THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1278

Testing macro-evolutionary models using incomplete molecular phylogenies

Oliver G. Pybus* **and Paul H. Harvey**

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Phylogenies reconstructed from gene sequences can be used to investigate the tempo and mode of species diversification. Here we develop and use new statistical methods to infer past patterns of speciation and extinction from molecular phylogenies. Specifically, we test the null hypothesis that per-lineage speciation and extinction rates have remained constant through time. Rejection of this hypothesis may provide evidence for evolutionary events such as adaptive radiations or key adaptations. In contrast to previous approaches, our methods are robust to incomplete taxon sampling and are conservative with respect to extinction. Using simulation we investigate, first, the adverse effects of failing to take incomplete sampling into account and, second, the power and reliability of our tests. When applied to published phylogenies our tests suggest that, in some cases, speciation rates have decreased through time.

Keywords: phylogeny; speciation; extinction; primates; *Dendroica*; *Sorex*

1. INTRODUCTION

Macro-evolutionary biology aims to uncover and understand the processes that have led to the biological diversity we observe today. We would like to document past changes in diversity and then test whether these changes are compatible with various models of diversification. The widespread and increasing availability of gene-sequence data makes molecular phylogenies an attractive source of information about historical patterns of diversity. In addition to showing past speciation events, phylogenies of extant taxa may also contain information about extinction (Nee *et al.* 1994*a*,*b*; Kubo & Iwasa 1995).

Several methods have been developed to infer speciation and extinction rates from reconstructed molecular phylogenies (reviewed by Sanderson & Donoghue 1996; Purvis 1996; Mooers & Heard 1997). Almost all of these methods have a common statistical framework, the generalized birth–death process. This process, which was initially applied to the fossil record (Raup *et al.* 1973), represents speciation and extinction as instantaneous events occurring in continuous time. During a short time interval, Δt, each extant lineage speciates with probability $b(t)\Delta t$ and goes extinct with probability $d(t)\Delta t$. Hence $b(t)$ and $d(t)$ are time-dependent per-lineage rates of speciation and extinction.

Birth–death methods can be used to answer two general questions. First, is speciation or extinction more likely to occur on some lineages than on others (Slowinski & Guyer 1989; Hey 1992; Nee *et al.* 1992, 1994*b*; Kirkpatrick & Slatkin 1993; Sanderson & Donoghue 1994; Paradis 1998*a*)? Second, supposing that speciation and extinction do not vary among lineages, what are the rates of these processes and do they change through time (Hey 1992; Harvey *et al.* 1994; Nee *et al.* 1994*a*,*b*; Kubo & Iwasa 1995; Zink & Slowinski 1995; Wollenberg *et al.* 1996; Paradis 1997, 1998*a*,*b*)? In this paper we concentrate on the latter question, but note that it can only be asked once the former has been answered.

Nee *et al.* (1994*b*) suggested that a birth–death process in which speciation and extinction rates remain constant through time is an appropriate null model for diversification (i.e. $b(t) = b$ and $d(t) = d$, where b and d are constants). Consequently, they developed a maximum-likelihood method for estimating b and d from reconstructed phylogenies. Obviously, it is essential to ascertain whether speciation and extinction rates have remained constant before using such a method. Furthermore, deviation from this null model may suggest the presence of evolutionary and ecological processes such as adaptive radiations or key adaptations. Unfortunately, all previous methods for testing the null hypothesis of constant speciation and extinction rates have limiting assumptions. Paradis (1997, 1998*a*,*b*) developed methods based on survival analysis that assume that the distribution of node heights in a reconstructed phylogeny depends only on $b(t) - d(t)$. This is only true if there is no extinction ($d(t) = 0$) (Nee *et al.* 1994*b*; Paradis 1998*b*). Wollenberg *et al.* (1996) reported a 'null model of stochastic lineage bifurcation and extinction' that has been used to infer changes in cladogenesis through time. However, Wollenberg *et al.*'s (1996) method tests whether $b = d$, not whether b and d are constant through time (Paradis 1998*b*). Zink & Slowinski (1995) developed a test that, like Paradis' (1997, 1998*a*,*b*) and Wollenberg *et al.*'s (1996) methods, is limited by the assumption of complete sampling: the reconstructed phylogeny must contain all members of the investigated clade.

