Chapter 1
Ordination Methods and the Evaluation of Ediacaran Communities

Matthew E. Clapham

Contents

Abstract Analysis of paleocommunity data poses a challenge because of its multivariate nature, containing counts of many species in many samples. Comparison of the abundance of a single species among all samples provides only incomplete information, whereas attempting to consider every species is impractical. Ordination methods are analytical techniques that reduce the original multivariate dataset to a few important components by creating new synthetic variables designed to explain the maximum amount of original data variability. The ultimate goal is to order the samples along ecologically or environmentally meaningful gradients in order to interpret differences in community structure. This chapter describes three of the most widely-used ordination methods, principal components analysis (PCA), detrended correspondence analysis (DCA), and non-metric multidimensional scaling (NMDS), explaining the methodology of each and outlining their strengths and weaknesses for analysis of paleoecological data. The techniques are illustrated using Ediacaran paleocommunity data from Mistaken Point, Newfoundland. PCA relies on assumptions that are inappropriate for ecological data, such as the requirement

---

M.E. Clapham (✉)
Department of Earth and Planetary Sciences, University of California,
Santa Cruz, CA 95060, USA
e-mail: mclapham@ucsc.edu

M. Laflamme et al. (eds.), *Quantifying the Evolution of Early Life*, Topics in Geobiology 36,
DOI 10.1007/978-94-007-0680-4_1,

that species abundances change in a linear fashion along the environmental gradient, and is not well suited for community ordination. In contrast, DCA and NMDS both perform well with ecological data; DCA incorporates a more ecologically-realistic measure of distance between samples but some of the detrending methods have been criticized, whereas NMDS only assumes a monotonic relationship between compositional similarity and gradient distance. The two methods also have complementary strengths, with DCA typically better at extracting the primary gradient and NMDS better at resolving the overall pattern.

Keywords [AU1]

1.1 Introduction

Ecologists and paleoecologists are commonly faced with large datasets containing many variables from many samples (multivariate data), with the ultimate goal of revealing and interpreting ecologically significant patterns in those data. In paleoecology these datasets often take the form of an occurrence matrix of presence/absence or relative-abundance counts of perhaps 10 to >100 taxa (variables) from as many as 10–100s of collection sites. Analysis of between-sample similarities in a single variable (e.g., the abundance of a single taxon) is inadequate at capturing the full range of variability in the original dataset. However, the original dataset also contains a great deal of redundant information as, for example, several taxa can respond in a similar fashion to the same environmental gradient (Palmer 1993). Thus, it is almost always desirable to reduce the number of variables, typically to the two or three most important components, allowing the data to be displayed and interpreted more easily than by examining the dozens of variables in the original data.

A wide range of different analytical methods have been devised for reducing complex multivariate datasets so that they can be displayed in two or three dimensions (James and McCulloch 1990; Shi 1993). They are collectively called ordination techniques and in ecology and paleoecology they typically are used for gradient analysis – identifying environmental gradients and placing species in their correct position along those gradients (Bray and Curtis 1957). Environmental gradients, such as depth (more precisely, depth-related variables such as energy and substrate) in marine communities, exert a fundamental control on the spatial distribution of species and, as a result, are the primary influence on community composition (Cisne and Rabe 1978; Holland et al. 2001; Holland 2005). Other important gradients may include biogeographic variation, other environmental changes such as salinity (in estuarine environments, for example), temporal changes in community structure due to origination/immigration and extinction/emigration, or ecological succession as argued by Clapham et al. (2003).

