On the Probability of Ancestors in the Fossil Record

Mike Foote

Paleobiology, Vol. 22, No. 2 (Spring, 1996), 141-151.

Paleobiology, 22(2), 1996, pp. 141–151

On the probability of ancestors in the fossil record

Mike Foote

Abstract.—Three homogeneous models of species origination and extinction are used to assess the probability that ancestor-descendant pairs are preserved in the fossil record. In the model of cladogenetic budding, a species can persist after it branches and can therefore have multiple direct descendants. In the bifurcation model, a species branches to give rise to two distinct direct descendants, itself terminating in the process. In the model of phyletic transformation, a species gives rise to a single direct descendant without branching, itself terminating in the process. Assuming homogeneous preservation, even under pessimistic assumptions regarding the completeness of the fossil record, the probability of finding fossil ancestor-descendant pairs is not negligible. Even if all species of Phanerozoic marine invertebrates in the paleontologically important taxa had the same probability of preservation, on the order of 1%–10% or more of the known fossil species would be directly ancestral to other known fossil species. However, this is likely to be an underestimate, since the probability of finding ancestor-descendant pairs is enhanced by taxonomic, temporal, and spatial heterogeneities in preservation probability. Moreover, indirect genealogical relationships substantially increase the probability of finding ancestor-descendant pairs. The model of budding, the only one in which an ancestor can persist after a branching event, predicts that half or more of extant species have ancestors that are also extant. Thus, the question of how to recognize ancestor-descendant pairs must be carefully considered.

Mike Foote. *Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637*

Accepted: November 28, 1995

General Problems

Considering that the fossil record is incomplete, how likely are we to find pairs of species related, directly or indirectly, as ancestor and descendant? This question is of fundamental importance to a number of areas in evolutionary paleobiology, including speciation and evolutionary trends as well as genealogy. Recent discussions (e.g., Fisher 1994; Huelsenbeck 1994; Smith 1994; Alroy 1995) attest to the timeliness of the issue. Yet, the recognition of ancestor-descendant (A-D) pairs in the fossil record remains a major difficulty (Engelmann and Wiley 1977; Paul 1992; Fisher 1994; Smith 1994). We could safely tolerate or neglect this problem if the fossil record were so incomplete as to make the recovery of an A-D pair very unlikely (Hull 1979: p. 429; Schoch 1986: pp. 166–167). On the other hand, if the probability of recovering A-D pairs were high, then we would feel compelled to modify phylogenetic methods in order to recognize them with greater confidence.

By the *ancestor* of a species, I mean another species including the organisms from which descended the species in question. If part of the evolving lineage between the ancestor and descendant is considered, by whatever species concept one adopts, to be a species distinct from both the ancestor and descendant, then an A-D relationship is said to be *indirect.* Otherwise, it is *direct* (see Fig. 1). These definitions essentially agree with those of Paul (1992). Although a more restricted concept of ancestors is sometimes demanded, for example, one specifying the actual deme(s) or individual(s) from which the descendant arose, the broader view is perhaps more often in line with common evolutionary usage. I emphasize that my goal is not to specify how to recognize A-D pairs, but rather to assess the probability that we actually have such pairs in our various fossil collections, whether or not we are aware of it. How to recognize such pairs remains an important challenge for phylogenetic analysis (Alroy 1995).

Fortey and Jefferies (1982) used simulated degradation of evolutionary trees to explore the effect of incompleteness on the accuracy of cladograms and the probability of recovering A-D pairs. I will extend this approach by using homogeneous birth-death models (Kendall 1948; Raup 1985) with homogeneous

0096-8373/96/2202-0002/$1.00

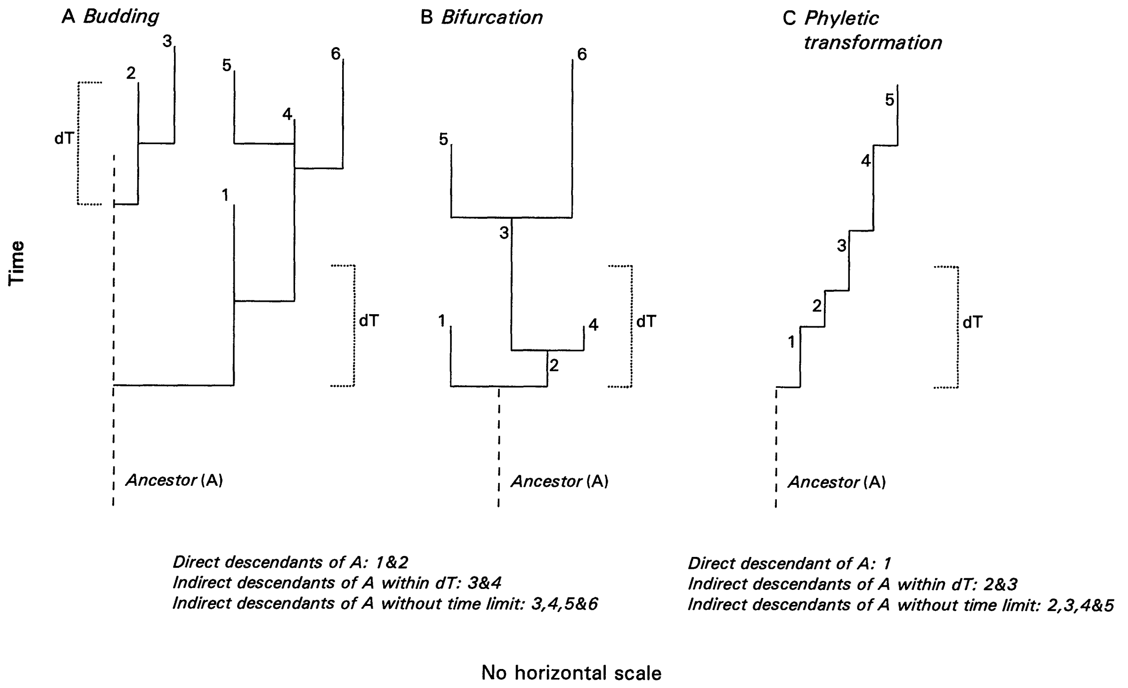

FIGURE 1. Birth-death models. A, Budding cladogenesis. A species may persist after it leaves a descendant. B, Bifurcation. A species is considered to terminate and leave two descendants at the point of branching. C, Phyletic transformation. A species either changes into a new species (pseudoextinction; transitions to species 1–5) or the lineage truly terminates (extinction; end of species 5). Horizontal lines are not meant to imply sudden morphological transitions; models in A and C should not be construed as punctuated equilibrium and punctuated anagenesis.