Here we investigate the effect of incomplete sampling on methods that test the null hypothesis of constant per-lineage speciation and extinction rates. Using Monte Carlo simulation we develop a new test that assumes neither a complete sample nor zero extinction, and apply this method to several reconstructed phylogenies. Our results suggest that incorrect inferences are made when incomplete sampling is not specifically taken into account.

*Author for correspondence (oliver.pybus@zoo.ox.ac.uk).

Received 7 June 2000 *Accepted* 30 August 2000

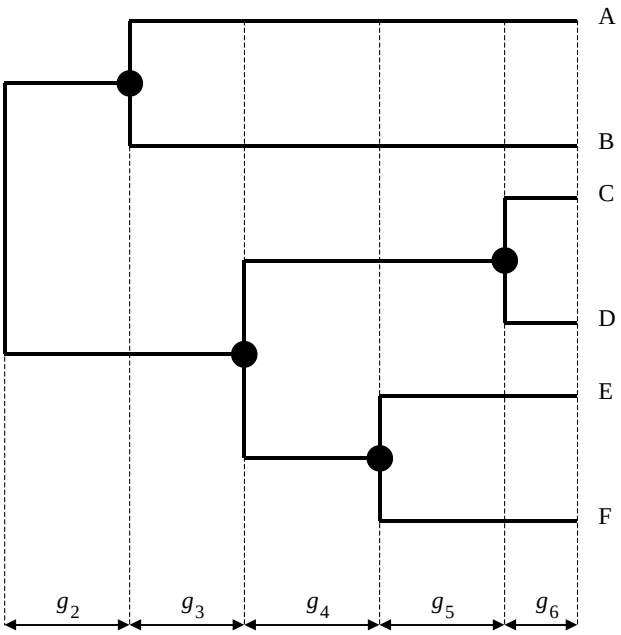

Figure 1. A reconstructed phylogeny containing six extant species (A–F). The internode distances of the phylogeny are denoted by $g_2, g_3, \ldots, g_6$. Note that the subscripts refer to the number of lineages present in the phylogeny during each internode interval. The filled circles show the positions of internal nodes.

2. MODEL

We first define four common birth–death processes: first, the generalized birth–death model, in which $b(t)$ and $d(t)$ can vary through time; second, the constant-rates birth–death model, in which both $b(t) = b$ and $d(t) = d$ are constant and $b \geqslant d$; third, the generalized birth model, in which $b(t)$ varies through time and $d(t) = 0$; and fourth, the pure birth model, in which $b(t) = b$ is constant and $d(t) = 0$.

Suppose that a birth–death process, starting with one lineage, gives rise to n taxa at the present day. We call a tree containing all n extant taxa a 'complete reconstructed phylogeny'. The properties of such phylogenies are well described (Harvey *et al.* 1994; Nee *et al.* 1994*a*,*b*; Kubo & Iwasa 1995). If $b(t)$ and $d(t)$ do not vary among lineages then the branching structure (topology) of the reconstructed phylogeny is independent of the underlying birth–death process. However, the birth–death process does determine the internode distances of the reconstructed phylogeny (figure 1).

Consider a complete reconstructed phylogeny that is the result of a constant-rates birth–death process with parameters b and d. If d is large then the phylogeny's internal nodes will tend to occur near its tips. This effect, known as 'the pull of the present', occurs because lineages that arose in the recent past are less likely to be removed by extinction than lineages that arose in the distant past. The strength of the pull depends on d/b, and increases as d/b increases (Nee *et al.* 1994*a*). Therefore, we might expect that the internal nodes of the phylogeny are, on average, closest to the root when $d/b = 0$ (the pure birth model). This observation suggests that we can reject the constant-rates model if the internal nodes of a reconstructed phylogeny are closer to the root than expected under the pure birth model.