This chapter will focus on three techniques for reconstructing environmental gradients and analyzing community structure: (1) Principal Components Analysis (PCA), (2) Correspondence Analysis (CA) and its relative Detrended Correspondence Analysis (DCA), and (3) Non-Metric Multidimensional

Scaling (NMDS). Other methods (Polar Ordination and Principal Coordinates Analysis) will be discussed briefly but are less commonly applied by paleoecologists. These are all indirect gradient analysis methods, as they use the distribution of species among the sites to infer the original environmental gradients that structured those sites (Minchin 1987; Palmer 1993). In contrast, direct gradient analysis techniques (e.g., Canonical Correspondence Analysis) explicitly relate species composition to measured environmental parameters (Palmer 1993), but are less applicable to paleontological data where environmental parameters typically are not directly measured or even estimated. The key differences between these three indirect gradient analysis methods are the mechanisms by which they quantify sample dissimilarity and the assumptions inherent in relating dissimilarity to separation along the environmental gradient (Faith et al. 1987). The implications of those differences will be discussed in greater detail below, with the ultimate goal of answering the question: which ordination method should you use to analyze your data?

1.2 Dataset Summary

This chapter will use a dataset of relative abundances from Ediacaran communities at Mistaken Point (Table 1.1), previously used by Clapham et al. (2003), to explore the procedures, assumptions, and application of different multivariate techniques (see Chap. 3 for a more detailed description of the fossil sites). The occurrence matrix contains abundance counts of 16 taxa from nine fossiliferous bedding planes, including three outcrops of the well-known "E" surface. All abundances

Table 1.1 Occurrence matrix showing abundance of 16 taxa in nine Mistaken Point samples

Sample/Taxon	PC	BC	LMP	D	Ey	Eq	Ewc	G	SH
Aspidella	0	0	0	1	3	0	1	0	1
Bradgatia	0	1	0	76	226	15	14	55	10
Charnia	18	0	152	20	84	5	14	4	1
"Charnia" B	0	26	58	0	0	0	0	0	0
Charniodiscus	0	0	13	0	1,398	52	35	49	3
"Dusters"	0	0	4	8	612	28	35	5	0
Fractofusus "Spindle"	0	76	4	1,169	1,497	71	100	0	0
Hiemalora	0	0	1	0	21	1	0	0	0
"Holdfast Frond"	0	0	0	0	0	0	0	22	0
Ivesheadia ("*Ivesia*")	79	1	1	7	77	5	3	8	27
"Lobate Disc"	0	0	0	0	142	3	10	0	0
Hapsidophyllas "Network"	0	0	0	1	1	0	2	0	0
"Ostrich Feather"	0	0	62	0	0	0	0	0	0
Pectinifrons "Pectinate"	0	0	0	175	0	0	0	0	304
"Spoon Frond"	0	0	0	0	1	1	1	0	0
Thectardis "Triangle"	140	0	0	0	25	3	7	0	0
Total	237	104	295	1,457	4,087	184	222	143	346

were standardized to percent abundance by normalizing by the total number of specimens per sample, as many of the analyses are sensitive to variations in absolute abundance among sites that reflect sampling intensity rather than an actual ecological parameter. Most taxa are identified to genus level but several are polyphyletic form taxa: the "dusters" category likely encompasses two or more undescribed genera, while *Charnia* counts also include the recently-described genus *Beothukis*, which was not differentiated in the original data collection of Clapham et al. (2003). All of the ordination techniques will be performed on the same occurrence matrix to facilitate understanding of their methodology and enable comparison of their results. Most analyses shown here were performed with the free software package PAST (Hammer et al. 2001) but can be done with a variety of statistical programs.

1.3 Data Standardization

The first important choice to be made, even before deciding on the appropriate ordination technique, is whether to analyze the raw abundance counts or whether to standardize them. All of the analyses in this chapter will be performed after converting each raw abundance value to percent abundance, normalizing the data to eliminate the influence of variations in total sample size. Otherwise, it is possible that two samples may be erroneously grouped together because rare species in the large sample have a similar raw abundance to common species in the small collection. This standardization was chosen because sample size in the Mistaken Point dataset is a reflection of sampling intensity, rather than having ecological significance. This is the case in many, but not all, paleoecological datasets, and investigators will need to consider whether variations in total abundance have biological significance (in which case they should not convert the data to proportions) or reflect sampling intensity.

