

# Successive approximations of diversity curves: Ten more years in the library

John Alroy National Center for Ecological Analysis and Synthesis, 735 State Street, Santa Barbara, California 93101, USA

## ABSTRACT

**Traditional paleontological diversity curves are based on tallies of all taxa appearing in formally defined time units. These tallies are thought to be robust to further data collection. Thus, they supposedly do not reflect nuisance factors like variable time unit lengths and sampling intensity biases. A comparison of a decade-old North American Cenozoic mammal diversity curve and a newer database shows major differences. At least three major factors differentiate the two: use of shorter, fixed-length time intervals; restriction of counts to taxa that cross boundaries between intervals; and correction for variation in sampling intensity. The difference between genus- and species-level data also was examined, but appears to be minor by comparison. Because at best only one pattern can be close to the true historical trajectory, the analyses suggest that the new data and protocols together have yielded a curve that is converging on an accurate signal.**

**Keywords:** Cenozoic, diversity curves, macroevolution, mammals, sampling.

## INTRODUCTION

One of paleobiology's most crucial roles is to present and analyze diversity curves showing how taxonomic richness evolves through geological time. Major data sets like those of Sepkoski (1978, 1979, 1984) serve as a gold standard against which all other biological trends in Earth history are contrasted. These meticulous compilations intuitively seem like the best possible approximations of true historical patterns. However, major questions about the veracity of these curves have existed for three decades. For example, the number of taxa appearing in each time interval appears to be strongly correlated with indicators of sampling intensity (Raup, 1972, 1976). These skeptical concerns have been countered with two key arguments: that apparently independent data sets yield qualitatively similar patterns (Sepkoski et al., 1981), and that large amounts of additional compilation only result in small changes (Sepkoski, 1993).

Over the past few years, the question of bias has once again become a major concern. Variation through time in the number of studied fossils demonstrably has major effects on diversity curves (Alroy, 1996; Miller and Foote, 1996; Markwick, 1998), as do secular variation in the nature of the rock record (Holland, 1995) and variation in the lengths of time units used to compute turnover rates (Foote, 1994).

Here I compare two major data sets that seemingly pertain to exactly the same phenomenon: the Cenozoic radiation of North American mammals. Within the terrestrial realm, the North American mammal record is equaled only by the European mammal record (Savage and Russell, 1983; Hunter and Jernvall, 1995). Taxonomic problems are relatively few, geochronological calibration points are numerous, and stratigraphic correlations are relatively well understood (Woodburne, 1987).

Stucky (1990) summarized this large body of literature in the first paper to rigorously analyze complete genus-level data resolved down to the level of North American land-mammal ages (NALMAs), which average between 3 and 4 m.y. in duration. Stucky personally collected many of his Paleogene faunal samples, and he extensively vetted the taxonomic nomenclature. His data are a very reasonable summary of the state of the art in mammal paleontology one decade ago.

I have built and continually updated a taxonomically standardized, locality-level faunal database summarizing the Cretaceous and Cenozoic record of North American mammals (Alroy, 1992, 1996, 1998, 2000b). Appearance event ordination (Alroy, 1992, 1994) was used to infer a complete, relatively ordered sequence of first and last appearance events, and geochronological age estimates were used to calibrate the sequence. Traditional NALMAs were abandoned; instead, diversity data were computed for evenly spaced, 1.0-m.y.-long intervals. Diversity counts were based on taxa crossing the boundaries between intervals, instead of all taxa appearing anywhere in each interval (the traditional method). Because the record is heavily dominated by localities in the western United States and Canada, faunal lists from other regions were excluded. The data were standardized to uniform sampling intensity levels using a randomized subsampling protocol. Thus, the final data show what the fossil record would look like at discrete, evenly spaced moments in geological time if sampling were completely uniform.

