W. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* 53:732–744.

DONOGHUE, M. J. AND ACKERLY, D. D. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philosophical Transactions of the Royal Society of London B* 351:1241–1249.

DOYLE, J. A. AND DONOGHUE, M. J. 1993. Phylogenies and angiosperm diversification. *Paleobiology* 19:141–167.

EDGINGTON, E. S. 1972a. An additive method for combining probability values from independent experiments. Journal of Psychology 80:351–363.

EDGINGTON, E. S. 1972b. A normal curve method for combining probability values from independent experiments. *Journal of Psychology* 82:85–89.

EDGINGTON, E. S. AND HALLER, O. 1984. Combining probabilities from discrete probability distributions. *Educational and Psychological Measurement* 44: 265–274.

ERIKSSON, O. AND BREMER, B. 1991. Fruit characteristics, life forms, and species richness in the plant family Rubiaceae. *American Naturalist* 138:751–761.

ERIKSSON, O. AND BREMER, B. 1992. Pollination systems, dispersal modes, life forms, and diversification rates in angiosperm families. *Evolution* 46:258–256.

FARRELL, B. D. 1998. "Inordinate fondness" explained: why are there so many beetles? *Science* 281:555–559.

FARRELL, B. D. AND MITTER, C. 1998. The timing of insect/plant diversification: might *Tetraopes* (Coleoptera: Cerambycidae) and *Asclepias* (Asclepiadaceae) have co-evolved? *Biological Journal of the Linnean Society* 63:553–577.

FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: Inference and reliability. *Annual Review of Genetics* 22:521–565.

FELSENSTEIN, J. 1989. PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164–166. (http://evolution.genetics.washington.edu/phylip.html)

FISHER, R. A. 1932. *Statistical Methods for Research Workers*. 4th edition. Oliver and Boyd, Edinburgh.

FURNAS, G. W. 1984. The generation of random, binary unordered trees. *Journal of Classification* 1:187–233.

FUSCO, G. AND CRONK, Q. C. B. 1995. A new method for evaluating the shape of large phylogenies. *Journal of Theoretical Biology* 175:235–243.

GOUDET, J. 1999. An improved procedure for testing the effects of key innovations on rate of speciation. *American Naturalist* 153:549–555.

GOULD, S. J., RAUP, D. M., SEPOWSKI, J. J., SCHOPF, T. J. M., AND SIMBERLOFF, D. S. 1977. The shape of evolution: a comparison of real and random clades. *Paleobiology* 3:23–40.

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

GUYER, C. AND SLOWINSKI, J. B. 1993. Adaptive radiations and the topology of large phylogenies. *Evolution* 47:253–263.

HARRIS, T. E. 1964. *The Theory of Branching Processes*. Springer-Verlag, Berlin.

HARVEY, P. H., NEE, S., MOOERS, A. Ø., AND PARTRIDGE, L. 1991. These hierarchical views of life: phylogenies and metapopulations. In R. J. Berry, T. J. Cranford, and G. M. Hewitt (eds), *Genes in Ecology*, pp. 123–137. Blackwell Scientific, Oxford.

HARVEY, P. H. AND NEE, S. 1993. New uses for new phylogenies. *European Review* 1:11–19.

HARVEY, P. H. AND NEE, S. 1994. Comparing real with expected patterns from molecular phylogenies. In P. Eggleton and R. I. Vane-Wright (eds), *Phylogenetics and Ecology*, pp. 219–231. Academic Press, London.

HARVEY, P. H., HOLMES, E. C., MOOERS, A. Ø., AND NEE, S. 1994a. Inferring evolutionary processes from molecular phylogenies. In R. W. Scotland, D. J. Siebert, and D. M. Williams (eds), *Models in Phylogeny Reconstruction*, pp. 313–333. Clarendon Press, Oxford.

HARVEY, P. H., MAY, R. M., AND NEE, S. 1994b. Phylogenies without fossils. *Evolution* 48:523–529.

HARVEY, P. H., RAMBAUT, A., AND NEE, S. 1996. New computer packages for analysing phylogenetic tree structure. In J. Colbert and R. Barbault (eds), *Aspects of the Genesis and Maintenance of Biological Diversity*, pp. 60–68. Oxford University Press, Oxford.

HARDING, E. F. 1971. The probabilities of rooted tree-shapes generated by random bifurcation. *Advances in Applied Probability* 3:44–77.