It is also possible to normalize species to their maximum abundance so that the species relative abundances all sum to a value of 1 (or 100%), or to perform a double standardization by both site and species totals (Faith et al. 1987). Standardization by species total abundance strongly weights rare species and reduces the influence of abundant taxa, which may not be appropriate unless the species occur at similar abundance or different trophic levels are present and the investigator wishes to account for the rarity of carnivores relative to primary producers, for example Jongman et al. (1995).

1.4 Ordination Methods

1.4.1 Principal Components Analysis (PCA)

Principal Components Analysis is one of the oldest techniques used in multivariate data ordination and was the first to be applied to ecological data (Goodall 1954).

It still enjoys some use in paleoecology (e.g., Rodriguez 2004; Botquelen et al. 2006; McKinney and Hageman 2006) despite having several attributes that are not ideally suited to certain types of ecological data (Gauch and Whittaker 1972; Fasham 1977; Minchin 1987). The most severe issue, discussed in more detail below, is the distortion induced by fitting a linear (Euclidean) distance model to non-linear ecological gradients (Gauch and Whittaker 1972; Minchin 1987). As a result, PCA should not be used for analysis of ecological (species count) data. It is instructive, however, to examine the analytical procedure used in PCA because it illustrates the general principles of ordination, in particular eigenanalysis-based multivariate ordination. PCA, like all ordination, seeks to summarize the original multivariate dataset, which can often contain dozens of species at multiple sites, in a smaller number of new, synthetic variables – the "principal components" in the method's name (James and McCulloch 1990). It is based on linear algebra, using a technique called eigenanalysis to rigidly rotate the original data matrix (i.e., changing the coordinate system without altering the relative position of the points) so that the new principal component axes account for the maximum amount of variability in the original data (Quinn and Keough 2002). The underlying mathematical basis is complex but the basics of PCA and other eigenanalysis methods can be explained graphically by considering a simplified example with only two variables (Fig. 1.1).

In this simple example, analysis of ecological data begins with a bivariate plot in which the axes are the abundance of species 1 and species 2, and the data points represent each sample in the study (Fig. 1.1a). The data are then "centered" by subtracting the mean species abundance from the abundance in each sample; this has the effect of moving the points so that the graph axes run through the mean of each species (Fig. 1.1b). The final step rotates the axes so that one axis, principal component 1, is aligned with the maximum variance of the sampled points (Fig. 1.1c). In fact, the new axes are simply obtained by a linear regression through the data points. Principal component 2 is created in a similar fashion and accounts for the maximum amount of remaining variance with the constraint that it is perpendicular to axis 1. This procedure can easily be conceptually generalized to

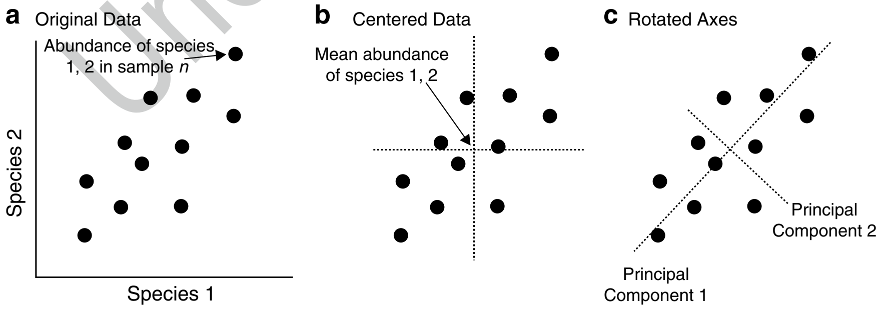

Fig. 1.1 Simplified two-variable example conceptually illustrating the process involved in PCA. (**a**) The abundance of each species is plotted on a bivariate (or multivariate) plot. (**b**) Data points are centered so that the axes pass through the mean point of each species. (**c**) Axes are rotated to align with the direction of maximum variability in the original data

multivariate datasets: principal component 1 is the direction of maximum variance in the multidimensional cloud of points and additional axes are generated to explain the maximum remaining variance while remaining perpendicular to all other axes.

