

Community inertia of Quaternary small mammal assemblages in North America

Brian J. McGill*[†], Elizabeth A. Hadly[‡], and Brian A. Maurer[§]

*Department of Biology, McGill University, Stewart Biology Building, Montreal, QC, Canada H3A 1B1; [†]Department of Biological Sciences, Stanford University, 371 Serra Mall, Stanford, CA 94305-5020; and [§]Department of Fisheries and Wildlife, Michigan State University, Natural Resources Building, Room 13, East Lansing, MI 48824

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved September 27, 2005 (received for review May 21, 2005)

One of the longest running debates in ecology is whether chance or determinism structures biotic communities, and this question is often studied by looking for the presence or absence of community inertia (lack of change) over time or space. Results have been equivocal. We adopted three tactics for a fresh approach: (i) allowing the answer to vary with the geographic, temporal, and taxonomic scale of study, (ii) using appropriate reference points for the amount of inertia in random biological systems, and (iii) using a robust approach for measurement of inertia. We examined fossil assemblages of mammalian communities across almost 1,000,000 years and at sites spanning $\approx 3,500$ km. We showed that in general there is good evidence for inertia but that the results change in a quantifiable fashion with taxonomic, spatial, and temporal scales. By using neutral theory we place a reference point on the degree of inertia and demonstrate that empirical mammalian communities show greater inertia than neutral communities over time scales $> 3,000$ year. Although our results do not specifically reveal mechanism, they emphasize that deterministic forces are at work in structuring communities over millennia.

neutral theory | paleoecology

One of the oldest debates in ecology (1) is between the view that there are deterministic processes that lead to predictable patterns in the structure of communities and the idea that random or at least unpredictable factors dominate that structure. A specific debate within this broad question is whether or not communities retain a coherent structure as abiotic forcing conditions change across space and time. We will call this putative coherency “community inertia.” Although this is an important and old question (2, 3), it is still largely unresolved and hotly debated today. Traditionally, community inertia has been quantified by studying whether relative abundances of taxa change across time or space. Neoecologists have made claims of inertia (4–6) and lack of inertia (7) in space as well as inertia (4) and lack thereof (8) in time. Paleoecologists have primarily focused on time and have also produced claims of inertia (9–16) and lack of inertia (17–21). Much of the debate has been of the unproductive form of “there’s inertia in my system” vs. “there’s not inertia in my system” (but see refs. 11, 14, and 15). Perhaps Simberloff (22) is correct in suggesting that communities are too contingent and context-dependent for general rules to be found. However, we suggest that improved methods for analyzing community dynamics may yet yield general rules. Full and careful consideration of three challenges can begin to provide a resolution to this debate. We elaborate these three factors in the next three paragraphs.

One obvious challenge is that the patterns obtained depend on (and may well change with) temporal, spatial, and taxonomic scale (23, 24). Yet rarely have there been attempts to make statements specific to scale (but see ref. 14). To fully address the question in the context of scale is a four-dimensional problem (inertia is a function of spatial, temporal, and taxonomic scale), which may further change for different groups of organisms. Such concerns are not realistically addressable by one data set,

researcher, or manuscript. But it is crucial to begin to include these dimensions in studies of community inertia. In this article, we choose to explore the effects of scale along the time dimension with a strong secondary emphasis on taxonomic scale and a minor emphasis on space. We also focus on a single taxonomic group (small mammals), but our approach can be adopted for other taxonomic groups.

The second challenge is that it is impossible to give a simple yes/no answer to the question of whether there is community inertia or not. The correct answer is “both;” there is always some inertia and some change. There are no purely deterministic, noise-free systems and no truly random (white noise) systems. Suppose we measure community inertia on a scale of 0 (complete randomness) to 1 (full inertia). We call this a community inertia index (CII). Measurements of CII are invariably intermediate (between 0 and 1). Whether to call this result a sign of inertia has been largely determined by subjective assessments of what qualifies as a little or a lot of inertia (but see ref. 19). In practice, it would be useful if we could place reference points on this scale of 0–1 that would allow us to say that a system has more or less inertia than a meaningful reference point. But it has been very difficult to select an appropriate model of stochasticity in community ecology (25). There is now an obvious reference point to help calibrate the scale (15), the neutral theory of biogeography (26, 27), which explicitly incorporates a random dynamic process based on ecologically identical species. This stochastic process results in a drift of abundance within communities that can act as a reference point for the measurement of community inertia. The neutral model also includes the idea of ongoing immigration and reassembly from a regional pool (metacommunity), which has been considered an appropriate null model by paleontologists for analyzing community inertia (19).