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

HENNIG, W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana, Illinois.

HEY, J. 1992. Using phylogenetic trees to study speciation and extinction. *Evolution* 46:627–640.

HUELSENBECK, J. P., LARGET, B., AND SWOFFORD, D. 2000a. A compound Poisson process for relaxing the molecular clock. *Genetics* 154:1879–1892.

HUELSENBECK, J. P., RANNALA, B., AND MASLY, J. P. 2000b. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* 288:2349–2350.

HULBERT, R. C. 1993. Taxonomic evolution in North American Neogene horses (subfamily Equinae): the rise and fall of an adaptive radiation. *Paleobiology* 19:216–234.

JENSEN, J. S. 1990. Plausibility and testability: Assessing the consequences of evolutionary innovation. In M. H. Nitecki (ed.), *Evolutionary Innovations*, pp. 171–190. University of Chicago Press, Chicago.

JOBSON, R. W. AND ALBERT, V. A. 2002. Molecular rates parallel diversification contrasts between carnivorous plant sister lineages. *Cladistics* 18:127–136.

JONES, K. E., PURVIS, A., MACLARNON, A., BININDA-EMONDS, O. R. P., AND SIMMONS, N. B. 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biological Reviews* 77:223–259.

JUDD, W. S., SANDERS, R. W., AND DONOGHUE, M. J. 1994. Angiosperm family pairs: preliminary phylogenetic analyses. *Harvard Papers in Botany* 1:1–51.

KELLEY S. T. AND FARRELL, B. D. 1998. Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). *Evolution* 52:1731–1743.

KENDALL, D. G. 1948. On the generalized birth-and-death process. *Annals of Mathematical Statistics* 19:1–15.

KENNEDY, M. AND PAGE, R. D. M. 2002. Seabird supertrees: combining partial estimates of procellariiform phylogeny. *The Auk* 119:88–108.

KIRKPATRICK, M. AND SLATKIN, M. 1993. Searching for evolutionary patterns in the shape of a phylogenetic tree. *Evolution* 47:1171–1181.

KISHINO, H., THORNE, J. L., AND BRUNO, W. J. 2001. Performance of divergence time estimation methods under a probabilistic model of rate evolution. *Molecular Biology and Evolution* 18: 352–361.

KUBO, T. AND IWASA, Y. 1995. Inferring rates of branching and extinction from molecular phylogenies. *Evolution* 49:694–704.

LAPOINTE, F.-J. AND CUCUMEL, G. 1997. The average consensus procedure: combination of weighted trees containing identical or overlapping sets of taxa. *Systematic Biology* 46:306–312.

LAPOINTE, F.-J. AND LEVASSEUR, C. 2004. Everything you always wanted to know about the average consensus, and more. In O. R. P. Bininda-Emonds (ed.), *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 87–105. Kluwer Academic, Dordrecht, the Netherlands.

LEE, M. S. Y. 1999. Molecular clock calibrations and Metazoan divergence dates. *Journal of Molecular Evolution* 49:385–391.

LOSOS, J. B. AND ADLER, F. R. 1995. Stumped by trees? A generalized null model for patterns of organismal diversity. *American Naturalist* 145:329–342.

MADDISON, W. P. 1989. Reconstructing character evolution on polytomous cladograms. *Cladistics* 5:365–377.

MAGALLÓN, S. AND SANDERSON, M. J. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.

MCKENZIE, A. AND STEEL, M. 2000. Distributions of cherries for two models of trees. *Mathematical Biosciences* 164:81–92.

MINDELL, D. P., SITES, J. W., JR., AND GRAUR, D. 1989. Speciational evolution: a phylogenetic test with allozymes in *Sceloporus* (Reptilia). *Cladistics* 5:49–61.

MOOERS, A. Ø. 1995. Tree balance and tree completeness. *Evolution* 49:379–384.

MOOERS, A. Ø. AND HEARD, S. B. 1997. Inferring evolutionary process from phylogenetic tree shape. *Quarterly Review of Biology* 72:31–54.

NEE, S. 2001. Inferring speciation rates from phylogenies. *Evolution* 55:661–668.

NEE, S., MOOERS, A. Ø., AND HARVEY, P. H. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America* 89:8322–8326.

NEE, S., R. AND HARVEY, P. H. 1994. Getting to the root of flowering plant diversity. *Science* 264:1549–1550.

