

Tempo and mode of evolution revealed from molecular phylogenies

SEAN NEE*, ARNE Ø. MOOERS*, AND PAUL H. HARVEY*†

*Agriculture and Food Research Council Unit of Ecology and Behaviour, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom; and †Natural Environment Research Council Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berks, SL5 7PY, United Kingdom

Communicated by Richard Southwood, May 27, 1992

ABSTRACT The analysis of the tempo and mode of evolution has a strong tradition in paleontology. Recent advances in molecular phylogenetic reconstruction make it possible to complement this work by using data from extant species.

Cladogenesis is the division of one evolutionary lineage into two. How often do lineages undergo cladogenesis to give daughter lineages that will survive for long periods of evolutionary time? And what characteristics of species determine rates of successful cladogenesis? Such questions are tackled traditionally by using the fossil record (1–8), the quality of which limits the accuracy of the answers obtained (9). However, paleontological data could be usefully complemented with information from extant species if the dates when pairs of species last shared a common ancestor were known (10). The advent of more complete molecular phylogenies for many extant species, together with the existence of molecular clocks, is beginning to provide such data. Here we analyze Sibley and Ahlquist's (11) phylogeny of the birds derived from DNA-DNA hybridization studies, to show what molecular phylogenies can tell us about macroevolution. We use those data because they are the most extensive yet available, and we caution that the validity of the results we report is limited, like paleontological analyses, by the quality of the data used. As molecular phylogenies become more accurate, so will the strength of the conclusions that can be drawn from them.

Density-Dependent Cladogenesis in Birds?

Have rates of cladogenesis changed through time and, if so, why? In this section, we tackle these questions using the phylogenetic tree and branch lengths that Sibley and Ahlquist (11) derived by subjecting their DNA-DNA hybridization data to Sokal and Michener's (12) UPGMA (unweighted pair group method using arithmetic averages) clustering procedure. Details of the species sampled, the genetic distance measure used, and an assessment of the assumptions made in the phylogenetic reconstruction are given by Sibley and Ahlquist (11). The analysis is based on hybridization data from a taxonomically widespread sample of some 1700 of the ca. 9700 species of living birds (13). We have restricted our analyses to that portion of the phylogenetic tree embraced by the earliest node and the 121st node. This portion includes all lineages whose extant species are separated by 10 or more of Sibley and Ahlquist's genetic distance units (species separated at their family level differ by between 9 and 11 units). We did not include more of the tree, because the species data set is incomplete, so that many of the more recent nodes are not recorded. Nevertheless, the data set contains representative species from 171 of the 174 families recognized by Wetmore (14) and, since we are using only the first 122 lineages, we feel sure that very few (if any) nodes leading to

extant lineages are missing from the time frame chosen. For the moment, we assume that, on average, there is a linear relationship between the genetic distance separating two taxa and the time when they last shared a common ancestor. The effect of relaxing this assumption is discussed below.

Molecular data provide information about the *effective* rate of cladogenesis, defined as the rate at which lineages that are still extant gave rise to new lineages that are also extant. Fossil data provide, in addition, information about lineages with no contemporary descendants, highlighting the complementarity of the two approaches to the study of the tempo and mode of evolution. The effective rate of cladogenesis has a simple qualitative relationship to the *instantaneous* rate of cladogenesis, defined as the lineage birth rate minus the death rate, which is a natural statistic for paleontological data. Suppose, for illustration, that birth and death rates were constant. Then the effective rate of cladogenesis would appear to *increase* close to the present because lineages that arose more recently are less likely to have gone extinct. So a constant or decreasing rate of effective cladogenesis implies a decreasing instantaneous rate.

Fig. 1 line A is a semilogarithmic plot of the number of lineages against time since the first bifurcation. As there is uncertainty about the calibration of the molecular clock pertaining to these data (15, 16), we measure time in arbitrary units since the time of the first bifurcation. The slope of this curve reflects the *per-lineage* rate of effective cladogenesis and, so, would appear to be a straight line, with stochastic wiggling, if the rate were constant through time. Instead, the rate appears to decrease quite smoothly over time. As the legend of Fig. 1 describes, the data are incompatible with a constant-rate model.

