XI. Sampling Models, 2: Some applications of homogeneous sampling models (forward problems)


1.1 Rationale: Conventional cladistic methodology makes recognition of ancestors problematic.
Is probability of finding sampled ancestor-descendant pairs high enough for this to be a practical concern?

1.2 Explore alternative models of descent:

1.2.1 Bifurcation (ancestor gives rise to two distinct descendants)

1.2.2 Budding (ancestor persists and gives rise to one descendant)

1.2.3 Phyletic transformation (ancestor gives rise to one descendant, does not persist)

1.3 Use homogeneous branching and sampling models in discrete time to predict the proportion of species with at least one preserved descendant, direct or indirect.

Comment: Study uses discrete-time approximation in certain places (for convenience) where continuous-time model would have been preferable.
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.
**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 \((P_r)\) and the proportion of these that have at least one direct descendant preserved \((P_d)\). 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 \(1-(1-R)^T\). From this it follows that the proportion of species preserved is given by (Foote and Raup 1996: eq. 6)

\[
P_r = \sum_{T=1}^{\infty} [P_d(T)[1 - (1 - R)^T]].
\]  

(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 \(1 - (1 - P_r)^N\). 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 \([1 - (1 - R)^T]P_d(T)/P_r\). Putting all this together, we have the probability that a preserved species has at least one preserved direct descendant:

\[
P_A = \sum_{T=1}^{\infty} \left( [1 - (1 - R)^T]P_d(T)/P_r \right)
\]

\[
\times \sum_{N=1}^{\infty} [P(N)[1 - (1 - P_r)^N]].
\]  

(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 \(P_r\). Therefore, \(P_A\) can be calibrated directly against \(P_r\).
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):

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

and

\[
P(N) = [e^{-qT}(pT)^{N-1}] / N! \quad \text{(Feller 1968: eq. 6.1)}
\]

Substituting into equation (2), it is easy to show that \( P_A = 2(p/q)P_p \) 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

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

Substituting into equation (2), it is easy to show that \( P_A = 2[p/(p + q)]P_p \) 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

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

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)

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

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

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

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

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

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

\[
P_{z,t} = \sum_{n=1}^{\infty} P_{z,n,t}.
\]

We would like the probability that there are a total of \( N_{tot} \) originations (i.e., a total progeny of \( M = 1 + N_{tot} \)) 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 \( A = P_{i1}P_{j2}P_{k3} \) (from eq. 4) if \( i + j + k = N_{tot} \) 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 $M = 1 + N_{tot}$) is equal to

$$Q^*_M = \sum_{t=1}^{T-1} \left[ \left( e^{-qt(1-1)} - e^{-qt} \right) \left( e^{-wT} (wT)^{M-1} \right) / (M - 1)! \right] + \left[ e^{-q(T-1)} \right] \left( e^{-wT} (wT)^{M-1} \right) / (M - 1)!$$

The summation reflects the probability distribution of the number of instances of phyletic speciation if true extinction occurs in the first through $(T - 1)^{th}$ 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.
1.4 Calibrate empirically to estimate the probability of finding ancestors given $p$, $q$ and $R$. 

![Diagram showing probability of finding ancestor-descendant pairs ($P_e$) 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 $P_e$ is equal to 0.5 when all species are preserved. A, Budding. $P_e$ 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, Phylectic 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 cladogegetic 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 phylectic 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.
1.5 Probability that ancestor of extant lineage is also extant

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 \( T_{\text{max}} \). 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 \( T^n \) time increment in the past (see Pease 1987: p. 486) is given by

\[
P_N(T) = e^{-q(T-t)} - e^{-qT}
\]

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

\[
P(A_k) = \sum_{T=1}^{T_{\text{max}}} [e^{-q(T-1)} - e^{-qT}] e^{-qT}
\]

(5a)

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

\[
P(A_k) = \int_{T=0}^{T_{\text{max}}} pe^{-qT} e^{-qT} dT.
\]

(5b)

This is equal to

\[
P(A_k) = [p/(p+q)][1 - e^{-q+qT_{\text{max}}}],
\]

(5c)

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

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

\[
P(A_k) = [p/(p+q)][1 - e^{-q+qT_{\text{max}}} + e^{-q+qT_{\text{max}}}
\]

(5d)

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

2.1 Rationale: Molecular clocks often imply long gap between group origin and first fossil find. Is sampling poor enough that this gap is plausible?