The third challenge is the method of quantification of similarity of communities (i.e., the CII). Dozens of measures have been proposed (28). At one extreme are measures that ignore abundance and only use presence/absence, such as the Jaccard index. At the other extreme are measures that place heavy emphasis on small differences in rare species, such as the squared chord distance (SCD), because a change in abundance from one individual to two represents a 100% increase (29). It has been argued, for example, that the identification of nonanalog (non-inertial) communities relies on these properties of the SCD (29). Our approach to this problem is based on the idea of robustness. We focus on results that hold true across a variety of CIIs.

Methods

Because we want to emphasize the time dimension, we necessarily chose a paleontological system. Neoecological time series

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CII, community inertia index; ybp, years before present.

[†]To whom correspondence should be addressed. E-mail: mail@brianmcgill.org.

© 2005 by The National Academy of Sciences of the USA

lower taxa were randomly reassigned to the higher taxa (reshuffled) while keeping the distributions of counts of species among genera, etc. For example, there is one genus, *Sorex*, containing eight species and an “unidentified *Sorex*” category, so in the reshuffling one genus received nine random species. This reshuffling destroyed any taxonomic associations and provides a null for the effect of aggregation when analyzing higher taxonomic levels.

One issue that must be addressed is the question of time averaging. Paleontological data are assigned to a layer or stratigraphic unit, which represents a span of time (typically tens to thousands of years for our data), and it is known that averaging abundances over time can change the apparent structure of a community (39). This averaging could potentially be a problem in comparing empirical data with different degrees of time averaging or comparing empirical data with data from neutral model simulations that have no time averaging. In particular, our third type of comparison (whole-site) does have an unusually high degree of time averaging relative to the first two types of comparisons and might suffer from these problems to a greater degree, but we include this analysis because of its conceptual simplicity and confirm any results obtained in this analysis with the other types of comparisons. For the first two types of analyses, we believe that time averaging did not affect our results because (i) time-averaged woodrat data such as we mostly used have been shown to be a good predictor of modern day (i.e., not time-averaged) communities (34, 40); (ii) time averaging introduces distortions to the abundances of rare species that varies significantly with the amount of time averaging (such as between our layers of greatly different duration) (39) and is therefore primarily a source of increased noise in the empirical data and thereby conservatively should create lower CII and favor neutral (unaveraged) data; and (iii) the degree of difference found between neutral and empirical data was large enough so it seems unlikely that a factor such as time averaging could affect this result, especially over just a few hundreds or thousands of years.

Results

Whole-site comparisons indicated that abundances of taxa among sites were positively correlated among all pairs of sites (Fig. 1 and Table 3, which is published as supporting information on the PNAS web site). Three of the comparisons were statistically significant. The odds of achieving three significant tests of six by chance (for $\alpha = 0.05$) are minimal ($P < 0.0001$ binomial test). This finding provides evidence for community inertia, but the role of historical contingency (41) was also clear in the empirical data. The three least similar comparisons all involve Baker’s Bluff. On these comparisons, *Ochotona* (pikas), *Thomomys* (pocket gophers), and *Clethrionomys* (red-backed voles) were outliers, being reasonably common at one site and not present at the other. *Ochotona* and *Thomomys* are taxa only found in the western United States and thus have no possibility of being present in Baker’s Bluff. A post hoc removal of these two species causes the Lamar/Baker’s Bluff correlation to rise to $r = 0.52$ with $P < 0.05$. Likewise, *Clethrionomys* arrived from Eurasia after the deposition in Porcupine Cave so additionally removing this genus from the analysis causes the Porcupine/Baker’s Bluff comparison to become significant ($r = 0.51$, $P < 0.05$). Hence, the cause of the three nonsignificant but positive correlations between communities involved historical/biogeographic contingency involving these three species.

