Online Supplemental Information

Time-specific occupancy results–There is temporal variation in the stratigraphic distribution of fossils. To demonstrate this, we analyzed separately the results for all packages that initiate in the Neogene, Permian, Ordovician, and Cambrian. Each sediment cohort has a different occupancy pattern for fossil collections that reflects a combination of variably complete sampling and environmental differences. Culling the data in this fashion greatly reduces the number of packages and fossil collections available for analysis and, therefore, greatly increases the error associated with the estimates (null distributions not shown, but they generally encompass most of the occupancy curves).
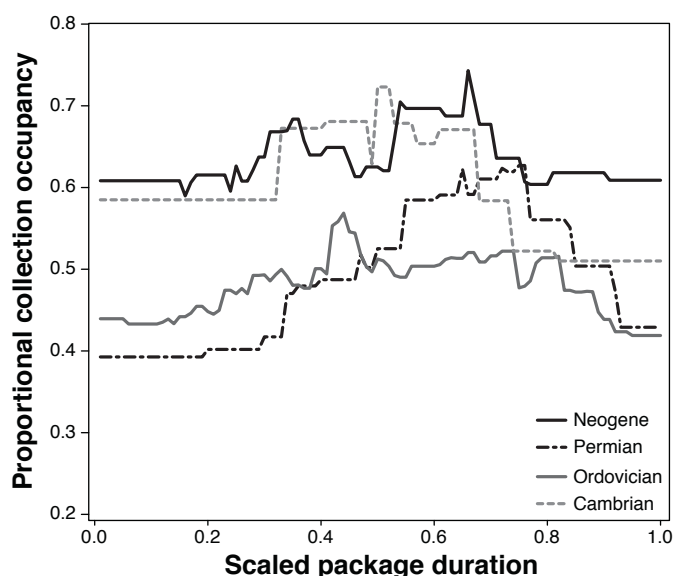

Figure DR1. Fossil collection occupancy for four sediment package initiation cohorts.

Exclusion of Late Cretaceous–The empirical distribution of fossils within sedimentary packages reflects both the actual stratigraphic distribution of fossils and sampling. Time intervals and/or stratigraphic units that have a disproportionately high level of interest can, therefore,

strongly influence our results. Peters and Heim (2010) document the geological completeness of paleontological sampling in North America and note that the Late Cretaceous is disproportionately well sampled. The potential impact of this sampling bias on our results is large because many marine sedimentary packages also truncate in the Late Cretaceous. Here, we present results similar to those in the main body of the paper, but that exclude all sediment packages that truncate in the Late Cretaceous, thereby eliminating any potential influence of disproportionately complete sampling during this time interval.
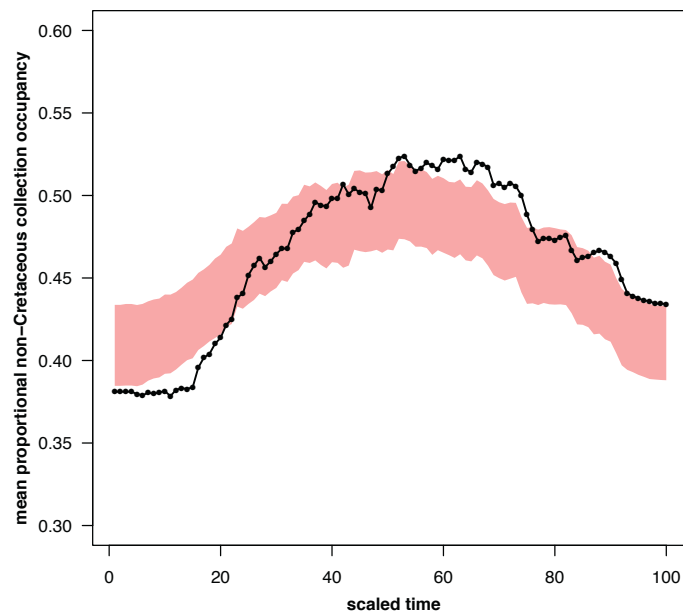

