THE ROYAL
SOCIETY

Received 26 May 2003
Accepted 8 July 2003

Analysis of diversification: combining phylogenetic and taxonomic data

Emmanuel Paradis

Laboratoire de Paléontologie, Paléobiologie & Phylogénie, Institut des Sciences de l'Évolution, Université Montpellier II, 34095 Montpellier cédex 05, France (paradis@isem.univ-montp2.fr)

The estimation of diversification rates using phylogenetic data has attracted a lot of attention in the past decade. In this context, the analysis of incomplete phylogenies (e.g. phylogenies resolved at the family level but unresolved at the species level) has remained difficult. I present here a likelihood-based method to combine partly resolved phylogenies with taxonomic (species-richness) data to estimate speciation and extinction rates. This method is based on fitting a birth-and-death model to both phylogenetic and taxonomic data. Some examples of the method are presented with data on birds and on mammals. The method is compared with existing approaches that deal with incomplete phylogenies. Some applications and generalizations of the approach introduced in this paper are further discussed.

Keywords: diversity; estimation; extinction; maximum likelihood; phylogeny; speciation

1. INTRODUCTION

Evolutionary biology deals with changes in species diversity through time and their mechanisms. Such changes are brought about by speciation and extinction events, and quantifying the rates of these events contributes to a better understanding of the biological mechanisms of evolution. During the past decade, there has been a lot of interest in using phylogenetic data to estimate speciation and extinction rates. This follows from the increasing number of phylogenetic studies using molecular techniques, which have resolved the relationships among more and more taxa. Most of the methods to study diversification with phylogenetic data consider phylogenies at the species level, and thus require that most (if not all) species of the studied lineages be included in the reconstructed phylogenies (Hey 1992; Nee *et al.* 1994*a*; Sanderson & Donoghue 1996; Paradis 1997, 1998*a*; Pybus & Harvey 2000). However, quite often phylogenetic data are available at a higher level than the species, for instance giving the relationships between families even though the relationships between species within the families are unknown although the numbers of species in these groups are known. Examples of such situations are provided by the phylogeny of birds, where almost all 144 families are included but only about 1000 species out of 9646 are represented (Sibley & Ahlquist 1990), and the phylogeny of insects, where the relationships among the 30 orders have been inferred but the relationships among the more than a million species cannot be deduced (Mayhew 2002).

In this paper, I present a method to estimate speciation and extinction rates, which combines phylogenetic and species-richness data. Some examples using birds and mammals are presented.

2. MATERIAL AND METHODS

The approach used here is an extension of the birth–death model developed by Nee *et al.* (1994*a*). They used a birth–death process to model phylogenetic diversification and to estimate speciation and extinction rates from the branching times of an observed phylogenetic tree. Basically, this method requires that all species belonging to the studied clade are included in the reconstructed tree (Nee *et al.* 1994*a*,*b*).

The present approach assumes that two kinds of information are available. First, the phylogenetic relationships (tree topology and branch lengths) of a subset of all species belonging to the studied clade are known; this is referred to as the phylogenetic data in this paper. It is further assumed that the reconstructed phylogenetic tree is ultrametric, that is that the branch lengths are linearly related to time. Second, the number of species belonging to each tip of this reconstructed phylogeny is also known; this is called the taxonomic data. This is likely to be a very common situation. For instance, it is usual in a phylogenetic study to resolve the relationships among higher level taxa (e.g. among orders within a class, or among families within an order), and the number of species in each of these taxa is known. It is assumed below that both the phylogenetic and the taxonomic data are known without error.

Let N be the number of tips in the reconstructed phylogeny, and n_i be the number of species belonging to the ith tip ($i = 1,\ldots,N$).

Let us assume that a living species has an instantaneous probability of splitting into two species called the rate of speciation and denoted by λ, and an instantaneous probability of disappearing called the rate of extinction and denoted by μ. We further assume that both rates are constant through time and across lineages. Under this 'birth-and-death' model, we can formulate the probabilities of the observed data using formulae in Kendall (1948) and in Nee *et al.* (1994*a*).

The phylogenetic data comprise a set of N lineages that originated from the nodes of the reconstructed phylogeny and have survived until the present. The probability of a lineage originating at time t not going extinct before time T is (Nee *et al.* 1994*a*, eqn 2):

$$P(t,T) = \frac{\lambda - \mu}{\lambda - \mu e^{-(\lambda - \mu)(T - t)}}. \tag{2.1}$$

The probability density of the branching events in the observed tree is proportional to the number of extant lineages multiplied by λ. It is thus possible to compute the likelihood of the phylogenetic data by multiplying together the probabilities of the observed branching events and the probabilities that the lineages survived to the present. This likelihood is

$$L_{\mathrm{P}} = (N-1)!\lambda^{N-2}\prod_{i=3}^{N} P^2(t_i,T)P^2(t_2,T). \tag{2.2}$$