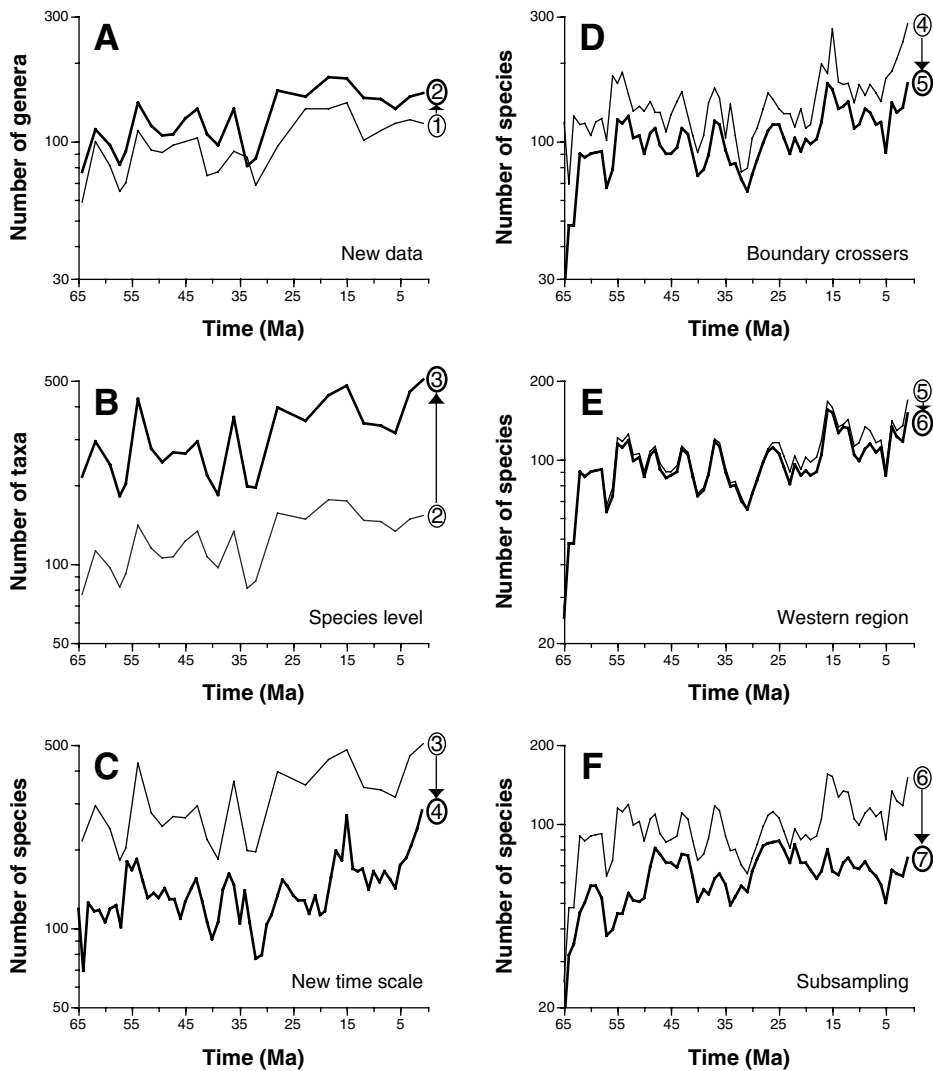
This paper focuses on two key issues. First, are the primary data in these two studies comparable? To answer this question, the new data are binned into the same time intervals as Stucky's and at first subjected to no further treatment. This question of whether new compilations yield different results has figured in previous

analyses of diversity patterns (foraminiferans: Sepkoski, 1993; plants: Niklas and Tiffney, 1994; dinoflagellates: MacRae et al., 1996), origination and extinction rates (brachiopods: Williams, 1957), geographic ranges (gastropods: Vermeij, 1999), and multivariate morphometrics (echinoderms and trilobites: Foote, 1997). Second, do the additional steps that I advocated (Alroy, 1996, 1998, 2000b) really make any difference? The importance of each factor is tested by quantifying the similarity between Stucky's curve and each successively refined version of the new curve, and by showing how each successive approximation differs from the last one. In addition to computing simple correlations, I perform additional analyses in which the data are detrended and autocorrelation is removed.

## DATA

Stucky's data set (curve 1 in Fig. 1) consists of counts of genera within NALMAs. His original data were consistently subdivided into numbered subages (e.g., Puercan 1) throughout the Paleocene and Eocene, but divisions were broader in the Oligocene and Neogene: there were 32 Puercan-Chadronian intervals (about 1/1.0 m.y.), but only 11 for the Orellan-Hemphillian (about 1/2.6 m.y.). Even the few Neogene NALMA subdivisions were contentious; the Irvingtonian zones are an example (Bell and Repenning, 1999). All of the Paleocene NALMAs, the end-Eocene Chadronian, and the very short early Oligocene Orellan and Whitneyan also appear to have been oversplit.