Following the failure of the one-parameter constant-rate model, we fitted a two-parameter density-dependent model, in which the *per-lineage* rate of cladogenesis is a function of the form p/N^α , where N is the number of lineages and α is a constant. The compatibility of this model with the data is good (Fig. 1), not only because there is an extra parameter to fit but also because of its qualitative form (the possession of a positive second derivative). This contrasts with the relatively poor performance of the two-parameter logistic model of density dependence, $p(1 - N/K)$, which is commonly used in population biology.

It has been suggested that the degree of molecular divergence between taxa is not related to time since divergence in a linear fashion, as we have assumed here, but that the genetic divergence between taxa was more rapid early in the history of an adaptive radiation (17). If this controversial (18, 19) suggestion is true, the deceleration in cladogenesis was even more rapid than we have estimated (i.e., a larger α). So our conclusions are qualitatively robust to this form of departure from linearity. More recently, Wayne *et al.* (20) have looked for departures from linearity by regressing logarithmically transformed molecular divergence data against logarithmically transformed fossil-dated temporal divergence data for two groups, primates and carnivores, that have relatively good fossil records. For the DNA-DNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

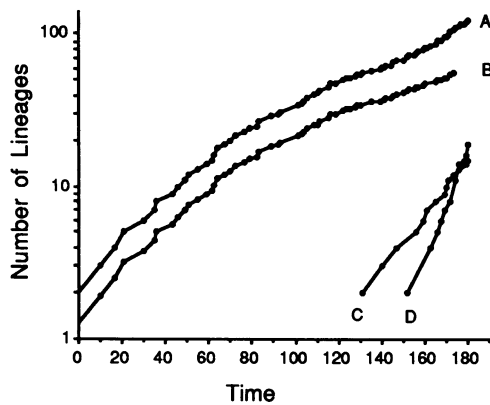


FIG. 1. Line plots of the number of lineages against time. Each point corresponds to a change in the number of lineages. Line A, the pattern of origination of all 122 lineages; line B, same as A, but without the Passeri (line C) and the Ciconiiformes (line D). Line B has been shifted downward to aid visual comparison. For convenience, time is measured in increments of 1 from time 0, when the first effective branching of the ancestral lineage occurred, to time 180. The original genetic distance scale goes from 28 to 10 in increments of 0.1, and each increment corresponds very approximately to 0.45 million years according to Sibley and Ahlquist (11). The phylogenetic tree consists entirely of dichotomous branches, so a lineage cannot give rise to more than one new lineage in one time unit (the single exception, a trichotomy at time 130, was treated as two dichotomies with nodes separated by 0 time units). The following statistical analysis was performed on the full data set of 122 lineages. If N_t is the number of lineages existing at time t , then $N_t - N_{t-1}$, denoted as X_t , is the number of lineages born between t and $t - 1$. Given the sequence N_t , the 180 X_t are modeled as independent binomial random variables with parameters N_{t-1} and p , where p is the per-lineage probability of effective cladogenesis. The maximum likelihood estimate for a constant p is 0.018268. With this estimate of p , the variables Z_t , where $Z_t = (X_t - \bar{X}_t)/[\text{Var}(X_t)]^{1/2}$, were constructed, each, by hypothesis, having a mean of 0 and a variance of 1. The variables Z_t are highly skewed, so the sum, $\sum Z_t^2$, which we will denote as ϕ^2 , is not expected to follow a χ^2 distribution. The distribution of ϕ^2 was determined by simulating the hypothesized Z_t 5000 times and we reject the hypothesis of a constant probability of cladogenesis ($\phi_{179}^2 = 245.75$, $P < 0.05$). The following analysis supports the visual impression that the failure of the constant-rate model is due to a slowdown in the rate of cladogenesis. Consider the linear model, $Z_t = \alpha + \beta(t - \bar{t}) + \epsilon_t$. The constant-rate model asserts that α and β are both 0. The least-squares estimate of β , b , is unbiased and, to a good approximation (confirmed by simulation), its sampling distribution is normal with a variance of $1/\sum(t - \bar{t})^2$. The observed slope, $b = -0.002827$, is significantly negative ($P < 0.025$). Finally, for each cladogenesis event early in the tree, we can compare the observed and expected times until the next event using a conventional χ^2 test. This analysis also excludes the constant probability model ($\chi_{13}^2 = 44.16$, $P < 0.005$). The parameters p and α in the p/N^α model of per-lineage cladogenesis probabilities were also estimated by the maximum-likelihood method: $p = 0.040305$ and $\alpha = 0.2028$. This model passes all the tests failed by the constant-probability model: $\phi_{178}^2 = 200.97$, $P < 0.2$; $b = -0.000369$, $P < 0.4$; $\chi_{12}^2 = 14.15$, $P < 0.9$. The performance of the constant-probability model on the reduced data set corresponding to line B is worse, and the performance of the density-dependent model is even better. The Passeri and Ciconiiformes radiations are, as they appear (lines C and D), quite compatible with a constant-probability model.