This is actually a likelihood from Nee *et al.* (1994*a*, eqn 20) modified to take into account the fact that the lineages survived with one or more species until the present. Let us rewrite equation (2.1) as

$$P(t,T) = \frac{1-a}{1-a\mathrm{e}^{-rx}}, \tag{2.3}$$

where $a = \mu/\lambda$, $r = \lambda - \mu$ and $x = T - t$. If T is the present, x is in fact the branching time. Some algebra leads us to rewrite equation (2.2) as

$$L_{\mathrm{P}} = (N-1)!r^{N-2}(1-a)^N\exp\left(2r\sum_{i=2}^{N}x_i\right)\prod_{i=2}^{N}(\mathrm{e}^{rx_i}-a)^{-2}. \tag{2.4}$$

A logarithmic transformation of the above equation leads to an expression that is easier to manipulate:

$$\begin{aligned}\ln L_{\mathrm{P}} = {} & \ln(N-1)! + (N-2)\ln r + N\ln(1-a) \\ & + 2r\sum_{i=2}^{N}x_i - 2\sum_{i=2}^{N}\ln(\mathrm{e}^{rx_i}-a).\end{aligned} \tag{2.5}$$

The taxonomic data consist of a set of monophyletic groups of species for which numbers and dates of origin (the lengths of the terminal branches of the reconstructed tree) are both known. The probability of a lineage originating from a single species having n species after a time t is (Kendall 1948, eqn 16)

$$\frac{\mu(\mathrm{e}^{(\lambda-\mu)t}-1)}{\lambda(\mathrm{e}^{(\lambda-\mu)t}-\mu)} \quad \text{if } n_t = 0, \tag{2.6}$$

$$\left[1-\frac{\mu(\mathrm{e}^{(\lambda-\mu)t}-1)}{\lambda(\mathrm{e}^{(\lambda-\mu)t}-\mu)}\right](1-\zeta_t)\zeta_t^{n-1} \quad \text{if } n_t > 0, \tag{2.7}$$

where

$$\zeta_t = \frac{\lambda(\mathrm{e}^{(\lambda-\mu)t}-1)}{\lambda\mathrm{e}^{(\lambda-\mu)t}-\mu}. \tag{2.8}$$

The same reparameterization used above for the phylogenetic data ($a = \mu/\lambda$, $r = \lambda - \mu$) and an algebraic development lead us to rewrite equation (2.7) as

$$P(n_t = n) = \frac{(1-a)^2\mathrm{e}^{rt}}{(\mathrm{e}^{rt}-a)^{n+1}}(\mathrm{e}^{rt}-1)^{n-1} \quad \text{for } n \geqslant 1. \tag{2.9}$$

Multiplying together the corresponding probabilities for the N observed lineages gives the likelihood of the taxonomic data

$$L_{\mathrm{T}} = (1-a)^{2N}\prod_{i=1}^{N}\mathrm{e}^{rt_i}(\mathrm{e}^{rt_i}-1)^{n_i-1}(\mathrm{e}^{rt_i}-a)^{-(n_i+1)}, \tag{2.10}$$

and after logarithmic transformation

$$\begin{aligned}\ln L_{\mathrm{T}} = {} & 2N\ln(1-a) + r\sum_{i=1}^{N}t_i + \sum_{i=1}^{N}(n_i-1)\ln(\mathrm{e}^{rt_i}-1) \\ & - \sum_{i=1}^{N}(n_i+1)\ln(\mathrm{e}^{rt_i}-a).\end{aligned} \tag{2.11}$$

The log-likelihood of the combined phylogenetic and taxonomic data (denoted by $\ln L$) is $\ln L_{\mathrm{P}} + \ln L_{\mathrm{T}}$. Given some data, finding the maximum of $\ln L$ over a and r gives their respective maximum-likelihood estimates (MLEs) denoted $\hat{a}$ and $\hat{r}$. Under the assumption that $\ln L$ is approximately normally distributed around its maximum, it is possible to estimate the standard errors of the MLEs using its second derivatives.

Practically, maximizing $\ln L$ would require us to find its first partial derivatives with respect to a and r; however, these expressions are too complex to be solved analytically. I used instead a numerical minimization method. The likelihood was transformed as the deviance (equal to $-2\ln L$); finding the maximum likelihood was equivalent to minimizing the deviance. This was done with the nonlinear minimization function of *R* (Ihaka & Gentleman 1996), which uses a Newton-type algorithm for unconstrained minimization (Schnabel *et al.* 1985). This method allows one to find the minimum of a function even when its partial and second derivatives are unknown. However, these derivatives can be computed numerically, allowing one to calculate the standard errors of the MLEs.

If the normal approximation of the likelihood surface is not satisfied (e.g. owing to asymmetries), the standard errors of the MLEs based on the second derivatives of the log-likelihood may be incorrect. An alternative approach is to use the profile likelihood (Hudson 1971): this consists of looking at the likelihood surface around its maximum. As a difference between two log-likelihoods of less than or equal to 1.92 is not significant at $p > 0.05$ (Edwards 1992), the set of all likelihood values greater than or equal to the maximum likelihood minus 1.92 define a 95% confidence region for both parameters a and r. This is best visualized by plotting the likelihood surface for different values of the two parameters. The confidence intervals of a and r can then be transformed into confidence intervals of λ and μ by developing the inequalities defining these intervals (see Appendix A).