In practice, PCA solutions are not derived graphically but are instead calculated using a linear algebra technique involving spectral decomposition of an association matrix to extract the eigenvectors and eigenvalues (these two terms will be explained below). The association matrix defines the interrelationships between the samples based on the covariance or correlation between variables (covariance and correlation are mathematical metrics of the degree to which changes in two variables are associated) (Quinn and Keough 2002). It is important to note that PCA with the covariance matrix will often give significantly different results from PCA with a correlation matrix; however, because PCA is not appropriate for ecological data analysis, the details are not important for this chapter. In brief, the correlation matrix is calculated from standardized data, removing the effects of differences in the variances of the original variables (Quinn and Keough 2002). As such, it is useful for analyzing variables measured in different units or with different scales, whereas the covariance matrix (sensitive to differences in variance) is more appropriate when differences in the variances have biological significance.

Regardless of the choice of association matrix, PCA will produce a final plot with several principal component axes ("eigenvectors") each explaining a certain amount of the variance present in the initial data (quantified by the axis "eigenvalue"). Because the principal component axes are linear regressions through the data cloud, each axis can be defined by a linear combination of the analyzed variables. For example,

$$\mathrm{PC1} = \mathrm{c}_1\mathrm{X}_1 + \mathrm{c}_2\mathrm{X}_2 + \ldots + \mathrm{c}_n\mathrm{X}_n$$

The coefficients c_n are termed "loadings" and quantify the importance of each variable in influencing the sample's position along the principal component axis. In the case of community analysis, they indicate the degree to which a sample's position is influenced by the abundance of each species (X_n). When the loading of a species on a particular principal component is high, variations in its abundance are a strong control on the distribution of samples along that axis.

Two PCA ordinations of Mistaken Point assemblages, one using a covariance matrix and one using a correlation matrix, are shown in Fig. 1.2 with the eigenvalues (given the Greek letter λ) displayed by each principal component axis and the loadings of each species displayed as lines radiating from the origin. The graphs are presented solely to demonstrate the interpretation of ordination plots, even though it is likely that neither result accurately reflects the underlying ecological structure of Mistaken Point communities due to the inappropriate linear assumption inherent to PCA. The units on each axis depend on the scale of original measurements and are not of importance for data interpretation. Note that the amount of variance explained by each principal component is low (no more than 40%) – a result that is typical of ecological data ordination. Explained variance decreases with successive principal components so only the first two, or sometimes three, axes are typically

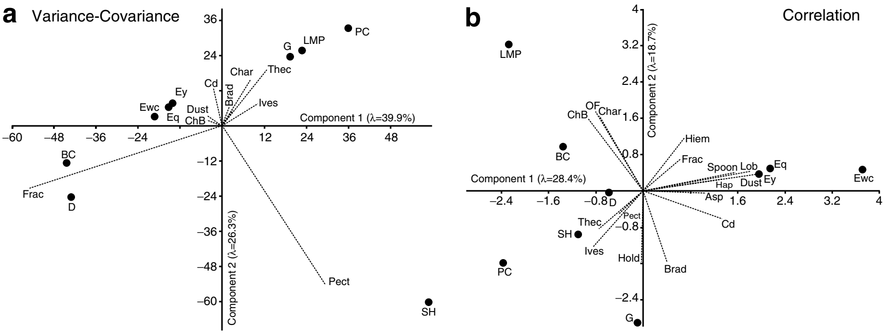

Fig. 1.2 Principal Components Analysis (PCA) ordination of Mistaken Point data (only PC1 and PC2 are shown for simplicity). (**a**) Ordination using the variance-covariance matrix. Samples are indicated by *black circles* and taxon loadings are indicated by *dotted lines* with the length of the *line* indicating the degree of loading on each axis. (**b**) Ordination using the correlation matrix