hybridization data, the primate slope of 1.19 did not differ significantly from 1.0, the expected value if rates of evolution do not change with divergence time. The carnivores, with a slope of 0.68, show rates of molecular divergence that appear to decrease the longer that lineages have been separated. Our conclusions for the bird data would not be qualitatively robust to this form of departure from linearity. However, Wayne *et al.*'s use of model 1 regression was inappropriate for the carnivore data, which had a relatively low correlation coefficient and an independent variable, fossil-dated diver-

gence time, which is subject to nontrivial error variance (21): the apparent carnivore slowdown may be artifactual (indeed, the rate of primate divergence may have sped up). For example, if the error variance in the carnivore relationship is attributed to fossil dating rather than molecular divergence, the slope of the double logarithmic plot is 0.94, which, for this data set, is statistically indistinguishable from 1.0. Clearly, the true estimate lies somewhere between 0.68 and 0.94, but whether the estimate differs significantly from 1.0 depends critically on the actual estimate that is used. The work of Wayne *et al.* is a very important beginning, and we await further developments before drawing firm conclusions.

Sibley and Ahlquist (11) suggest that rates of genetic divergence are more highly correlated with generation time than with real time. If they are correct, it is conceivable that the decrease in apparent rates of effective cladogenesis results from a (previously unsuspected) shortening of generation time, rather than a decrease in effective cladogenesis over the time frame examined. Either way, the pattern is of intrinsic biological interest.

Paleontologists have observed that instantaneous rates of cladogenesis are higher in the early history of adaptive radiations (22, 23), and it has been suggested that this slowdown is a result of niche filling (5, 23). As we have seen for the avian radiation, the effective rate of cladogenesis is higher earlier on as well. What is, perhaps, surprising is the way in which the effective rate decreases: in contrast to the logistic model, deceleration of the form $1/N^\alpha$ implies an initially rapid slowdown and then a leveling off, as opposed to, for example, a constant rate early on, generating exponential growth, and then a subsequent slowdown.

What could account for such a pattern? It is natural to think in terms of models such as those of Valentine (8), in which speciation and extinction occur in an ecological niche space. One may conjecture that instantaneous cladogenesis reflects two distinct processes: evolutionary processes such as sexual selection producing, from a coarse view, ecological equivalents, and ecological processes producing diversification into an adaptive landscape of distinct ecological niches or ways of life. A lineage that invades an empty niche, coarsely defined, may be assured of leaving descendants in the distant future, barring calamities such as comet impacts or niche usurpation by a quite different sort of beast. However, the subsequent history of cladogenesis *within* the coarsely defined niche may be well described by random speciation and extinction models in which the overall number of lineages is roughly constant (24). The form of density-dependent effective cladogenesis that we have presented can be interpreted in the light of such a niche-filling model. Perhaps, early in the adaptive radiation of the birds, there were a number of preexisting vacant niches waiting to be filled (such as "aerial predator of flying insects") and, as the diversification of the birds proceeded, new niches were then created (such as "aerial predator of birds").

