Statistical testing for paleocommunity recurrence: Are similar fossil assemblages ever the same?

J.B. Bennington a, R.K. Bambach b

a Department of Geology, 114 Hofstra University, Hempstead, NY 11550, USA
b Department of Geological Sciences, Virginia Tech, Blacksburg, VA 24060, USA

Received 15 July 1995; accepted 25 August 1995

Abstract

Observations of the recurrence of similar fossil assemblages through long intervals of geologic time punctuated by rapid changes in both the composition and structure of fossil assemblages has recently resulted in the concept of "coordinated stasis" to describe this pattern in the behavior of paleocommunities through time. Coordinated stasis may imply the existence of ecological mechanisms that actively maintain particular community structures and thus have a significant impact on the process of macroevolution. However, before such mechanisms can be invoked, it must be shown that ecological stasis as observed in the fossil record is more significant than the persistence of similar community types due to the repeated reinvansion of recurring habitats from a persistent species pool (the "null hypothesis" for paleocommunity recurrence). Analysis of the relationship between communities and paleocommunities and an expansion of the ecological hierarchy at the community/paleocommunity level allows the creation of a rigorous definition of the entities composing the local paleocommunity (the samples collected within a stratigraphic horizon at a single locality), paleocommunity (groups of samples shown to be statistically identical), and paleocommunity type (groups of samples that are similar but can be shown to be statistically different). These definitions permit the development of the null model for paleocommunity recurrence and establish a base-level of variability within a local paleocommunity that permits rigorous statistical comparisons to be made between paleoecological samples at larger temporal and geographic scales. Paleoecological samples from four marine tongues in the Middle Pennsylvanian Breathitt Formation are used to analyze species abundance variability at several spatial and temporal scales and to test for paleocommunity recurrence. Although recurrence of statistically identical local paleocommunities occurs at single localities and between localities within individual marine units, it is usually not detectable between marine tongues, suggesting that paleocommunity recurrence in the marine strata of the Breathitt Formation is the recurrence of paleocommunity types and does not falsify the null hypothesis.

Keywords: paleoecology; Breathitt Formation; coordinated stasis; analysis of variance; Pennsylvanian

1. Introduction

Entities identified as paleocommunities are normally recognized by the recurrence of species, usually with a characteristic pattern of relative abundances. The consistent basin-wide recurrence of such species patterns in the Appalachian Basin over intervals of 2–8 m.y. during the Silurian and Devonian, coupled with basin-wide breakup and reorganization of such patterns over shorter intervals, led Brett and Baird (Brett et al., 1990; Brett and Baird, 1992, 1995) to propose the idea
of coordinated stasis, which they suggest may be characteristic of the entire history of marine life. Brett and Baird claim that the same communities, with the same species membership and general relative abundances, recur in appropriate environments during each interval of coordinated stasis and that the majority of change in species and in patterns of community occurrence occurs during the short intervals between times of stasis.

It should be emphasized that the units of coordinated stasis noted by Brett and Baird, who regard them as subunits of Boucot’s Ecologic-Evolutionary Units (Boucot, 1983), are not functional ecological entities (as is also the case for the full-scale EEU’s). They are intervals of time during which a majority of the species in a region (such as the Appalachian Basin) persist with little change and during which the whole spectrum of paleocommunity types also persist, each paleocommunity type recurring when the appropriate environmental conditions are present.

In several important papers William Miller III (Miller, 1986, 1990, 1993a,b) has clarified different temporal scales and processes of community change, ranging from short term events, such as seasonal cycles and succession, up to long term patterns of community replacement mediated by environmental change. Niles Eldredge (Eldredge, 1989, p. 190) has observed that the phenomenon now termed coordinated stasis by Brett and Baird is a pattern of community replacement, but on a larger scale than that discussed by Miller. The matter of coordinated stasis focuses on an issue described as a macroevolutionary phenomenon by Eldredge (1989, pp. 190–191):

“The point of the Hamilton [mid-Devonian of the Appalachian Basin] situation is that while there are a large number of distinct community types, most of those communities remain recognizably the "same" throughout Hamilton time. The species found in the Hamilton seas in general persist, and for the most part display little change, throughout Hamilton time. … The critical issue then becomes: Do the community (ecosystem) types persist — and does the entire Hamilton economic system persist — because the species persist? Or do the species persist because the economic system persists? Or is it, somehow, both simultaneously?”

Here we have a major issue in macroevolution hanging on the observation that “most of the communities remain recognizably the same”. Eldredge puts quotation marks around “same”, suggesting that there is uncertainty about the precision implied by the word as it is used. In this paper we address the distinction of similarity versus being the same, both at the conceptual level of community definition and in testing whether the same communities regularly recur. We propose a subdivision of the general “community” level of the ecological hierarchy which permits more precise definition of paleocommunity entities, argue that the pattern of species occurrence and abundance in an environment produced by stochastic recruitment from the available species pool should be the null model against which proposed examples of stasis need to be tested, and provide a step-by-step analysis of a particularly good test case for examining the potential for paleocommunity recurrence.

1.1. Definitional issues

If we are to test the idea of coordinated stasis we must have clear definitions for the ecological units involved, yet agreement on what is meant by terms such as community and paleocommunity has never been achieved. One reason for this is that both ecological interactions and species occurrence patterns take place at a variety of scales. One worker may focus on a local patch, another on a widespread environment. Each may consider that the fauna and flora in the geographic region studied composes a community, yet they are not dealing with comparable entities. However, by using the concept of an ecologic hierarchy as advocated originally by Valentine (1968, 1973) and more recently by Miller (1986, 1990, 1991, 1993a,b) and Eldredge and his colleagues (Eldredge and Salthe, 1984; Eldredge, 1985, 1989; Eldredge and Grene, 1992), it is possible to formulate several community related concepts to which many of the varied ideas scattered through the literature can be related in a consistent fashion.
1.2. Community and paleocommunity comparability

Before going farther, however, we must make two distinctions of profound importance between neontological and paleontological samples. One concerns information loss and its implications, the other involves the effect of time-averaging. These taphonomic issues raise questions about the comparability between ecological and paleoecological units above the level of the individual.

Although there is a direct parallel between many neontological studies (on lizard communities, bird communities, soil arthropod communities, coral communities and the like) and paleontological studies (which are necessarily restricted to the preserved fauna and/or flora) that makes paleoecologic studies as legitimate as most neoeologic studies, information loss from decay and incomplete preservation in paleontological samples creates insurmountable differences between the interpretive potential of the two disciplines. Because of information losses, knowledge of the actual range of interactions and the documentation of their quantitative distributions for any organism, even if well preserved, is not possible in paleoecology. Analysis of community structure in paleoecology will never be comparable to that in neoeecology.

There are many ideas about the nature of communities in neontology, ranging from the Gleasonian view of chance aggregations of individuals living in an environment to an extreme Eltonian view of the community as a superorganism integrated by complex, interdependent, biotic interactions. However, all neontological concepts of community include the idea of an aggregate of local species populations (avatars) among which interactions can and, even in chance aggregations, will occur. Some aspect of contemporaneity of constituents underlies all formulations of the community concept. Paleontological samples from marine benthic habitats never fulfill this criterion. Time-averaging of marine benthic fossil assemblages makes them forever non-comparable to life assemblages. The recently dead and the recently reworked and exhumed shells of any age get included in fossil assemblages as they form, even in those preserved by sedimentation events equivalent to “Pompeian” style ashfalls. This means we can never determine the actual members of any avatar in a fossil representation of a community. Although paleontologists get the advantage of time-averaging smoothing some of the variation in occurrence and abundance of taxa over time at any one place, we don’t see the complete community, because of preservation failure, and, because of time-averaging, we cannot sort out which of the individuals in a fossil assemblage were, in fact, contemporaneous. Due to these factors, paleocommunities are never exactly comparable to neontological communities.

1.3. Expanding the ecological hierarchy

Table 1 presents a comparison of several major steps in the development of subdivisions of the ecological hierarchy. We recognize five major levels in the ecological hierarchy. These are (a) the structural level, for features included within individual organisms, (b) the individual level, (c) the taxonomic level, of which the avatar (the population of a species that exists and functions in a community) is the only point at which an entity (the species) in the taxonomic hierarchy is represented in the ecological hierarchy, (d) the ecosystem level, the level that includes community concepts, and (e) the biogeographic level, the most inclusive level. We are preparing a more detailed description of these levels and their subdivision, as well as those of the evolutionary hierarchy, for a comparison between neontological and paleontological categories to be published elsewhere. In this paper we will concentrate only on the ecosystem level.