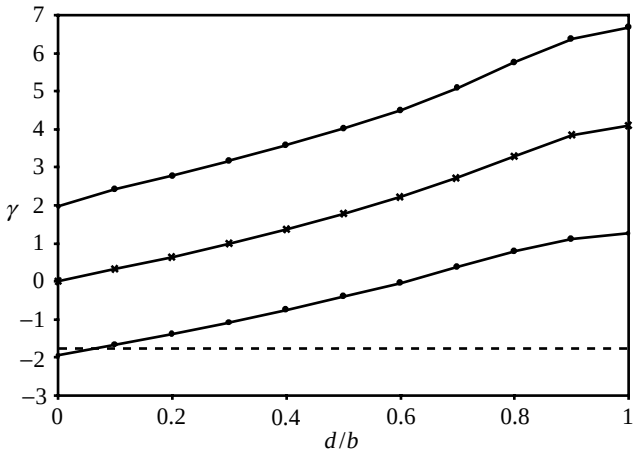

Figure 2. Distribution of the γ-statistic under the constant-rates birth–death model. For each value of d/b the distribution of γ was obtained from 50 000 simulated phylogenies with 30 tips. The curve marked with crosses is the mean of the γ-distribution. The curves marked with circles are the 0.975 and 0.025 percentiles of the γ-distribution. The dashed line ($\gamma = -1.645$) represents the critical value of the constant-rates test. The hypothesis of constant b and d is rejected when $\gamma < -1.645$. For the pure birth model $d/b = 0$.

(a) *The γ-statistic*

To implement this test a measure of the relative positions of internal nodes within a phylogeny is needed. Let $g_2, g_3, \ldots, g_n$ be the internode distances of a reconstructed phylogeny with n taxa (figure 1). The statistic we use is

$$\gamma = \frac{\left(\frac{1}{n-2}\sum_{i=2}^{n-1}\left(\sum_{k=2}^{i} k g_k\right)\right) - \left(\frac{T}{2}\right)}{T\sqrt{\frac{1}{12(n-2)}}}, \quad T = \left(\sum_{j=2}^{n} j g_j\right). \tag{1}$$

This statistic is modified from Cox & Lewis (1966, p. 47). Although complicated, γ has a useful property: under the pure birth process, γ-values of complete reconstructed phylogenies follow a standard normal distribution. If $\gamma > 0$ then a phylogeny's internal nodes are closer to its tips than expected under the pure birth process. Conversely, if $\gamma < 0$ then the internal nodes are closer to the root than expected under the pure birth model. Therefore, the null hypothesis of constant b and d is rejected at the 5% level if $\gamma < -1.645$ (one-tailed test). We call this the constant-rates (CR) test.

The above approach follows the same rationale as Zink & Slowinski (1995), who presented a statistic similar to γ. However, their statistic contained errors that have been fixed in equation (1) (J. Slowinski, personal communication). The correctness of equation (1) can be verified in two ways. First, the expected value of g_k under the pure birth process is $1/kb$. If this value is substituted into equation (1) then $\gamma = 0$, as expected for a standard normal distribution. Second, our Monte Carlo simulations (see § 2(b)–(d)) consistently show that under the pure birth model γ follows a standard normal distribution.