The programs implementing the method developed here are freely distributed in the package APE (analysis of phylogenetics and evolution; Paradis *et al.* 2003).

3. EXAMPLES

(a) *Birds*

The extensive study of Sibley & Ahlquist (1990) provided an estimate of the phylogenetic relationships for more than 1000 out of the 9646 living bird species based on DNA/DNA hybridization experiments (their famous 'tapestry'). This covered all avian orders, and almost all families. Sibley & Monroe (1990) reviewed the taxonomy of birds, and listed the number of species for all taxonomic levels, thus providing the numbers of species for the 23 orders and 144 families of living birds.

The goal of the present analysis was to estimate diversification rates for birds at the order and family levels. The data for 137 families were considered. The seven families not included here were groups with uncertain relationships (Sibley & Ahlquist 1990) and contained very few species: Brachypteraciidae (five species), Raphidae (three species), Mesitornithidae (three species), Philepittidae (four species), Callaeatidae (three species), Picathartidae (four species) and Hypocoliidae (one species). The resulting tree was not fully dichotomous and contained one trichotomy (see Sibley & Ahlquist 1990, fig. 356). This was resolved by assuming a null branch length below this node, so that the corresponding branching time was counted twice in the analysis. The original units of Sibley & Ahlquist (1990) ($\Delta T_{50}H$, a measure of the distance

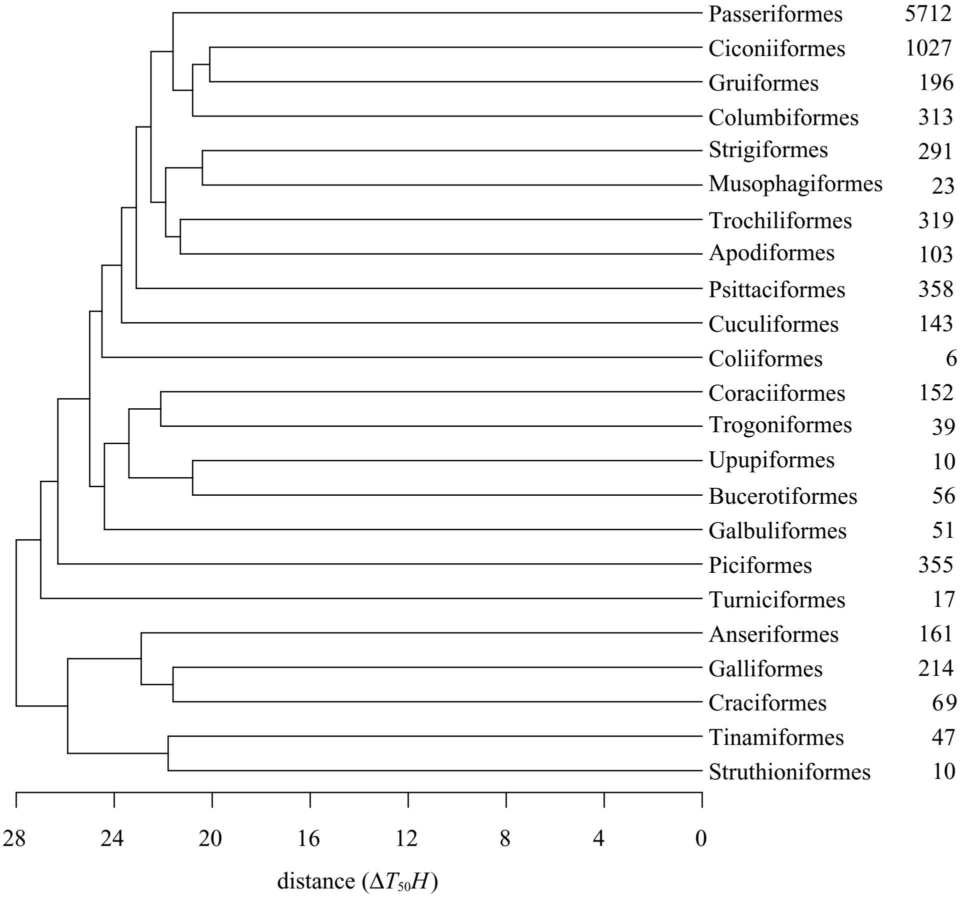

Figure 1. The phylogenetic relationships among the orders of birds according to Sibley & Ahlquist (1990). The numbers on the right-hand side are the number of species in each order according to Sibley & Monroe (1990).

between the tips of the tree) were used here for the branch lengths.

All 23 extant orders were considered: the resulting phylogeny was fully dichotomous (figure 1). As for families, the number of species were found in Sibley & Monroe (1990).