Valentine (1968) recognized the major levels of the ecological hierarchy, from the structural level within the level of the individual (characters) up to the biogeographic level (provinces and the biosphere), but in that pioneering exercise he did not seek to explicate the sublevels that could be defined within some of the levels. Eldredge and Saltz (1984) further reviewed the formal properties of hierarchies and defined both a genealogical (evolutionary) and an ecological hierarchy, noting that “no single nested hierarchy can include both”
Table 1
Subdividing the ecological hierarchy

<table>
<thead>
<tr>
<th>Entity group</th>
<th>Valentine, 1968</th>
<th>Eldredge and Salthe, 1984</th>
<th>Eldredge, 1989</th>
<th>This paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neontology</td>
<td>Paleontology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biogeographic</td>
<td>total planetary biosphere province</td>
<td>entire biosphere biotic region</td>
<td>biosphere m</td>
<td>biosphere realm</td>
</tr>
<tr>
<td>Ecosystem</td>
<td>community</td>
<td>local ecosystem</td>
<td>ecosystems</td>
<td>community type</td>
</tr>
<tr>
<td>Taxonomic</td>
<td>population</td>
<td>organism</td>
<td>avatar</td>
<td>organism</td>
</tr>
<tr>
<td>Structural</td>
<td>phenotype</td>
<td>character</td>
<td>organism</td>
<td>organism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cells</td>
<td>organs</td>
<td>organs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>enzymes</td>
<td>tissues</td>
<td>tissues</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cells</td>
<td>cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>proteins</td>
<td>proteins</td>
</tr>
</tbody>
</table>

Genealogic and ecological entities (Eldredge and Salthe, 1984, pp. 186–187). They identified and defined a series of levels in the ecological hierarchy similar to those specified by Valentine. Eldredge (1985, Eldredge, 1989) has written more about both the genealogical and ecological hierarchies, expanding on the lower levels in each and perceptively discussing the general nature of the higher levels, but he has not subdivided the higher levels in the ecological hierarchy in detail, primarily because his focus has been on evolution and on the individual and taxonomic levels. Damuth (1985) coined the term avatar for the group of individuals from a single species that actually participate in the “economy” of a community or local ecosystem (“the ’embodiment’ or ”representation“ of the species in the local community.” Damuth, 1985, p. 1137). Avatar is more ecologically precise than the term population, which also can carry the connotation of a deme, the group of individuals in a species that forms a local reproducing population. Eldredge (1985) discussed the use of the term avatar, rather than population, in the ecological hierarchy and has used it explicitly in later work. Eldredge (1985, 1989) also distinguished between community, the living association of avatars, and ecosystem, the organisms plus their environment. In later studies on economic systems he has preferred the term ecosystem, but in this paper we are referring to organisms and the remains of organisms in fossil assemblages, so we will use the term community rather than ecosystem.

In his 1989 book, Eldredge refers to ecosystems (plural). Because of the multiplicity of concepts incorporated under the umbrella of ecosystem (or, for organisms only, community) this is reasonable. We propose here to subdivide the ecosystem level of the ecological hierarchy in a way that, we hope, will lead toward a more consistent method of using terms for the varied entities at this general level in the ecological hierarchy. Because of the inescapable differences noted above between fossil assemblages and associations of living species, we will also propose roughly parallel, but different, terms for the paleoecological entities associated with the subdivisions of the ecological hierarchy at the ecosystem level.

Because the potential for interaction is inherent in neontological community concepts, contemporaneity and spatial proximity are both criteria for defining the smallest subdivision of the ecosystem level in the neoecological hierarchy. We suggest local community, a term parallel to local ecosystem as used by Miller (1986, 1990), for this fundamental group entity. A local community is the set of
avatars that do, in fact, exist together and can interact directly in space and time. Everyone can agree that this fulfills the basic community concept.

Because of time-averaging, however, it is not possible to identify exactly contemporaneous individuals, or groups of individuals, in a marine fossil assemblage. As the recent studies by Karl Flessa and his colleagues (Flessa et al., 1993; Flessa, 1993; Flessa and Kowalewski, 1994), as well as other work by Powell and Davies (1990) and Muray-Wallace and Belperio (1994), indicate, the shells being incorporated into single shell beds today have median ages of hundreds of years, with some ranging up to 3000 to over 9000 yr in age. Therefore, even a single fossil assemblage collected from a single horizon at one spot is not temporally the same as any neontological collection. Avatars are simply not identifiable in fossil collections, nor is it possible to observe the area over which those interacting populations would have ranged at any one time. Because we cannot see the full range over which the organisms that lived in a local community in the past interacted, the only way to recognize a paleoecological unit that might be roughly comparable to a local community is to restrict the suite of fossil remains included in a local paleocommunity to those that occur at a spatial scale over which we can be sure that they would have been derived from the avatars that did live in a local setting during the time interval represented by a fossiliferous horizon. A local paleocommunity, then, is defined as the assemblage collectable from a single bed at one outcrop, assuming that sedimentological and taphonomic interpretation indicate that the fossil deposit is generally untransported.

Because patches shift with time, time-averaging in a local paleocommunity will have incorporated representatives of many of the species that lived in the local habitat over time (fide Warme et al., 1976; Staff and Powell, 1988). Even though only a narrow geographic range can be “safely” included in a local paleocommunity, the species assemblage of the local paleocommunity will more closely approximate the total species assemblage that was present in the entire living local community than the sampling of live organisms at a single location would produce. By extending the concept of local paleocommunity to the scale of the outcrop, rather than restricting it to a single sample spot, it becomes possible to collect replicate samples laterally along a bed, or to sample at scattered points on a bedding surface, and so accumulate data to evaluate the variability of species occurrence in the local paleocommunity, something that is necessary if we are to compare local paleocommunities from different localities or different strata quantitatively.

How, then, does a community differ from the idea of a local community? This is where statistical analysis enters the picture; where we need to move beyond the methods of exploratory descriptive ecology into the area of rigorously reproducible results. While it is tempting to think of a community as a set of interacting, or potentially interacting, organisms, that is only spatially practical on the relatively small scale of the local community. But, if one finds the same species in the same abundance pattern at several places, then the form and structure of a local community has recurred, and it seems only reasonable that both places should be regarded as having the same ecological entity present. This should be true whether those places are connected by similarly inhabited territory or not, and regardless of whether the two places are inhabited by the same association at the same time or at different times. If, after accounting for variability within the scale of the local community, the species and their abundances in two local communities are not statistically significantly different from each other, then the two local communities belong to the same community. In a hierarchical ecological scheme a community is the aggregate of the local communities arrayed in space and time that are sufficiently similar to each other that they can not be shown to be significantly different statistically. Thus, the community concept becomes one of recurrence beyond the scale of the local community. A local community is a real community, but if it is a unique entity there is little to generalize about, whereas recurrence opens the way to further ecological analysis.

A paleocommunity, then, is defined as the aggregate of local paleocommunities that are not statistically significantly different from each other.
Paleocommunities are based on local paleocommunities, which, although at an equivalent position in the paleoecological hierarchy to local communities in the neoeocological hierarchy, are not spatially or temporally comparable entities, as noted above. In like manner, paleocommunities are not exactly comparable to communities, although paleocommunities are at the equivalent level in the paleoecological hierarchy as communities in the neoeocological hierarchy.

The case of similar, but statistically significantly different, communities and local communities remains. There can be a number of reasons why local communities might be similar but not be demonstrably members of the same community. We propose the more general concept of community type to include such apparently allied assemblages together. The community type is the aggregate of local communities and communities that have similar, but not identical, taxonomic membership and occur in similar, but not necessarily the same, environments. This is a level in the ecological hierarchy similar to the one that Jackson et al. (in press) calls the metacommunity. Similarly, the grouping of paleocommunities and local paleocommunities that occur in stratigraphically similar settings (and may be identified using analytical techniques such as cluster analysis) are termed paleocommunity types.

As Jackson et al. (in press) also conclude, the level of community type (metacommunity) is undoubtedly involved in coordinated stasis. Our question is whether coordinated stasis is the general case at the community level (aggregates of samples with statistically similar species abundances), as was initially claimed by Brett and Baird in formulating the concept. General descriptive observation suggests that the same community types do persist during an interval of coordinated stasis — but do the same communities persist?

1.4. The null model of community type recurrence

The ubiquitous feature associated with the recognition of recurrent ecologic entities on all time scales is the association of environment and adaptive type. For example, similarity of ecospace utilization persists through each of Sepkoski’s three evolutionary faunas (Bambach, 1983, 1985), Boucot claims that each of his Ecologic-Evolutionary Units is characterized by structurally similar communities (Boucot, 1983, 1990a,b), and Brett and Baird (Brett et al., 1990; Brett and Baird, 1995) point out that there are subunits of different stable paleocommunity patterns within Boucot’s units, which are the intervals of coordinated stasis they recognize.

Direct control by environmental conditions is obviously a major factor in determining which taxa can be successful in particular habitats. Rocky shores differ from sandy beaches — and different taxa dominate in each setting (barnacles and byssate mussels on rocks, rapid burrowing clams and mobile arthropods in sand). This obvious point is a basic determinant, however, in the recurrence of similar adaptive suites when habitats recur.

In the Recent, parallel communities are consistently associated with similar physical environments even if located in different biogeographic provinces. This is nicely illustrated in Thorson’s classic description of parallel communities from around the world (Thorson, 1957). He lists several suites of parallel communities, the communities in each suite characterized by different species from the same genus, and he also notes several other parallel community patterns in the same general environmental settings at different locations. The “homeomorphic” paleocommunities noted by Wallace (1978) in the Devonian of France and Iowa record a similar situation in the fossil record.

Similarly, congruent community structures developed by different faunas at vastly different times share similarity of environment as a selective force but do not share similarity of genetic heritage. This is seen in the suite of communities from Middle Ordovician and Lower Devonian carbonate rocks described by Walker and Laporte (1970). A series of paleocommunity types in each stratigraphic interval crossing the environmental spectrum from intertidal to protected subtidal show similar adaptive types and diversities, even though there are virtually no common taxa below the class level in the collections from the two intervals. The similarity of feeding depths of deposit feeders in Silurian and Recent settings
described by Levinton and Bambach (1975) is another case of congruent community structure persisting under appropriate environmental conditions.