A more sophisticated comparison takes into account the fact that community inertia may behave differently at different spatial, temporal, and taxonomic scales and avoids the large degree of time averaging in the previous analysis. We repeated this analysis by using the between-sites and within-sites and between-time comparisons, and the results are summarized in Fig. 2 and in Table 4, which is published as supporting infor-

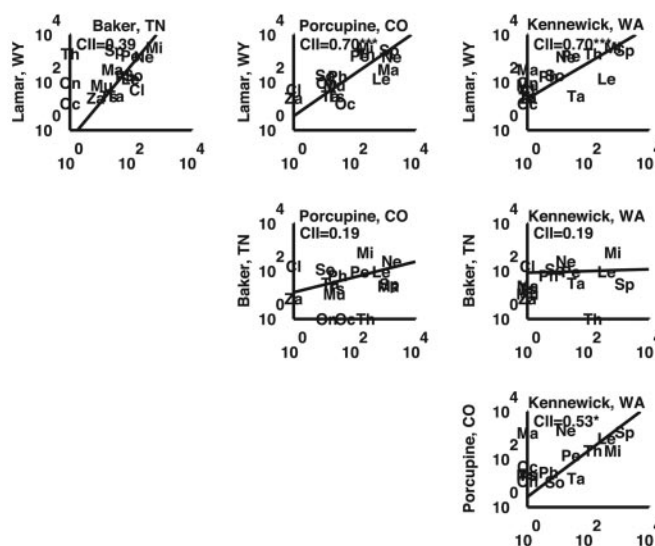


Fig. 1. Inertia in community structure. The similarity of community structure is plotted for all six pairwise combinations of the four sites. The axes are log₁₀-scale in units of number of identified bones. Each taxon is represented by its first two letters, except *Tamiascurius*, represented by Ts. The Pearson r is reported in the upper left corner, with * denoting $P < 0.05$ and *** denoting $P < 0.001$. The line is a ranged major axis regression (28). The abundance for a given site was summed across all levels. The distance in km and average separation in years between sites are summarized in Table 3.

mation on the PNAS web site. Again, some comparisons (Table 4) showed significant ($\alpha = 0.05$) differences from a null model of CII = 0 and some did not, but a binomial test suggested that the probability of the observed number of significant results by chance is vanishingly small ($P \ll 0.001$) for both types of comparison.

CII decreased with both temporal and spatial distance between communities. A Mantel test on matrices of distance and time between sites suggests that time and space are independent in the between-site comparisons ($r = -0.44$, $P = 0.79$, but the power is low). A visual examination of the locations of points in Fig. 2 *a–d* suggests that time and space are not strongly correlated with each other, but nor are they fully independent. An inspection of Fig. 2 shows that distance effects strongly dominated time effects. Therefore, to analyze the effect of time, within-site comparisons are much more accurate. Thus time and space both had statistically significant effects for species, genus, and family and are nonsignificant for order (although clearly trending significant for time) (Table 4).

An examination of Fig. 2 *e* and *f* suggests the possibility of some intriguing nonlinearities as well. For example, it appears (especially for species) that most of the decrease in CII happens exponentially in the first 700 km and slows to almost no decrease over the next 2,500+ km, but we did not have data points between 0 and 700 km to quantify this idea further. Similarly, the decrease with time appears largely linear, except that there is a rapid, steep drop-off in CII after 200,000 years. But this difference is caused by one data point, comparing the top and bottom layers at Kennewick, and may be an artifact; more data are needed in these time ranges.

CII also decays more slowly with increasing taxonomic level. An analysis of covariance showed that the slopes vary significantly by taxon (i.e., the interaction term for taxon level vs. the slope for CII over distance or time is significant, $P < 0.001$ for both comparison types). The graphs in Fig. 2 suggest that community similarity increases with taxonomic level, although using analysis of covariance with its assumption of independence