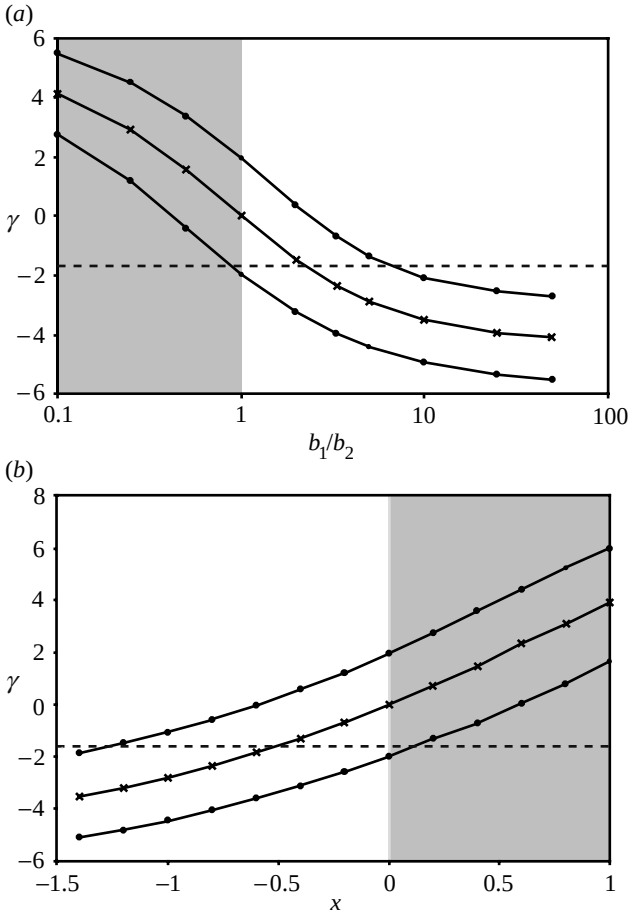

Figure 3. Distribution of the γ-statistic under two generalized birth models. The shaded region contains cases in which $b(t)$ increases through time and the unshaded region contains cases in which $b(t)$ decreases through time. The dashed line and marked curves are explained in figure 2. The γ-distributions were obtained from 50 000 simulated phylogenies with 30 tips. (*a*) Phylogenies were simulated from one to 15 lineages under $b(t) = b_1$, then simulated from 16 to 30 lineages under $b(t) = b_2$. The hypothesis of constant rates is only rejected when $b_1/b_2 > 1$, that is, when $b(t)$ decreases through time. For the pure birth process $b_1/b_2 = 1$. (*b*) Phylogenies were simulated under the model $b(t) = N^x$, where N is the number of extant lineages (N always increases through time). The hypothesis of constant rates is only rejected when $x < 0$, that is, when $b(t)$ decreases through time. For the pure birth process $x = 0$.

(b) *Monte Carlo simulation*

We investigated the behaviour of the γ-statistic using Monte Carlo simulation. Complete reconstructed phylogenies were simulated using the code implemented in the computer program Bi-De (Rambaut *et al.* 1996). Given a specified birth–death process, this program uses a Markov-chain algorithm to simulate a 'complete actual phylogeny', which contains both extinct and extant lineages. Once the algorithm has finished, a complete reconstructed phylogeny is obtained by removing all the extinct lineages from the actual phylogeny.

Figure 2 shows the behaviour of γ under the constant-rates birth–death process. Here, γ depends only on d/b, is smallest when $d/b = 0$ and increases as $d/b \rightarrow 1$. More extinction results in a stronger pull of the present, so internal nodes tend to occur closer to the tree's tips. The shallow gradient of figure 2 suggests that γ is a weak estimator of d/b. Figure 2 also shows that the type I error of the CR test, when applied to complete phylogenies, is never greater than 0.05 (at the 5% level).

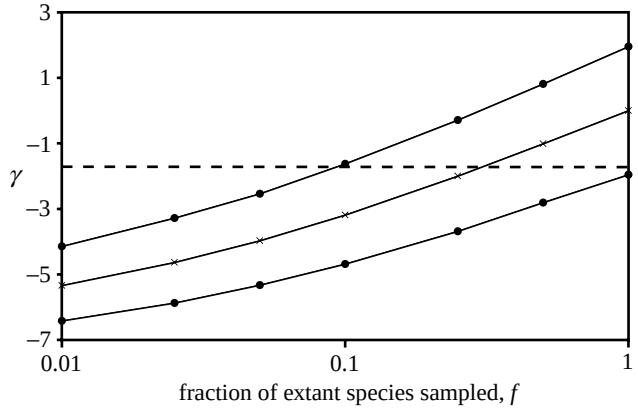

Figure 4. Effect of incomplete sampling on the constant-rates test. The dashed line and marked curves are explained in figure 2. For each value of f, 50 000 complete phylogenies with $(30/f)$ tips were simulated under the pure birth process. The γ-values were calculated from incomplete phylogenies that were obtained by randomly sampling 30 tips from the complete phylogenies. For the pure birth process $f = 1$. The type I error of the constant-rates test increases as f decreases.

(c) *Type II error*

To investigate the type II error of the CR test, reconstructed phylogenies were simulated using birth–death models in which speciation and extinction rates vary through time. Figure 3 shows the distribution of γ under two generalized birth models. In both cases, the null hypothesis of constant speciation and extinction rates is only rejected when $b(t)$ decreases through time. This seems reasonable: if $b(t)$ increases through time then more speciation events will occur near the present, resulting in a larger γ-value and a high type II error.