*Fig. 1.* Hypothetical illustration of the missing diversity problem. Species 1 through 5 comprise the extant part of the group of interest, whose outgroup is O. Solid lines show the known fossil record. (A) Relatively even distribution of branching events. (B) Clustering of some branching events, as is often thought to occur in the early stages of an evolutionary radiation (47). In both cases, the tree topology, the length of the known fossil record, and the age of the common ancestor to the outgroup and the group of interest are the same. In (B), however, the interval of missing history is shorter and the sum of missing species durations is lower.

2.2 Use homogeneous branching model in continuous time to predict $S$, the sum of missing lineage durations.

This depends on:

1. postulated extinction rate
2. gap between clade origin and first appearance in fossil record
3. minimum clade diversity at time of first fossil appearance
4. assumed diversification model
13. We start by assuming that species originations and extinctions within a taxonomic group occur at stochastically constant per capita rates, $p$ and $q$, per lineage-million-years. The hypothesized divergence times imply an interval of missing history of length $T$ and a minimal diversity of $N$ species at the end of this interval. We use lineage-million-years because each species has some chance of preservation throughout its lifetime, and so all species are multiplied by their durations to yield the sum of species durations, $S$, during the interval $T$. This is equivalent to

$$\int_0^T D_t \, dt$$

where $D_t$ is the expected species diversity at time $t$. Several diversification models are plausible. Exponential growth, in which $D_t = e^{p-qt}$, includes the chance of total group extinction before time $T$. With exponential growth, $S$ depends only on $(p-q)$; the same value of $S$ can thus result from higher rates, yielding many short-lived species, or lower rates, yielding fewer longer-lived species. A crucial aspect of the hypotheses we wish to test is a minimal diversity of $N$ species at time $T$, when fossil taxa are first observed. The most appropriate model tracks expected diversity at time $t$ conditioned upon a diversity at least equal to $N$ at time $T$. Let $t$ be the elapsed time since $t = 0$. Let $P(0, t, a)$ be the probability that a group is extinct at or before time $t$, given that its diversity is equal to a lineages at $t = 0$; $P(0, t, a) = \frac{p}{(1 + pt)}$ if $p = q$, and $P(0, t, a) = (qe^{p-qt} - 1)/(pe^{p-qt} - q)$ if $p \neq q$. Let $P(s, t, a)$ be the probability that a group survives at least until time $t$, given that its diversity is equal to a lineages at $t = 0$; $P(s, t, a) = 1 - P(0, t, a)$. Let $P(n, t, a)$ be the probability that a group’s diversity is exactly equal to $n$ lineages at time $t$, given that it is equal to a lineages at $t = 0$; for $a = 1, P(n, t, a) = (pt)^{a-1}/(1 + pt)^{a+1}$ if $p = q$ (9), and $P(n, t, a) = (1 - A)(1 - B)b^{a-1}$ if $p \neq q$ (9), where $A = P(0, t, 1)$ and $B = pA/q$. For $a > 1$,

$$P(n, t, a) = \left[ \frac{pt}{(1 + pt)} \right] a^a - n \sum_{j=1}^{\min(a,n)} \binom{a}{j} \binom{n-1}{j-1} (pt)^{-j} \right] \left( 1 - A \right) \left( 1 - B \right) b^{a-1}$$

and

$$P(n, t, a) = \sum_{j=1}^{\min(a,n)} \binom{a}{j} \binom{n-1}{j-1} (pt)^{-j} \right] \left( 1 - A \right) \left( 1 - B \right) b^{a-1}$$