interpreted in ecological studies, although it is possible that lower order axes may also contain important information. In the covariance analysis (Fig. 1.2a), *Fractofusus* (“spindles”) and *Pectinifrons* (“pectinates”) have high loadings (indicated by longer lines radiating from the origin) on axes 1 and 2, primarily because they account for 57% of the variance in the original data. This indicates that variations in their abundance are the primary controls on the position of samples along those axes – for example, principal component 1 is essentially the abundance of *Fractofusus*, from high abundance at negative PC1 scores to low abundance at positive PC1 scores. Variations in the abundance of rare species (*Aspidella*, lobate discs, *Hapsidophyllas* (“networks”), “holdfast fronds”, etc.) make essentially no contribution to the structure of the ordination plot. In contrast, all species contribute equally to the correlation PCA due to the effects of variable standardization (small variations in the length of the loadings are an artifact of displaying only two axes of a multidimensional diagram), which gives as much weight to the abundant *Fractofusus* as it does to *Hiemalora*, only present on three surfaces and never accounting for more than 0.5% of the community.

As mentioned above, it is likely that neither PCA ordination accurately represents the actual ecological structure of Mistaken Point communities. That error arises because PCA requires the assumption of an underlying linear relationship between variables (Euclidean distance) – a severe problem that results in a distorted representation of the original ecological gradient when applied to community data (Gauch and Whittaker 1972; Minchin 1987). Many studies have documented curved unimodal variations in taxon abundance along ecological gradients (depth, altitude, etc.) and a non-linear relationship between compositional similarity and distance between samples along the gradient, with the rate of decrease in similarity reduced at greater distances (Minchin 1987). Because PCA attempts to fit this nonlinear relationship to an underlying assumption of linearity, the resulting PCA ordination is distorted and likely does not reflect the actual ecological gradients. One common distortion is the characteristic “horseshoe” shape, in which the ends of the gradient curve inwards are depicted as having greater similarity than is actually the case. This contrasts to other ordination techniques that are not based on the assumption of a linear relationship, such as detrended correspondence analysis (discussed in more detail below), which do not exhibit the horseshoe effect. The effect is less pronounced when beta diversity (between-sample species difference) is low, but when multiple gradients are present in the original sample space the distortion is severe and undetectable, and the results cannot be interpreted (Fasham 1977; Minchin 1987). As a result, PCA should not be used for ordination of ecological data.

1.4.2 *Correspondence and Detrended Correspondence Analysis (CA/DCA)*

The recognition of faults in PCA methodology when applied to community data (e.g., Gauch and Whittaker 1972) led to the development of new techniques for ecological ordination. Correspondence analysis, the underlying mechanism for DCA, is an eigenanalysis technique similar to principal components analysis (CA can also be calculated through an iterative process called reciprocal averaging) (Hill and Gauch 1980). Whereas PCA conceptually seeks to rotate axes through the multidimensional cloud of sample points to maximize the variance explained, CA can be thought of as rotating the axes simultaneously through both species and sample points to maximize the correspondence between the two. In matrix algebra terminology, this simultaneous axis rotation means that eigenvectors for species and samples are extracted jointly and allows samples and species to be plotted directly on the same axis instead of separately determining loadings of species onto sample axes. Correspondence analysis performs the same matrix decomposition techniques as PCA but on the original data after data transformation to yield chi-square distances rather than on an association matrix of correlation or covariance between sites (Chardy et al. 1976; Faith et al. 1987). As a result, many features of the output of CA are analogous to those in PCA, only differing due to the use of chi-square distance rather than Euclidean distance (Kenkel and Orloci 1986; Faith et al. 1987). The sum of CA eigenvalues measures the lack of independence between species and samples in the original data set (the overall chi-squared statistic divided by total frequency – called “inertia”); in PCA they measure the total variance in the original data set. As in PCA, successive axes account for progressively less inertia, with the first axis explaining a relatively high proportion of the lack of independence between species and samples.