We performed further simulations in which extinction was added to the generalized birth models used in figure 3 (results not shown). No matter what form $b(t)$ takes, adding extinction always increases γ and hence increases the type II error. This result follows from Nee *et al.* (1994*b*, eq. 24) and was predicted by Zink & Slowinski (1995).

(d) *Incomplete sampling*

Simulations were performed to measure the effect of incomplete taxon sampling on the type I error of the CR test. Incomplete sampling was implemented by simulating a complete reconstructed phylogeny under the pure birth process, randomly selecting a fraction, f, of the extant taxa (the sampled taxa), then removing all non-sampled lineages from the reconstructed phylogeny. Figure 4 shows that the type I error of the CR test increases as f decreases. Internal nodes near the root leave more descendants and are more likely to be included in a small sample than nodes near the tips, resulting in lower γ-values (Nee *et al.* 1994*a*). Hence, incomplete sampling can lead us to conclude incorrectly that speciation and extinction rates vary through time.

We suggest a Monte Carlo solution to this problem. Suppose we have an incomplete reconstructed phylogeny

of x species, representing a random sample from a larger clade containing y species. Using the method outlined above, we can simulate the distribution of γ under the pure birth process when only x of y extant species are sampled. An appropriate critical value of γ is then obtained from the simulated distribution. Observed γ-values below the critical value are sufficient to reject the CR hypothesis. This test is named the Monte Carlo constant-rates (MCCR) test. By definition, the type I error of the MCCR test is never greater than 0.05 (at the 5% level), even when it is applied to incomplete phylogenies. The MCCR test is conservative to extinction, which has the effect of increasing γ and thus raising the type II error.

In summary, the CR and MCCR tests can reject the null hypothesis of constant b and d in favour of a scenario in which $b(t)$ decreases through time. Incomplete taxon sampling results in the CR test incorrectly rejecting the null hypothesis. In contrast, the MCCR test is robust to incomplete sampling. Both tests are conservative with respect to extinction and have high type II errors. Therefore rejection of the null hypothesis is a significant result but failure to reject should not be considered as evidence in favour of the CR model.

3. RESULTS

We have applied the CR and MCCR tests to several published phylogenies whose speciation and extinction rates have been investigated using other methods. In each case, the assumption that diversification occurs equally among lineages is investigated using two methods. First, the B_1 tree-asymmetry test (Kirkpatrick & Slatkin 1993) is used to test the null hypothesis that $b(t)$ and $d(t)$ are equal across lineages. Second, the relative cladogenesis statistic (Nee *et al.* 1994*b*; Rambaut *et al.* 1997) is used to locate lineages that have diversified unusually rapidly or slowly.

(a) Sorex *shrews*

Fumagalli *et al.* (1999) investigated the evolution of *Sorex*, a genus of red-toothed shrews found predominantly in the palaeartic and neartic. They obtained 1011 bp mitochondrial DNA sequences from 27 out of the 70 extant *Sorex* species. The authors noted that the lineages-through-time plot of their phylogeny was convex (Harvey *et al.* 1994), leading them to conclude that the rate of speciation in the *Sorex* genus had decreased through time. However, convex lineages-through-time plots also occur when there is incomplete sampling (Nee *et al.* 1994*a*).

We reconstructed a *Sorex* phylogeny (denoted *Sorex A*) using Fumagalli *et al.*'s (1999) sequence data. One species, *Sorex tundrensis*, was omitted because it appears to have a different rate of molecular evolution from the other species investigated (Fumagalli *et al.* 1999). The tree topology was estimated using neighbour joining with outgroup rooting, whilst branch lengths were estimated using maximum likelihood (REV-Γ substitution model, under the assumption of a molecular clock). The molecular clock could not be rejected using a likelihood ratio test (Felsenstein 1981). Using the B_1 test, we rejected the hypothesis of equal rates across lineages in *Sorex A* ($B_1 = 11.50$, $p < 0.05$). The relative cladogenesis statistic highlighted one clade (the *Sorex araneus–arcticus* group) with unusually rapid diversification (Fumagalli *et al.* 1999). This clade contained six out of eight *Sorex* species with an unusual chromosome system, suggesting that Fumagalli *et al.*'s (1999) data set is a non-random sample of the *Sorex* genus as a whole. One member of this group (*Sorex samniticus*) was subsequently removed, resulting in a second phylogeny (denoted *Sorex B*), which was not rejected under the B_1 test ($B_1 = 11.39$, $p > 0.05$).