The parameter estimates for the bird-family data were $\hat{a} = 0$ and $\hat{r} = 0.401$ (s.e. = 0.006). The maximum log-likelihood was -427.45. A plot of the log-likelihood around its maximum shows 95% confidence intervals of [0, 0.008] for a and [0.385, 0.418] for r (figure 2*a*). The 95% confidence intervals for the speciation and extinction rates (in $\Delta T_{50}H$ time units) were $0.385 \leqslant \lambda \leqslant 0.421$ and $0 \leqslant \mu \leqslant 0.003$, respectively.

The same estimates for the analysis at the order level were $\hat{a} = 0$ and $\hat{r} = 0.287$ (s.e. = 0.007); the maximum log-likelihood was -144.58. The log-likelihood surface indicates 95% confidence intervals of [0, 0.040] for a and [0.268, 0.308] for r (figure 2*b*). The 95% confidence intervals for the speciation and extinction rates (in $\Delta T_{50}H$ time units) were $0.268 \leqslant \lambda \leqslant 0.321$ and $0 \leqslant \mu \leqslant 0.013$, respectively. It is noteworthy that the confidence intervals for λ estimated from family and order data do not overlap. This is discussed in § 4.

(b) *Eutherian mammals*

Wilson & Reeder (1993) listed 26 orders of mammals together with a complete taxonomic listing at the species level. However, there is, for the moment, no complete phylogeny with branch lengths of the orders of mammals. Douzery *et al.* (2003) reconstructed a phylogeny with dated nodes for 18 orders of mammals based on molecular data for 42 species. They used a non-parametric method that does not assume a molecular clock (Sanderson 1997), to estimate the divergence dates of the tree. I discarded the data for the order Diprotodontia (represented by the kangaroo *Macropus*), which was the only non-Eutherian order in this study. I used data for 17 orders in which the species richness of Cetacea (78 species) was pooled with that of Artiodactyla (220 species), thus leaving 16 orders. I considered only Eutherian mammals since the divergence between this clade and the other mammals is not yet well dated (E. J. P. Douzery, personal communication).

The tree was not fully dichotomous and contained one trichotomy (figure 3). This was resolved in the same way as for the phylogeny of avian families. The unit of branch lengths was millions of years ago (Ma) calculated from the ages of the nodes reported by Douzery *et al.* (2003).

The parameter estimates for this mammalian phylogeny were $\hat{a} = 0$ and $\hat{r} = 0.080$ (s.e. = 0.003). The maximum log-likelihood was -102.21. A plot of the log-likelihood around its maximum shows 95% confidence intervals of [0, 0.050] for a and [0.074, 0.088] for r (figure 4). The 95% confidence intervals for the speciation and extinction rates are $0.074 \leqslant \lambda \leqslant 0.093$ and $0 \leqslant \mu \leqslant 0.005$, respectively.

4. DISCUSSION

There is great interest in using phylogenetic data to study macroevolutionary processes as shown by recent

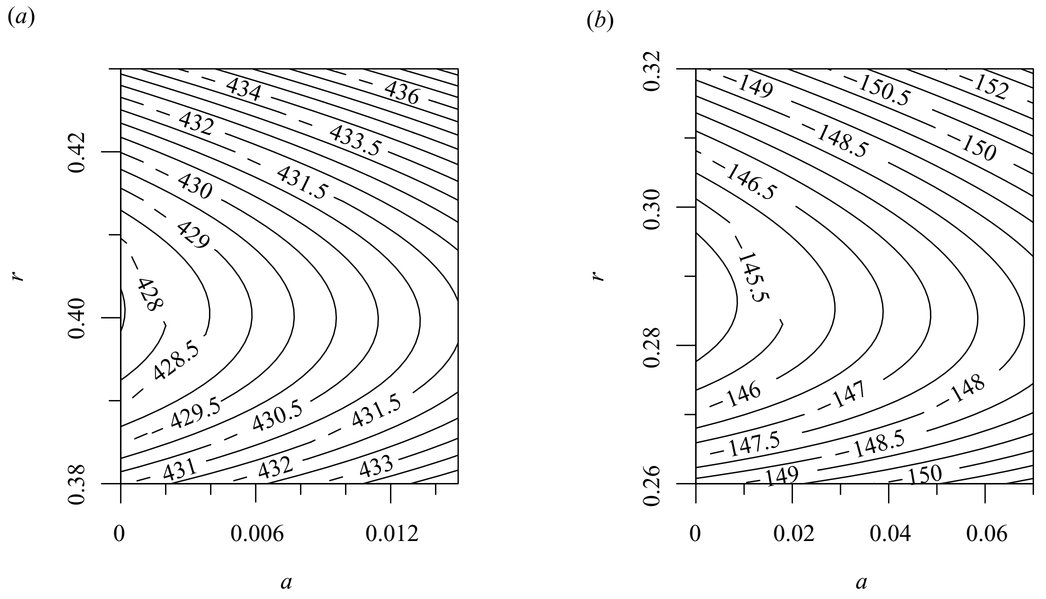

Figure 2. Likelihood surface of the diversification model described in § 2 for the analysis of the bird data. (*a*) Results at the family level. A 95% confidence region of the parameters is approximately given by the contour line −429.5. (*b*) Results at the order level. A 95% confidence region of the parameters is approximately given by the contour line −146.5.