preservation (Foote and Raup 1996) to predict analytically the probability of recovering A-D pairs from the fossil record. Specifically, I will address the question: *What is the probability that a preserved species has at least one descendant species that is also preserved?* (Note that by *preserved* I mean *preserved and recovered.*) I will address this question first for direct descendants, then for indirect descendants as well. Note that I am not assessing the probability that a preserved species' ancestor is also preserved. This is because, for genealogical analyses of fossil taxa, it is useful to know which taxa should be in an ancestral as opposed to a terminal position. I will argue that the probability of finding A-D pairs is not negligible. Therefore, the use of a homogeneous preservation model, which actually decreases the probability of recovering A-D pairs relative to more realistic models incorporating taxonomic, spatial, and temporal variation in preservation probability, is conservative. For simplicity, I will emphasize results for the case where speciation and extinction rates are equal to each other. When speciation rate is higher, as during an evolutionary radiation, the probability of recovering A-D pairs is even greater (Appendix 1).

Modeling

Two alternative models of the branching process are that in which an ancestral species persists after a descendant arises (Fig. 1A), and that in which speciation is considered to coincide with the termination of the ancestral species and the origin of two distinct descendants (Fig. 1B). I will refer to these as *budding* and *bifurcation,* respectively. (Wagner and Erwin [1995] referred to these as *cladogenesis* and *bifurcation,* but both represent what would generally be considered cladogenetic processes. As long as we describe the branching patterns explicitly, it is unimportant which names we choose to hang on them.) An additional model not involving branching is the origin (and termination) of species by *phyletic transformation* (Fig. 1C). (Note that the horizontal lines in Figure 1 do not imply sudden morphological shifts associated with the new species; rather, they simply denote a point in time, possibly arbitrary, at which we would recognize and name a new species if we had the relevant fossil record. Thus, for example, the models of Figures 1A and 1C should not be construed as equivalent to punctuated equilibrium and punctuated anagenesis, respectively.)

With budding cladogenesis, a species may have multiple descendants, and species that are longer-lived leave more descendants, on average. As developed here, this model implicitly assumes that the duration of a species is not greatly affected by its descendants. The most extreme violation of this assumption, in which a descendant immediately causes the extinction of its ancestor, yields a model effectively equivalent to that of phyletic transformation. Because a longer-lived species is, all else equal, more likely to be preserved, the fossil record should be biased toward species with more descendants. This in turn enhances the probability of finding A-D pairs. Note that this model considers two species to represent an A-D pair even if the ancestral species is preserved only after it has given rise to a descendant. The stricter requirement that an ancestor must be preserved before its descendant is implicit in the next two models.

With bifurcation, a species has either no descendants or two descendants, and the ratio of probabilities of these two outcomes is simply the ratio between the extinction and origination rates. Because species are terminated by extinction or origination, they are shorter-lived than in the budding model, and therefore less likely to be preserved.

In the extreme case of phyletic transformation without branching, a species has either no descendants or one descendant, and the ratio of the corresponding probabilities is the ratio between the rate of true lineage termination and the rate of pseudoextinction. It is unlikely that many major clades have evolved for substantial periods of time without branching (Raup 1985: eqs. A11–A14). However, consideration of the pure models of budding and phyletic transformation allows us to investigate end members of the birth-death process.

We can consider at least three concepts of what constitutes a descendant (Fig. 1): (1) We may focus on direct descendants only. (2) We may incorporate indirect descendants originating within some arbitrary time, dT, after the origination of direct descendants. (3) We may include indirect descendants without limit of time. Depending on the question, one of these concepts may be more relevant than the others, but it is instructive to consider all three.

The probability of finding A-D pairs (Appendix 1) is illustrated in Figure 2A, B for the case where speciation rate (p) and extinction rate (q) are equal to 0.25 per species per million years, corresponding to Raup's (1991) estimate of mean species duration for Phanerozoic marine animals. Figure 2C illustrates the case where true extinction rate (q) and pseudoextinction rate (w) are each equal to 0.25. The general form of the relationships in Figure 2 is the same for other values of these parameters. Other values are explored explicitly below.

In general, the greater the proportion (P_P) of species that are preserved and recovered (which I will refer to as *completeness*; see Valentine 1989), the greater the proportion (P_A) of these preserved species that are ancestral to *at least* one other preserved species (see Fig. 2A for an exception to this monotonic relationship). As explained in Appendix 1, P_A is derived from a particular birth-death model with its parameters and a particular probability of preservation (R) per species per time interval. Because, for a given birth-death model with its parameters, there is a one-to-one correspondence between values of R and values of P_P, it is possible to calibrate P_A directly against P_P rather than R. Moreover, because P_P may be easier to estimate than R (Durham 1967; Valentine 1970), and P_P, unlike R, is independent of the level of stratigraphic resolution (Foote and Raup 1996), this calibration of P_A against P_P can be rather convenient. It will therefore be used in what follows.