The communities distributed along ecotones in space and time are clearly related to continuous shifts in conditions along environmental gradients. In one case, Warne (1969, 1971), found correspondence of community type with major environmental setting in Mugu Lagoon, California, even using presence-absence data only. Death assemblages from sand channel settings form a separate cluster from the pond, marsh and eelgrass samples in a multivariate cluster analysis. But by using the relative abundances in the analysis, it is possible to extract even more detailed correspondence of these samples with their environmental setting. Quantitatively based clusters of samples are found to map along the salinity, topographic and vegetation gradients of the lagoon (Fig. 1). In a like manner, Levinton and Bambach (1975) demonstrated that in a shoaling sequence in the Silurian one suite of taxa waned and another increased. The series of paleocommunity types in that sequence matched the change in environment with time.

Why should this similarity of communities in particular environments be so consistent in time and space? Possibly because selection does its job on the functional corner of Seilacher’s triangle (Seilacher, 1970; Raup, 1972), despite the constraints of the structural and historical-phylogenetic corners. Early in the history of any bauplan, selection probably drives adaptation to the highest attainable efficacy given the limitations of heritage, biomechanics, biochemistry, and environmental variability. The appearance of aerodynamic flight feathers, asymmetrically placed shaft and all, in the earliest known bird, *Archaeopteryx*, is an example. Once well adapted organisms evolve to exploit any set of environmental parameters they should simply persist in doing so whenever they have access to the appropriate setting — as has been the case for deposit-feeding bivalves and worms for 420 m.y. These factors alone may be responsible for the apparent stability of communities because environments do not have any inherent tendency for secular change on the “steady-state”

Fig. 1. Sketch map of the eastern portion of Mugu Lagoon, California, as it was in the late 1960s. Numbers and shaded areas illustrate the gradient array of Q-mode clusters, based on abundance data, for samples collected by John Warne (Warne, 1971, table 1). The gradient extends from the inlet through the barrier beach eastward to the isolated end of the long tidal channel and landward (northward) from the tidal channel onto the tidal flats. Five subdivisions of the gradient are apparent from the distribution of samples in various clusters. Cluster 1 (dense stipple area) from sand channel near inlet, Cluster 2 (light stipple area) in belt just shoreward and up-channel from Cluster 1, Cluster 3 (upper left to lower right cross-hatched area) in first large pond in long tidal channel, Cluster 4 (unshaded) in tidal channels and ponds landward of first three clusters and in more remote eastern ponds of long tidal channel, and Cluster 5 (upper right–lower left cross-hatched area) in isolated area on landward side of far eastern end of long tidal channel.
earth. Climates have fluctuated, but not ranged very far from “normal” for long at any time throughout the Phanerozoic, and sedimentation has been about the same, too (Ronov et al., 1980; Tardy et al., 1989).

Community structures, however, are not fully self-sustaining. When local or regional environmental conditions change, communities commonly break up, rather than migrate as a unit. This has been demonstrated for both marine and terrestrial communities in relation to the major environmental perturbations of the Pleistocene and Holocene (Davis, 1986; Graham, 1986; Addicott, 1966; Thomsen and Vorren, 1986). Yet, during these changes many species have shifted geographic range to remain in the general conditions to which they are adapted rather than either go extinct or evolve different adaptive responses (Pielou, 1991). This has meant that general patterns of temperature change have been reliably tracked through change and recurrence in vegetation (Comboutrie-Nebout, 1993), beetle (Coope, 1987) or molluscan (Peacock, 1989) assemblages.

Although biotic interactions are present in all communities (even in new associations in which the colonizers are only chance migrants) and these interactions can contribute to community structure and pattern, the species that become dominant in any environment will obviously be the ones that have access to and are best suited for life in that environment. With recurrence of particular environmental conditions, the expected success of the well adapted species present in the available species pool alone may be responsible for the ubiquitous correlation of community type with environmental setting and the general similarity of community type whenever environments recur (for example, explaining the recurrence of Late Paleozoic plant communities — see DiMichele and Phillips, 1996). By available species pool we simply mean the suite of species with populations living in the geographic region. These would be the species that would have the most ready access to any local habitat and would be the species available to colonize any new locations as environmental conditions change. If the apparent stability of paleocommunities in time is to be shown to be anything more than environmental selection from the available pool of species, the application of analytic techniques that can distinguish the additional control on community structure caused by biotic interaction will be required.

The null model is that similar community types should recur whenever similar environmental conditions recur while the same general species pool is available from which to recruit organisms to populate a habitat. To demonstrate that the intervals of apparent coordinated stasis have unusual community stability it will be necessary to demonstrate statistically that the same paleocommunities, not just similar paleocommunities of the same type, recur. The latter is what we would expect from the null model of drawing adapted species from a continuing species pool whenever the same general environmental conditions recur.

1.5. A taphonomic null hypothesis

The null hypothesis presented above is not the only null model that might be applied to explaining the recurrence or lack thereof of paleocommunities or paleocommunity types. It may be that taphonomic processes are primarily responsible for generating the patterns of species occurrences and abundances seen in fossil assemblages. Miller (1993a,b) refers to this scenario as “the Devil’s own model, pointing out that taphonomic processes associated with recurring sedimentary facies might create the appearance of coordinated stasis by only allowing the preservation of a uniform subset of the skeletonized taxa in spite of geographic or temporal differences that might have existed in the compositions of the living communities.

That taphonomic processes could have generated similar fossil assemblages from dissimilar living communities seems unlikely. Taphonomic processes are deterministic, so that the same processes operating in two different places or at two different times on the same assemblage of shells would be expected to produce similar alterations in preservation of species and their abundances. Thus, we expect that similar living communities should have led to similar fossil assemblages in the face of the same taphonomic forces. Except in cases where all but a few exceptionally preserveable
species were eliminated, we also expect that different living communities should have produced different fossil assemblages. Taphonomy is generally considered to be a disordered process that would work against recognizing the existence of coordinated stasis. Far from worrying about artifactual similarity, most paleontologists are concerned that different taphonomic processes operated in different areas or at different times to create dissimilarity between fossil assemblages generated from originally similar living communities. The pertinent question then is not whether coordinated stasis is an artifact of taphonomy, but whether taphonomy is masking coordinated stasis at some level of the ecological hierarchy. In other words, if two local paleocommunities are statistically dissimilar from one another, is this because their original living communities were dissimilar, or because taphonomy has rendered them dissimilar? In this case, taphonomic alteration of the fossil record does create a “taphonomic” null hypothesis that is a viable alternative to the “biological” null hypothesis presented above. In the absence of information on the taphonomic processes affecting the fossil assemblages being compared it is impossible to distinguish between the two alternative null hypotheses; either could be reasonably inferred from the observation of statistical differences between local paleocommunities. However, if it can be shown that differential taphonomic process were not a factor in generating differences between fossil assemblages, or if differences due to taphonomic effects can be accounted for in the statistical tests, then the “biological” null hypothesis becomes the most reasonable working null hypothesis.

1.6. A setting to test for paleocommunity recurrence

An ideal setting to test whether we can find recurrence of paleocommunities is provided by the Middle Pennsylvanian transgressive marine shales and associated strata found in eastern Kentucky in the Appalachian basin (Fig. 2). These shales form tongues of marine strata (called “marine units”) that record periodic flooding of a subsiding foreland basin that hosted non-marine, fluvial and coastal swamp environments between the episodes of marine incursion (Chesnut, 1989). We have measured and sampled four of the most extensive marine units in the Breathitt Formation (in ascending stratigraphic sequence the Elkins Fork, Kendrick, Magoffin, and Stoney Fork — see
Fig. 2) over a large region of eastern Kentucky and have found that each records a sequence of marine environments with distinctive marine faunas that developed through transgression and subsequent regression (Bennington, 1991a, b, Bennington, in press; Bennington, in prep.; Bennington and Bambach, 1994). The major marine units we examined were separated in time by episodes of deltaic and coastal plain progradation that completely disrupted marine conditions throughout the central Appalachian basin on a temporal scale of from 400 ka to 2.5 Ma (Chesnut, 1989).

Paleoecological sampling and analysis (Bennington, in prep.) has shown that, although different marine units represent different magnitudes of marine transgression and not all identified marine environments were developed in every marine unit, the marine units do contain several recurring facies with very similar fossil assemblages. All of the fossiliferous facies sampled for quantitative analyses represent low-energy, mud-dominated environments where no significant transport of shelly material would have occurred (Kidwell and Bosence, 1991). Preservation is excellent, with both brachiopods and mollusks (encompassing the majority of species) represented as original hard parts in all samples. Although it cannot be stated that these fossil assemblages were unaffected by taphonomic processes beyond decay of soft part anatomy, it can be argued that these processes were similar in all of the environments sampled, and probably very similar within facies from which samples were selected for direct statistical comparison.

Using clustering and ordination methods we have recognized five major assemblage types that recur among the four different marine units (the Productid, Spiriferid, Chonetid–Productid, Chonetid–Mollusk, and Small Mollusk clusters, Fig. 3). The cluster analysis in Fig. 3 was done using species abundance (not just presence–absence) data. The same genera and, in most cases, the same species are found in very similar abundance associations in all four marine units. Each assemblage type is characterized by a dominant suite of taxa (Table 2), and each recurs in a characteristic general stratigraphic and lithological setting, indicating that we are dealing with a pattern of recurrence that is comparable in scale and style to the recurrence of fossil assemblages during the intervals of coordinated stasis observed by Brett et al. (1990) and Brett and Baird (1995). The assemblages within each cluster are similar from locality to locality and from one horizon to another (hence the clustering), but how similar are they? Do they recur as distinct paleocommunities with statistically identical composition and abundance, or is the recurrence one of paleocommunity type, as would be expected under the null hypothesis of environmental selection from a persistent species pool? This is the question we attempt to answer in this paper.