NEE, S., HOLMES, E. C., MAY, R. M., AND HARVEY, P. H. 1994a. Extinction rates can be estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society of London B* 344:77–82.

NEE, S., MAY, R. M., AND HARVEY, P. H. 1994b. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London B* 344:305–311.

NEE, S., HOLMES, E. C., MAY, R. M., AND HARVEY, P. H. 1995. Estimating extinction from molecular phylogenies. In J. H. Lawton and R. M. May (eds), *Extinction Rates*, pp. 164–182. Oxford University Press, Oxford.

NEE, S., BARRACLOUGH, T. G., AND HARVEY, P. H. 1996. Temporal changes in biodiversity: detecting patterns and identifying causes. In K. J. Gaston (ed.), *Biodiversity: a Biology of Numbers and Differences*, pp. 230–252. Blackwell Science, Oxford.

PAGE, R. D. M. 1993. On describing the shape of rooted and unrooted trees. *Cladistics* 9:93–99.

PARADIS, E. 1997. Assessing temporal variations in diversification rates from phylogenies: Estimation and hypothesis testing. *Proceedings of the Royal Society of London B* 264:1141–1147.

PARADIS, E. 1998a. Detecting shifts in diversification rates without fossils. *American Naturalist* 152:176–187.

PARADIS, E. 1998b. Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. *Molecular Biology and Evolution* 15:476–479.

PURVIS, A. 1995. A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London B* 348:405–421.

PURVIS, A. 1996. Using interspecies phylogenies to test macroevolutionary hypotheses. In K. J. Gaston (ed.), *Biodiversity: a Biology of Numbers and Differences*, pp. 151–168. Blackwell Science, Oxford.

PURVIS, A., NEE, S., AND HARVEY, P. H. 1995. Macroevolutionary inferences from primate phylogeny. *Proceedings of the Royal Society of London B* 260:329–333.

PURVIS, A., KATZOURAKIS, A, AND AGAPOW, P.-M. 2002. Evaluating phylogenetic tree shape: two modifications to Fusco and Cronk's method. *Journal of Theoretical Biology* 214:99–103.

PYBUS, O. G. AND HARVEY, P. H. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London B* 267:2267–2272.

PYBUS, O. G., RAMBAUT, A, HOLMES, E. C., AND HARVEY, P. H. 2002. New inferences from tree shape: numbers of missing taxa and population growth rates. *Systematic Biology* 51:881–888.

RAMBAUT, A., HARVEY, P. H., AND NEE, S. 1997. End-Epi: an application for inferring phylogenetic and population dynamical processes from molecular sequences. *Computer Applications in the Biosciences* 13:303–306.

RAMBAUT, A. AND BROMHAM, L. 1998. Estimating divergence dates from molecular sequences. *Molecular Biology and Evolution* 15:442–448.

RAUP, D. M., GOULD, S. J, SCHOPF, T. J. M., AND SIMBERLOFF, D. S. 1973. Stochastic models of phylogeny and the evolution of diversity. *Journal of Geology* 81:525–542.

RICKLEFS, R. E. AND RENNER, S. S. 1994. Species richness within families of flowering plants. *Evolution* 48:1619–1636.

ROGERS, J. S. 1993. Response of Colless's tree imbalance to number of terminal taxa. *Systematic Biology* 42:102–105.

ROGERS, J. S. 1994. Central moments and probability distribution of Colless' coefficient of tree imbalance. *Evolution* 48:2026–2036.

ROGERS, J. S. 1996. Central moments and probability distributions of three measures of phylogenetic tree imbalance. *Systematic Biology* 45:99–110.

RONQUIST, F., HUELSENBECK, J. P., AND BRITTON, T. 2004. Bayesian supertrees. In O. R. P. Bininda-Emonds (ed.), *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 193–224. Kluwer Academic, Dordrecht, the Netherlands.

SALAMIN, N., HODKINSON, T. R., AND SAVOLAINEN, V. 2002. Building supertrees: an empirical assessment using the grass family (Poaceae). *Systematic Biology* 51:136–150.

SANDERSON, M. J. 1994. Reconstructing the history of evolutionary processes using maximum likelihood. In D. M. Fambrough (ed.), *Molecular Evolution of Physiological Processes*, Society of General Physiologists Series 49:13–26. Rockefeller University Press, New York.

SANDERSON, M. J. 1997. A non-parametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14:1218–1231.

SANDERSON, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19:101–109.