This chi-square transformation of the original data matrix is a key aspect of CA and is one reason why it is more suitable for ordination of ecological data (Faith et al. 1987). Chi-square distance does not assume a linear relationship between sample dissimilarity and ecological distance, an assumption inherent in PCA that is rarely met by ecological data (Faith et al. 1987; Minchin 1987). As a result, CA correctly ordinates samples from the ends of the environmental gradient along the primary axis, avoiding the "horseshoe effect" common in PCA. The chi-square distance has other limitations, however, and is not robust to variations in the underlying structure of the ecological gradient when taxon abundances deviate from a unimodal normal distribution (Faith et al. 1987), a relatively common feature in natural ecosystems (Oksanen and Minchin 2002). It tends to overweight species with low abundance, exaggerating the distance of samples containing more rare taxa (Minchin 1987), although different software packages contain options for downweighting their importance (Hill and Gauch 1980; Holland et al. 2001; Scarponi and Kowalewski 2004). Chi-square distance also does not reach a constant maximum when samples with no compositional overlap are compared, instead varying depending on the relative occurrence of common and rare species in those samples (Minchin 1987).

Despite the limitations of the chi-square distance, correspondence analysis often performs well at reconstructing the primary ecological gradient (Gauch et al. 1981; Olszewski and West 1997), although there are two potential artifacts of the CA process that may hamper interpretation. First, samples are often compressed at either end of the gradient relative to the middle (Hill and Gauch 1980) because taxon breadths along the gradient are artificially truncated and narrowed at the edge of the study area (Peet et al. 1988). Second, and more significantly, CA does not produce additional axes that are independent of prior axes (although they must not be linearly correlated, they can have a non-linear relationship) and the second axis often appears as a quadratic function of the first (due to projection of a multidimensional dataset into two dimensions) – termed the "arch effect" (Hill and Gauch 1980). Two separate techniques have been devised to correct for these artifacts in CA: detrending to eliminate the arch effect on higher-order axes, and rescaling to reduce compression at the ends of axes. The resulting ordination is called detrended correspondence analysis (DCA) (Hill and Gauch 1980).

Detrending uses a running window to divide the primary axis into segments; all points within a segment are centered by subtracting the mean axis 2 score in that segment from each axis 2 value (Fig. 1.3) (Hill and Gauch 1980). This technique

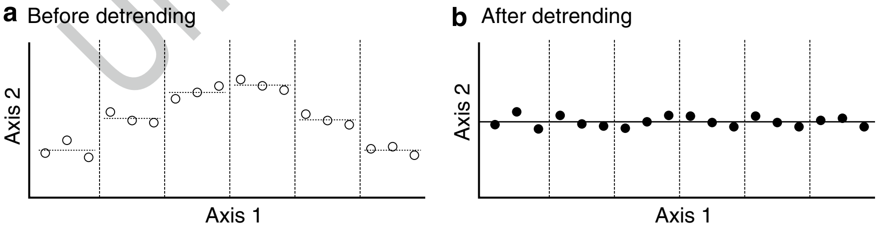

Fig. 1.3 Example illustrating the detrending procedure used in DCA. (**a**) Axis 1 is divided into segments (*dashed lines*) and the original points in each section (*open circles*) are centered by subtracting the mean axis 2 value for that segment (*dotted line*). (**b**) The resulting detrended points are shown by *filled circles*. The actual detrending procedure is more complicated because it uses overlapping running segments and is performed in conjunction with rescaling (Modified from Hill and Gauch (1980))

removes any arch effect (by distorting the ordination plot), regardless of whether the arch is an artifact or a real feature of the ecological structure (Kenkel and Orloci 1986; Minchin 1987; Wartenberg et al. 1987; Olszewski and West 1997). Detrending can also be sensitive to the number of segments in some cases (Jackson and Somers 1991), although the default value (26 segments) often improves the ordination (Peet et al. 1988; Knox 1989).