For these reasons, it was not possible to draw meaningful comparisons between all of Stucky's original 43 intervals and the new North American mammalian paleofaunal database time scale. Therefore, I was forced to convert his data into counts for entire NALMAs, except where a two-way subdivision could be determined easily and the entire NALMA had a duration of at least 3.0 m.y. (i.e., Tiffanian, Wasatchian, Bridgerian, Uintan, Duchesnean, Hemphillian). The less obvious splits of Stucky's NALMAs were: Tiffanian 2–4 versus 5; Wasatchian 1–4 versus 5; Bridgerian 1 versus 2 + 3; and Uintan 1 + 2 versus 3. Stucky's "Tiffanian 1" was included in the Torrejonian. The "Early Arikarean" (Geringian + Monroecreekian, 6.0 m.y.) and several relatively long Neogene NALMAs like the Hemingfordian (4.3 m.y.) were left unsplit because they were single units in Stucky's data. Stucky's RanchoLabrean and Recent data were excluded because of their short duration. The Lancian (latest Cretaceous) was excluded in order to confine the data set to the Cenozoic.



**Figure 1.** Changes in diversity curves resulting from collection of new data and corrections for artifacts. **A:** Data of Stucky (1990; curve 1) compared with new data (curve 2). **B:** New data for genera (curve 2) compared with species-lineage data (curve 3). **C:** Species-lineage data for traditional, North American land-mammal age-based time intervals (curve 3) compared with same data partitioned into 1.0-m.y.-long sampling bins (curve 4). **D:** Counts of species in sampling bins (curve 4) compared with counts of species crossing boundaries going into bins (curve 5). **E:** Counts of boundary-crossing species across North America (curve 5) compared with counts restricted to western North American region (curve 6). **F:** Counts of western North American species (curve 6) compared with sampling-standardized counts (curve 7).

A final set of 25 intervals was employed. Although the variance of the interval lengths is substantial (logged data: mean = 0.870,  $1 \sigma = 0.481$ ), the mean and standard deviation would have been equal instead if boundaries had been generated by a random Poisson process. Furthermore, the interval lengths do not correlate with the midpoint ages of the intervals ( $r = -0.299$ ,  $t = 1.274$ , n.s.).

The methods used to prepare the North American paleofaunal database were described previously (Alroy, 2000b, and earlier papers; Wing et al., 1995). This new data set includes 4713 faunal lists, 29,795 taxonomic records, 1223 genera, and 3207 species, and is based on 2760 publications (Alroy, 2000a). The lists were subjected to maximum likelihood appearance event

ordination (Alroy, 2000b), which considers such factors as stratigraphic superposition and differing probabilities of preservation among taxa. The appearance event sequence was calibrated using 106 high-quality (i.e.,  $^{40}\text{Ar}/^{39}\text{Ar}$ , paleomagnetic, and uranium series) geochronologic age estimates. Instead of the earlier hinge interpolation method (Alroy, 1996, 1998, 1999), this analysis employs the shrink-wrap nonparametric interpolation method (Alroy, 2000b), which makes use of as many calibration points as possible. The calibration explains 99.91% of the variance in the age estimates. Sampling standardization was performed using the smoothed version of the records-squared list subsampling algorithm (Alroy, 2000b), in contrast to the records-tallied method used earlier (Alroy, 1996, 1998, 1999).

Several other, earlier data sets are not comparable: Romer (1966) defined only 16 distinct subepoch level time units for the entire Cenozoic; Lillegraven (1972) published family and order data only; much of the Oligocene and Neogene data of Savage and Russell (1983) were repeated by Stucky; Van Valkenburgh and Janis (1993) published latest Eocene and Neogene data for large mammals only; and Hunter and Jernvall (1995) based their data on Savage and Russell (1983). Thus, the current comparisons focus exclusively on Stucky's data.