Radiations Versus Background Cladogenesis

The previous section showed how Sibley and Ahlquist's phylogeny is compatible with a model in which, at a given point in time, all lineages have the same probability of cladogenesis per unit time, at least to the degree of discrimination of the statistical tests described in Fig. 1. To look for radiations, defined as clades that expanded to leave extant representatives at exceptionally high rates, we performed the following analysis. Considering the lineages in existence at time 80, what is the distribution of their progeny number by time 130, and what is the distribution of *their* progeny number by time 180? The analysis, described in Fig. 2, which seeks outliers from the expected geometric distributions, identifies the Passeri (the oscines or songbirds) and Ciconiiformes

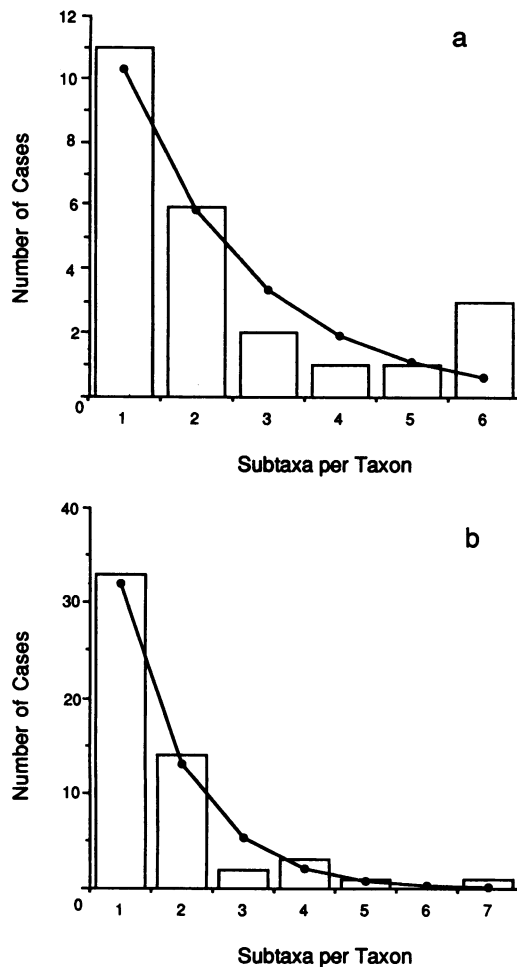


FIG. 2. Inspection of Fig. 1 suggests that over each time interval of 50 units, from 80 to 130 and from 130 to 180, we can treat the per-lineage cladogenesis rate as roughly a constant. As an excellent approximation for the following analysis, we treat effective cladogenesis as a birth process in *continuous* time over these intervals, in spite of the discrete time nature of the data being analyzed. As before, we suppose that each lineage has the same probability per unit time of giving rise to a new lineage. Under this hypothesis, the progeny distributions are expected to be geometric with parameter q (Eq. 1). q is related to p , the branching probability per time unit, by the formula $q \approx e^{-50p}$. For each time window, the inverse of the average number of progeny lineages is the maximum-likelihood estimate of q (note that a lineage which does not give rise to any new lineages over the time period is considered to give rise to one progeny lineage—*itself*). For the time window from 80 to 130 (a), the fit to the geometric distribution is excellent ($\chi^2 = 0.1376$, not significant). For the time period from 130 to 180, the probability of finding any lineage giving rise to more than 14 lineages is less than 0.005 (see below). This means that the Passeri and Ciconiiformes, with 15 and 19 subtaxa respectively, are anomalous. When we remove these taxa (b), this time window is also well described by the geometric series ($\chi^2 = 2.67$, not significant). The fitted parameters are as follows: 80 to 130, $q = 0.429$ ($p \approx 0.017$); 130 to 180 (Passeri and Ciconiiformes removed), $q = 0.593$ ($p \approx 0.010$); 130 to 180 (Passeri and Ciconiiformes included), $q = 0.448$ ($p \approx 0.016$). As expected from Fig. 1, q is greater for the second than the first time period, whereas p is less. The statistical identification of the Passeri and Ciconiiformes as outliers in the geometric distribution is based on the fact that if progeny numbers are geometrically distributed, then all vectors of progeny number are equally probable, as long as the elements sum to the observed total number of progeny. So, in this case, the statistics of progeny number distribution can be determined by randomly breaking a stick 125 units long into 56 fragments (where 125 is the total number of progeny and 56 is the number of ancestral lineages) and allowing breakages to occur only at unit boundaries. The lengths of the fragments correspond to progeny numbers and each broken stick