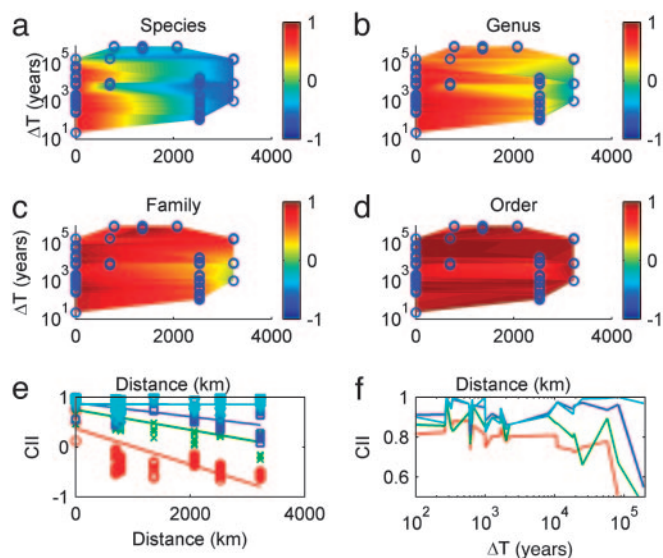


Fig. 2. How temporal and spatial scales affect community inertia. (a–d) Plots of the CII as a function of the time between communities (log scale) and the spatial distance between communities for four taxonomic scales. The CII ranges from -1 (complete reversal) through 0 (random reshuffling) to $+1$ (identical abundance proportions). The surface is a simple linear extrapolation between these points. (e) The between-sites comparison is used to look at how CII varies with distance. The four lines show different taxonomic levels (red = species, green = genus, blue = family, and cyan = order). (f) The within-sites, between-times comparison is used to look at how CII varies with time (lines represent different taxonomic levels as in e).

to calculate the magnitude and significance of the effects is inappropriate because of the issues related to aggregation of data mentioned in *Methods*.

We used random reshuffles of taxonomic associations as a null model to evaluate the effects of analysis at different taxonomic levels vs. a null model of aggregation in a randomized fashion (Table 1). These results suggest that at least for small mammals of the Quaternary the taxonomic level of analysis has no effect on analysis of community inertia within sites (time dimension

only), but has a significant effect on analysis of community inertia between sites. Between sites, analyzing at higher levels (genus-order) increases estimates of community inertia by ≈ 0.25 – 0.30 , which is highly significant statistically. Some caution is needed in the analysis of species-level data because the older sites contain fewer bones identified to the species level. Thus the significant effect of taxonomic level on rate of decrease of inertia with space and time (Fig. 2 and analysis of covariance presented earlier) is entirely caused by the nonbiological effect of aggregating data for comparisons across time within a site, but between sites it appears to be largely caused by some biological effect of taxonomic aggregation.

Having addressed the effects of temporal, spatial, and taxonomic scale on the question of inertia, we now turn to the use of neutral theory as a context for making comparative statements (there is/is not a lot of inertia relative to a drift model). A typical example of the simulated neutral community is shown in Fig. 3. Two facts about the drift of community structure in the neutral model are obvious. First, the neutral model loses rare species at a very high rate (see also Fig. 5, which is published as supporting information on the PNAS web site). Second, the relative abundances of species in a neutral community drift significantly, with moderately rare species becoming common and vice versa. For example after $t = 7,500$ years, the second and fourth most common species have gone extinct, whereas two moderately rare species are common (contrast Fig. 5 with Fig. 1).

Results averaged across 20 local communities and also across 20 Monte Carlo replicates are summarized in Fig. 4 and confirm that the rapid loss of species and high degree of drift in community structure is a general result of the neutral model. Thus, even for the highest migration rates ($m = 1.0$), half the species are lost from the average local community in ≈ 300 years and two-thirds of all species are lost in 1,000 years (Fig. 5), leaving only three to five of the more common species in the community. Thus a major finding is that whereas extinctions can occur over the time scales explored, the neutral model causes both rare and moderately abundant species to go extinct at extremely high rates relative to the empirical data (42) (in our empirical data, none of the small mammal species experience extinction). This result is related to the unusually high rate of speciation necessary to make the neutral theory fit empirical data (42). Real communities apparently possess an additional mechanism not found in neutral communities that allows rare species to persist over time and space.