## ANALYSES

In addition to curve 1, representing Stucky's data, six other diversity curves were prepared from the new data set (Fig. 1, Table 1). Curves 2 and 3 also use the coarse, NALMA-based time scale. Curve 2 (new genus-level data) is consistently higher than curve 1 (Fig. 1A), probably because of newly described taxa and age-range extensions: 702 publications in the database (25.4%) appeared during the 1990s, and 116 new genera were described (9.3%). Only one bin (Orellan) shows lower diversity in the new curve. This case may relate to extensive synonymizing of such mid-Tertiary large-mammal taxa as oreodonts (e.g., Janis et al., 1998).

Curve 3 counts species instead of genera. The counts are augmented by assuming that if a genus ranges through a bin where no particular species of this genus has been identified, at least one species must be present (i.e., species data are transformed to species lineages; Alroy, 1996, 1998). Curve 3 is consistently two to three times higher than curve 2 (Fig. 1B), which results from the data set's overall 2.6:1 species to genus ratio.

The other four species-level curves use the equally spaced 1.0-m.y.-long sampling bins. The shorter bins separate many species in the same NALMA that never actually coexisted, so curve 4 is much lower (Fig. 1C). Curve 5 is lower still (Fig. 1D) because the counts represent tallies of species crossing boundaries between bins instead of ranging anywhere into a bin. Because all species crossing one point in time must have coexisted at that time, curve 4 represents bona fide faunal cohorts. Curve 6 restricts the data to the western region of the United States and Canada; because most data already come from this region, almost no change results (Fig. 1E). The sampling-standardized curve 7 lowers the data one final time (Fig. 1F).

Although most pairs of curves have very different means, what really matters is not these offsets but similarities and differences in shape. The easiest way to quantify shape similarity is to compute correlation coefficients. The data are first-log transformed, which reduces skewness. To compare curves using different time scales, data from the more highly resolved data set are boiled down by computing the geometric mean of all diversity estimates falling in each NALMA-based time interval.

All but two of the correlations between consecutive versions of the curves (Last [raw], Table 1) are at or below +0.895, meaning that pairs of curves share only about four-fifths of their variance. Thus, a fifth or more of the signal in the curve is transformed at most steps. The exceptions involve the highly correlated genus- and species-level data (curves 2 and 3) and the essentially pro forma geographic restriction of the data set (curves 5 and 6).

The correlations are mostly lower for comparisons of all curves to Stucky's and to the sampling-standardized version. Oddly, some curves (e.g., curve 5) are just as similar to computationally related curves as they are to computationally distant curves. This pattern may be attributable to the statistical perils of cross-correlating time series that show secular trends and autocorrelation.

Generalized differencing (McKinney, 1990) helps to remove these effects. It works by first detrending the data (by taking residuals of the diversity values regressed against numerical time), and then removing autocorrelation (by taking differences between neighboring values). It is important that these differences are modulated by the strength of the correlation between the neighbors (Lagged, Table 1). When autocorrelation is very weak, this weighting leaves the curve unchanged; when it is strong, the generalized differences converge on simple first differences.

All of the curves show a strong increase from the Paleocene to the Pleistocene (Time, Table 1). Therefore, linear detrending has a major impact (detrended values, Table 1). All of the correlations drop, although the ones involving neighboring curves are still moderate. Correlations between curve 1 and the others are substantially weakened. In particular, the curve 1 versus curve 2 comparison yields an  $r^2$  of just 0.66. In other words, 34% of the variance in Stucky's data is not replicated by the new data set, even though the curves were prepared in exactly the same way. Correlations involving the sampling-standardized curve are even lower, and in one case just under zero (curve 4 vs. curve 7, Table 1).

Generalized differencing has a more subtle effect (differenced values, Table 1). All but one of the correlations involving curve 1 fall, and now none of them implies an  $r^2$  higher than 0.59. Parametric significance tests are now appropriate because the data retain no autocorrelation, but only the first two comparisons are significant (C1 [differenced], Table 1). So all the curves go up through time, but the interval-by-interval pattern of change in Stucky's data is different from those seen in the new curves employing a highly resolved time scale.

Quite to the contrary, correlations involving last-step curves or curve 7 often stay the same or even increase slightly with generalized differencing. None of the NALMA-based curves are significantly correlated with the standardized data.