(including the sandgrouse, plovers, gulls, herons, flamingos, ibises, New World vultures, and storks) as radiations. As can be seen in Fig. 1, the Passeri radiation occurred at about the same rate as the original radiation of the birds, whereas the ciconiiform radiation exhibited a much higher rate. It will be interesting to see whether these radiations also exhibit the same pattern of deceleration as the total data set when the analysis can be reliably extended further toward the present.

The analysis of cladogenesis within specified time frames provides an alternative approach to the well-established problem of defining the distribution of subtaxa per taxon (species per genus or families per order, for example). In the first half of the century there was considerable interest in the fact that the distributions of many phenomena can be described by power laws—i.e., relationships of the form $Y = AX^{-d}$. Examples include the numbers of species in genera (26, 27), the population sizes of cities, incomes, and word use [Kendall (28) provides a general discussion]. Interest in power laws appears to be reviving partly because the exponent of the power law can be called a “fractal dimension.” Recently, Burlando (29) presented a large amount of evidence from many taxa that the number of species in genera follows a power law—that is, a relationship of the form $G = AN^{-d}$, where G is the number of genera with N species (see also ref. 30). The exponent d is usually between 1 and 2, so most genera have very few species, but some genera have a large number of species: the taxonomic “diversity” is dominated by a small number of genera (31). Plots of $\log G$ against $\log N$ typically look like Fig. 3.

The question to be put to such observations is whether we can infer anything from the exponent d about the evolutionary process that produced the contemporary taxa. To decide this, we need a model of how power laws arise. Over the century, several people have generated power-law distributions from birth–death stochastic processes (27, 32–34). These models are particular cases of the simple and general model we shall now derive. We can summarise the following by saying that multiplicative growth processes give rise to the one-parameter geometric distribution, and if this parameter is a variable, rather than a constant, the result is a power-law distribution. To make the presentation concrete, we will refer to the thing undergoing multiplicative growth as a “lineage.”

Consider a single lineage at time 0 that has a constant probability per unit time of either giving birth to a new lineage or dying. Each of its progeny lineages has the same parameters. Given that the lineage has *some* descendants at time t , the probability distribution of the number of daughter lineages, N , at time t is geometric (32); that is,

$$P(N = n) = q(1 - q)^{n-1}. \quad [1]$$

The geometric distribution arises not only from multiplicative processes, as discussed here, but also from divisional processes, such as the branch diameter reduction in successive “generations” of tree branching, for example.

The parameter q ($0 < q < 1$) is a function of t and the birth and death probabilities. Now suppose that q is itself a random variable with the two-parameter Beta density

corresponds to an equiprobable vector of progeny numbers. It is straightforward to repeatedly break sticks on a computer and so determine whatever statistics are of interest. We note that Slowinski and Guyer (25) derived a special case of the result that, for a random birth–death process, all vectors of progeny number are equally probable, the special case being two ancestral lineages. The general result, for an arbitrary number of ancestral lineages, arises immediately as a property of the geometric distribution.

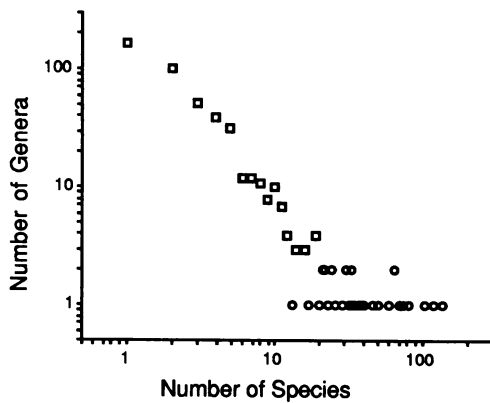


FIG. 3. Power-law simulation of the number of genera consisting of 1, 2, 3, . . . species. Each genus out of a total of 500 was assigned a number of species drawn at random from the Yule distribution, $B(n + 2, 2.5)$. This plot shows the usual features of real data [e.g., Burlando (29)], such as the long tail and the slight curvature. Following Burlando's procedure of fitting a model 1 regression line through only the square symbols (more than two genera) yields an exponent of -1.5 .