The last equation is a correction of equation (A18) from (9). Let $P(\geq n, t, a)$ be the probability that diversity is greater than or equal to $n$ lineages at time $t$, given that it is equal to a lineages at $t = 0$. Then

$$P(\geq n, t, a) = 1 - P(0, t, a) - \sum_{m=1}^{n-1} P(m, t, a)$$

Let $t$ and $T$ be two points in time such that $t < T$, and let the group have diversity equal to one lineage at $t = 0$. Let $P(n, t, s, T)$ be the probability that diversity is exactly equal to $n$ lineages at time $t$, given that the group survives at least until time $T$. Let $P(n, t, N, T)$ be the probability that diversity is exactly equal to $n$ lineages at time $t$, given that diversity is exactly equal to $N$ lineages at time $T$. And let $P(n, t, \geq N, T)$ be the probability that diversity is exactly equal to $n$ lineages at time $t$, given that diversity is greater than or equal to $N$ lineages at time $T$. Then, by the rules of conditional probability,

$$P(n, t, s, T) = P(n, t, 1) \cdot P(s, T - t, n)P(s, t, 1);$$
$$P(n, t, N, T) = P(n, t, 1) \cdot P(N, T - t, n)P(N, t, 1);$$
and

$$P(n, t, \geq N, T) = P(n, t, 1) \cdot P(\geq N, T - t, n)P(\geq N, T, 1)$$

The corresponding expected (mean) diversities at time $t$ are given by

$$D_t = E(n, t, s, T) = \sum_{m=1}^{N} m \cdot P(m, t, s, T);$$
$$D_t = E(n, t, N, T) = \sum_{m=1}^{N} m \cdot P(m, t, N, T);$$
and

$$D_t = E(n, t, \geq N, T) = \sum_{m=1}^{N} m \cdot P(m, t, \geq N, T)$$

NB: Comment that equation is correction of (A18) of Raup (1985) is wrong! The two look different but are algebraically identical.
2.3 Use homogeneous sampling model in continuous time to calculate the probability that $S$ could go unobserved for assumed value of sampling rate $r$: $P_0 = Pr(\text{nofossils}) = e^{-rS}$.

2.4 Choose critical value of $P_0$ (say 0.5) and find $r_{\text{crit}} = -\ln (P_0/S)$. If empirical $r$ is greater than this, then probability of missing $S$ is less than $P_0$.

2.5 Use observed stratigraphic ranges to estimate $r$ empirically, and compare this with value required for missing clade history to be plausible.

2.5.1 $\hat{q} = m/\sum_i t_i$, where $m$ is the number of non-singleton taxa and $t_i$ are their observed ranges.

2.5.2 $\hat{r} = \hat{q}(1 - f_0)/f_0$, where $f_0$ is the observed proportion of species that are singletons.

Note the importance of conditioning probabilities on survival of clade to some specified time. Standard exponential growth model is an expectation (mean) which averages together zero values for numerous clades that do not survive. Much higher diversities are expected, especially early in clade history, if we condition on survival of the clade. In general, it is important to think carefully about whether conditioning is appropriate for a given question.
Fig. 3. Analysis of the hypothesis that many lineages of modern eutherians originated before the Tertiary (4, 13, 17). $T$ is taken to be 64 My, $N$ is taken to be 9 species, and $q$ is taken to be 0.25 Lmy$^{-1}$ (see text). (A to C) Exponential diversity model; (D to F) diversity before time $T$ conditioned upon minimal diversity of $N$ at time $T$; other diversity models yield results between these extremes. (A and D) Variation in $T$ with $N = 9$ species and $q = 0.25$ Lmy$^{-1}$. (B and E) Variation in $N$ with $T = 64$ My and $q = 0.25$ Lmy$^{-1}$. (C and F) Variation in $q$ with $T = 64$ My and $N = 9$ species.