Considering only direct descendants in the model of budding, P_A is roughly twice P_P when P_P is small and thereafter increases more slowly than P_P. Since P_A is the probability that a preserved species, not an arbitrarily chosen species, is an ancestor to another pre-

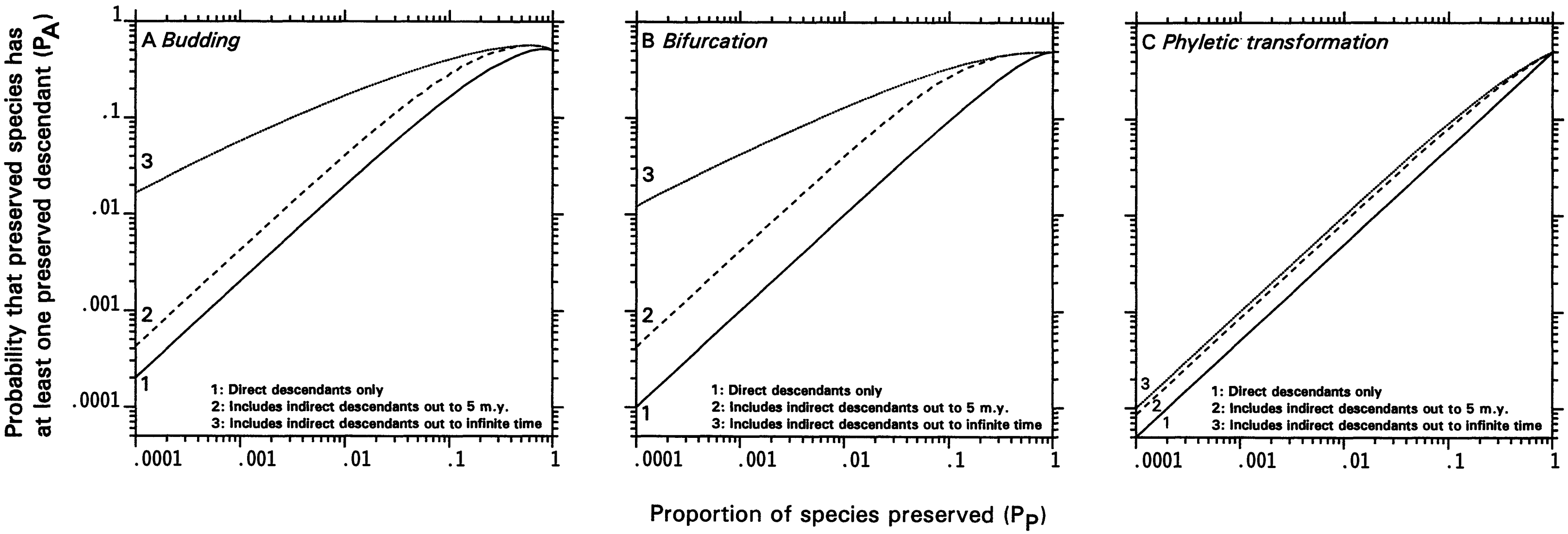

FIGURE 2. Probability of finding ancestor-descendant pairs (P_A) as a function of the proportion of species preserved at least once in the fossil record (P_P). In A and B, stochastically constant origination and extinction rates are each equal to 0.25 per species per million years (see Raup 1991), and in C the pseudoextinction and extinction rates are each equal to 0.25. Similar curves are obtained with other turnover rates. Although every species has an ancestor, not all species leave descendants. When origination and extinction rates are equal, half of all species die without issue; therefore P_A is equal to 0.5 when all species are preserved. A, Budding. P_A rises above 0.5 at values of P_P less than unity, reflecting bias of fossil record in favor of longer-lived species with more descendants when preservation is not complete. B, Bifurcation. Because fewer direct descendants can be produced than with budding, curve 1 is lower than in A. C, Phyletic transformation. Even fewer direct descendants are possible, so curve 1 is lower still. Because there is no branching in this extreme model, the incorporation of indirect descendants has a much smaller effect than it does in the cladogenetic models of budding and bifurcation. Note that for the probability of finding direct ancestor-descendant pairs the maximal difference between the models of budding and phyletic transformation is a factor of four. Models A and C represent end-members of the birth-death process; combinations of the three models therefore predict intermediate results relative to these extremes.

served species, P_A need not be less than or equal to P_P. It may seen counterintuitive that, with the budding model, the maximal value of P_A occurs when P_P is less than unity. However, this result makes sense given the bias of the fossil record in favor of longer-lived species with more descendants under this model (see above). Considering only direct descendants in the bifurcation model, P_A is roughly equal to P_P when P_P is small and thereafter increases more slowly than P_P. Considering only direct descendants in the model of phyletic transformation, P_A is equal to one-half P_P (Appendix 1).

Note that for direct descendants the maximal difference in P_A among the three models is a factor of four; this differences diminishes as the fossil record becomes more complete. It may seem counterintutive that, in all three models, P_A is equal to 0.5 when all species are preserved. This is simply because while each species has an ancestor not all species leave descendants. In randomly branching trees, the proportion of branches that terminate without issue is equal, in the limit of infinite time, to $q/(p+q)$, which is equal to 0.5 when $p = q$. Likewise, with true extinction and pseudoextinction occurring randomly in time, the proportion of species that terminate before transforming phyletically to a new species is equal to $q/(w+q)$, which equals 0.5 when $w = q$.

Thus, in the worst, unrealistic case, i.e., adopting a model of phyletic transformation with no branching and counting only direct descendants, the probability that a preserved species is ancestral to another preserved species is simply equal to one-half the probability that the species is preserved in the first place. For example, if P_P were equal to 1%, then, of each 20,000 species that ever existed, some 10,000 of them would in fact be ancestral to others, but only 200 species would be preserved and, on average, only one of these preserved species (0.5%) would be directly ancestral to another preserved species. Including indirect descendants of course increases the probability that a preserved species will be ancestral to at least one other preserved species (Fig. 2). The effect of including indirect descendants diminishes as completeness of the fossil record increases.