TABLE 1. CORRELATIONS BETWEEN DIVERSITY CURVES

	C1	C2	C3	C4	C5	C6	C7
Last (raw)	N/A	+0.895	+0.928	+0.708	<b>+0.698</b>	<b>+0.990</b>	<b>+0.730</b>
Last (detrended)	N/A	+0.814	+0.881	+0.531	<b>+0.572</b>	<b>+0.990</b>	<b>+0.617</b>
Last (differenced)	N/A	+0.765*	+0.910*	+0.517	<b>+0.350</b>	<b>+0.989*</b>	<b>+0.707*</b>
C1 (raw)	N/A	+0.895	+0.857	+0.665	+0.796	+0.773	+0.691
C1 (detrended)	N/A	+0.814	+0.762	+0.443	+0.629	+0.606	+0.487
C1 (differenced)	N/A	+0.765*	+0.765*	+0.440	+0.582	+0.569	+0.379
C7 (raw)	+0.691	+0.706	+0.540	<b>+0.251</b>	<b>+0.730</b>	<b>+0.730</b>	N/A
C7 (detrended)	+0.487	+0.507	+0.263	<b>-0.039</b>	<b>+0.598</b>	<b>+0.617</b>	N/A
C7 (differenced)	+0.379	+0.505	+0.350	<b>+0.003</b>	<b>+0.711*</b>	<b>+0.707*</b>	N/A
Time	-0.655*	-0.669*	-0.618*	<b>-0.505*</b>	<b>-0.596*</b>	<b>-0.535*</b>	<b>-0.553*</b>
Lagged	+0.510	+0.514	+0.382	<b>+0.674*</b>	<b>+0.777*</b>	<b>+0.732*</b>	<b>+0.842*</b>

*Note:* Diversity values are logged before all analyses. Correlations in bold have a sample size of 65 because both time series employ the new time scale; sample size is 25 for other correlations. Because other comparisons involve pairs of autocorrelated data sets, significance tests are only performed for correlations against time; lagged correlations; and correlations involving differenced data. Correlations are based on Pearson's equation, which assumes normality, but rank-order correlations (not shown) are similar. First three rows give data contrasting each curve with the preceding curve in Figure 1; next three, contrasting curve 1 and others; last three, contrasting curve 7 and others. Each set of three rows gives results for raw, untransformed data, detrended data, and data after generalized differencing. C1–C7 = curves 1–7 (Fig. 1, Table 1); Last = correlation against preceding variable; Time = correlation against time in Ma; Lagged = correlation of values for time T against values of the same variable for time T – 1. N/A = not applicable because the two variables are the same.

\*Value of  $p < 0.001$ .

Despite this, all but two neighboring pairs do show substantial and significant correlations. Thus, each separate step has a constrained, but still notable, effect on the shape of the curve. The cumulative result after several steps often is a completely altered pattern.

## CONCLUSIONS

These simple analyses imply four major conclusions.

First, all of the curves climb, making it inevitable that most neighboring pairs of curves will show a moderate correlation. Nonetheless, these raw, uncorrected cross-correlations are mostly below +0.9 or even +0.8. Thus, 20%, 30%, or more of the variance in each curve must relate strictly to methodology. Because curves 2–7 are based on exactly the same raw data and underlying sequence of appearance events, the differences cannot be attributed to random sampling error.

Second, the more rigorous time series analyses show again that most analytical steps do have a major impact (Last [differenced]; Table 1). These key steps are: (1) switching from a traditional, NALMA-based time scale to a uniformly split, highly precise, quantitative biochronology-based time scale (curves 3 vs. 4); (2) switching from traditional counts of all species falling in bins to counts of species crossing between bins (curves 4 vs. 5); and (3) standardizing the data for variation in sampling intensity by means of randomized subsampling (curve 6 vs. 7). In each case, the shared variance between detrended shapes of neighboring curves is 50% or less. The corrections have the cumulative effect that none of the NALMA-based curves shows any significant shape similarity to the final, most refined curve.