SANDERSON, M. J. AND BHARATHAN, G. 1993. Does cladistic information affect inferences about branching rates? *Systematic Biology* 42:1–17.

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

SANDERSON, M. J. AND DONOGHUE, M. J. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends in Ecology and Evolution* 11:15–20.

SANDERSON, M. J. AND WOJCIECHOWSKI, M. F. 1996. Diversification rates in a temperate legume clade: are there "so many species" of *Astragalus* (Fabaceae)? *American Journal of Botany* 83:1488–1502.

SANMARTÍN, I., ENGHOF, H., AND RONQUIST, F. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society* 73:345–390.

SARICH, V. AND WILSON, A. C. 1967. Rates of albumin evolution in primates. *Proceedings of the National Academy of Sciences of the United States of America* 58:142–148.

SAVOLAINEN, V. AND GOUDET, J. 1998. Rate of gene sequence evolution and species diversification in flowering plants: a re-evaluation. *Proceedings of the Royal Society of London B* 265:603–607.

SHAO, K.-T. AND SOKAL, R. R. 1990. Tree balance. *Systematic Zoology* 39:266–276.

SIMMS, H. J. AND MCCONWAY, K. J. 2003. Nonstochastic variation of species-level diversification rates within angiosperms. *Evolution* 57:460–479.

SLOWINSKI, J. B. 1990. Probabilities of *n*-trees under two models: Demonstration that asymmetrical interior nodes are not improbable. *Systematic Zoology* 39:89–94.

SLOWINSKI, J. B. AND GUYER, C. 1989a. Testing the stochasticity of patterns of organismal diversity: an improved null model. *American Naturalist* 134:907–921.

SLOWINSKI, J. B. AND GUYER, C. 1989b. Testing null models in questions of evolutionary success. *Systematic Zoology* 38:189–191.

SLOWINSKI, J. B. AND GUYER, C. 1993. Testing whether certain traits have caused amplified diversification: an improved method based on a model of random speciation and extinction. *American Naturalist* 142:1019–1024.

STANLEY, S. M. 1979. *Macroevolution: Pattern and Process*. W. H. Freeman, San Francisco.

STONE, J. AND REPKA, J. 1998. Using a nonrecursive formula to determine cladogram probabilities. *Systematic Biology* 47:617–624.

STONER, C. J., BININDA-EMONDS, O. R. P, AND CARO, T. 2003. The adaptive significance of coloration in lagomorphs. *Biological Journal of the Linnean Society* 79:309–328.

TAKEZAKI, N., RZHETSKY, A, AND NEI, M. 1995. Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution* 12:823–833.

THORNE, J. L., KISHINO, H, AND PAINTER, I. S. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15:1647–1657.

THORNE, J. L. AND KISHINO, H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689–702.

TIFFNEY, B. H. AND MAZER, S. J. 1995. Angiosperm growth habit, dispersal and diversification reconsidered. *Evolutionary Ecology* 9:93–117.

VOS, R. A. AND MOOERS, A. Ø. 2004. Reconstructing divergence times for supertrees. In O. R. P. Bininda-Emonds (ed.), *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pp. 281–299. Kluwer Academic, Dordrecht, the Netherlands.

WALLIS, W. A. 1942. Compounding probabilities from independent significance tests. *Econometrica* 10:229–248.

WILEY, E. O. 1981. *Phylogenetics: the Theory and Practice of Phylogenetic Systematics*. Wiley and Sons, New York.

WOJCIECHOWSKI, M. F., SANDERSON, M. J., STEEL, K. P., AND LISTON, A. 2000. Molecular phylogeny of the "temperate herbaceous tribes" of papilionoid legumes: a supertree approach. In P. Herendeen and A. Bruneau (eds), *Advances in Legume Systematics* 9:277–298. Royal Botanic Garden, Kew.

WU, C.-I. AND LI, W.-H. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of the National Academy of Sciences of the United States of America* 82:1741–1745.

YODER, A. D. AND YANG, Z. H. 2000. Estimation of speciation dates using local molecular clocks. *Molecular Biology and Evolution* 17:1081–1090.

YODER, A. D. AND YANG, Z. H. In press. Divergence dates for Malagasy lemurs estimated from multiple gene loci: fit with climatological events and speciation models. *Molecular Ecology*

YULE, G. U. 1924. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis. *Philosophical Transactions of the Royal Society of London B* 213:21–87.