2. Methods (distribution tests)

2.1. Testing for recurrence

2.1.1. Direct comparisons of species abundance distributions

To test for recurrence of paleocommunities we first need to find a reasonable method for making statistical comparisons of paleocommunity samples. The data most commonly used by paleoecologists to quantitatively describe a paleocommunity are the kinds of species present and their abundances (the species abundance distribution) in a bulk sample taken from a particular locality (the local paleocommunity). Each bulk sample has a particular species abundance distribution, so that to compare local paleocommunities requires that species abundance distributions be compared and tested for statistically significant differences. If no significant differences can be shown, then the samples can be considered statistically identical, and this identity is the basis for judging them to have been drawn from the same paleocommunity. However, before such a test for identity can be made, we must first ask how different species abundances can be expected to be, even if sampled from the same paleocommunity.

2.2. Sources of variability

2.2.1. Sampling error

It would be naive to expect species abundance distributions between any two samples to be
Fig. 3. Q-mode cluster analysis dendrogram showing similarity of quantitative paleoecological samples from the four major marine units of the Breathitt Formation. Sample groups with similar fossil assemblages identified by the analysis are labeled according to the dominant fauna. Samples are identified by marine unit as E = Elkins Fork, K = Kendrick, M = Magoffin, S = Stoney Fork.

exactly the same. Even if samples are drawn from the same underlying species abundance distribution, it is expected that they will be variable to some degree due to sampling error. Because of this, statistical tests that compare distributions must ignore differences attributable to sampling error when determining if two samples could have been drawn from the same underlying distribution.

2.2.2. Patchiness

A second potential source of variability within paleocommunities results from the fact that organisms are not evenly distributed throughout a community. Local variations in habitat and stochastic variations in the distributions of organisms appear to be present at many spatial scales, resulting in a mosaic of varied concentrations of organisms within a particular community called patchiness (Springer and Miller, 1990). Two samples taken from the same community are likely to encompass different patches and thus have different species abundance distributions, adding another source of expected variability to the sample comparison. Patchiness, however, may be less significant in paleocommunities because they are time-averaged aggregates of the living community (Kidwell and Bosence, 1991). It is possible that most paleocommunities are sufficiently time averaged that patch differences in the species abundance distributions have been “smeared out” over time, resulting in uniformly similar species distributions throughout the range of the paleocommunity.

Patchiness in the preserved fossil assemblage could also result from taphonomic disruption of
Table 2
Percent abundance of important taxa in Breathitt sample clusters

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Cluster</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Productid</td>
<td>Spiriferid</td>
<td>Chonetid–Productid</td>
<td>Chonetid–Mollusk</td>
<td>Small Mollusk</td>
</tr>
<tr>
<td>Desmosinius</td>
<td>50.75</td>
<td>4.41</td>
<td>18.95</td>
<td>16.95</td>
<td>0.36</td>
</tr>
<tr>
<td>Linoproduc tus</td>
<td>13.08</td>
<td>1.59</td>
<td>0.80</td>
<td>0.59</td>
<td>0.02</td>
</tr>
<tr>
<td>Juresania</td>
<td>5.30</td>
<td>0.48</td>
<td>0.52</td>
<td>1.04</td>
<td>–</td>
</tr>
<tr>
<td>Antiquitonia</td>
<td>2.08</td>
<td>9.86</td>
<td>3.94</td>
<td>0.52</td>
<td>–</td>
</tr>
<tr>
<td>Chonetids</td>
<td>7.13</td>
<td>3.72</td>
<td>13.80</td>
<td>34.42</td>
<td>5.75</td>
</tr>
<tr>
<td>Plicocconiotes</td>
<td>–</td>
<td>0.14</td>
<td>10.03</td>
<td>–</td>
<td>0.02</td>
</tr>
<tr>
<td>Derbya</td>
<td>3.49</td>
<td>8.00</td>
<td>6.80</td>
<td>0.97</td>
<td>0.34</td>
</tr>
<tr>
<td>Anthracospirifer</td>
<td>1.90</td>
<td>35.86</td>
<td>7.00</td>
<td>1.12</td>
<td>0.02</td>
</tr>
<tr>
<td>Punctospirifer</td>
<td>1.43</td>
<td>1.59</td>
<td>5.58</td>
<td>1.41</td>
<td>0.14</td>
</tr>
<tr>
<td>Crurithyris</td>
<td>0.39</td>
<td>–</td>
<td>0.47</td>
<td>2.75</td>
<td>2.01</td>
</tr>
<tr>
<td>Hustedla</td>
<td>0.42</td>
<td>0.07</td>
<td>4.08</td>
<td>0.97</td>
<td>0.05</td>
</tr>
<tr>
<td>Composita</td>
<td>1.32</td>
<td>15.17</td>
<td>4.79</td>
<td>1.12</td>
<td>0.05</td>
</tr>
<tr>
<td>Orbiculoidea</td>
<td>0.65</td>
<td>0.21</td>
<td>0.19</td>
<td>3.35</td>
<td>0.67</td>
</tr>
<tr>
<td>Septimyalina</td>
<td>0.49</td>
<td>5.24</td>
<td>0.39</td>
<td>0.15</td>
<td>–</td>
</tr>
<tr>
<td>Arculopecten</td>
<td>1.18</td>
<td>2.28</td>
<td>0.06</td>
<td>1.12</td>
<td>0.19</td>
</tr>
<tr>
<td>Parallelodon</td>
<td>1.55</td>
<td>2.34</td>
<td>4.82</td>
<td>0.97</td>
<td>0.02</td>
</tr>
<tr>
<td>Astarrella</td>
<td>1.11</td>
<td>2.41</td>
<td>3.36</td>
<td>3.42</td>
<td>0.12</td>
</tr>
<tr>
<td>Nuculoidea</td>
<td>0.09</td>
<td>0.21</td>
<td>0.58</td>
<td>0.89</td>
<td>23.81</td>
</tr>
<tr>
<td>Nuculopsis</td>
<td>0.35</td>
<td>0.07</td>
<td>0.96</td>
<td>1.19</td>
<td>13.30</td>
</tr>
<tr>
<td>Phesia</td>
<td>1.06</td>
<td>0.14</td>
<td>0.83</td>
<td>1.26</td>
<td>14.40</td>
</tr>
<tr>
<td>Bellerophon</td>
<td>0.37</td>
<td>0.48</td>
<td>0.17</td>
<td>2.53</td>
<td>0.41</td>
</tr>
<tr>
<td>Euphemites</td>
<td>0.05</td>
<td>0.07</td>
<td>0.25</td>
<td>0.59</td>
<td>1.63</td>
</tr>
<tr>
<td>Glabrocingulum</td>
<td>1.37</td>
<td>–</td>
<td>0.14</td>
<td>3.20</td>
<td>11.93</td>
</tr>
<tr>
<td>Trepospira</td>
<td>0.07</td>
<td>–</td>
<td>0.14</td>
<td>8.25</td>
<td>10.06</td>
</tr>
<tr>
<td>Ananais</td>
<td>–</td>
<td>–</td>
<td>0.08</td>
<td>–</td>
<td>12.46</td>
</tr>
</tbody>
</table>

All taxa listed are among the ten most abundant in at least one cluster (shown by bold face numbers). Plain numbers are percent abundance for taxa when not one of ten most abundant in that particular cluster. Dash indicates taxon does not occur in collections grouped in that cluster. Note that 17 of the 25 taxa occur in all five clusters.

the species abundance distribution of the living community if the distribution or magnitude of taphonomic forces (within habitat transport, shell dissolution, shell burial) was itself patchy within a habitat. This “taphonomic patchiness” would be indistinguishable from patchiness in the original distribution of living organisms. Also, like community patchiness, taphonomic patchiness would be expected to be mitigated with increasing amounts of time-averaging as the distribution and magnitude of taphonomic forces changed within a habitat.

2.3. Testing for patchiness in paleocommunities

If paleocommunities are not patchy, then each can be characterized by a particular species abundance distribution. In this case, making a statistically rigorous comparison of two or more local paleocommunities requires that samples from each local paleocommunity be tested to see if they could have been sampled from the same underlying distribution, given an expected level of sampling error. If, however, time-averaging does not erase patchiness from paleocommunities, then tests based on a direct comparison of species abundance distributions cannot be used to test for paleocommunity identity between samples. This is because there is no way to disprove the possibility that two statistically dissimilar samples could have been taken from different patches within the same paleocommunity. A test that finds no significant difference between samples would demonstrate identity, but a test that finds a significant difference could
not disprove identity. Therefore, before distribution-based tests can be used to test for paleocommunity identity, it must first be shown that samples from the same paleocommunity have statistically similar species abundance distributions. To do this, we can test samples drawn from the same local paleocommunity, i.e. replicate samples taken from the same stratigraphic horizon at the same outcrop locality. Because of their spatial and stratigraphic proximity, these replicate samples are, by definition, samples of the same local paleocommunity.

2.3.1. Sampling

Fossiliferous horizons within the four marine units studied were sampled at a number of localities throughout eastern Kentucky, West Virginia, and Virginia. Bulk samples of 5–10 kg of fossiliferous mudstones were disaggregated to exhaustively sample the enclosed fauna. Collections were made from each sample, and individuals were identified to species and counted. These data were analyzed using cluster analysis and principal components analysis to identify groups of similar samples representing paleocommunity types. A subset of localities encompassing a variety of paleocommunity types were selected to be revisited and replicate sampled to permit statistical testing of the data. For each locality resampled, three replicate samples of approximately 300 specimens each were made from the same bed originally sampled, with a spacing of 1–5 m between replicates. To test for patchiness in paleocommunities, the species abundance distributions of the replicate samples taken from the same stratigraphic horizon at the same outcrop locality were statistically compared using two tests that directly compare species abundance distributions.