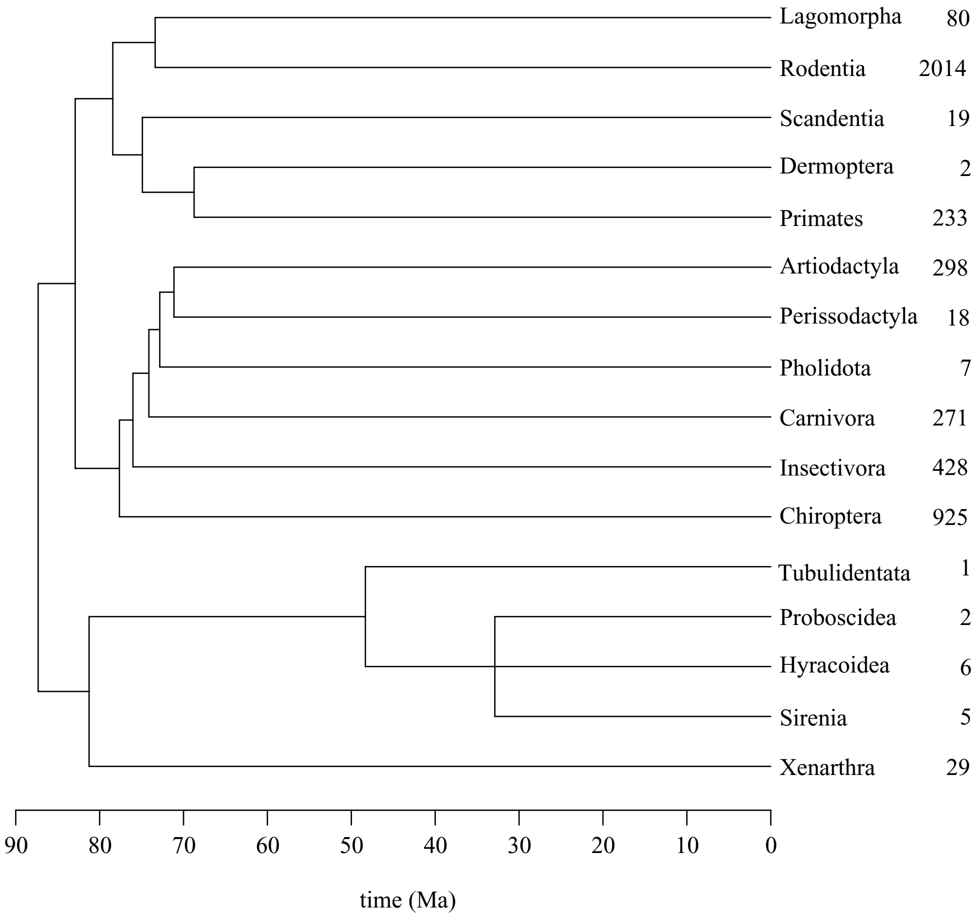

Figure 3. The phylogenetic relationships among the orders of Eutherian mammals according to Douzery *et al.* (2003). The numbers on the right-hand side are the number of species in each order according to Wilson & Reeder (1993).

reviews (Sanderson & Donoghue 1996; Barraclough & Nee 2001). Significant progress has been achieved when the available phylogeny is complete (i.e. has been reconstructed with all species; Slowinski & Guyer 1993; Harvey *et al.* 1994). However, the development of approaches for analysing incomplete phylogenies has been more difficult (Paradis 1997, 1998*b*; Pybus & Harvey 2000). The method presented in this paper is an attempt to combine