Table 1. Effect of empirical aggregation vs. randomized aggregation

Data set	Between sites/times ($n = 105$)				Within sites ($n = 25$)			
	Species	Genus	Family	Order	Species	Genus	Family	Order
MeanAct	-0.120	0.474	0.697	0.856	0.766	0.836	0.918	0.936
MeanRnd	-0.114	0.231	0.444	0.549	0.722	0.828	0.878	0.930
MeanDelt	-0.006	0.243	0.254	0.308	0.044	0.008	0.040	0.007
SD	0.103	0.301	0.385	0.484	0.112	0.097	0.110	0.115
SE	0.020	0.058	0.074	0.093	0.044	0.038	0.043	0.045
<i>P</i>	0.5838	<0.0001	<0.0001	<0.0001	0.0619	0.6855	0.0855	0.7737

The effects of taxonomic aggregation are summarized. The data are for the between-sites comparison and the within-sites comparison. MeanAct gives the average CII for the comparison at the taxonomic level indicated. MeanRnd gives the average CII if the same taxonomic level is used but the 72 detailed taxa are randomly reshuffled with respect to which higher taxa they are part of. MeanDelt gives the “effect size” or true increase in CII caused by studying at a higher level after removing the increase caused purely by using aggregated data (i.e. MeanDelt = MeanAct – MeanRnd). SD and SE give the standard deviation and standard error at each level, respectively and *P* gives the *P* value for MeanDelt being significantly different from zero (paired *t* test). Repeated randomizations with different starting values for the random seed suggest that these results are robust, and in particular show that the three levels for the between-site values are highly significant (and with approximately similar effect sizes), but the species and family within-site values do not trend near significance (usually have much higher *P* values than in the run reported).

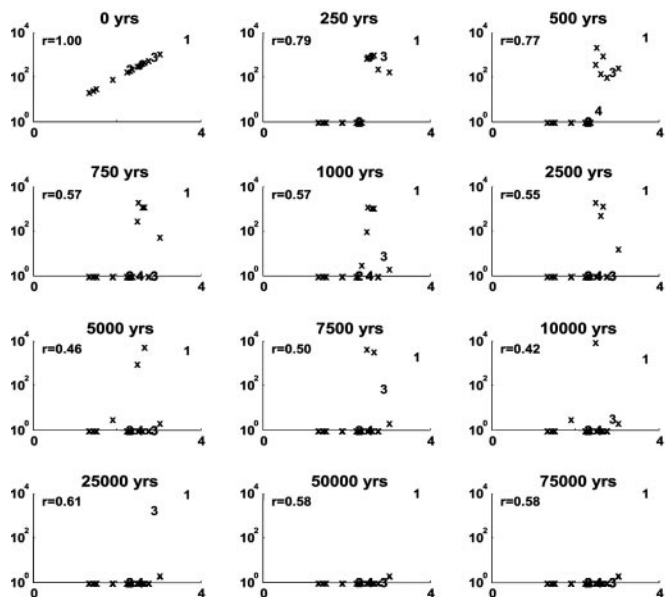


Fig. 3. Drift in community structure between a neutral community and itself over time. Shown is one Monte Carlo simulation comparing a community with itself for inertia or drift as the time between the points of comparison increases. Thus a single local community was selected at random and the community structure at various points in the future is compared with the initial community structure. Abundances in the initial community are plotted on the horizontal axis, and abundances at the specified amount of time later in the same community are plotted on the vertical axis. As a result, initially rare species appear on the left, and initially common species appear on the right. The four most common species in the metacommunity are numbered 1–4, and the rest are denoted by *x*. As a species becomes more common within the local community it moves up, and if it crosses the $y = x$ line it is more common than it was initially in the community. Similarly, a species moves down as it becomes rarer, hitting the x axis when it becomes locally extinct. CII (Pearson r) values are reported as in Fig. 1. Species are randomly assigned to a genus and analyzed at the genus level to compare with Fig. 1. Notice how most taxa, including all but the most common, go extinct. Model values are: $\theta = 6$, $M = 1$ ($m = 1/J$), $J = 1e4$. See note in *Supporting Text* on artificially high r values for later years.