Left-hand ordinate (solid line), expected sum of species durations, $S$ (13). Right-hand ordinate (dashed line), preservation rate required to yield a probability of complete nonpreservation exactly equal to 0.5 (17). Shaded area beneath corresponds to probabilities of nonpreservation greater than 0.5, and thus to combinations of preservation rate and value of abscissa for which the corresponding amount of missing diversity is plausible. For example, in (D), $q = 0.25$ Lmy$^{-1}$ and $N = 9$ species. If $T = 40$ My, then $S = 918$ Lmy. This value of $T$ implies that $r_{max} = -\ln(0.5)/918 = 0.0008$ Lmy$^{-1}$. For this value of $r$, any value of $T$ less than 40 My yields a probability of nonpreservation of the group greater than 0.5 (the shaded region), and, for this value of $T$, any value of $r$ less than $r_{max}$ yields a probability of nonpreservation of the group greater than 0.5. As $T$ increases so does $S$, and thus an ever smaller value of $r$ is required to make group nonpreservation likely. The same is true for an increase in $N$ with $T$ and $q$ fixed (B and E) or an increase in $q$ with $T$ and $N$ fixed (C and F).
Table 1. Analysis of the hypothesis that many modern eutherian lineages arose before the Tertiary (4, 13, 17). N is the minimal number of species present at the time the group is first found in the fossil record. T is the time between the postulated origin of the group and its first fossil appearance, q is the extinction rate, S is the expected summed species durations, \( r_{\text{sum}} \) is the preservation rate that yields a probability of 0.5 that S will completely escape preservation if \( r = 0.03 \text{ Lmy}^{-1} \) (see text) (that is, \( p = e^{-r T} \)). For the bottom part of the table, S is the grand sum of summed species durations of the individual lineages. See (13) for exponential (E) and conditional (C) diversity models; the latter conditions upon minimal diversity of N at time T. Results of the other models (13) are within these extremes. We do not consider the “star phylogeny” model, in which all extant lineages diverge at the origin of the group, because it is inconsistent with the hypothesis we are testing. That model yields even greater summed species durations than those we present.

<table>
<thead>
<tr>
<th>N</th>
<th>T (My)</th>
<th>Model</th>
<th>S (Lmy)</th>
<th>( q = 0.0 \text{ Lmy}^{-1} )</th>
<th>( q = 0.1 \text{ Lmy}^{-1} )</th>
<th>( q = 0.5 \text{ Lmy}^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( r_{\text{sum}} )</td>
<td>( r_{\text{sum}} )</td>
<td>P</td>
<td>( S (\text{Lmy}) )</td>
</tr>
<tr>
<td>9*</td>
<td>64†</td>
<td>C</td>
<td>401</td>
<td>17</td>
<td>( 6 \times 10^{-1} )</td>
<td>939</td>
</tr>
<tr>
<td>9*</td>
<td>64†</td>
<td>E</td>
<td>233</td>
<td>30</td>
<td>( 9 \times 10^{-1} )</td>
<td>233</td>
</tr>
<tr>
<td>14‡</td>
<td>64†</td>
<td>C</td>
<td>566</td>
<td>12</td>
<td>( 4 \times 10^{-10} )</td>
<td>1216</td>
</tr>
<tr>
<td>14‡</td>
<td>64†</td>
<td>E</td>
<td>315</td>
<td>22</td>
<td>( 1 \times 10^{-5} )</td>
<td>315</td>
</tr>
<tr>
<td>7§</td>
<td>46‡</td>
<td>C</td>
<td>227</td>
<td>31</td>
<td>( 1 \times 10^{-5} )</td>
<td>454</td>
</tr>
<tr>
<td>10</td>
<td>44‡</td>
<td>C</td>
<td>330</td>
<td>23</td>
<td>( 1 \times 10^{-5} )</td>
<td>562</td>
</tr>
<tr>
<td>10</td>
<td>44‡</td>
<td>E</td>
<td>172</td>
<td>40</td>
<td>( 6 \times 10^{-10} )</td>
<td>172</td>
</tr>
</tbody>
</table>

Modern eutherians considered as a whole

<table>
<thead>
<tr>
<th>N</th>
<th>T (My)</th>
<th>Model</th>
<th>S (Lmy)</th>
<th>( q = 0.0 \text{ Lmy}^{-1} )</th>
<th>( q = 0.1 \text{ Lmy}^{-1} )</th>
<th>( q = 0.5 \text{ Lmy}^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9*</td>
<td>64†</td>
<td>C</td>
<td>346</td>
<td>20</td>
<td>( 3 \times 10^{-1} )</td>
<td>1337</td>
</tr>
<tr>
<td>9*</td>
<td>64†</td>
<td>E</td>
<td>346</td>
<td>20</td>
<td>( 3 \times 10^{-1} )</td>
<td>346</td>
</tr>
<tr>
<td>14‡</td>
<td>64†</td>
<td>C</td>
<td>487</td>
<td>14</td>
<td>( 5 \times 10^{-7} )</td>
<td>1784</td>
</tr>
<tr>
<td>14‡</td>
<td>64†</td>
<td>E</td>
<td>487</td>
<td>14</td>
<td>( 5 \times 10^{-7} )</td>
<td>487</td>
</tr>
</tbody>
</table>