Empirical Scaling

Estimating the probability of finding A-D pairs requires some estimate of completeness. A common approach to estimating completeness at the global scale is to compare the number of known fossil species to some estimate of the total progeny (Kendall 1948), i.e., the number of species that have lived in some group (Simpson 1952; Newell 1959; Durham 1967; Valentine 1970;). If we focus on the paleontologically important marine groups, we have on the order of 200,000 described fossil species (Raup and Stanley 1978) and 150,000 species extant today (Valentine 1970) (although the latter number may need to be increased by a factor of two or more [May 1994]). We can estimate the total number of species that have lived on the basis of some diversity model and some turnover rate, both empirically scaled (Valentine 1970). Because incompleteness prevents our knowing the true diversity curve, it is useful to consider some reasonable extremes. Raup (1991) has estimated species-level extinction for marine animals on the whole at 25% per million years, on average. However, incomplete preservation truncates stratigraphic ranges and tends to exaggerate extinction rate. It is therefore worth considering a number of different extinction rates for each diversity model (Table 1). The estimates in Table 1 are based on the budding model (Fig. 1A), but these numbers could easily be adjusted to reflect other birth-death models (Fig. 2; Appendix 1).

Because standing diversity has almost surely increased since the Cambrian (Bambach 1977; Sepkoski et al. 1981), an extreme diversity model would assume constant standing diversity throughout the Phanerozoic, with a fixed proportion of species becoming extinct every million years, immediately replaced by others. This model is fully deterministic, with no stochastic variation. The assumption of constant diversity almost certainly overestimates the number of species that have lived and therefore underestimates completeness. An alternative model, perhaps extreme in the other direction, is based on a stochastic, exponential increase in diversity over the Phanerozoic (see Table 1, note 1). In this model, an

TABLE 1. Estimated total progeny (M), proportion of species preserved (P_P), and probability of finding direct ancestor-descendant pairs (P_A) for three different models of global Phanerozoic diversity of the paleontologically important marine groups, assuming budding birth-death model and homogeneous preservation with probability corresponding with the given value of P_P.

Extinction rate (per species per m.y.)	Median species duration (m.y.)	Model								
		Deterministic, constant diversity			Deterministic, linear diversity increase			Stochastic, exponential diversity increase[1]		
		M	P_P	P_A	M	P_P	P_A	M[2]	P_P	P_A
0.05	13.8	4.1×10^6	0.049	0.081	2.0×10^6	0.098	0.15	5.2×10^5	0.38	0.41
0.10	6.9	8.1×10^6	0.025	0.043	4.1×10^6	0.049	0.081	9.3×10^5	0.22	0.37
0.25	2.8	2.0×10^7	0.010	0.019	1.0×10^7	0.020	0.036	2.4×10^6	0.083	0.14
0.50	1.4	4.1×10^7	0.0049	0.0091	2.0×10^7	0.0098	0.019	4.9×10^6	0.041	0.070
1.0	0.7	8.1×10^7	0.0025	0.0048	4.1×10^7	0.0049	0.0091	1.1×10^7	0.018	0.033

[1] Strictly speaking, this diversity model is not exponential. For a given origination and extinction rate, the expected diversity over time follows an exponential trajectory if this expectation includes the possibility of prior extinction of the clade in question (see Kendall 1948; Raup 1985). The expected diversity of life over the Phanerozoic, *assuming survival of life to the Recent*, resembles exponential growth somewhat but increases extremely rapidly early on. This is intuitively reasonable, as clades that persist for a long time are likely to be those that get off to a good start.

[2] The expected total progeny of this stochastic process is based on Raup (1985: eq. A30). This equation is problematic, as the rationale behind it is not immediately obvious (D. M. Raup personal communication 1994). I have found, by comparing results of the equation to simulation results, that the equation is accurate when origination and extinction rates are roughly equivalent. However, with large differences between origination and extinction rates, the equation predicts values of total progeny that disagree with simulation results. The reader should therefore be cautioned against using this equation when rates of origination and extinction differ substantially from each other. With high-speed computers, it is a simple task to use simulation to determine whether the equation is greatly biased for given values of origination rate, extinction rate, and clade age.

extinction rate is assumed and a corresponding origination rate is chosen to yield an expected increase in standing diversity from 1 to 150,000 over the Phanerozoic, *assuming survival of marine animals to the Recent* (Raup 1985: eq. A25). An intermediate model is based on a deterministic, linear increase in diversity from the start of the Phanerozoic to the Recent.

Table 1 shows the probability of finding A-D pairs for each of these three models, for a number of different extinction rates. The estimates of completeness corresponding to the extinction rate of 0.25 are concordant with prior estimates (e.g., Durham 1967; Valentine 1970). Even with rather pessimistic assumptions, the proportion of preserved species directly ancestral to other preserved species is on the order of about 1% or higher.

Basing P_A on global completeness and on the assumption of homogeneous preservation clearly underestimates the probability of finding A-D pairs. There is reason to believe that completeness is higher on a local scale than on a global one, i.e., that the fossil record is rather complete where it occurs, while incompleteness stems largely from the loss or absence of fossiliferous rock rather than from the failure of species to enter the fossil record in the first place (Schopf 1978; Valentine 1989; Foote and Raup 1996). If ancestors and descendants are more likely to occur in the same geographic region than would be expected from a chance distribution of species, then the higher completeness at smaller geographic scales implies a greater probability of finding ancestor-descendant pairs. Finally, the fossil record is biased toward the better-preserved taxa, and the structural, chemical, and behavioral bases of preservation potential suggest that preservability is heritable. Thus, the average preservation potential is effectively concentrated in the better preserved groups, which therefore have a higher probability of containing preserved ancestor-descendant pairs.

The importance of spatial and taxonomic heterogeneity in the probability of preservation and recovery is suggested by comparing Table 1 to estimates of completeness for well-preserved taxa on smaller spatial scales (Table 2). These completeness values are based on an approach that enables preservation probability and temporal durations of taxa to be estimated from the frequency distribution of stratigraphic ranges (Foote and Raup 1996). These estimates suggest that, for well preserved groups in particularly well-documented geographic areas—which constitute a large part of the known fossil record—a considerable number of species belong in an ancestral rather than terminal position on their respective evolutionary trees.