$$f(q) = \frac{1}{B(a, b)} q^{a-1} (1-q)^{b-1}, \quad [2]$$

where B is the Beta function. [The Beta density is a very flexible density on $(0,1)$, so this choice underpins our claims to generality.] Then the distribution of N becomes

$$P(N = n) \propto B(n + a - 1, b + 1). \quad [3]$$

Simon (33) derived distribution (3) from a pure birth process for the special case of $a = 1$ and named it the Yule distribution. We will use this name for the more general distribution (3) as well.

The Yule distribution exhibits power-law behavior (e.g., Fig. 3). For large n , $B(n + a - 1, b + 1) \propto 1/n^{b+1}$. We cannot fully discuss the behavior of stochastic realizations of the Yule distribution here, but, as a rule of thumb, the observed exponent, d , determined, for example, by regression, is typically smaller than $b + 1$. Depending on how rapidly convergence to the asymptotic exponent of $b + 1$ occurs, and on whether or not the data are sufficiently numerous for large values of n to be frequently seen, this may sometimes give the appearance of curvature to the double logarithmic plots.

Once the parameters a and b have been fitted by an appropriate procedure, such as maximum likelihood, the question arises as to what the implied distribution of q actually means. As far as phylogenies are concerned, the main question to be addressed about q is the extent to which the distribution reflects variation in t (i.e., how long ago the lineages giving rise to distinct taxonomic groups arose) versus variation in the birth and death rates among lineages that arose at the same time. The difficulty of answering this question may limit the usefulness of this sort of analysis of taxonomic structure. Furthermore, studies of nonmolecular taxonomies must also contend with the fact that these taxonomies capture an *interaction* of taxonomist psychology and evolutionary history.

Species Selection

Smaller-bodied taxa often contain more species than their larger-bodied relatives (35–41). Does this mean that small body size promotes successful cladogenesis? We calculated representative body weight values for 135 of Sibley and Ahlquist's bird families, using procedures described else-

where that do not lead to bias from speciose subtaxa (21, 42), and there is a highly significant negative correlation between body weight and the number of species in a family (Spearman rank correlation coefficient = -0.374 , $P = 0.0001$, $N = 135$; see also ref. 41). If this pattern results from clades of lighter birds having higher rates of effective cladogenesis, those clades should be found on shorter branches of the phylogenetic tree. Ancestral estimates for body weight were assigned to nodes above the family level of the Sibley and Ahlquist tree by calculating the mean of the daughter nodes, weighted by branch lengths [following Felsenstein (43)]. We tested whether the shorter daughter branch derived from each node ended in the lighter or heavier daughter clade. This method yielded 124 independent comparisons. The shorter daughter branch was as likely to end in the lighter daughter clade as in the heavier daughter clade, comparing all nodes (62 lighter/62 heavier) or specified subsets (Passeriformes, 17/21; Passeri, 15/14; Ciconiiformes, 16/11; 62 nodes nearest root, 27/35; none of the subsets differed from the null expectation of random allocation of the shorter daughter branch to the lighter versus the heavier daughter clade).

We also examined the quantitative differences between the lengths of sister branches leading to the lighter and heavier daughter clades. Given a random model of branch production, we transformed and standardized the differences so that they became samples drawn from a normal distribution with mean 0 and standard deviation 1. If we assume that, to a good approximation, the branch lengths to the lighter and heavier clades are drawn from an exponential distribution, then the variance of the logarithmically transformed branch lengths is $\psi^{(1)}(1) (= \pi^2/6)$, where $\psi^{(1)}$ is the trigamma function (44). (The assumption of an exponential distribution requires the ancillary assumption that the rate of effective cladogenesis does not change very much over the relevant time scale.) Under the null hypothesis that the branch lengths leading to the lighter and heavier daughter clades have the same expectation, the difference in the logarithmically transformed lengths has an expectation of 0 and a variance of $2\psi^{(1)}(1)$. The sum of N such differences, where N is large, is approximately normally distributed with an expectation of 0 and variance of $2N\psi^{(1)}(1)$. Transformation now allows a Z test to be performed. Notice that the null hypothesis does not assume that the expected branch lengths are the same all over the tree, merely the same for the two branches derived from each node. For all nodes and for a number of subsets, there were no significant differences between the lengths of branches leading to lighter and heavier daughter clades (all $N = 124$, $Z = 0.48$; Passeriformes, $N = 38$, $Z = -0.83$; Passeri, $N = 29$, $Z = 0.15$; Ciconiiformes, $N = 27$, $Z = 1.14$; 62 nodes nearest root, $N = 62$, $Z = -1.80$).