Modern eutherians considered as individual lineages

*Modern placental lineages that are implied by hypothesis (4) to have diverged during the Late Cretaceous and to have passed unreviewed into the Paleocene are Eutheria, Squamata, Haplodonta, Paenungulata, Cetartiodactyla, Perissodactyla, Carnivora, Lagomorpha, and Scandentia. T is the difference between the postulated divergence time of Eutheria (129 million years ago [Ma]) (4) and the oldest fossil record of undisputed modern placental in the early Paleocene (21) at ~65 Ma (42). [Except as otherwise noted, we use the phylogenetic relationships in Kumar and Hedges’s (46) hypothesis in order to test their postulated divergence times. The hypothesis in question omits a number of modern orders. Some of these can be inferred to extend into the Cretaceous if we accept Kumar and Hedges’s divergence times and suprageneric groupings while incorporating additional interordinal relationships expressed in the widely cited, morphologically based phylogeny of Novacek (19). In this way, at least five additional lineages of modern placental are inferred.]

†The calculations in the top part of this table use only the oldest postulated divergence time. Here we treat the minimal number of lineages individually, each constrained with its own amount of missing time and with a minimal diversity of one species at the end of this time. The missing time, T, for an individual lineage is the difference between its postulated divergence time and the time of its oldest fossil record. For lineages added on the basis of Novacek’s (19) calculations, the divergence time is the youngest time consistent with his phylogeny and with Kumar and Hedges’s postulated divergence times (4). The divergence times, oldest fossil occurrences, and estimated absolute ages (42) of oldest fossil are: Eutheria, 129 Ma; late Paleocene (21); 60.5 Ma; Squamata, 112 Ma, late Paleocene (21); 60.3 Ma; Haplodonta, 109 Ma; early Eocene (21); 56.5 Ma; Paenungulata, 105 Ma; late Paleocene (43); 68.5 Ma; Ferungulata (sensu Kumar and Hedges), 92 Ma; early Paleocene (39, 44); 65 Ma; Cetartiodactyla, 83 Ma; late Paleocene (21); 74 Ma; Perissodactyla, 74 Ma; early Eocene (21); 56.5 Ma; Lagomorpha (excluding problematic zalambdalestids), 90.8 Ma; early Paleocene (21); 65 Ma; Scandentia, 85.9 Ma; middle Eocene (21); 50 Ma; Macroscelidea, 112 Ma, early Paleocene (27); 54 Ma; Lipotyphla, 74 Ma; early Eocene (21); 65 Ma; Primates, 85.9 Ma, early Paleocene (21); 65 Ma; Dermoptera (Chiroptera), 85.9 Ma, late Paleocene (21); 60.5 Ma; Tubulidentata, 74 Ma, early Oligocene (21); 35.4 Ma. A number of these groups (Macroscelidea, Lipotyphla, Primates, and Dermoptera (Chiroptera)) may have unequivocal first-fossil appearances that postdate those used here. Our acceptance of claims of earlier occurrences is conservative, because we thereby minimize implied gaps and thus favor the hypothesis of early origins of modern eutherians. In addition to the Perissodactyla and Cetartiodactyla, there must have been at least one continuous lineage of Ferungulata (sensu Kumar and Hedges, including Carnivora) that extends from the origin of this suprageneric grouping to the Paleocene. The gaps are minimized in the foregoing tabulation by taking the gap of this lineage to end at the first appearance of Carnivora. Because the ~85-Ma Asian fossils in question are thought to be stem members of one or two modern eutherian lineages (4, 26) if they belong to these lineages at all, the number and duration of gaps for individual lineages are affected minimally by the interpretation of these fossils. Separate analyses are therefore not presented.

3.1 Rationale: It is often stated that abundance of monotypic taxa is artifact of taxonomic practice. What is the expected proportion of monotypic taxa under homogeneous branching and sampling model?