Third, a major exception involves switching from genus-level data to species-level data. These curves share an impressive 83% of shape variance. Each type of data has failings. Workers are more hesitant to name new genera on marginally distinctive and poorly preserved material, but genera may be polyphyletic even when all species are valid (and so may inflate diversity), or they may collapse biologically distinct, coeval lineages of species (and so may deflate diversity). The substantial shape similarity between the differenced curves suggests that workers should think carefully when choosing between genus- and species-level data.

Fourth, no curve employing species-level data carries over any meaningful signal from the data set of Stucky (1990). Thus, the old and new data agree that diversity has increased since the Cretaceous-Tertiary boundary mass extinction, but not on much else. For example, a marked Hemphillian drop in diversity is only visible in the species-level data. The 1990 data even fail as mere indicators of sampling trends. The enormous differences between curves 6 and 7 do correctly indicate where peaks resulting from rich sampling should lie (e.g., Wasatchian, Barstovian; see also Alroy, 1996, 1998, 2000b), but these peaks are not clearly visible in the older data.

One false conclusion that could be drawn is that the work of one researcher or the other was somehow negligent. This skepticism is belied by Stucky's (1990) obviously painstaking research, and by the present study's use of all the same older literature plus hundreds of new papers. If one really wanted to dismiss either data set, then one would also have to cast doubt on virtually every published paleontological diversity curve.

Another misapprehension would be the idea that if analytical steps make major differences to the shapes of curves, then the most refined data sets must be meaningless. To the contrary, every one of these steps is intended to draw out original biological signals in the fossil record. That is particularly true of the key steps: refining correlations, counting coeval cohorts of species, and standardizing sampling. The importance of sampling standardization is only emphasized by the substantial offset between the two independent raw data compilations (C1, C2). Clearly, then, additional sampling can and will have major impacts, so removing sampling artifacts is crucial. Standardized data may not be perfect, but by virtue of the very fact that they differ systematically from data that must suffer from strong artifacts, they are likely to be converging on the truth.

Thus, although we see enormous differences between diversity curves that all represent the same historical pattern, this does not suggest that there are insoluble problems with diversity analysis. Instead, the results vindicate two key conclusions: properly preparing diversity data is remarkably important, and even one decade of study can lead to substantial, progressive changes in a continent- and era-scale data set for a major taxonomic group.

#### ACKNOWLEDGMENTS

I thank M. Foote, P. Koch, D. Raup, and R. Stucky for helpful comments on the manuscript. This work was conducted at the National Center for Ecological Analysis and Synthesis, a center funded by the National Science Foundation (grant DEB-94-21535), the University of California, Santa Barbara, the California Resources Agency, and the California Environmental Protection Agency.

#### REFERENCES CITED

Alroy, J., 1992, Conjunction among taxonomic distributions and the Miocene mammalian biochronology of the Great Plains: *Paleobiology*, v. 18, p. 326–343.

Alroy, J., 1994, Appearance event ordination: A new biochronologic method: *Paleobiology*, v. 20, p. 191–207.

Alroy, J., 1996, Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals: *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 127, p. 285–311.

Alroy, J., 1998, Equilibrial diversity dynamics in North American mammals, in McKinney, M.L., and Drake, J., eds., *Biodiversity dynamics: Turnover of populations, taxa and communities*: New York, Columbia University Press, p. 232–287.

Alroy, J., 1999, The fossil record of North American mammals: Evidence for a Paleocene evolutionary radiation: *Systematic Biology*, v. 48, p. 107–118.

Alroy, J., 2000a, North American mammalian paleofaunal database: <http://www.nceas.ucsb.edu/~alroy/nampfd.html> (accessed August 2000).

Alroy, J., 2000b, New methods for quantifying macroevolutionary patterns and processes: *Paleobiology*, v. 26 (in press).

Bell, C.J., and Repenning, C.A., 1999, Observations on dental variation in *Microtus* from the Cudahy Ash Pit fauna, Meade County, Kansas, and implications for Irvingtonian microtine rodent biochronology: *Journal of Vertebrate Paleontology*, v. 19, p. 757–766.

Foote, M., 1994, Temporal variation in extinction risk and temporal scaling of extinction metrics: *Paleobiology*, v. 20, p. 424–444.

Foote, M., 1997, Sampling, taxonomic description, and our evolving knowledge of morphological diversity: *Paleobiology*, v. 23, p. 181–206.