2.3.2. Constancy analysis

The first test is an analysis of constancy — the number or percentage of samples of a community that contain a particular species (McIntosh, 1985). Constancy is used as a means of assessing how uniform the species membership is in a group of samples thought to represent the same paleocommunity. Generally, one expects numerically abundant species to exhibit high constancy within a group of samples drawn from the same paleocommunity. Abundant species that are absent from some samples raise the suspicion that those samples are not from the same paleocommunity. An overall measure of the constancy of all species within a group of samples can be obtained by calculating a coefficient of constancy ($C_c$). Because it is more likely, due to sampling error, that rare species will be absent from some samples than will common species, the coefficient of constancy weights each species relative to its overall abundance in the group. For a group of $n$ samples with $r$ species, the $C_c$ is calculated by Eq. 1:

$$C_c = \left( \sum_{j=1}^{r} \frac{P_j \times k_j}{n} \right)$$

where

- $P_j$ is the percent of species $j$ in the pooled sample group
- $k_j$ is the number of samples containing at least one specimen of species $j$.

The maximum value of $C_c$ is 1.0, obtained if all samples contain the identical number of and kinds of species. The minimum value of $C_c$ is equal to $1/n$, where $n$ equals the number of samples being used to generate the coefficient.

Although there is no statistical significance associated with the constancy coefficient itself, it can be bootstrapped to determine if the actual coefficient obtained for a group of samples is within the range expected or is lower than would be expected if the samples had been drawn from the same underlying distribution. In a bootstrap analysis, the samples are pooled to generate an estimate of the underlying species abundance distribution. The estimated distribution is then randomly sampled to regenerate the original number and size of samples, from which a new, bootstrapped coefficient of constancy is calculated. This process is repeated numerous times until a distribution of bootstrapped coefficients is obtained. The coefficient distribution shows the range of constancy values that would be expected had the original samples been drawn from the same distribution. The actual coefficient of constancy is compared to the distribution of coefficients. If it is greater than
the bootstrapped values in the lower 5% of the bootstrapped distribution, the original samples are judged similar enough have been drawn from the same species abundance distribution. For a detailed explanation of how a bootstrap analysis is done, see Gilinsky and Bennington (1994).

2.3.3. Heterogeneity chi-square

A test statistic used to compare several multinomial populations suspected to have been drawn from the same underlying distribution is the heterogeneity chi-square (Ostle and Mensing, 1975). This statistic is designed for discrete data and tests the hypothesis:

\[ H: p_{1,j} = p_{2,j} = \ldots = p_{i,j} \quad j = 1, \ldots, k \quad i = 1, \ldots, r \]

The chi-square statistic is calculated by Eq. 2:

\[ \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \]  

\[ (2) \]

where

\[ O_{ij} = \text{observed number in the j-th species in the i-th sample} \]
\[ E_{ij} = \frac{R_i C_j}{n} = \text{the expected number in the j-th species in the i-th sample} \]
\[ R_i = \text{the row total for sample I} \]
\[ C_j = \text{the column total for species j} \]
\[ n = \text{the number of samples} \]

(Ostle and Mensing, 1975).

To test for statistical identity among replicate samples, observed and expected values were calculated for 20 species using all available replicate samples at each horizon and locality tested. Because expected values less than 5 result in a biased chi-square statistic (Ludwig and Reynolds, 1988), if any expected value obtained was less than 5, expected and observed values were recalculated removing the least abundant species. This step was repeated until all expected values were greater than or equal to 5, after which the chi-square statistic was calculated. The value of the chi-square obtained was compared to tabulated values of the chi-square statistic at \( v = (r - 1)(k - 1) \) degrees of freedom. If the obtained chi-square was larger than 95% of the chi-square distribution, the species abundance distributions of the replicate samples were judged to be statistically dissimilar.

3. Results and interpretation (distribution tests)

Chi-square and coefficient of constancy bootstrap test results are presented in Table 3. Replicate samples from 18 localities were tested. The chi-square test detected significant differences between replicate samples at the 95% or higher level at 16 of the 18 localities tested, with most tests showing highly significant differences. The coefficient of constancy bootstrap demonstrated a less than 5% chance of obtaining the observed coefficient from the same underlying abundance distribution in 10 of the 18 tests. It is not surprising that the coefficient of constancy test failed to find significant differences more often than the heterogeneity chi-square test. The constancy test is a more permissive test than the chi-square because the criteria of similarity in abundance distributions is relaxed relative to the chi-square test. Only one individual from each species is needed to establish that species in a sample and show similarity in the constancy test. In this way, the coefficient of constancy test is a test based on presence–absence, rather than numerical abundance. However, the probability of obtaining at least one individual of each species in a sample is proportional to the abundance of that species in the pooled distribution, so that the coefficient is weighted to reflect those probabilities. Also, time-averaging has the effect of greatly increasing the probability that rare species will be incorporated into a fossil assemblage (Kidwell and Bosence, 1991), increasing the tendency for the constancy test to commit a type I error by concluding that two dissimilar fossil assemblages have the same suite of species.

Both tests agree that, in a majority of the localities tested, replicate samples were drawn from different underlying species abundance distributions. This shows that time averaging does not sufficiently erase the spatial variability in species occurrences within a local paleocommunity to eliminate the patchiness that was inherent in the living community or that was caused by within-habitat differences in taphonomic processes. Because of this, local paleocommunities cannot be characterized by a unique species abundance distribution and tests that directly compare species abundance distributions cannot be used to
Table 3
Results from the chi-square and coefficient of constancy ($C_e$) bootstrap tests

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Sample no.</th>
<th>Spec.no.</th>
<th>DF</th>
<th>Chi$^2$</th>
<th>Sig. level</th>
<th>$C_e$</th>
<th>% boot dist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S55 4.5–5.25'</td>
<td>4</td>
<td>9</td>
<td>24</td>
<td>86.34</td>
<td>≪99.95</td>
<td>0.98</td>
<td>08.0</td>
</tr>
<tr>
<td>S51 1.5–2.5'</td>
<td>4</td>
<td>15</td>
<td>15</td>
<td>60.62</td>
<td>&gt;99.95</td>
<td>0.95</td>
<td>08.6</td>
</tr>
<tr>
<td>S39 2–2.5'</td>
<td>4</td>
<td>8</td>
<td>21</td>
<td>51.90</td>
<td>≪99.95</td>
<td>0.94</td>
<td>00.0</td>
</tr>
<tr>
<td>S31 3.5'</td>
<td>4</td>
<td>7</td>
<td>18</td>
<td>48.25</td>
<td>&gt;99.95</td>
<td>0.87</td>
<td>00.0</td>
</tr>
<tr>
<td>S31 2.5'</td>
<td>4</td>
<td>8</td>
<td>21</td>
<td>53.37</td>
<td>&gt;99.95</td>
<td>0.95</td>
<td>00.0</td>
</tr>
<tr>
<td>S29 19–23'</td>
<td>5</td>
<td>4</td>
<td>12</td>
<td>17.39</td>
<td>90–95</td>
<td>0.92</td>
<td>05.9</td>
</tr>
<tr>
<td>S29 0–10'</td>
<td>6</td>
<td>5</td>
<td>20</td>
<td>75.47</td>
<td>&gt;99.95</td>
<td>0.97</td>
<td>04.1</td>
</tr>
<tr>
<td>M45 5–10'</td>
<td>3</td>
<td>11</td>
<td>20</td>
<td>56.49</td>
<td>&gt;99.95</td>
<td>0.99</td>
<td>26.7</td>
</tr>
<tr>
<td>M37 6–6.5'</td>
<td>4</td>
<td>4</td>
<td>9</td>
<td>18.17</td>
<td>95–97.5</td>
<td>0.98</td>
<td>06.5</td>
</tr>
<tr>
<td>M37 4.25–4.75'</td>
<td>4</td>
<td>10</td>
<td>27</td>
<td>64.76</td>
<td>&gt;99.95</td>
<td>0.98</td>
<td>05.4</td>
</tr>
<tr>
<td>M37 21–36'</td>
<td>5</td>
<td>16</td>
<td>60</td>
<td>285.30</td>
<td>≪99.95</td>
<td>0.94</td>
<td>00.0</td>
</tr>
<tr>
<td>M28 21'</td>
<td>4</td>
<td>11</td>
<td>30</td>
<td>129.36</td>
<td>≪99.95</td>
<td>0.98</td>
<td>02.1</td>
</tr>
<tr>
<td>M26 5'</td>
<td>4</td>
<td>8</td>
<td>21</td>
<td>106.30</td>
<td>&gt;99.95</td>
<td>0.99</td>
<td>33.2</td>
</tr>
<tr>
<td>K66 9.5–10.5'</td>
<td>4</td>
<td>7</td>
<td>18</td>
<td>24.68</td>
<td>80–90</td>
<td>0.98</td>
<td>10.1</td>
</tr>
<tr>
<td>K66 4.5–5'</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>30.86</td>
<td>99.5–99.95</td>
<td>0.97</td>
<td>04.4</td>
</tr>
<tr>
<td>K35 57–58'</td>
<td>4</td>
<td>12</td>
<td>33</td>
<td>58.71</td>
<td>99–99.5</td>
<td>0.98</td>
<td>00.0</td>
</tr>
<tr>
<td>E61 8–14'</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>32.55</td>
<td>99.5–99.95</td>
<td>0.97</td>
<td>00.2</td>
</tr>
<tr>
<td>E36 11–15'</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>14.79</td>
<td>97.5–99</td>
<td>0.96</td>
<td>04.2</td>
</tr>
</tbody>
</table>

DF = degrees of freedom for chi-square. Results that are not statistically significant are noted in bold.

demonstrate that two or more paleoecological samples come from the same paleocommunity.