Discussion and Summary

The approach outlined above does not attempt to suggest how A-D pairs should be rec-

TABLE 2. Estimated median species duration, proportion of species preserved, and probability of finding direct ancestor-descendant pairs for three fossil groups, assuming budding birth-death model and homogeneous preservation with probability corresponding with the given value of P_P (see Foote and Raup 1996).

Data	Source	Median duration (m.y.)	P_P	P_A
Cambro-Ordovician trilobites from Oklahoma	Stitt 1977	0.15	0.9	0.5
Cenozoic mammals from North America	J. Alroy pers. comm. 1995; after Alroy 1994	1.7	0.6	0.5
European Jurassic bivalves	Hallam 1976	8.0	0.9	0.5

ognized, only to estimate the probability that we have them in hand, whether or not we know it. This approach also ignores the stratigraphic order in which taxa are found. Clearly, in the budding model, there is a diminished probability that an ancestral species and a descendant species are preserved, *and that the former occurs stratigraphically below the latter.* However, whether species are preserved in the "wrong" order does not affect the facts of their genealogical relationships, which we must attempt to reconstruct regardless of where we find the species stratigraphically. (Of course, in the models of bifurcation and phyletic transformation, the very nature of ancestors and descendants requires that the former occur stratigraphically below the latter.)

A question related to the order of preservation concerns the probability that a direct A-D pair coexists today, or at any other arbitrary point in time, a question that makes sense only in the context of the budding model. This can be assessed by considering the probability that an extant species originated at some time in the past and the probability that its ancestor, which must also have existed at the time of origination, is also extant (Fig. 3; Appendix 2). When origination and extinction rates are equal, and when the clade in question is old, half of all extant species, on average, are expected to have ancestors within the clade that are also extant (Fig. 3A). This proportion is even higher when origination rate exceeds extinction rate, as in diversifying clades. When clades are young, there is a substantial probability that the founding species of the clade in question is still extant. We must therefore look outside the clade in question for the ancestor of this founding species, but the youth of this founding species enhances the proba-

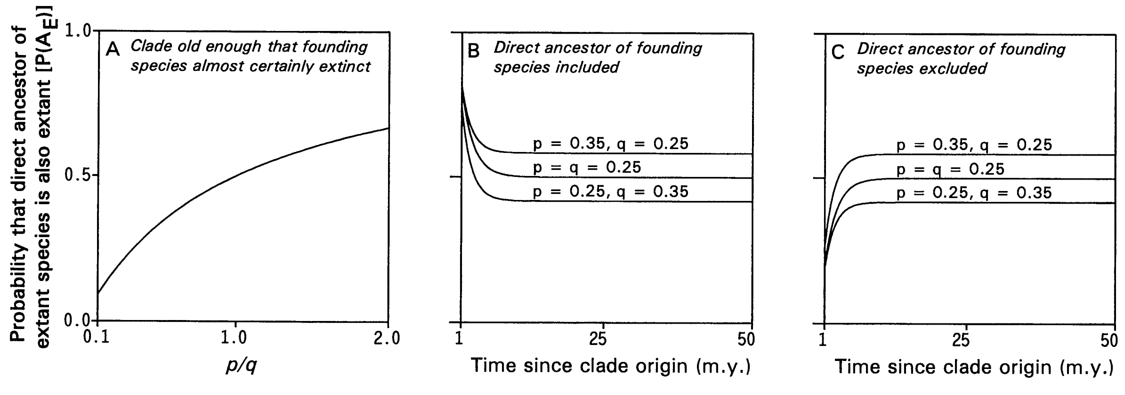

FIGURE 3. Probability that an extant lineage has a direct ancestor that is also extant. This probability makes sense only in the context of the budding model. Origination and extinction rates are denoted p and q, respectively. A. This probability, which is equal to $p/(p+q)$ when the clade is question is old, increases as the ratio of origination to extinction rates increases (Appendix 2). When the clade in question is young (B,C), there is an appreciable probability that the founding species of the clade is still extant. If we sample this species, whether we also sample its direct ancestor depends on whether we look outside the clade for it (B) or not (C). If we do, then a young clade, because of the youth of its founding species and the correspondingly high probability that the ancestor of this founding species is still extant, has an enhanced probability of containing extant species whose ancestors are also extant.

bility that its ancestor is extant (Fig. 3B,C). Thus, if the budding model frequently holds, *the world today should be teeming with species that are ancestral to other living species* (see Hennig 1966: pp. 60–61; Eldredge and Cracraft 1980: p. 131). This will prove to be less the case if descendant species commonly have a large, negative effect on the durations of their ancestors.

The approach developed here does not address the probability that each member of an A-D pair is preserved *during the period of transition.* If speciation is rapid relative to the duration of species (Gould and Eldredge 1993), then the chance of capturing transitions, especially major transitions among classes and orders, is substantially diminished. Given estimates of completeness, median species duration, the time required for evolutionary transitions, and the number of ordinal- or higher-level transitions, we could obtain an estimate of the number of major transitions we should expect to see in the fossil record. This would have important implications for whether the small number of documented major transitions provides strong evidence against evolution (Valentine 1986; D. M. Raup personal communication 1994).

It is worth repeating that heterogeneous preservation enhances the probability of finding A-D pairs. As discussed above, preservation is concentrated in certain taxa and regions, whose species, therefore, have a higher probability of standing in a position ancestral to other preserved species. Temporal variation in preservation probability has a similar effect. Because A-D pairs are more likely to coexist in the same interval of time than at randomly chosen times, concentration of preservation into episodes enhances the probability of finding A-D pairs (see Fig. 3; Appendix 2). To what extent heterogeneous preservation increases the chances of finding A-D pairs depends on assumptions about how preservation probability is distributed. In any event, simulated degradation of evolutionary trees (not presented), comparing homogeneous to inhomogeneous preservation models, supports the intuition that the more realistic assumptions of taxonomically, temporally, and spatially inhomogeneous preservation yield a higher probability of finding A-D pairs.