We applied the MCCR test to the *Sorex A* and *B* phylogenies. In both cases the hypothesis of constant b and d was not rejected, so it is not possible to draw any firm conclusions about *Sorex* diversification using the MCCR test at present (table 1). Interpretation is further complicated by the probable non-random sampling of the *S. araneus–arcticus* clade.

(b) Dendroica *warblers*

Dendroica is a morphologically conservative but behaviourally diverse genus of North American wood warblers that has played an important role in the study of species coexistence (MacArthur 1958). Lovette & Bermingham (1999) presented a phylogeny of the '*Dendroica* radiation' reconstructed from 3639 bp mitochondrial DNA sequences (denoted *Dendroica A*). The tree contained 24 out of 27 extant *Dendroica* species, plus *Setophaga ruticilla* and *Catharopeza bishopi*. Lovette & Bermingham (1999) analysed *Dendroica A* using the method of Wollenberg *et al.* (1996) and concluded that there was 'a significant burst of speciation occurring early in the history of the genus'. Wollenberg *et al.*'s (1996) method actually tests the null hypothesis that $b = d$ (Paradis 1998*b*) and consequently is not directly comparable with our tests: deviation from the hypothesis that $b = d$ is not evidence that per-lineage speciation and extinction rates have changed through time. Furthermore, a birth–death process with $b = d$ is prone to rapid extinction, so contemporary clades that have grown under this model are unlikely to exist. For example, if $b = d$ then 96.7% of simulated phylogenies go extinct before reaching 30 extant lineages.

The B_1 test indicates that diversification in *Dendroica A* is not equal among lineages ($B_1 = 10.5$, $p < 0.05$). The relative cladogenesis statistic suggests that this is mainly due to an outgroup of six lineages with unusually low diversification rates (*Dendroica pharetra*, *C. bishopi*, *Dendroica angelae*, *Dendroica plumbea*, *S. ruticilla* and *Dendroica tigrina*). We removed this out group, creating a second phylogeny (denoted *Dendroica B*), which was not rejected under the B_1 test ($B_1 = 9.0$, $p > 0.05$).

The MCCR test was applied to the *Dendroica A* and *B* phylogenies. In both cases the hypothesis of constant b and d is strongly rejected (table 1). Hence, there is good phylogenetic evidence for a decrease in *Dendroica* speciation rate through time. Furthermore, this conclusion is unaffected by the removal of several basal lineages.

(c) *Primates*

A composite estimate of the phylogeny of all extant primates was presented by Purvis (1995). Previous studies of this phylogeny have failed to detect clades in which speciation and extinction rates vary through time (Purvis *et al.* 1995; Paradis 1998*a*). However, the tests used in these studies assumed that the extinction rate is zero. Here, we follow Purvis *et al.* (1995) and investigate the *Strepsirhini* of Madagascar and the *Platyrrhini* of South America.

Table 1. *Results of the CR and MCCR tests*

phylogeny	number of sampled species (x)	true number of species (y)	γ	critical value of γ (at 5% level)[a]	significance level[a]
Sorex A	26	70	−2.678	−2.792	$0.06 < p < 0.07$
Sorex B	25	70	−2.491	−2.805	$p \sim 0.1$
Dendroica A	26	29	−4.171	−1.773	$p < 0.0001$
Dendroica B	20	29	−3.403	−1.996	$p < 0.001$
Madagascar *Strepsirhini*	24	24	−2.281	−1.645	$0.01 < p < 0.02$
South American *Platyrrhini*	66	66	−0.918	−1.645	$p \sim 0.2$

[a] Distribution obtained from 50 000 replicates.