4. Methods II (analysis of variance tests)

4.1. The problem of variation at different scales

The difficulty in formulating a statistically rigorous test for paleocommunity identity comes from the fact that variability in species distributions due to patchiness, taphonomy, and habitat gradients makes it difficult to identify or define discrete patterns of species abundance that can be used to make statistical comparisons. This is particularly a problem with paleoecological sampling because time averaging can blur the boundaries that may have existed between relatively discrete living communities (see the study by Miller, 1988). The key problem is to determine how much variation in species abundances is to be expected within a paleocommunity and how much variation is indicative of change to a structurally different paleocommunity. Whether paleocommunities are continuously intergrading over space and time or are divisible into relatively discrete entities, the problem of comparing them becomes a problem of determining boundary conditions within which a paleocommunity can be defined and sampled at a consistent and meaningful scale. Consistent implies that paleocommunities being compared have been defined using the same criteria of scale, whereas meaningful implies that this scale encompasses within-habitat variation due to patchiness without also including variation due to the mixing of samples from different habitats occupied by different paleocommunities. Finding this ideal scale is not a simple matter. Indeed, in a purely Gleasonian world where species are distributed individualistically throughout continuously intergrading habitat, there is no ideal scale — any subdivision of such a continuum would necessarily be arbitrary. However, there are natural discontinuities in habitat and species distributions at many spatial and temporal scales that break-up the continuum, and change along a continuum may itself be profitably measured provided that the continuum is subdivided into units of consistent size. As noted by Wiens (1986), what is important is that the scale of communities (and therefore paleocommunities) defined for analysis be appropriate to the questions being addressed, that the scale be
made explicit when reporting the study, and that the conclusions reached only be applied at the scale from which they were derived.

4.2. Defining scale in paleocommunity analysis

As explained above, we propose that a reasonable base scale for encompassing within-paleocommunity variation is the scale of the local paleocommunity as sampled at a particular stratigraphic level in a single outcrop. At the outcrop scale (from 1 to 10’s of meters) individual stratigraphic horizons are usually easily traced, replicate samples needed to quantify local paleocommunities and their internal variability can be made, and, at least for benthic marine invertebrates, this scale is localized enough to preclude the likelihood of mixing different paleocommunities. From this base of within-paleocommunity variability at the outcrop scale, local paleocommunities can be compared at larger spatial and temporal scales. For example, two outcrop localities separated by a wide covered interval might be compared. If they were found to be statistically similar, they could be grouped to quantify within-paleocommunity variability at the scale of the two outcrops and then used to make comparisons between other outcrop groups at the larger scale. In this way, the spatial scale occupied by a paleocommunity can be determined and hypotheses of paleocommunity sameness can be tested.

It must be emphasized that replicate samples at the outcrop scale are the best objective quantification of a paleocommunity because of their spatial proximity within a local paleocommunity. For the purposes of drawing robust conclusions as to the pattern of paleocommunity stability over time, we strongly suggest that 1) paleocommunity identity be determined on the basis of statistical similarity using replicate samples at the locality scale to define within-paleocommunity variability, and 2) that paleocommunity identity only be extended to larger spatial scales when the criteria of statistical identity is met after the variability within local paleocommunities has been accounted for. In the absence of statistically defined identity, sample groups with similar (but not the same) species abundance distributions should be considered to be examples of a paleocommunity type rather than a paleocommunity.

Even if shown to be statistically dissimilar, an investigator might still group samples from two different localities given some other criteria on which to conclude that the locality differences were still the result of variability within a paleocommunity. If this is done, the rational for grouping the localities, and the spatial scale and range of variation must be made explicit so that inferences and comparisons made using this sample group can be related back to their underlying assumptions. However, at larger spatial scales, grouping samples within the same paleocommunity becomes increasingly arbitrary and subjective in the absence of statistical testing based on the level of variation at the scale of the local paleocommunity.

4.3. Incorporating variability into testing for paleocommunity recurrence

Testing for statistically significant differences between groups of samples can be done by comparing the variance in the means between the different groups to the variance between samples within the different groups using an Analysis of Variance (ANOVA) or related techniques. For paleoecological data sets containing multiple species, separate ANOVAs can be performed on individual species and an overall multivariate analysis (MANOVA) can be performed that compares multiple species simultaneously. Nested (M)ANOVAs can also be performed that permit the decomposition of variability into several treatment levels, allowing variability at several scales to be assessed in the same analysis. Although it is often assumed that paleontological data are inappropriate for parametric statistical analyses, we have chosen to use parametric tests because non-parametric analyses based on rank abundance of species are prone to type II errors (in our case, concluding that fossil assemblages were different when, in fact, they were the same, thus failing to correctly reject the null hypothesis). This happens because sampling error and taphonomy can easily alter species counts by a few specimens, changing the rank order of rare taxa as well as abundant taxa that occur in comparable numbers (Kidwell
and Bosence, 1991). Parametric statistical tests based directly on abundance data are impacted much less by such small changes in species abundances.

4.3.1. Analysis of variance tests

Replicate samples of species abundance data from 17 different localities representing 4 different marine units were collected as described above. In most cases, from among four or more replicate samples available, three were chosen for analysis that were the closest in sample size. To make species abundance data amenable to analysis using the ANOVA, the data were transformed using the Freeman-Tukey variant of the arc-sine data transformation (Bishop et al., 1975):

\[
a = \frac{1}{2} \left( \sin^{-1} \left( \frac{x_i}{n_i + 1} \right) + \sin^{-1} \left( \frac{x_i + 1}{n_i + 1} \right) \right)
\]

(3)

where \( x_i \) is a particular species’ abundance in sample \( i \) and \( n_i \) is the total abundance of sample \( i \).

This transformation has the dual effect of rendering the data proportional (as in a percent transformation) and stabilizing the sample variances. Achieving uniformity of sample variances is important for performing ANOVA and MANOVA tests because type I error (deciding that two groups are different when in fact they are not) is strongly affected by variance heterogeneity. A second assumption underlying ANOVA and MANOVA tests is that observations are normally distributed on the dependent variable(s) in each group. This assumption is usually violated by ecological data (Day and Quinn, 1989). However, studies of the effects of non-normality indicate that for both ANOVA and MANOVA, even a large deviation from normality has only a small effect on type I error, usually within .02 of the target significance level, provided that the variances are stable (Stevens, 1992). ANOVA and MANOVA tests were performed on transformed data using the SAS statistical analysis program’s General Linear Models procedure (SAS Institute, 1982). Analyses were weighted by sample size to reflect the fact that larger samples contain more data and thus should have a proportionally greater influence on the test.

For each test comparing sample groups, both the multivariate MANOVA and individual univariate ANOVA tests were performed. For the MANOVA, the maximum number of species was used, in decreasing order of abundance in the pooled sample distribution, that would allow the SAS program to construct the partial correlation matrix needed to make the test. Increasing the number of localities or replicate samples increases the amount of data available to the SAS program to construct the partial correlation matrix. Likewise, increasing the number of variables (species) the test must incorporate increases the amount of data needed to construct the matrix. For the purposes of comparing the results of tests using different numbers of samples and localities, we decided to always use the maximum number of species that would allow SAS to make the test, in order to keep the power of each MANOVA approximately comparable. MANOVA sample groups were judged to be significantly different at a \( p \) value of 0.05. For the individual univariate ANOVA tests we used the six or eight most abundant species from the pooled sample distribution. Because multiple univariate significance tests were performed on the same sample group, the significance level of 0.05 was adjusted to control for the inflated type I error rate associated with multiple tests. We used the Bonferroni procedure (Bray and Maxwell, 1985) which divides the single test significance level by the number of tests performed to yield an adjusted significance level. The adjusted significance level is 0.008 for tests on 6 species and 0.006 for tests on 8 species.

4.3.2. Choosing localities for testing

To prevent inflation of the type I error rate, only a limited number of planned comparisons between localities were made out of all possible comparisons. Localities were chosen for comparison based on the results of a two-way cluster analysis of all Breathitt marine unit samples (Fig. 4). The two-way cluster analysis arranges samples in an \( X-Y \) plane based on their position in both a Q-mode cluster based on species abundance data and a Q-mode cluster based on species presence–absence. The cluster analyses were done using the Unweighted Pair-Group Method
Fig. 4. Two-way cluster analysis diagram showing the positions of all localities from which replicate samples were taken. Locality pairs tested for identity using ANOVA and MANOVA are connected by thin ovals. Multiple localities tested using nested ANOVA and MANOVA are grouped within rectangles. First letter of locality name indicates marine unit (E = Elkins Fork, K = Kendrick, M = Magoffin, S = Stoney Fork). Names at bottom of diagram refer to characteristic fauna of the community type identified by quantitative cluster analysis.

(UPGMA) and the quantified Czekanowski's similarity coefficient (for species abundance clustering) and Jaccard's similarity coefficient (for presence-absence clustering). These cluster analysis methods have been used previously in paleoecological analyses and are discussed in Miller (1988) and Springer and Bambach (1985).

