B: Formulae used in the calculation of diversification rates and 95% confidence intervals of expected diversity according to the age of a group.

To estimate the rate of diversification, *r*, of a Cape radiations we followed Magallón & Sanderson's (2001) method-of-moments estimator based on the standing diversity of the clade, *n*, and it's age, *t*. Relative extinction rate, ε, is also incorporated, and is defined as:

$$\varepsilon = \frac{\mu}{\lambda} \quad (1)$$

where λ and μ are constant speciation and extinction rates respectively.

The estimated rate for a dated stem group is given by:

$$\hat{r}_\varepsilon = \frac{1}{t}\log[n(1-\varepsilon)+\varepsilon] \quad (2)$$

and for a dated crown group by:

$$\hat{r}_\varepsilon = \frac{1}{t}\left\{\log\left[\frac{1}{2}n\left(1-\varepsilon^2\right)+2\varepsilon+\frac{1}{2}(1-\varepsilon)\sqrt{n(n\varepsilon^2-8\varepsilon+2n\varepsilon+n)}\right]-\log 2\right\}. \quad (3)$$

Minimum duration of Cape presence of putative relicts was estimated from the timing of divergence of these lineages from their closest non-Cape relatives, corresponding to their stem node ages. For a dated stem group, Strathmann & Slatkin (1983) showed that the combined probability of having *k* or more species at time *t* is

$$P\left[N(t) \geq k\right] = \beta_t^{k-1} \quad (4)$$

and the combined probability of having fewer than *k* species at *t* is

$$P\left[N(t) < k\right] = 1-\beta_t^{k-1} \quad (5)$$

where β_t is a parameter which depends on time (Magallon & Sanderson 2001):

$$\beta_t = \frac{e^{rt}-1}{e^{rt}-\varepsilon}\,. \quad (6)$$

The sum of these two combined probabilities is one. Equations (4) and (5) were used to obtain critical values of k above and below which 5% of replicates of the stochastic process are expected to fall. To obtain a 95% confidence interval on a clade at time t with diversity $N(t)$ given a diversification rate, r, and relative extinction rate, ε, an upper boundary value, k_u, was calculated at or below which 97.5% of the results of the replicates of the stochastic process will fall, and a lower boundary value k_l, above which 97.5% of the results of the replicates of the stochastic process will fall. Values of k_u for every 0.1 Ma time interval from t = 0 to t = 100 Ma were obtained by fixing equation (4) to be equal to 0.025. Likewise, values for k_l were obtained by fixing equation (5) to be equal to 0.025. For the relative extinction rate, ε, a value of zero, implying no extinction forms the lower bound. We follow Magallón & Sanderson (2001) in somewhat arbitrarily choosing 0.9 as an upper bound on the relative extinction rate since at values above $\varepsilon = 0.9$, "estimated values of λ and μ begin to exceed 1.0 events per million years, which is approximately the upper limit estimated from real data from a variety of taxa..".

Magallón, S. & Sanderson, M. J. 2001 Absolute diversification rates in angiosperm clades. *Evolution* **55**, 1762-1780.

Strathmann, R. R. & Slatkin, M. 1983 The improbability of animal phyla with few species. *Paleobiology* **9**, 97-106.