Holland, S.M., 1995, The stratigraphic distribution of fossils: *Paleobiology*, v. 21, p. 92–109.

Hunter, J.P., and Jernvall, J., 1995, The hypocone as a key innovation in mammalian evolution: *National Academy of Sciences Proceedings*, v. 92, p. 10,718–10,722.

Janis, C.M., Scott, K.M., and Jacobs, L.L., 1998, *Evolution of Tertiary mammals of North America, Volume 1: Terrestrial carnivores, ungulates, and ungulatelike mammals*: Cambridge, UK, University of Cambridge Press, 691 p.

Lillegraven, J.A., 1972, Ordinal and familial diversity of Cenozoic mammals: *Taxon*, v. 21, p. 261–274.

MacRae, R.A., Fensome, R.A., and Williams, G.L., 1996, Fossil dinoflagellate diversity, originations, and extinctions and their significance: *Canadian Journal of Botany*, v. 74, p. 1687–1694.

Markwick, P.J., 1998, Crocodylian diversity in space and time: The role of climate change in paleoecology and its implication for understanding K/T extinctions: *Paleobiology*, v. 24, p. 470–497.

McKinney, M.L., 1990, Classifying and analyzing evolutionary trends, in McNamara, K.J., ed., *Evolutionary trends*: Tucson, University of Arizona Press, p. 28–58.

Miller, A.I., and Foote, M., 1996, Calibrating the Ordovician radiation of marine life: Implications for Phanerozoic diversity trends: *Paleobiology*, v. 22, p. 304–309.

Niklas, K.J., and Tiffney, B.H., 1994, The quantification of plant biodiversity through time: *Royal Society of London Philosophical Transactions, ser. B*, v. 345, p. 35–44.

Raup, D.M., 1972, Taxonomic diversity during the Phanerozoic: *Science*, v. 177, p. 1065–1071.

Raup, D.M., 1976, Species diversity in the Phanerozoic: An interpretation: *Paleobiology*, v. 2, p. 289–297.

Romer, A.S., 1966, *Vertebrate paleontology*: Chicago, Illinois, University of Chicago Press, 468 p.

Savage, D.E., and Russell, D.E., 1983, *Mammalian paleofaunas of the world*: Reading, Massachusetts, Addison-Wesley, 432 p.

Sepkoski, J.J., Jr., 1978, A kinetic model of Phanerozoic taxonomic diversity. I. Analysis of marine orders: *Paleobiology*, v. 4, p. 223–251.

Sepkoski, J.J., Jr., 1979, A kinetic model of Phanerozoic taxonomic diversity. II. Early Phanerozoic families and multiple equilibria: *Paleobiology*, v. 5, p. 222–251.

Sepkoski, J.J., Jr., 1984, A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions: *Paleobiology*, v. 10, p. 246–267.

Sepkoski, J.J., Jr., 1993, Ten years in the library: New data confirm paleontological patterns: *Paleobiology*, v. 19, p. 43–51.

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

Stucky, R.K., 1990, Evolution of land mammal diversity in North America during the Cenozoic: *Current Mammalogy*, v. 2, p. 375–431.

Van Valkenburgh, B., and Janis, C.M., 1993, Historical diversity patterns in North American large herbivores and carnivores, in Ricklefs, R.E., and Schluter, D., eds., *Species diversity in ecological communities: Historical and geographical perspectives*: Chicago, Illinois, University of Chicago Press, 414 p.

Vermeij, G., 1999, The accumulation of taxonomic knowledge: The history of species descriptions of some predatory gastropods: *Malacologia*, v. 41, p. 147–150.

Williams, A., 1957, Evolutionary rates of brachiopods: *Geological Magazine*, v. 94, p. 201–211.

Wing, S.L., Alroy, J., and Hickey, L.J., 1995, Plant and mammal diversity in the Paleocene to early Eocene of the Bighorn Basin: *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 115, p. 117–155.

Woodburne, M.O., 1987, *Cenozoic mammals of North America: Geochronology and biostratigraphy*: Berkeley, University of California Press, 336 p.

Manuscript received March 13, 2000

Revised manuscript received August 18, 2000

Manuscript accepted August 25, 2000