With pessimistic assumptions, a lower bound on the proportion of known fossil species that are directly ancestral to other known fossil species is on the order of 1%. With more realistic assumptions, this proportion is on the order of perhaps 10% or more, and there are even more preserved A-D pairs if we consider indirect relationships (see Fig. 2). These figures suggest that the problem of how to recognize ancestor-descendant pairs must remain a major priority. If existing phylogenetic methods have trouble recognizing A-D pairs, yet we have reason to believe that their frequency in the fossil record is not negligible, contrary to some assessments (e.g., Schoch 1986: pp. 166–167), then clearly some enhancement of methodology is needed (Hull 1979; Paul 1992). Stratophenetics (Gingerich 1979), stratocladistics (Fisher 1994), and continuous track analysis (Alroy 1995) are among the possibilities that should move us toward a fuller resolution of genealogical relationships.

Acknowledgments

For helpful discussions and comments I thank J. Alroy, B. E. Bodenbender, B. Chernoff, T. M. Collins, D. H. Erwin, D. C. Fisher, P. D. Gingerich, D. Jablonski, M. LaBarbera, C. R. Marshall, D. J. Miller, A. I. Miller, M. E. Patzkowsky, D. M. Raup, and P. J. Wagner. J. Alroy also generously supplied the mammalian data used in Table 2 and suggested the term *budding*. Supported by the National Science Foundation (DEB-9207577, DEB-9496348, and EAR-9506568).

Literature Cited

Alroy, J. 1994. Quantitative mammalian biochronology and biogeography of North America. Unpublished Ph.D. dissertation, University of Chicago, Chicago.

———. 1995. Continuous track analysis: a new phylogenetic and biogeographic method. Systematic Biology 44:152–178.

Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic. Paleobiology 3:152–167.

Durham, J. W. 1967. The incompleteness of our knowledge of the fossil record. Journal of Paleontology 41:559–565.

Eldredge, N., and J. Cracraft. 1980. Phylogenetic patterns and the evolutionary process. Columbia University Press, New York.

Engelmann, G. F., and E. O. Wiley. 1977. The place of ancestor-descendant relationships in phylogeny. Systematic Zoology 26:1–11.

Feller, W. 1968. An introduction to probability theory and its applications, Vol. I, 3d ed. Wiley, New York.

Fisher, D. C. 1994. Stratocladistics: morphological and temporal patterns and their relation to phylogenetic process. Pp. 133–171 *in* L. Grande and O. Rieppel, eds. Interpreting the hierarchy of nature: from systematic patterns to evolutionary process theories. Academic Press, San Diego.

Foote, M. 1988. Survivorship analysis of Cambrian and Ordovician trilobites. Paleobiology 14:258–271.

Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.

Fortey, R. A., and R. P. S. Jefferies. 1982. Fossils and phylogeny—a compromise approach. Pp. 197–234 *in* K. A. Joysey and A. E. Friday, eds. Problems of phylogenetic reconstruction (Systematics Association Special Volume No. 21). Academic Press, London.

Gingerich, P. D. 1979. Stratophenetic approach to phylogeny reconstruction in vertebrate paleontology. Pp. 41–77 *in* J. Cracraft and N. Eldredge, eds. Phylogenetic analysis and paleontology. Columbia University Press, New York.

Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. Nature 366:223–227.

Hallam, A. 1976. Stratigraphic distribution and ecology of European Jurassic bivalves. Lethaia 9:245–259.

Hennig, W. 1966. Phylogenetic systematics [transl. D. D. Davis and R. Zangerl]. University of Illinois Press, Urbana.

Huelsenbeck, J. P. 1994. Comparing the stratigraphic record to estimates of phylogeny. Paleobiology 20:470–483.

Hull, D. L. 1979. The limits of cladism. Systematic Zoology 28: 416–440.

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

May, R. M. 1994. Biological diversity: differences between land and sea. Philosophical Transactions of the Royal Society of London, **B** 343:105–111.

Newell, N. D. 1959. Adequacy of the fossil record. Journal of Paleontology 33:488–499.

Patzkowsky, M. E. 1995. A hierarchical branching model of evolutionary radiations. Paleobiology 21:440–460.

Paul, C. R. C. 1992. The recognition of ancestors. Historical Biology 6:239–250.

Pease, C. M. 1987. Lyellian curves and mean taxonomic durations. Paleobiology 13:484–487.

Raup, D. M. 1978. Cohort analysis of generic survivorship. Paleobiology 4:1–15.

———. 1985. Mathematical models of cladogenesis. Paleobiology 11:42–52.

———. 1991. A kill curve for Phanerozoic marine species. Paleobiology 17:37–48.

Raup, D. M., and S. M. Stanley. 1978. Principles of paleontology, 2d ed. W. H. Freeman, San Francisco.

Schoch, R. M. 1986. Phylogeny reconstruction in paleontology. Van Nostrand Reinhold, New York.

Schopf, T. J. M. 1978. Fossilization potential of an intertidal fauna: Friday Harbor, Washington. Paleobiology 3:261–271.

Sepkoski, J. J., Jr., R. K. Bambach, D. M. Raup, and J. W. Valentine. 1981. Phanerozoic marine diversity and the fossil record. Nature 293:435–437.

Simpson, G. G. 1952. How many species? Evolution 6:342.

Smith, A. B. 1994. Systematics and the fossil record: documenting evolutionary patterns. Blackwell Scientific, Oxford.

Stitt, J. H. 1977. Late Cambrian and earliest Ordovician trilobites, Wichita Mountains Area, Oklahoma. Oklahoma Geological Survey Bulletin 124:1–79.