Figure DR2. Mean Phanerozoic distribution of fossil collections within gap-bound sedimentary packages excluding all packages terminating in the Late Cretaceous. Shaded region show 95% confidence limits around the expected mean occupancy based on 1000 randomizations of fossil collections within gap-bound sedimentary packages.
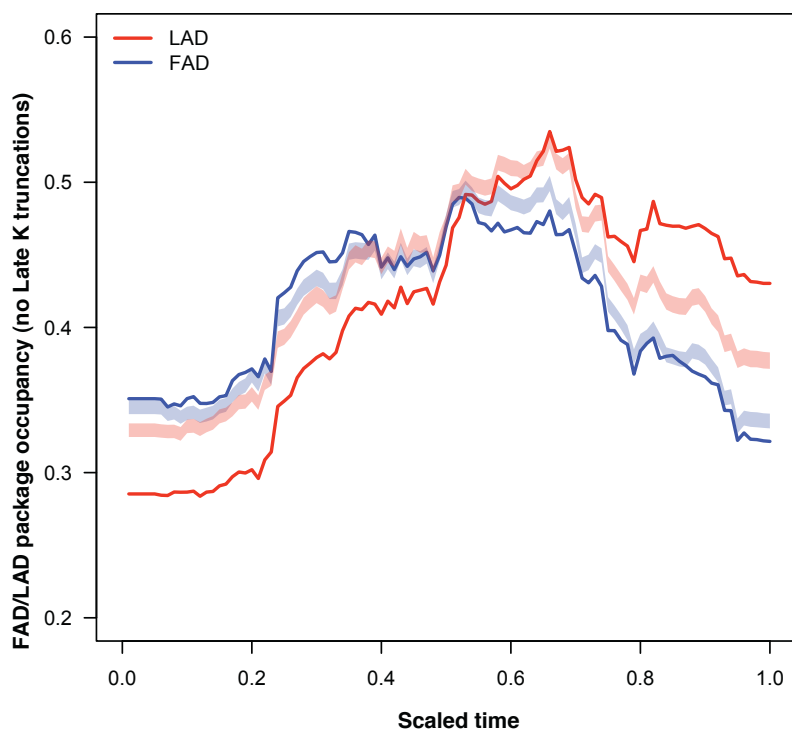
Figure DR3. Mean stratigraphic distribution of generic first and last appearance datums in North American marine sediment packages excluding all packages terminating in the Late Cretaceous. Shaded regions encompass all outcomes for 100 randomizations of FADs and LADs among collections within packages. Null distributions differ because packages that contain FAD- and LAD-defining collections are a partially overlapping subset of all fossil-bearing packages.

Randomization of individual collections–Our primary interest was in testing for significant patterns in the distribution of fossil collections within the ranges of sedimentary packages, not for the evenness (i.e., clustering) of the distribution of fossil collections within units. The results we present in the main body of the text accomplishes this test by shuffling the identify of the stratigraphic units for entire groups of fossil collections assigned to the same unit, rather than

randomly assigning each collection to a randomly chosen rock unit. Here, we present the results

for this later randomization.

The observed occupancy of fossil collections and genera (from Fig. 2) is entirely below

the null distribution obtained by randomizing the stratigraphic unit of each fossil collection indi-

vidually. This indicates that fossil collections in the PaleoDB are more "clustered" within lithos-

tratigraphic units than expected due to chance alone (clustering results in single sharp-peaked

occupancy curves surrounded by many zero values, whereas packages with evenly distributed

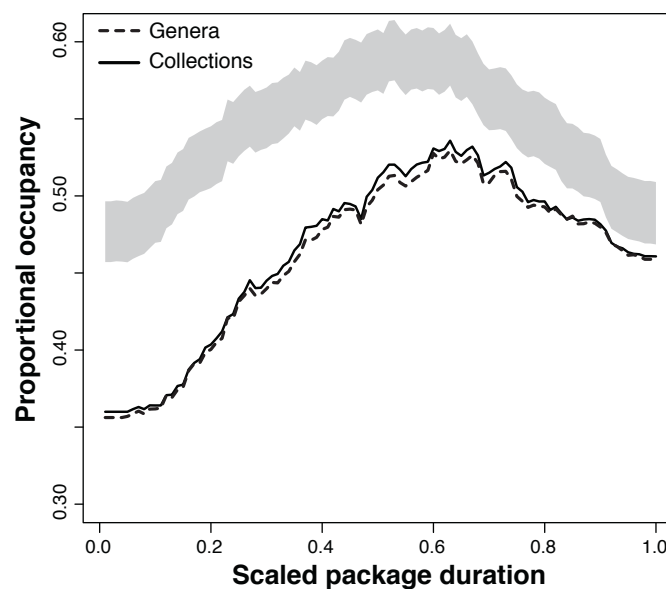collections have few zero values and a higher average occupancy).

Fig. DR4. Observed distribution of fossil genera and collections within packages and 95% confi-

dence limits around expected distribution based on the randomization of the stratigraphic unit of

individual fossil collections. Compare null distribution here to randomization in Fig. 3, where

groups of collections assigned to the same unit were shuffled rather than individual collections.