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

1 The probability of ancestors in the fossil record (see Foote, 1996, *Paleobiology* 22:141-151).

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.

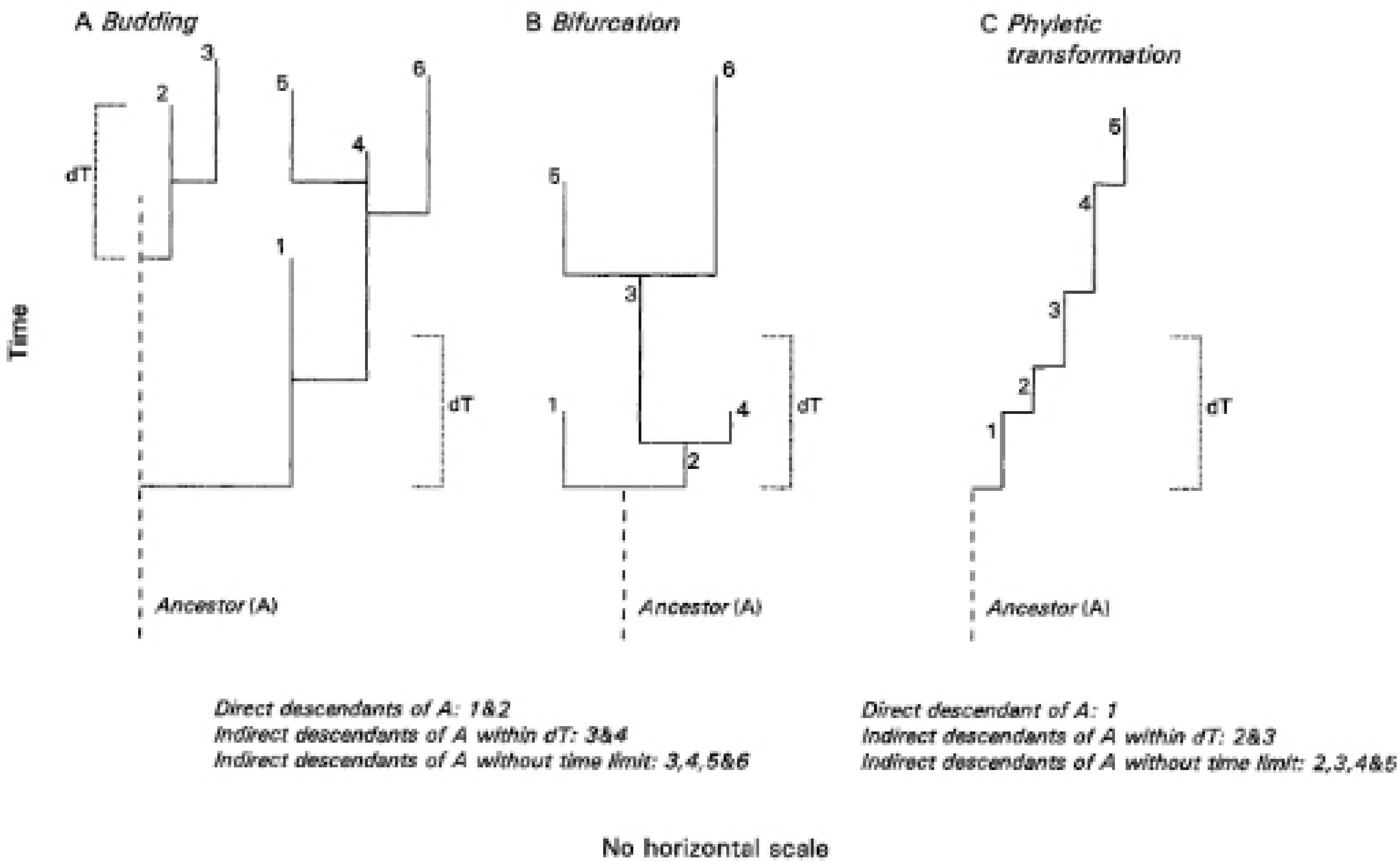

FIGURE 1. Birth-death models. A, Budding cladogenesis. A species may persist after it leaves a descendant. B, Bifurcation. A species is considered to terminate and leave two descendants at the point of branching. C, Phyletic transformation. A species either changes into a new species (pseudoextinction; 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_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-(1-R)^T. From this it follows that the proportion of species preserved is given by (Foote and Raup 1996: eq. 6)

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

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:

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

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.