We were unable to calibrate migration rate with distance for the neutral model. To do this calibration would require a precise estimate of how M varies with distance in small mammal communities, which is difficult to obtain. Thus we can only compare our within-site, between-time analysis with the neutral model. In the neutral model, the local community structure is compared against the same local community at a different point in time. The results are shown in Fig. 4, where CII (Pearson's r) is plotted vs. time. The empirical data show a much greater degree of community inertia than does the neutral drift model over extended periods of time (with $r \approx 0.9$ for time periods $< 10,000$ years and $r \approx 0.7–0.9$ out to $100,000$ years). For time periods of $< 1,000$ ybp (Lamar Cave only), there is considerable noise and the empirical data and neutral model overlap each other as the drift begins. For the time scale between $1,000$ and $3,000$ ybp, the empirical data fall in the same range as the neutral model for $M \gg 4$, but well above the range for $M \leq 4$. $M \gg 4$ would have to be considered to be a much higher migration rate than the rate of $M < 1$ most small mammals experience (43, 44), especially when the biases of the F_{ST} estimation method (45) are taken into account. For time periods $> 3,000$ years, the empirical data are all well above the neutral simulations, even for the highly unrealistic scenario of $M = J$ ($m = 1.0$, complete replacement of the community by migrants each year or total panmixia). When the time span exceeds $3,000$ years (allowing for

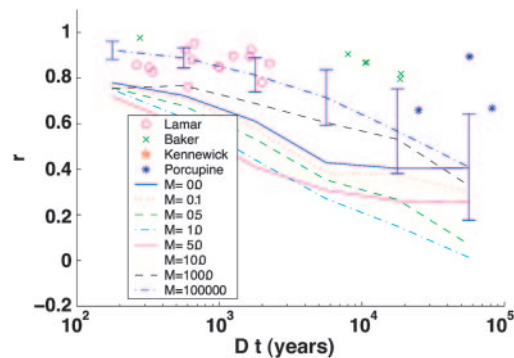


Fig. 4. Comparison of empirical data with the neutral model. CII is plotted vs. time. For empirical and neutral data the CII for all possible comparisons was calculated and plotted along the ordinate as the difference in time between the pair of communities. Symbols represent the within-site, between-time empirical comparisons. For neutral data only the line through the means is plotted with a 95% confidence interval based on ± 2 SE added to the $m = 1.0$ line (95% range of observed points gave very similar intervals). After $\approx 2,500$ years, the empirical data are always above the neutral model even for the case of $M = J$ or total panmixia. The Kennewick case for difference in time = $190,000$ years, $r = 0.46$, is left off the plot although it is above the $M = 1$ line for five Monte Carlo replicates extended out to $200,000$ years. Model parameters are $\theta = 6$, $M = \text{various}$, $J = 1e4$, $C = 20$.

initial transients in the neutral model), the empirical data always show more inertia than the neutral model (a binomial test gives $P \ll 0.001$; seven of eight empirical points are above the 95% confidence interval for the neutral $M = J$ or $m = 1.0$ line). These results are conservative because the empirical data also contains measurement error, decreasing the CII (Pearson r), whereas the neutral does not. The neutral model further shows (Fig. 4 vs. Fig. 6, which is published as supporting information on the PNAS web site) that decay occurs faster across space than time (at least for the scales studied here), just as was found in the empirical data.

A good confirmation of the accuracy of our neutral-model simulations was given by the fact that, as expected from analytical results (38), the neutral model displays three distinct groups depending on $M \gg 4$, $M \approx 4$, or $M \ll 4$ (data not shown but the three groups decay to CII's of about $r = 0.8$, 0.6 , and 0.4 , respectively over the first $10,000$ years, see also comparison between sites in Fig. 6).

Finally, we address our third challenge: that of choice of CII. We reanalyzed the within-site, between-time comparisons for both empirical and neutral data and the between sites for empirical data (Figs. 2 and 4 and Table 4 compared with Tables 5 and 6 and Fig. 7, which are published as supporting information on the PNAS web site) by using the Jaccard, Whittaker, squared chord distance, and our own percentage swap measures of similarity or distance as CII. Only minor changes were found (see details in *Supporting Text*). This finding demonstrates that our results are independent of the choice of CII. We suggest that this robustness to the exact measure of CII should be demonstrated before making claims for or against community inertia.

Discussion

All our results are concordant and support a nonrandom view of communities through space and time. It appears that between sites at the level of species change occurs that is so extensive that it is nonrandom (CII < 0), although some caution is needed for interpretation of our results at the species level. This one exception is probably caused by (i) the historical contingencies discussed earlier and (ii) the nonrandom, systematic replacement of species by congeners between sites (and to a much lesser degree replacements of genera by confamilial genera). In