The original strong negative relationship that we described among families is, in fact, a result of taxonomic nonindependence. The pattern results from two monophyletic taxa—the five species-poor, generally large-bodied families of Eoaves (ostriches, rheas, cassowaries and emus, kiwis, and tinamous) and the 40 species-rich, small-bodied families of Passeriformes (or passerines, an order containing the suboscines and the oscines). When the Eoaves and the passerines are removed from the analysis, the relationship between body size and number of species in a family is nonsignificant (Spearman rank correlation coefficient = -0.19 , $P = 0.08$, $N = 90$). Body mass alone therefore does not seem to be a general correlate of effective cladogenesis within and among bird families.

There has been discussion about whether small body size was in some way responsible for the radiation of the passerines (41, 45–48). No resolution was possible, in part because analytical procedures were not available to cope with "the nonindependence of data points drawn from a hierarchically structured system" (p. 68 of ref. 41). The tests were per-

formed above, which were designed to deal with the problem of taxonomic nonindependence, reveal no correlation between rates of effective cladogenesis and body size. Those same tests can, however, be used to investigate which factors, if any, do correlate with rates of cladogenesis.

As more accurate molecular phylogenies become available (and they are getting better all the time), it should be possible to identify the causes of the different components of evolutionary success outlined in this article. The links that are being forged between paleontologists and molecular biologists promise a synergistic response that will not only provide answers to old questions in evolutionary biology but will also allow us to pose new questions for the future.

A. Purvis helped with data analysis for the section on species selection. We are grateful to Prof. J. Felsenstein, Prof. J. R. Krebs, Prof. R. M. May, Prof. J. Maynard Smith, Prof. C. G. Sibley, Dr. A. Grafen, and Dr. A. F. Read for help, advice, and criticism. This work was funded in part by European Economic Community Contract SCI*.0814.C.