Valentine, J. W. 1970. How many marine invertebrate fossil species? A new approximation. Journal of Paleontology 44:410–415.

———. 1986. Fossil record of the origin of Baupläne and its implications. Pp. 209–222 *in* D. M. Raup and D. Jablonski, eds. Patterns and processes in the history of life. Springer, Berlin.

———. 1989. How good was the fossil record? Clues from the Californian Pleistocene. Paleobiology 15:83–94.

Wagner, P. J., and D. H. Erwin. 1995. Phylogenetic patterns as tests of speciation models. Pp. 87–122 *in* D. H. Erwin and R. L. Anstey, eds. New approaches to studying speciation in the fossil record. Columbia University Press, New York.

Appendix 1

Probability of finding fossil ancestor-descendant pairs

Direct Descendants

Given a probability of preservation per taxon per unit time, and a distribution of taxonomic durations, we want to predict the proportion of species preserved (P_P) and the proportion of these that have *at least* one direct descendant preserved (P_A). Adopting a discrete-time convention, let R be the probability that a species is preserved at least once during a time interval in which it exists, and let $P_D(T)$ be the probability that the duration of a species is T time units. The probability that a species with duration T is not preserved at all is equal to $(1-R)^T$; therefore, the probability that it is preserved at least once is equal to $1\text{-}(1\text{-}R)^T$. From this it follows that the proportion of species preserved is given by (Foote and Raup 1996: eq. 6)

$$P_P = \sum_{T=1}^{\infty} \{P_D(T)[1 - (1 - R)^T]\}. \tag{1}$$

Let $P(N)$ be the probability that a species with duration T has exactly N direct descendants. Then the probability that at least one of these is preserved is equal to $1 - (1 - P_P)^N$. Since we are interested only in species that are in fact preserved, we modify $P_D(T)$ to be the probability that a species has duration T, *given that it is preserved.* This is equal to $[1 - (1 - R)^T]P_D(T)/P_P$. Putting all this together, we have the probability that a preserved species has at least one preserved direct descendant:

$$P_A = \sum_{T=1}^{\infty} \left(\{[1 - (1 - R)^T]P_D(T)/P_P\} \times \sum_{N=1}^{\infty} \{P(N)[1 - (1 - P_P)^N]\} \right). \tag{2}$$

Note that this equation requires R as an input, but in fact, for each birth-death model, each value of R along with the parameters of the model (see below) clearly predicts a unique value of P_P. Therefore, P_A can be calibrated directly against P_P.

Budding.—Let p be the origination (branching) rate per species per time unit, and let q be the extinction rate. Note that, in contrast to some other works (e.g., Raup 1978; Foote 1988; Patzkowsky 1995), p here includes all branching events, both those leading to new species within the same higher taxon and those leading to what would be considered new higher taxa. In this model (Fig. 1A), a species is terminated by extinction only. If branching can occur at any time during a species' duration, then the number of direct descendants of a species of a given duration follows a Poisson distribution with parameter pT (Feller 1968: eq. 6.1). Therefore, we make the following substitutions into equations (1) and (2):

$$P_D(T) = e^{-q(T-1)} - e^{-qT} \qquad \text{(Raup 1985: eq. A2)}$$

and

$$P(N) = [e^{-pT}(pT)^N]/N! \qquad \text{(Feller 1968: eq. 6.1)}$$

Substituting into equation (2), it is easy to show that $P_A \approx 2(p/q)P_P$ when $R << 1$ and time steps are small enough that $p << 1$.

Bifurcation.—Let p and q be as in the budding model. In the bifurcation model (Fig. 1B), a species terminates by extinction or origination, and the number of descendants can be equal only to 0 or 2. Therefore, we make the following substitutions into equations (1) and (2):

$$P_D(T) = e^{-(q+p)(T-1)} - e^{-(q+p)T}$$

and

$$\begin{aligned} P(N) &= q/(p + q) && \text{if } N = 0 \\ &= p/(p + q) && \text{if } N = 2 \\ &= 0 && \text{otherwise.} \end{aligned}$$

Substituting into equation (2), it is easy to show that $P_A \approx 2[p/(p + q)]P_P$ when $R << 1$.

Phyletic transformation.—Let the rates of true extinction and pseudoextinction (species origination) be q and w, respectively. A species terminates by true extinction or pseudoextinction, and the number of descendants can be equal only to 0 or 1. We therefore make the following substitutions into equations (1) and (2):

$$P_D(T) = e^{-(q+w)(T-1)} - e^{-(q+w)T}$$

and

$$\begin{aligned} P(N) &= q/(q + w) && \text{if } N = 0 \\ &= w/(q + w) && \text{if } N = 1 \\ &= 0 && \text{otherwise.} \end{aligned}$$

Substituting into equation (2), it is easy to show that $P_A = [w/(q + w)]P_P$.

Indirect Descendants

Budding.—to modify equation (2) to incorporate indirect descendants, consider each direct descendant as a clade having a total progeny (including itself) of one or more species. If at least one of the total progeny is preserved, then the ancestral species in question has at least one preserved (in)direct descendant. Kendall (1948: eqs. 52 and 53) gives the probability distribution of total progeny in the limit of infinite time when $p \leq q$. Note that, in contrast to the total progeny calculations in Table 1, the relevant calculations here allow for the extinction of the clade (i.e., the direct descendant plus its direct and indirect descendants). Using Kendall's notation, the probability Q_M that the total progeny of a direct descendant will be equal to M is given by (Kendall 1948: eqs. 52 and 53)

$$Q_M = \frac{p + q}{2p} \cdot \frac{(2M)!}{2^{2M}(M!)^2} \cdot \frac{x^M}{2M - 1}, \qquad p \leq q$$

where $x = (4pq)/[(p + q)^2]$.