5. Results and interpretation II (analysis of variance tests)

5.1. Localities compared

A Q-mode cluster analysis of samples from the Breathitt marine units grouped samples in broadly similar clusters characterized by dominant macroinvertebrate types (identified in quotes at the base of Fig. 4 and listed in Table 2). These clusters represent one or more paleocommunity types which encompass an undetermined number of paleocommunities. Localities to be replicate sampled were chosen so that statistical comparisons could be made at a variety of spatial scales (single and multiple localities within individual marine units), at a variety of temporal scales (adjacent stratigraphic levels, between different marine units), and within several of the sample clusters. For tests of similarity between marine units, localities from different marine units were chosen for comparison that were grouped in the same cluster and appeared from the cluster analyses to be very similar. Thus, from the range of samples, localities and community types represented in the Breathitt marine units we made comparisons that maximized the chance that we would find recurrence of the same paleocommunity in different marine units. The results of these comparisons are presented below in a sequence from replicate sample groups taken in close geographic and temporal proximity to those taken at increasingly divergent spatial and temporal scales.

5.2. Tests of adjacent stratigraphic horizons at a locality

At two localities replicate samples were taken at adjacent stratigraphic horizons. Locality S31 was replicate sampled at two levels within a continuously fossiliferous stratigraphic interval of uniform lithology. Both samples clustered in the Chonetid-Productid cluster and the ANOVA and MANOVA were unable to detect significant differences in the species abundances between the two levels (Fig. 5A). At locality M37 replicate samples that clustered in the Chonetid-Productid cluster were taken from two proximal stratigraphic horizons with slightly different lithologies (the lower horizon, M37 21–36' was more calcareous than the upper horizon). The MANOVA test failed to find a significant overall difference between the two horizons, although the ANOVA detected significant differences in the distributions of the second and third most abundant species, Rugosochonetes delicatus and Punctospirifer kentuckyensis (Fig. 5B).

The results from testing at adjacent stratigraphic horizons show that, given the persistence of the same habitat (shown by the constant lithology over 0.3 meters of section from S31 2.5' to S31
3.5) it can be demonstrated that statistically identical local paleocommunities can recur at a locality, with no significant stratigraphic variability added to the local variability that exists due to patchiness. In this case, both local paleocommunities should be regarded as representatives of the same paleocommunity. Furthermore, at M37 we see that, although a slight change in lithology results in the addition of significant stratigraphic variation in two species beyond the variation seen locally at one stratigraphic level, no overall significant difference in the two successive local paleocommunities is shown by the MANOVA. This suggests that some component species of the precursor community responded to changes in habitat that occurred through time at locality M37, but that overall the two local paleocommunities are not statistically demonstrably different. In this case, with no significant difference in the abundances of 6 of the 8 species and with the MANOVA test showing no significant difference, it is also reasonable to regard both local paleocommunities as belonging to the same paleocommunity.

5.3. Tests of the same horizon at different localities within the same marine unit

We next compared the two stratigraphic horizons at locality S31 to a horizon within the same fossiliferous interval at locality S39 (app. 6 km distant) that also clusters in the Chonetid–Productid cluster. Results from the MANOVA test are equivocal; one test statistic (Wilks) found no significant overall difference with an F value just over the 0.05 level, while two other test statistics suggested differences at a significance just below the 0.05 level (Fig. 5C). The ANOVA failed to find significant differences between localities in all species except for Punctospirifer kentuckiensis, the second most abundant species in the pooled
distribution, which was highly significantly different between localities. In this case it is shown that increasing geographic and temporal distance (temporal distance because individual beds are often time-transgressive over distance) resulted in a detectable amount of variability between two local paleocommunities. Whether or not these local paleocommunities should be considered to be representatives of the same paleocommunity depends on the strictness with which the definition of sameness is applied. If one is conservative and feels that all the MANOVA test statistics should show significance before accepting that two local paleocommunities should be grouped in the same community, then the two localities would not be considered the same. On the other hand, with seven of the eight species abundances being statistically indistinguishable and one of the MANOVA tests giving the same result, both localities might very well be regarded as representing the same paleocommunity. What is certain, however, is that the tests were able to detect a major difference in the abundance distribution of one of the characterizing species (P. kentuckiensis) between the two localities.

Four additional tests for differences between replicate sample groups from a single fossiliferous interval at different localities in the same marine unit are presented in Fig. 6. Between two of the locality groups (E36/E61 from the Productid cluster, Fig. 6A and M37/M28 from the Chonetid–Productid cluster, Fig. 6C) both ANOVA and MANOVA tests failed to show significant differences in the species abundances of any of the tested species between localities separated by 22 km and 37 km, respectively. At a spatial scale of several tens of kilometers, these localities show no more variability in their species compositions than is seen at adjacent horizons within the same fossiliferous interval or at a single horizon at a single locality. Thus, we interpret these tests as demonstrating two examples (each from a different community type) in which two local paleocommunities are found to represent the same statistically defined paleocommunity. For the remaining two locality pairs (M45/M28 from the Small Mollusk cluster, Fig. 6B and S51/S55 from the Productid cluster, 6D) the MANOVA test detected significant overall differences and the ANOVA tests detected significant differences in the distributions of several numerically abundant species. Thus, in spite of great overall similarity in species composition and abundances (revealed by the close association of samples from the locality pairs in cluster and ordination analyses), these sample groups cannot be attributed to the same statistically defined paleocommunity. Instead, they are examples of a community type — representatives of similar, but not identical, paleocommunities.

At this point it should be mentioned that, in addition to testing for identity within community types as defined by clusters in the cluster analysis, we have also tested locality pairs from different
clusters, representing different community types (Bennington, in prep.). Although we do not present the results of those analyses here, we wish to note that they show a significant difference between localities tested, with a range of difference that overlaps the range of difference seen in the tests between localities from within clusters. This suggests that when two localities within a cluster are found to represent two different paleocommunities, the differences between them are not trivial and are, in fact, often as great as the differences found between localities suggested by other multivariate analyses to be different paleocommunity types. This implies that community types, while displaying enough faunal and environmental similarity that they make good clusters in cluster analyses and show stratigraphic consistency in their distribution, have no particular structural or organizational integration as aggregate entities beyond the adaptive appropriateness of their component species to the environment.

5.4. Tests of locality pairs in different marine units

Two locality pairs from different marine units with similar fossil assemblages in lithologically similar stratigraphic intervals were tested for paleocommunity recurrence. Comparing replicate sample groups from localities M37/S29 from the Productid cluster, the MANOVA detected a significant overall difference at a type I error probability of 0.03 (Fig. 7A), but the ANOVA tests detected a significant difference in the distribution of only one species, Derbyia crassa, the second most abundant species in the pooled distribution. For the locality pair S29/K66 from the Chonetid–Mollusk cluster, the MANOVA could not detect a significant overall difference, although the ANOVA tests show significant differences in the distributions of the three least abundant species of the six tested (Fig. 7B).

The results of tests on these two pairs of localities suggest that the local paleocommunities that did form are very similar to one-another, albeit in two different ways, and that the decision to recognize paleocommunity recurrence is made from a continuum of possible results, not at a natural discontinuity. In the M37/S29 pair a majority of species argue for identity, with only one relatively abundant species being highly dissimilar between the paleocommunity types in the Magoffin and Stoney Fork marine units, yet the MANOVA test recognizes an overall dissimilarity. In the K66/S29 locality pair, the MANOVA test permits both local paleocommunities to be regarded as the same paleocommunity. The three species that compose 84 percent of the pooled sample distribution also are statistically similar between marine units, whereas the remaining three species tested are not. In this case the MANOVA and the ANOVAs on the abundant species argue for identity, whereas the rare species appear to be dissimilarly distributed. A liberal interpretation could be used to argue both cases are recurrence of the same community, but in each the point of justification, the similarity of 5 of 6 species abundances in one instance, the agreement of the MANOVA and the abundances of the top three species in the other, would have to be explicitly stated. A conservative approach could, likewise, be used to reject either case, the MANOVA and the one dissimilar species arguing for dissimilarity in the first, and, in the second, the differences in the rare species could be considered compelling enough to define the two localities as two different paleocommunities. In the second case, the decision would rest on whether rare species are considered to be important in defining community structure.
5.5. Tests of multiple localities in different marine units

As discussed above, the results of tests between locality pairs M28/M37 (Fig. 6C) and S39/S31 (Fig. 5C) show that these locality pairs each represent a particular paleocommunity. Samples from both of these locality pairs group within the Chonetid–Productid cluster in the quantitative Q-mode cluster analysis (Fig. 4) and so probably represent members of a single paleocommunity type. We can now use both locality pairs to test for recurrence of the same paleocommunity in the Magoffin and Stoney Fork marine units. Results from the MANOVA (Fig. 8A) show that the test failed to find significant differences between localities within the two marine units and between the two marine units. This suggests that not only does the same paleocommunity recur within both the Magoffin and Stoney Fork marine units (as concluded from the separate analyses discussed above) but that there is no overall detectable difference between species distributions from one marine unit to the next and the same paleocommunity recurs in the Magoffin and in the Stoney Fork. The individual ANOVA tests, however, do show some detectable species differences both between marine units and within the marine units. Desmoinesia muricatina is significantly different between localities within one of the marine units (the test does not indicate which one) but not between the marine units. This suggests that although D. muricatina is somewhat variable within at least one of the paleocommunities within a marine unit, it does not show any additional significant variability between the two marine units. Punctospirifer kentuckyensis is highly significantly different both within and between the two marine units, whereas Hustedia miser is not significantly variable within the marine units, but does have detectable differences in its distribution between marine units. Overall, the results of this test show that although the MANOVA suggests that there is statistically definable recurrence of this paleocommunity from the Magoffin to the Stoney Fork, it is recurrence with some detectable variability in species abundances.