DCA also uses a mathematical rescaling process to remove the effects of gradient compression near the ends of axes, assumed to be an artifact (Hill and Gauch 1980). It requires that the abundances of all taxa are normally distributed along the gradient with equal variances (i.e., that species composition changes at the same rate along the gradient) and adjusts the position of species along the axis to maintain a constant within-sample variance of species scores at each point (Hill and Gauch 1980; Peet et al. 1988). This has the benefit of rescaling the axis so that distance can be directly interpreted in terms of compositional change (Peet et al. 1988) but has been criticized because of its potentially unjustified assumption of constant rates of taxon turnover (Wartenberg et al. 1987). Species abundances along a gradient are not always Gaussian (Oksanen and Minchin 2002) and different species response curves may have different variances (Holland 2005).

The choice of whether or not to detrend the data has been intensely debated (Minchin 1987; Wartenberg et al. 1987; Peet et al. 1988; Knox 1989) and CA (Cisne and Rabe 1978; Olszewski and West 1997) and DCA (Holland et al. 2001; Scarponi and Kowalewski 2004; Holland 2005; Zuschin et al. 2007; Clapham and James 2008) have both been used in paleoecological data analysis. Despite the arbitrary and perhaps unfounded assumptions inherent to DCA (Wartenberg et al. 1987), the detrending and rescaling process may yield more accurate and interpretable ordination results, especially of long ecological gradients (Peet et al. 1988; Knox 1989) or in the case of rapid ecological turnover (Kenkel and Orloci 1986), although in other cases they produce distortions in the underlying gradient (Kenkel and Orloci 1986; Minchin 1987). In particular, DCA ordination is often twisted so that residual variation is actually spread over axes 2 and 3, resulting in the characteristic wedge shape, where variation on axis 2 is maximal near the mid-point of axis 1, seen in many DCA plots (e.g., Scarponi and Kowalewski 2004; Tomasovych and Siblik 2007; Zuschin et al. 2007; Clapham and James 2008). In analyses of simulated data, CA and DCA typically produce accurate ordinations of the primary axis (Gauch et al. 1981) but the structure along axes 2 and 3 may be distorted, regardless of the effects of detrending and rescaling. In Mistaken Point data, the detrending/rescaling procedure modifies the spacing between points, increasing the distance between E, G, and BC samples while bringing SH and PC samples closer to the main cluster, but does not greatly alter the relative position of the samples (Fig. 1.4). The main difference is in the position of the PC sample, which is likely difficult to ordinate due to its low species count (only three taxa). Overall, CA and DCA both appear to result in a reasonable ordination of Mistaken Point surfaces.

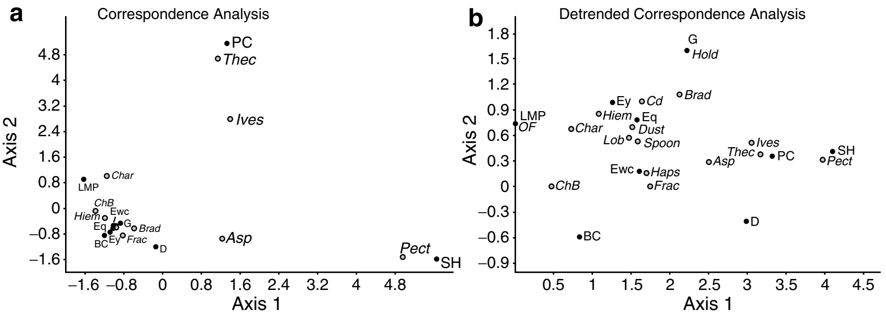

Fig. 1.4 Correspondence Analysis (**a**) and Detrended Correspondence Analysis (**b**) ordinations of Mistaken Point data (only the first two axes are shown for simplicity). Sample positions are shown by *black filled circles* and species positions by *gray filled circles*

1.4.3 Non-Metric Multidimensional Scaling (NMDS)

Eigenanalysis ordination methods such as PCA or DCA are metric techniques that stipulate a proportional relationship between compositional change and distance along the gradient (Kenkel and Orloci 1986; Minchin 1987). As discussed above, ordination results may be distorted when the actual relationship deviates from the idealized assumptions. Another family of techniques compares between-sample similarity using quantitative distance metrics (discussed below) rather than taking an eigenanalysis approach. This section will focus primarily on non-metric multidimensional scaling, the most widespread of these techniques, but it is useful to briefly mention two other distance-based ordination methods: polar ordination and principal coordinates analysis (PCoA).