I know of no corresponding equation for the case where $p > q$, a situation that may commonly hold. However, assuming, as I have, that $p = q$, when in fact $p > q$, leads to an underestimate of P_A and is therefore conservative for the present purposes. From the foregoing equation, the probability of preservation of at least one of the total progeny of a direct descendant, including itself, is obtained as

$$P' = \sum_{M=1}^{\infty} \{Q_M[1 - (1 - P_P)^M]\}. \tag{3}$$

This value is then substituted for the second P_P in equation (2) to yield the probability that a preserved species has at least one preserved descendant, direct or indirect, where the indirect descendants are followed infinitely into the future.

We may also be interested in following indirect descendants out to some limited time, dT, after the branching of direct descendants (Fig. 1). We would like to know the probability distribution of the total progeny up to some time y. Let $P_{z,n,t}$ be the probability that standing diversity is equal to n and that there are exactly z originations ($z \leq n$) during the t^{th} time increment. This is given by the Poisson distribution with parameter pn (Feller 1968: eq. 6.1) as

$$P_{z,n,t} = P_{n,t}[e^{-pn}(pn)^z]/z!$$

where $P_{n,t}$ is the probability that standing diversity is equal to n (Raup 1985: eqs. A15 and A17). From this, the probability that there are exactly z originations during time increment t is equal to

$$P_{z,t} = \sum_{n=1}^{\infty} P_{z,n,t}. \tag{4}$$

We would like the probability that there are a total of N_{tot} originations (i.e., a total progeny of $M = 1 + N_{\text{tot}}$) over all time increments. To do so, we sum the probabilities corresponding to all pathways that yield this total. For example, following indirect descendants out to $y = 3$, we have the following sum:

$$Q'_M = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \{A\},$$

where $A = P_{i,1}P_{j,2}P_{k,3}$ (from eq. 4) if $(i + j + k) = N_{\text{tot}}$ and $A = 0$ otherwise. Q'_M is then simply substituted for Q_M in equation (3) above. The number of summations is equal to the number of time steps, and so the computation becomes prohibitive rather quickly. Using a reasonably fast (as of 1995!) desktop computer, it can take a few days to calculate the probability distribution of total progeny, even if the process is followed to only ten time steps or less. For the intermediate curves in Figure 2, indirect descendants are followed out to $y = 5$ million years.

Bifurcation.—Topologically, this model is the same as the budding model. We can therefore use Q_M and Q'_M as above, provided that for each value of R we use the value of P_P corresponding to the budding model.

Phyletic transformation.—In the absence of branching, indirect descendants of a species can continue to be generated provided that true extinction does not occur. The probability that the total number of originations within a span of time T is exactly equal to N_{tot} (i.e., that the total progeny $M = 1 + N_{\text{tot}}$) is equal to

$$\begin{aligned} Q^*_M = &\sum_{t=1}^{T-1} \{[e^{-q(t-1)} - e^{-qt}][e^{-wt}(wt)^{M-1}]/(M - 1)!\} \\ &+ [e^{-q(T-1)}][e^{-wT}(wT)^{M-1}]/(M - 1)! \end{aligned}$$

The summation reflects the probability distribution of the number of instances of phyletic speciation if true extinction occurs in the first through $(T - 1)^{\text{th}}$ intervals, while the second expression reflects the distribution of the number of phyletic speciation events given survival into the T^{th} interval. This same equation can be used for following indirect descendants for a finite period of time as well as infinitely into the future. Q^*_M is substituted for Q_M in equation (3) above.

Appendix 2

Probability that direct ancestor of extant species is also extant

The question of whether the ancestor of an extant species is itself extant makes sense only in the context of the budding model (Fig. 1A). Let p and q be the rates of branching and extinction, respectively. Consider a clade and a species within that clade that is extant at the Recent or any other arbitrary point in time, where time zero is the arbitrary point, and time increases into the past until the origin of the clade at time $T_{\max}$. As this species is followed backward in time, there is an increasing probability that it will ultimately "terminate" by merging (topologically) with its ancestral species. Just as the probability of species durations followed forward in time is exponentially distributed, the probability that a species extant at time zero originated during the T^{th} time increment in the past (see Pease 1987: p. 486) is given by

$$P_{O}(T) = e^{-p(T-1)} - e^{-pT}$$

The origination of a species at time T implies that its ancestor was extant at time T. The probability of survival of the ancestor to time zero is therefore simply equal to e^{-qT} (Raup 1985: eq. A1). Therefore, in a clade sufficiently old that all extant species originated after the origin of the clade, the probability that the ancestor of a species is extant, given that the species in question is extant, is given by

$$P(A_{\mathrm{E}}) = \sum_{T=1}^{T_{\max}} \{[e^{-p(T-1)} - e^{-pT}]e^{-qT}\} \tag{5a}$$

If the time steps are sufficiently small, then $e^{-p(T-1)} - e^{-pT}$ is approximately equal to pe^{-pT} (see Pease 1987: p. 486), and we have

$$P(A_{\mathrm{E}}) = \int_{T=0}^{T_{\max}} pe^{-pT}e^{-qT}\, dT. \tag{5b}$$

This is equal to

$$P(A_{\mathrm{E}}) = [p/(p + q)][1 - e^{-(p+q)T_{\max}}], \tag{5c}$$

which is approximately equal to $p/(p + q)$ when $T_{\max}$ is large.

There is a certain probability that the origination of an extant species did not occur between time zero and time $T_{\max}$, i.e., that an extant species is in fact the founding species of the clade in question. The corresponding probability is simply $1 - \int_{T=0}^{T_{\max}} pe^{-pT}\, dT$, which is equal to $e^{-pT_{\max}}$. The origin of this founding species coincides with the origin of the clade; thus, the probability that the ancestor of the founding species is extant is simply equal to $e^{-qT_{\max}}$. Combining this with equation (5c) we have

$$P(A_{\mathrm{E}}) = [p/(p + q)][1 - e^{-(p+q)T_{\max}}] + e^{-(p+q)T_{\max}} \tag{5d}$$

if the ancestor of the founding species is taken into consideration, and equation (5c) if it is not.