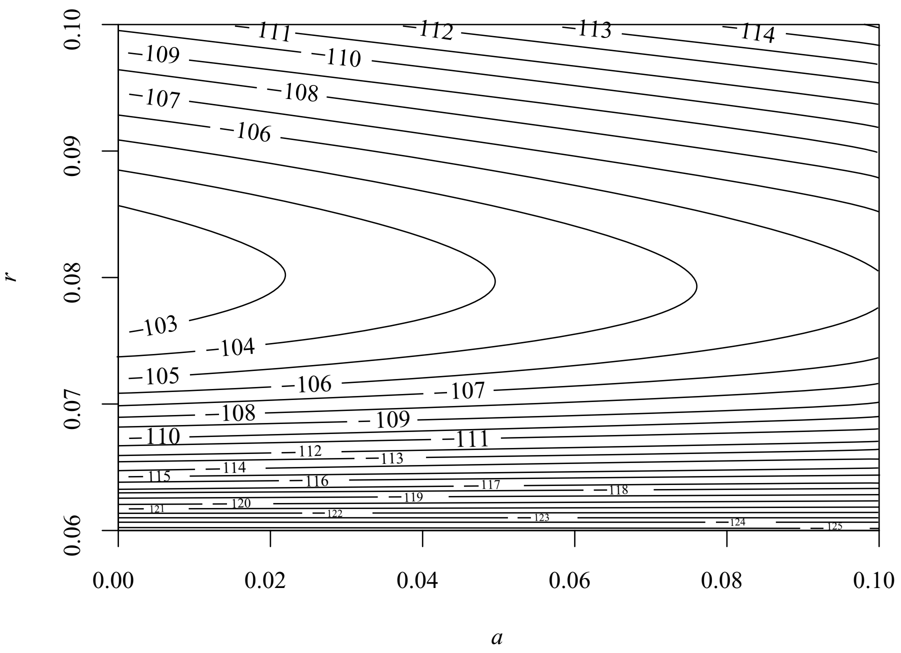

Figure 4. Likelihood surface of the diversification model described in § 2 for the analysis of the mammal data. A 95% confidence region of the parameters is approximately given by the contour line −104.

phylogenetic and taxonomic data to estimate speciation and extinction rates, and thus can be used to analyse incomplete phylogenies provided some data on species richness are available.

The present method can be viewed as an extension of the method introduced by Nee *et al.* (1994*a*). Both methods are based on the birth-and-death process studied by Kendall (1948). The main distinction between the two methods, particularly apparent in the computation of their respective likelihoods, is that Nee *et al.* (1994*a*) consider that all surviving lineages have one species living at present, whereas the present method simply assumes that these lineages survive (with one or more species).

It is interesting to point out that both approaches aim to estimate parameters rather than to test hypotheses. The method of Nee *et al.* (1994*a*) can be extended to more general birth–death models (e.g. assuming time dependence in both parameters, see Nee *et al.* 1994*a*) but, as far as I know, this has not been used in practice. Whether the approach presented here could be extended to more complex birth–death models requires more work. It would be interesting in the future to compare extensively the estimates from the method of Nee *et al.* (1994*a*) and those from the present one.

In a previous work, I developed a method to take into account missing species in phylogenetic data when estimating diversification rates (Paradis 1997, 1998*b*). This is based on making an analogy between branching times in a phylogeny and survival events: thus missing branching times can be treated as censored data (i.e. when time of death is not known precisely but it is known that death occurred after a certain time). This approach has the advantage of making available likelihood-based fitting of alternative models, and thus allows testing of alternative hypotheses on variation in diversification rates (Paradis 1997, 1998*b*). However, this has the inconvenient property that when the proportion of missing data is too large there is not enough information to allow efficient estimation. Thus, the approach based on survival models cannot be applied to the kind of data analysed in the above examples.

Though the approach presented here aims to analyse incomplete phylogenies, the data considered (branch lengths, tree topology and number of species) are assumed to be known without error. This assumption is likely to be untrue in some situations since the tree is usually an estimate and species numbers in some groups may be inaccurate, particularly in speciose or poorly known groups. The maximum-likelihood approach developed here could be, at least in theory, extended to take into account such uncertainties but this would certainly require complex mathematical developments. Furthermore, it is likely that in most situations the uncertainty in the data reflects conflicts between a limited number of scenarios (such as the attribution of a sub-group to a particular subclade) and thus it is possible to consider all of them alternatively.

Pybus & Harvey (2000) proposed a method to assess the impact of incompleteness in phylogenetic data on testing for constant diversification. This is based on simulating phylogenies with constant rates, then sampling randomly a subset of the species ‘living’ at the end of the simulation. This Monte Carlo approach makes it possible to estimate the distribution of their test statistic (the γ-statistic) in the presence of missing data.

Nee *et al.* (1994*b*) used simulations to assess the effect of missing data on analysis of diversification with a graphical method: the lineages-through-time plot (Harvey *et al.* 1994; Nee *et al.* 1995). They showed that the analysis of incomplete phylogenies may indicate a spurious decline through time of diversification (Nee *et al.* 1994*b*). There is still an obvious need for methods that would separate the effects of incomplete sampling from those of temporal variation in phylogenetic diversification rates.

The analyses presented in this paper have a mainly illustrative purpose, but the results obviously call for a few comments. I will first convert the rates estimated for birds into meaningful units (Ma^{-1}). The depth of the bird tree of Sibley & Ahlquist (1990) is 28 in $\Delta T_{50}H$ units. Paton *et al.* (2002), using various methods, dated the origin of the modern bird orders to between 110 and 130 Ma ago, depending on the method. We need thus, roughly, to multiply the time-scale of Sibley & Ahlquist (1990) by four to obtain a time-scale in Ma. Therefore, the corresponding rates in Ma^{-1} (after dividing the speciation and extinction rates by four) are $0.096 \leqslant \lambda \leqslant 0.105$ and $0 \leqslant \mu \leqslant 0.001$ for the family-based analysis, and $0.067 \leqslant \lambda \leqslant 0.080$ and $0 \leqslant \mu \leqslant 0.003$ for the order-based one.