1. Simpson, G. G. (1944) *Tempo and Mode in Evolution* (Columbia Univ. Press, New York).
2. Simpson, G. G. (1953) *The Major Features of Evolution* (Columbia Univ. Press, New York).
3. Stanley, S. M. (1979) *Macroevolution: Pattern and Process* (Freeman, San Francisco).
4. Raup, D. M. & Jablonski, D., eds. (1986) *Patterns and Processes in the History of Life* (Springer, Berlin).
5. Hallam, A., ed. (1977) *Patterns of Evolution, As Illustrated by the Fossil Record* (Elsevier, Amsterdam).
6. Valentine, J. W., ed. (1985) *Phanerozoic Diversity Patterns: Profiles in Macroevolution* (Princeton Univ. Press, Princeton, NJ).
7. Sepkoski, J. J. (1988) *Paleobiology* 14, 221–234.
8. Valentine, J. W. (1990) in *Population Regulation and Dynamics*, eds. Hassell, M. P. & May, R. M. (The Royal Society, London), pp. 261–268.
9. Raup, D. M. (1979) *Carnegie Mus. Nat. Hist. Bull.* 13, 85–91.
10. Harvey, P. H., Nee, S., Mooers, A. Ø. & Partridge, L. (1992) in *Genes in Ecology*, eds. Berry, R. J. & Crawford, T. J. (Blackwell, Oxford), in press.
11. Sibley, C. G. & Ahlquist, J. E. (1990) *Phylogeny and Classification of Birds* (Yale Univ. Press, New Haven, CT).
12. Sokal, R. R. & Michener, C. D. (1958) *Univ. Kansas Sci. Bull.* 38, 1409–1438.
13. Sibley, C. G. & Monroe, B. L. (1990) *Distribution and Taxonomy of Birds of the World* (Yale Univ. Press, New Haven, CT).
14. Wetmore, A. (1960) *Smithson. Misc. Coll.* 139, 1–37.
15. Cracraft, J. (1987) *Evol. Biol.* 21, 47–96.
16. Sibley, C. G., Ahlquist, J. E. & Sheldon, F. H. (1987) *Evol. Biol.* 21, 97–125.
17. Goodman, M. (1981) *Prog. Biophys. Mol. Biol.* 37, 105–164.
18. Runnegar, B. (1982) *Lethaia* 15, 199–205.
19. Raup, D. M. (1987) in *Rates of Evolution*, eds. Campbell, K. S. W. & Day, M. F. (Allen and Unwin, London), pp. 1–14.
20. Wayne, R. K., Van Valkenburgh, B. & O'Brien, S. J. (1991) *Mol. Biol. Evol.* 8, 297–319.
21. Harvey, P. H. & Mace, G. M. (1982) in *Current Problems in Sociobiology*, ed. Kings College Sociobiology Group (Cambridge Univ. Press, Cambridge, U.K.), pp. 343–361.
22. Gould, S. J., Raup, D. M., Sepkoski, J. J., Schopf, T. J. M. & Simberloff, D. S. (1977) *Paleobiology* 3, 23–40.
23. Stanley, S. M. (1973) *Syst. Zool.* 22, 486–506.
24. Raup, D. M., Gould, S. J., Schopf, T. J. M. & Simberloff, D. S. (1973) *J. Geol.* 81, 525–542.
25. Slowinski, J. B. & Guyer, C. (1989) *Am. Nat.* 134, 907–921.
26. Willis, C. J. (1922) *Age and Area: A Study in Geographical Distribution and Origin of Species* (Cambridge Univ. Press, Cambridge, U.K.).
27. Yule, G. U. (1924) *Philos. Trans. R. Soc. London A* 213, 21–87.
28. Kendall, M. G. (1961) *J. R. Stat. Soc. A* 124, 1–16.
29. Burlando, B. (1990) *J. Theor. Biol.* 146, 99–114.
30. Williams, C. B. (1964) *Patterns in the Balance of Nature and Related Problems in Quantitative Ecology* (Academic, London).
31. Dial, K. P. & Marzluff, J. M. (1989) *Syst. Zool.* 38, 26–37.
32. Kendall, D. G. (1948) *Biometrika* 35, 6–15.
33. Simon, H. A. (1955) *Biometrika* 42, 425–440.
34. Anderson, R. M., Gordon, D. M., Crawley, M. J. & Hassell, M. P. (1982) *Nature (London)* 296, 245–248.
35. Hutchinson, G. E. & MacArthur, R. H. (1959) *Am. Nat.* 93, 117–125.
36. Fleming, T. H. (1973) *Ecology* 54, 555–563.
37. Stanley, S. M. (1973) *Evolution* 27, 1–26.
38. Van Valen, L. (1973) *Evolution* 27, 27–35.
39. May, R. M. (1986) *Ecology* 67, 1115–1126.
40. Dial, K. P. & Marzluff, J. M. (1988) *Ecology* 69, 1620–1624.
41. Kochmer, J. P. & Wagner, R. H. (1988) *Syst. Zool.* 37, 68–69.
42. Harvey, P. H. & Pagel, M. D. (1991) *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, Oxford).
43. Felsenstein, J. (1985) *Am. Nat.* 125, 1–15.
44. Cox, D. R. & Lewis, P. A. A. (1966) *The Statistical Analysis of Series of Events* (Methuen, London).
45. Raikow, R. J. (1986) *Syst. Zool.* 35, 255–259.
46. Vermeij, G. J. (1988) *Syst. Zool.* 37, 69–71.
47. Fitzpatrick, J. W. (1988) *Syst. Zool.* 37, 71–76.
48. Raikow, R. J. (1988) *Syst. Zool.* 37, 76–79.