In this instance the B_1 test is inappropriate because the phylogenies contain a number of polytomies (Kirkpatrick & Slatkin 1993). Fortunately the relative cladogenesis statistic detects no anomalous lineages in the *Strepsirhini* phylogeny and only one weakly significant lineage ($p = 0.04$) in the *Platyrrhini* phylogeny (Purvis *et al.* 1995), so it is reasonable to assume that there is no significant variation in speciation and extinction rates among lineages.

Using the CR test, the null hypothesis of constant speciation and extinction rates is rejected for the Madagascan *Strepsirhini* but not for the *Platyrrhini* (table 1). One interpretation of this result is that the restricted geographical area of Madagascar has led to density-dependent cladogenesis in the *Strepsirhini* (Purvis *et al.* 1995).

4. DISCUSSION

In common with all methods for inferring evolutionary processes from phylogenies, the tests presented here make a number of important assumptions.

(a) *No phylogenetic error*

The CR and MCCR tests are applied to reconstructed phylogenies that are assumed to be correct. This assumption can always be called into question, justifiably so if the sequences used are uninformative. Since the CR and MCCR tests use internode-distance information, they are expected to be more sensitive to branch-length errors than topology errors. In particular, biases that do not apply equally to all branches could lead to systematic bias in γ. For example, large genetic distances are more seriously underestimated than small distances when among-nucleotide substitution rate variation is ignored (Yang 1996).

Phylogenetic error could be incorporated into our tests by integrating across all possible topologies and branch lengths. Sanderson & Donoghue (1994) used this approach to investigate diversification among three lineages. Larger phylogenies could be investigated using Markov-chain Monte Carlo methods (*sensu* Yang & Rannala 1997).

(b) *Molecular clock*

The γ-statistic requires that observed node heights represent relative times. To satisfy this condition molecular phylogenies are usually reconstructed under the assumption of a molecular clock. However, the clock is not obligatory: recent work has demonstrated that speciation events can be dated even when the rate of molecular evolution is not uniform across the tree (Sanderson 1997; Thorne *et al.* 1998; Huelsenbeck *et al.* 2000).

(c) *Random taxon sampling*

The MCCR test assumes that incomplete phylogenies contain taxa that are sampled randomly with respect to the complete phylogeny. Non-random sampling may be overdispersed, for example when one representative of each subgenus is used, or underdispersed, if taxa are excluded because they exhibit a trait (such as rarity or geographical location) that is correlated with phylogeny. Overdispersed sampling will raise the type I error of the MCCR test, whilst underdispersed sampling will raise the type II error.

(d) *Equal rates across lineages*

Most methods, including our tests and Nee *et al.*'s (1994*b*) framework, assume that speciation and extinction occur equally on all lineages. At present, the effect of lineage-specific diversification on such methods is not known. We chose to locate and remove anomalous lineages, then note if this *ad hoc* procedure qualitatively affected our results. The optimal solution would be to develop null models that incorporate both lineage- and time-dependent speciation and extinction.

(e) *Summary*

We have shown that it is possible to make useful inferences about species diversification in the presence of confounding factors such as extinction and incomplete sampling. The continuing development of such methods should lead to a better understanding of macroevolutionary and biogeographical processes. Furthermore, we have shown that it is possible to quantify phylogenetic information about incomplete sampling. We are currently investigating the possibility of using such information to provide phylogenetic estimates of taxon under-sampling and species richness.

We thank Andrew Rambaut for program code and advice, and Walter Jetz for the ecological input. This work was funded by the Wellcome Trust (grant no. 050275), the Biotechnology and Biological Sciences Research Council and The Royal Society.

REFERENCES

Cox, D. R. & Lewis, P. A. W. 1966 *The statistical analysis of series of events*. London: Methuen.

Felsenstein, J. 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376.

Fumagalli, L., Taberlet, P., Stewart, D. T., Gielly, L., Hausser, J. & Vogel, P. 1999 Molecular phylogeny and evolution of *Sorex* shrews inferred from mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* **11**, 222–235.

Harvey, P. H., May, R. M. & Nee, S. 1994 Phylogenies without fossils. *Evolution* **48**, 523–529.