Concerning birds, the analysis at the level of orders yielded smaller estimates and larger confidence intervals than the same analysis at the level of families. There could be two explanations for this discrepancy. First, the family-based data may be more accurate than the order-based data. Consequently, the analysis of the former may give more accurate estimates (with narrower confidence intervals) than that of the latter. Second, heterogeneity in rates may affect the present estimates. The present method assumes that speciation and extinction rates are constant through time and the same for all lineages. This is certainly not true for the present data since species richness is extremely uneven among bird families, with some passerines being extremely diverse (e.g. Fringillidae). It could be that at the level of orders the effect of heterogeneity was lessened since the heterogeneity among orders may be smaller than that among families. In other words, the family-based estimates may have been 'pushed up' by the many diversified families of passerines.

The estimates for mammals are close to those for birds, particularly for the order-based analysis of the latter group. This is in agreement with the slightly younger origin of the modern orders of Eutherian mammals (*ca.* 90 Ma) compared with birds (*ca.* 110 Ma). The Eutherian mammals have 4260 species (mammals total 4628 species according to Wilson & Reeder 1993) against 9623 for the birds. However, we need to keep in mind that rates are certainly heterogeneous in both groups: rodents, like passerines for birds, have surely diversified at a higher rate than other lineages of mammals (see Purvis *et al.* (1995) and Paradis (1998*b*) for analyses of heterogeneity in diversification rates among lineages of primates, and Mooers & Heard 1997 for a review). However, the present analyses suggest that, on average, birds and mammals have diversified at similar rates.

The three analyses reported in this paper yielded estimates of a equal to zero. It is obvious that some extinction events occurred during the history of mammals and birds, as evidenced by the fossil record. For instance, Purvis *et al.* (1995) reported estimates of a equal to zero in primate lineages in two cases out of four, and Nee *et al.* (1994*b*) reported similar instances in other groups. The method presented in this paper can, in theory, yield estimates of a greater than zero (Appendix B). One referee of the present paper pointed out that there is currently some confusion in the literature with respect to the issue of estimating extinction rates from phylogenies without fossils (e.g. Sims & McConway 2003). In the absence of an extensive theoretical treatment of this issue, it seems sound to treat the estimates of extinction rates inferred from phylogenies without fossils (complete or not) with great care. An interesting point would be to assess whether the confidence interval for the extinction rate inferred from the profile likelihood correctly covers the true value of μ.

The method presented in this paper surely has a wide range of potential applications. Future research will need to address the issues of temporal variation and heterogeneity in rates. Ultimately, linking the methods dealing with phylogenetic diversification and the methods dealing with the evolution of characters, such as the various phylogenetic comparative methods, should result in significant progress in evolutionary analysis.

The author is grateful to three anonymous referees for their constructive comments on an earlier version of this paper, and to E. Douzery for sending me his results on divergence dates for mammals. This research was financially supported by the Institut Français de la Biodiversité and the Centre National de la Recherche Scientifique. This is publication 2003-052 of the Institut des Sciences de l'Évolution (Unité Mixte de Recherche 5554 du Centre National de la Recherche Scientifique).

APPENDIX A

This appendix shows how the confidence intervals of the parameters a and r are transformed to give confidence intervals of the speciation and extinction rates (λ and μ, respectively). Suppose the profile likelihood of the model gives the two 95% confidence intervals

$$a_1 \leqslant a \leqslant a_2, \tag{A 1}$$

$$r_1 \leqslant r \leqslant r_2. \tag{A 2}$$

As we have $a = \mu/\lambda$ and $r = \lambda - \mu$, and under the condition that λ is positive and non-zero, we can rewrite both inequalities as

$$a_1\lambda \leqslant \mu \leqslant a_2\lambda, \tag{A 3}$$

$$r_1 + \mu \leqslant \lambda \leqslant r_2 + \mu. \tag{A 4}$$

Thus, we have for λ the inequalities

$$r_1 + a_1\lambda \leqslant \lambda \leqslant r_2 + a_2\lambda, \tag{A 5}$$

which can be written separately as

$$r_1 + a_1\lambda - \lambda \leqslant 0, \qquad 0 \leqslant r_2 + a_2\lambda - \lambda. \tag{A 6}$$

Developing both inequalities algebraically gives (under the condition that both a_1 and a_2 are less than 1)

$$\frac{r_1}{1 - a_1} \leqslant \lambda, \qquad \frac{r_2}{1 - a_2} \geqslant \lambda, \tag{A 7}$$

giving the 95% confidence interval for λ as

$$\frac{r_1}{1 - a_1} \leqslant \lambda \leqslant \frac{r_2}{1 - a_2}. \tag{A 8}$$

Substituting in the above inequalities of μ gives its 95% confidence interval as

$$\frac{a_1 r_1}{1 - a_1} \leqslant \mu \leqslant \frac{a_2 r_2}{1 - a_2}. \tag{A 9}$$

APPENDIX B

This appendix shows that the method presented in this paper can yield estimates of a of greater than 0. This can be shown by examining the first partial derivative of the log-likelihood function $\ln L$. As $\ln L = \ln L_{\mathrm{P}} + \ln L_{\mathrm{T}}$, we have

$$\frac{\partial \ln L}{\partial a} = \frac{\partial \ln L_{\mathrm{P}}}{\partial a} + \frac{\partial \ln L_{\mathrm{T}}}{\partial a}. \qquad \text{(B 1)}$$