3.2 NB: For complete sampling, probability of being monotypic (in limit of infinite time) is simply the probability that founding member of clade becomes extinct before it leaves any descendants: This is equal to \( q/(p + q) \).

3.3 Simulate branching history of clade in continuous time (using approach similar to the R code given earlier). Add rules for generating new higher taxa analogous to model of Patzkowsky (1995). Repeat numerous times.

3.4 Degrade stratigraphic ranges of species within higher taxa by dropping sampling events randomly in time.

3.5 Compare size-frequency distributions of pristine and degraded clades.

3.6 Principal results (with respect to monotypy):

3.6.1 Most truly monotypic taxa are not sampled at all.

3.6.2 Most apparently monotypic taxa were not in fact monotypic before sampling.

3.6.3 However, the entire size-frequency distribution is affected in such a way that the overall proportion of taxa that are monotypic is similar in the pristine and degraded distributions.

3.6.4 Likewise, the proportion of species in the largest genus is also accurately represented (although this proportion is highly variable in both the raw and sampled simulations).

Note similarity to situation with stratigraphic ranges: Many taxa lost, entire distribution changed, but certain properties maintained.
### Table 1

The effect of sampling on $M$ and $L$, the proportion of monotypic genera and the proportion of species in the largest genus, respectively. Species diversity equilibrium is set to equal the number of extant species in Chiroptera, namely 974. Results are given for two types of sampling, exponential and logistic, of two average intensities, 2.5% and 10%. The 95% confidence intervals are given for the actual $M$ and the actual $L$, which refer respectively to the proportion of monotypic genera and the proportion of species in the largest genus under perfect sampling. The 95% confidence intervals for the apparent $M$ and the apparent $L$ indicate the proportions as observed after sampling. $X$ is the proportion of genera sampled to be monotypic that are actually monotypic. $Y$ is the proportion of actual monotypic genera that are sampled at least once.

<table>
<thead>
<tr>
<th>Sampling</th>
<th>Actual $M$</th>
<th>Apparent $M$</th>
<th>Actual $L$</th>
<th>Apparent $L$</th>
<th>$X$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exponential Growth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5% uniform</td>
<td>0.48–0.58</td>
<td>0.51–0.76</td>
<td>0.10–0.44</td>
<td>0.12–0.50</td>
<td>0.06–0.31</td>
<td>0.01–0.06</td>
</tr>
<tr>
<td>10% uniform</td>
<td>0.48–0.58</td>
<td>0.47–0.65</td>
<td>0.10–0.44</td>
<td>0.11–0.48</td>
<td>0.22–0.47</td>
<td>0.07–0.18</td>
</tr>
<tr>
<td>2.5% exponential</td>
<td>0.45–0.59</td>
<td>0.51–0.77</td>
<td>0.10–0.44</td>
<td>0.11–0.52</td>
<td>0.04–0.32</td>
<td>0.01–0.07</td>
</tr>
<tr>
<td>10% exponential</td>
<td>0.48–0.58</td>
<td>0.46–0.66</td>
<td>0.10–0.43</td>
<td>0.11–0.49</td>
<td>0.19–0.49</td>
<td>0.05–0.23</td>
</tr>
<tr>
<td><strong>Logistic Growth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5% uniform</td>
<td>0.49–0.55</td>
<td>0.49–0.64</td>
<td>0.11–0.43</td>
<td>0.11–0.46</td>
<td>0.12–0.26</td>
<td>0.03–0.07</td>
</tr>
<tr>
<td>10% uniform</td>
<td>0.49–0.55</td>
<td>0.45–0.55</td>
<td>0.11–0.42</td>
<td>0.11–0.44</td>
<td>0.33–0.48</td>
<td>0.13–0.20</td>
</tr>
<tr>
<td>2.5% exponential</td>
<td>0.49–0.55</td>
<td>0.49–0.63</td>
<td>0.11–0.44</td>
<td>0.11–0.46</td>
<td>0.11–0.27</td>
<td>0.02–0.07</td>
</tr>
<tr>
<td>10% exponential</td>
<td>0.49–0.55</td>
<td>0.45–0.55</td>
<td>0.10–0.45</td>
<td>0.10–0.47</td>
<td>0.30–0.50</td>
<td>0.12–0.22</td>
</tr>
</tbody>
</table>