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

Huelsenbeck, J. P., Larget, B. & Swofford, D. 2000 A compound Poisson process for relaxing the molecular clock. *Genetics* **154**, 1879–1892.

Kirkpatrick, M. & Slatkin, M. 1993 Searching for evolutionary pattern in the shape of a phylogenetic tree. *Evolution* **47**, 1171–1181.

Kubo, T. & Iwasa, Y. 1995 Inferring the rates of branching and extinction from molecular phylogenies. *Evolution* **49**, 694–704.

Lovette, I. J. & Bermingham, E. 1999 Explosive speciation in the New World *Dendroica* warblers. *Proc. R. Soc. Lond.* B **266**, 1629–1636.

MacArthur, R. H. 1958 Population ecology of some warblers of northeastern coniferous forests. *Ecology* **39**, 599–619.

Mooers, A. & Heard, S. B. 1997 Inferring evolutionary process from phylogenetic tree shape. *Q. Rev. Biol.* **72**, 31–54.

Nee, S., Mooers, A. & Harvey, P. H. 1992 Tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl Acad. Sci. USA* **89**, 8322–8326.

Nee, S., Holmes, E. C., May, R. M. & Harvey, P. H. 1994*a* Extinction rates can be estimated from molecular phylogenies. *Phil. Trans. R. Soc. Lond.* B **344**, 77–82.

Nee, S. May, R. M. & Harvey, P. H. 1994*b* The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond.* B **344**, 305–311.

Paradis, E. 1997 Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. *Proc. R. Soc. Lond.* B **264**, 1141–1147.

Paradis, E. 1998*a* Detecting shifts in diversification rates without fossils. *Am. Nat.* **152**, 176–187.

Paradis, E. 1998*b* Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. *Mol. Biol. Evol.* **15**, 476–479.

Purvis, A. 1995 A composite estimate of primate phylogeny. *Phil. Trans. R. Soc. Lond.* B **348**, 405–421.

Purvis, A. 1996 Testing macroevolutionary hypotheses. In *New uses for new phylogenies* (ed. P. H. Harvey, A. J. Leigh Brown, J. Maynard Smith & S. Nee), pp. 153–168. Oxford University Press.

Purvis, A., Nee, S. & Harvey, P. H. 1995 Macroevolutionary inferences from primate phylogeny. *Proc. R. Soc. Lond.* B **260**, 329–333.

Rambaut, A., Grassly, N. C., Nee, S. & Harvey, P. H. 1996 Bi-De: an application for simulating phylogenetic processes. *Comput. Appl. Biosci.* **12**, 469–471.

Rambaut, A., Harvey, P. H. & Nee, S. 1997 End-Epi: an application for inferring phylogenetic and population dynamical processes from molecular sequences. *Comput. Appl. Biosci.* **13**, 303–306.

Raup, D. M., Gould, S. J., Schopf, T. J. M. & Simberloff, D. 1973 Stochastic models of phylogeny and the evolution of diversity. *J. Geol.* **81**, 525–542.

Sanderson, M. J. 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1232.

Sanderson, M. J. & Donoghue, M. J. 1994 Shifts in diversification rate with the origin of the angiosperms. *Science* **264**, 1590–1593.

Sanderson, M. J. & Donoghue, M. J. 1996 Reconstructing shifts in diversification rates on phylogenetic trees. *Trends Ecol. Evol.* **11**, 15–20.

Slowinski, J. & Guyer, C. 1989 Testing the stochasticity of patterns of organismal diversity: an improved null model. *Am. Nat.* **134**, 907–921.

Thorne, J. L., Kishino, H. & Painter, I. S. 1998 Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657.

Wollenberg, K., Arnold, J. & Avise, J. C. 1996 Recognising the forest for the trees: testing temporal patterns of cladogenesis using a null model of stochastic diversification. *Mol. Biol. Evol.* **13**, 833–849.

Yang, Z. 1996 Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* **11**, 367–372.

Yang, Z. & Rannala, B. 1997 Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol. Biol. Evol.* **14**, 717–724.

Zink, R. M. & Slowinski, J. B. 1995 Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. *Proc. Natl Acad. Sci. USA* **92**, 5832–5835.