Partially differentiating equations (2.5) and (2.11) with respect to a is straightforward

$$\frac{\partial \ln L_{\mathrm{P}}}{\partial a} = 2\sum_{i=2}^{N} \frac{1}{\mathrm{e}^{rx_i} - a} - \frac{N}{1-a}; \qquad \text{(B 2)}$$

$$\frac{\partial \ln L_{\mathrm{T}}}{\partial a} = \sum_{i=1}^{N} \frac{n_i + 1}{\mathrm{e}^{rt_i} - a} - \frac{2N}{1-a}; \qquad \text{(B 3)}$$

giving for the log-likelihood function

$$\frac{\partial \ln L}{\partial a} = 2\sum_{i=2}^{N} \frac{1}{\mathrm{e}^{rx_i} - a} + \sum_{i=1}^{N} \frac{n_i + 1}{\mathrm{e}^{rt_i} - a} - \frac{3N}{1-a}. \qquad \text{(B 4)}$$

If the partial derivative of $\ln L$ can be positive at $a = 0$, then $\ln L$ can have a maximum for positive values of a. Fixing $a = 0$ in equation (B 4) gives

$$2\sum_{i=2}^{N} \frac{1}{\mathrm{e}^{rx_i}} + \sum_{i=1}^{N} \frac{n_i + 1}{\mathrm{e}^{rt_i}} - 3N. \qquad \text{(B 5)}$$

The first two terms in equation (B 5) are always positive, whereas the third one is always negative. Thus, the above expression can be positive depending on the values of r, x_i, t_i and N. Consequently, the log-likelihood can potentially achieve a maximum for positive values of a.

REFERENCES

Barraclough, T. G. & Nee, S. 2001 Phylogenetics and speciation. *Trends Ecol. Evol.* **16**, 391–399.

Douzery, E. J. P., Delsue, F., Stanhope, M. J. & Duchon, H. 2003 Local molecular clocks in three nuclear genes: divergence times for rodents and other mammals, and incompatibility among fossil calibrations. *J. Mol. Evol.* (In the press.)

Edwards, A. W. F. 1992 *Likelihood*, expanded edn. Baltimore, MD: Johns Hopkins University Press.

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.

Hudson, D. J. 1971 Interval estimation from the likelihood function. *J. R. Statist. Soc.* B **33**, 256–262.

Ihaka, R. & Gentleman, R. 1996 R: a language for data analysis and graphics. *J. Comput. Graphical Statist.* **5**, 299–314.

Kendall, D. G. 1948 On the generalized "birth-and-death" process. *Ann. Math. Stat.* **19**, 1–15.

Mayhew, P. J. 2002 Shifts in hexapod diversification and what Haldane could have said. *Proc. R. Soc. Lond.* B **269**, 969–974. (DOI 10.1098/rspb.2002.1957.)

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

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

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

Nee, S., Holmes, E. C., Rambaut, A. & Harvey, P. H. 1995 Inferring population history from molecular phylogenies. *Phil. Trans. R. Soc. Lond.* B **349**, 25–31.

Paradis, E. 1997 Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. *Proc. R. Soc. Lond.* B **264**, 1141–1147. (DOI 10.1098/rspb.1997.0158.)

Paradis, E. 1998*a* 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.

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

Paradis, E., Claude, J. & Strimmer, K. 2003 APE: an R package for analyses of phylogenetics and evolution. *Bioinformatics* (In the press.)

Paton, T., Haddrath, O. & Baker, A. J. 2002 Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *Proc. R. Soc. Lond.* B **269**, 839–846. (DOI 10.1098/rspb.2002.1961.)

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

Pybus, O. G. & Harvey, P. H. 2000 Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. Lond.* B **267**, 2267–2272. (DOI 10.1098/rspb.2000.1278.)

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

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

Schnabel, R. B., Koontz, J. E. & Weiss, B. E. 1985 A modular system of algorithms for unconstrained minimization. *ACM Trans. Math. Software* **11**, 419–440.

Sibley, C. G. & Ahlquist, J. E. 1990 *Phylogeny and classification of birds: a study in molecular evolution*. New Haven, CT: Yale University Press.

Sibley, C. G. & Monroe Jr, B. L. 1990 *Distribution and taxonomy of birds of the world*. New Haven, CT: Yale University Press.

Sims, H. J. & McConway, K. J. 2003 Nonstochastic variation of species-level diversification rates within angiosperms. *Evolution* **57**, 460–479.

Slowinski, J. B. & Guyer, C. 1993 Testing whether certain traits have caused amplified diversification: an improved method based on a model of random speciation and extinction. *Am. Nat.* **142**, 1019–1024.

Wilson, D. E. & Reeder, D. M. (eds) 1993 *Mammal species of the world. A taxonomic and geographic reference*, 2nd edn. Washington, DC: Smithsonian Institution Press.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.