### Table 2

The effect of sampling of $M$ and $L$, the proportion of monotypic families and the proportion of species in the largest family, respectively. Species diversity equilibrium is set to equal the number of extant species in Chiroptera, namely 974. Results are given for two types of sampling, exponential and logistic, of two average intensities, 2.5% and 10%. Polytypic origins are modeled by starting the simulation with four families each of two species. The 95% confidence intervals are given for the actual $M$ and the actual $L$, which refer respectively to the proportion of monotypic families and the proportion of species in the largest family under perfect sampling. The 95% confidence intervals for the apparent $M$ and the apparent $L$ indicate the proportions as observed after sampling. $X$ is the proportion of families sampled to be monotypic that are actually monotypic. $Y$ is the proportion of actual monotypic families that are sampled at least once.

<table>
<thead>
<tr>
<th>Sampling</th>
<th>Actual $M$</th>
<th>Apparent $M$</th>
<th>Actual $L$</th>
<th>Apparent $L$</th>
<th>$X$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exponential Growth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5% uniform</td>
<td>0.31–0.66</td>
<td>0–0.83</td>
<td>0.53–0.98</td>
<td>0.54–0.99</td>
<td>0–1</td>
<td>0–0.14</td>
</tr>
<tr>
<td>10% uniform</td>
<td>0.31–0.66</td>
<td>0.14–0.75</td>
<td>0.52–0.98</td>
<td>0.53–0.99</td>
<td>0–1</td>
<td>0–0.33</td>
</tr>
<tr>
<td>2.5% exponential</td>
<td>0.30–0.68</td>
<td>0–0.80</td>
<td>0.53–0.97</td>
<td>0.54–0.99</td>
<td>0–1</td>
<td>0–0.14</td>
</tr>
<tr>
<td>10% exponential</td>
<td>0.30–0.68</td>
<td>0.15–0.75</td>
<td>0.54–0.97</td>
<td>0.55–0.98</td>
<td>0–1</td>
<td>0–0.32</td>
</tr>
<tr>
<td>2.5% uniform + polyphyly</td>
<td>0.26–0.59</td>
<td>0–0.71</td>
<td>0.33–0.95</td>
<td>0.33–0.97</td>
<td>0–1</td>
<td>0–0.14</td>
</tr>
<tr>
<td>10% uniform + polyphyly</td>
<td>0.26–0.59</td>
<td>0.14–0.67</td>
<td>0.33–0.95</td>
<td>0.33–0.96</td>
<td>0–1</td>
<td>0–0.32</td>
</tr>
<tr>
<td><strong>Logistic Growth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5% uniform</td>
<td>0.37–0.62</td>
<td>0.25–0.71</td>
<td>0.59–0.97</td>
<td>0.60–0.98</td>
<td>0–0.50</td>
<td>0–0.13</td>
</tr>
<tr>
<td>10% uniform</td>
<td>0.37–0.61</td>
<td>0.29–0.63</td>
<td>0.62–0.97</td>
<td>0.62–0.98</td>
<td>0.13–0.67</td>
<td>0.04–0.30</td>
</tr>
<tr>
<td>2.5% exponential</td>
<td>0.38–0.62</td>
<td>0.25–0.70</td>
<td>0.58–0.97</td>
<td>0.58–0.98</td>
<td>0–0.50</td>
<td>0–0.13</td>
</tr>
<tr>
<td>10% exponential</td>
<td>0.37–0.61</td>
<td>0.26–0.64</td>
<td>0.64–0.97</td>
<td>0.63–0.98</td>
<td>0.13–0.71</td>
<td>0.04–0.32</td>
</tr>
<tr>
<td>2.5% uniform + polyphyly</td>
<td>0.35–0.58</td>
<td>0.21–0.68</td>
<td>0.34–0.96</td>
<td>0.34–0.97</td>
<td>0–0.50</td>
<td>0–0.14</td>
</tr>
<tr>
<td>10% uniform + polyphyly</td>
<td>0.34–0.59</td>
<td>0.25–0.59</td>
<td>0.34–0.96</td>
<td>0.35–0.96</td>
<td>0.1–0.67</td>
<td>0.04–0.30</td>
</tr>
</tbody>
</table>