A final test was performed to determine if the same paleocommunity occurred in all four marine units. For this test, localities were chosen from the Productid cluster (Fig. 4), the only cluster that contained samples from all four marine units. Replicate samples were made of the same stratigraphic horizon at two localities within each marine unit (except for the Magoffin, in which only one locality with a Productid-dominated assemblage was found). Results of both the

---

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Locality</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species</td>
<td>% Pool</td>
<td>F Value</td>
</tr>
<tr>
<td>Rugosochonetes</td>
<td>34.21</td>
<td>7.79</td>
</tr>
<tr>
<td>Desmoinesia</td>
<td>13.15</td>
<td>15.27</td>
</tr>
<tr>
<td>Derbyia</td>
<td>9.26</td>
<td>1.05</td>
</tr>
<tr>
<td>Punctospirifer</td>
<td>6.38</td>
<td>6.25</td>
</tr>
<tr>
<td>Antiquatoria</td>
<td>5.97</td>
<td>0.54</td>
</tr>
<tr>
<td>Anthracospirifer</td>
<td>5.76</td>
<td>5.61</td>
</tr>
<tr>
<td>Composita</td>
<td>5.01</td>
<td>9.44</td>
</tr>
<tr>
<td>Hustedia</td>
<td>3.46</td>
<td>6.82</td>
</tr>
</tbody>
</table>

α = .006 α = .006

---

**Fig. 8.** Results of ANOVA and MANOVA tests for samples from multiple localities in different Breathitt marine units. All else as in Fig. 5.
MANOVA and ANOVA tests (Fig. 8B) show that, with one exception, all species tested showed significant differences in distribution both within localities and between marine units. In no way could the local paleocommunities from the four marine units be grouped within the same paleocommunity. Instead they are shown by the statistical tests to be broadly similar local paleocommunities representative of a paleocommunity type whose membership is specified by the cluster analysis.

6. Discussion

We can now return to the question that was posed at the beginning of this paper and ask, do the four Breathitt marine units studied show recurrence, and if so, what is recurring? Collections from the four marine units show that the same species (or very similar species within genera) recurred in each marine unit. Each marine unit itself is a relatively thin package of strata deposited through a single cycle of transgression and regression. Although the Magoffin marine unit contains a wider variety of habitats and paleocommunity types (all of the Small Mollusk and most of the Spiriferid cluster samples, Fig. 3; see also Bennington, in press), most marine facies observed in the Breathitt Formation and three of the paleocommunity types defined by cluster analysis (Chonetid–Mollusk, Productid–Chonetid, and Productid, Fig. 3) are well represented by samples from at least two different marine units. Furthermore, there is an impressive amount of redundancy in species membership through the different community types (Table 2) with distinctions between community types created more by changes in species abundance than by changes in paleocommunity membership. Thus, not only does it appear that the same species pool persisted through the time during which the four marine units formed, but also that the same general species associations recurred, and that most species were present throughout the range of marine habitats that recurred from marine unit to marine unit. All of the evidence indicates that the Breathitt marine units are a good example of the repeated reconstruction and recolonization of a spectrum of similar marine habitats from a relatively stable and persistently available pool of species (the “source-sink” network of Miller, 1993a,b). The recurrence in different marine units of similar species assemblages characterized by similar numerically dominant taxa demonstrates at the very least the recurrence of paleocommunity types through the Breathitt marine units.

Within the Breathitt paleocommunity types there is variability in the species abundance structure of individual samples at all scales. Variability due to patchiness appears to be ubiquitous between replicate samples at the scale of the local community (outcrop). At close stratigraphic horizons within a continuously fossiliferous interval, very little additional variation beyond that attributable to patchiness was noted. At larger scales between outcrops within single marine units, a significant increase in variability beyond that attributable to patchiness at the outcrop scale can be demonstrated in some cases, but in others no significant difference between localities could be found. Between marine units, two examples of possible paleocommunity recurrence were noted, but in both cases some significant variability between marine units also was detected in one or more species. In general, as the temporal and spatial scale of the comparisons increased, so did the probability of finding some significant variability. This suggests that paleocommunities are continuously variable in time and space and that, although general species composition is conserved or recurs, the detailed species composition and the abundances of species in paleocommunities recur with less and less fidelity at increasing spatial and temporal distance.

It must be emphasized that we have constructed our tests to maximize the chance of finding statistical recurrence of paleocommunities. We have done this by choosing to make statistical tests only when other multivariate analyses show that the local paleocommunities being compared are very similar. Many more comparisons of local paleocommunities between marine units could have been made, but these would not have compared fossil assemblages with similar species abundances. We find it striking that, in spite of the overall similarity
in facies sequence, facies types, and species occurrences in all of the four marine units studied, detailed analyses of species composition and species abundance reveal unique aspects of the samples of the paleocommunity types found in each marine unit. It should also be emphasized that without collecting and numerically analyzing large paleontological samples, no paleontologist (ourselves included) would have noted the uniqueness of many of the paleocommunities in the different marine units.

When interpreting the results of the statistical tests presented here, it should also be kept in mind that our tests allowed for a great amount of variability between local paleocommunities that was not directly a result of significant differences between the original living communities. By using replicate samples from a bed at a single outcrop as the basis for the (M)ANOVA tests, we were able to eliminate the influence of patchiness due to local differences in species distribution and differences in taphonomic effects. Furthermore, because the statistical comparisons made were between local paleocommunities from low energy facies with very similar lithologies and taphonomic characteristics, it is unlikely that significant differences would have existed in taphonomic regimes from locality to locality or from one marine unit to the next, beyond those that might have existed from place to place within a locality. Because of this, we argue that the taphonomic null hypothesis can be discounted as an explanation for the differences detected between local paleocommunities by the statistical analyses.

The null model we apply here to coordinated stasis in paleocommunities is analogous to a null model for morphological stasis in evolutionary lineages developed by Bookstein (1988). Bookstein observed that the phenomenon of “stasis” was assumed to be operating on a lineage if it did not show an obvious anagenetic trend. After noting the presence of stasis in many lineages, workers began to formulate hypotheses based on such ideas as stabilizing selection to explain stasis. Bookstein noted that a more appropriate null hypothesis explained stasis as a random walk, with apparent stability caused by nothing more than random evolutionary change through time. In this study, we suggest that coordinated stasis (analogous to morphologic stasis) only requires explanation with recourse to stabilizing phenomenon such as “ecological locking” (analogous to stabilizing selection) if it can first be demonstrated that paleocommunities recur in the fossil record with any more fidelity than would be expected from the reinvasion of recurring habitats by suitably adapted species.

7. Conclusions

Can we reject the null hypothesis of recurrence simply due to recruitment into similar habitats from a persistent species pool (in effect, recurrence of community type) to explain recurrence in the Breathitt marine units? The results of this analysis suggest that we cannot, and that scale does play a major role in the likelihood of the recurrence of entities we can regard as being statistically the same. Recurrence of the same statistically defined paleocommunity could not be unambiguously falsified in the tests that compared samples from a cluster taken within a single marine unit at a single locality, and the same can be said for half of the comparisons made between localities within a marine unit and for those between marine units. However, half the cases incorporating separate localities in a marine unit or several marine units were significantly dissimilar, and all comparisons made between marine units contained some statistically significant variability. Although particular paleocommunities do sometimes recur in time and space at small scales, this is not the general case at larger temporal scales. The general pattern of recurrence on a large temporal scale (such as from one major marine transgression to the next) is recurrence of community type.

This study examines the possibility of stasis within a time interval equivalent to one of Brett and Baird’s intervals of stasis in the Silurian and Devonian and does not incorporate intervals of paleocommunity reorganization such as those that bound Brett and Baird’s intervals of coordinated stasis. Did this study fail to detect coordinated stasis in the Breathitt marine units? Although stasis is observed in the sense that some paleocommunity types recur through the Breathitt marine units, we
see little evidence for the widespread conservation of particular species abundance relationships beyond the local scale. The recurrence of statistically identical paleocommunities from one marine unit to the next is not the general case. However, our interest in presenting this study is not only to show that, statistically, coordinated stasis was not detected regularly in the Breathitt marine units below the level of paleocommunity type, but to also develop a reasonable null hypothesis against which explanations for coordinated stasis can be tested. As with any null hypothesis, the burden of proof should be shifted away from this null hypothesis and onto whatever alternative hypotheses for the cause of coordinated stasis might be presented. To demonstrate the operation of any process that would actively conserve the structure of paleocommunities over long intervals of time (a coordinating mechanism, for example, the hypothesis of “Ecological Locking” outlined by Morris et al., 1992) it must first be determined how that process would affect community structure. Second, it must be shown that the effect of the coordinating mechanism on the community would produce a structure in paleocommunities different from that expected from the operation of the null model. Finally, the observed pattern of paleocommunity structure must be tested against the pattern expected from the operation of the coordinating mechanism. Only then can detection of the coordinating mechanism in the fossil record be established. Our aim in this paper has been to help this process along by outlining the null hypothesis for paleocommunity recurrence and presenting a methodology for making statistically rigorous and meaningful comparisons between paleoecological samples.

Acknowledgements

Portions of this research were supported by grants from the Geological Society of America and Sigma Xi to J.B. Bennington, and National Science Foundation grant EAR-9105469 to R.K. Bambach. We also would like to thank Dr. George Terrel and Yan Jia of the Department of Statistics at Virginia Tech for their assistance in developing the analyses used in this study. Peter Wagner and Doug Erwin provided thoughtful and constructive reviews of the original manuscript that contributed greatly to the quality of this paper. Finally, we thank Linda Ivany and Ken Schopf for organizing this theme issue on faunal stability and for their continued enthusiasm for discussing issues involving community recurrence and coordinated stasis.

References


Bennington, J.B., in press. Stratigraphic and biofacies patterns in the Middle Pennsylvania Magoffin marine unit in the Appalachian basin, U.S.A. Int. J. Coal Geol.