Polar ordination (also called Bray-Curtis ordination) is so named because it projects all samples onto an axis using their relationship to two selected reference points ("poles"), as calculated by the chosen distance metric (Bray and Curtis 1957; Cisne and Rabe 1978; Beals 1984). In the original formulation of Bray and Curtis (1957), the two data points with the maximum separation were chosen as the reference points; later workers proposed a variety of methods for selecting reference points (Beals 1984). Higher-order axes can be constructed from a matrix of residual distances (i.e., the remaining distance not accounted for by the first axis), producing axes that are perpendicular and uncorrelated to the primary axis (Beals 1984). The comparison of all samples to only two reference points, rather than analyzing all intersample differences, may seem like a limiting constraint but polar ordination appears to perform well at ecological ordination (Gauch and Whittaker 1972; Beals 1984).

Principal Coordinates Analysis (PCoA) attempts to maximize the linear correlation between actual intersample distances (using a user-chosen distance metric) and those distances projected onto a two- or three-dimensional space (Gower 1966). It is sometimes called "metric multidimensional scaling" because it is similar to non-metric multidimensional scaling but uses a parametric (or "metric") linear correlation. Although PCoA is able to use an ecologically-appropriate distance metric, the linear correlation imparts similar distortion to that seen in PCA (Podani and Miklós 2002) and PCoA generally performs poorly at ordinating ecological data (Minchin 1987).

Non-metric multidimensional scaling differs markedly from those metric approaches by assuming only that the relationship between distance and sample dissimilarity is monotonic (i.e., increasing rank-order distance corresponds to increased rank-order dissimilarity) rather than proportional (Kruskal 1964; Fasham 1977; Kenkel and Orloci 1986; Clarke 1993). In this manner, rank-order NMDS is analogous to non-parametric classical statistics. NMDS uses this rank-order agreement between the calculated solution and the original dataset, a value called "stress," to assess the quality of ordination results. Analysis begins with a random configuration of points (or a user-specified configuration, such as the results of another ordination method) and performs an iterative optimization procedure to refine the solution until the stress reaches a minimum (Clarke 1993). The algorithm may reach a local stress minimum rather than identifying the globally best solution (Fasham 1977; Kenkel and Orloci 1986; Minchin 1987; Clarke 1993), but this is rare when the starting configuration is random and can generally be overcome by selecting several random starting points and comparing the final stress to confirm that the optimal solution has been reached (Clarke 1993). The lowest stress indicates the best rank-order agreement between calculated solution and original matrix, with a stress of zero indicating perfect rank-order agreement and stress values less than 0.1 corresponding to a good representation that can reliably be interpreted (Clarke 1993). Higher stress values can still be interpreted, but with greater caution, although stresses greater than 0.2 indicate a poor to nearly random ordering of points (Clarke 1993).

Another fundamental difference of NMDS is the ability to select from a variety of similarity or distance coefficients, whereas PCA is based on Euclidean distance and DCA is constrained to use chi-square distance to calculate sample dissimilarity (Faith et al. 1987). This is a significant benefit because Euclidean (and perhaps also chi-square) distance often perform poorly with ecological data (Faith et al. 1987). A myriad of similarity or distance metrics have been proposed, each with benefits and drawbacks for specific data types (Faith et al. 1987; Shi 1993). Studies with various simulated ecological gradients have suggested that the Bray-Curtis, Kulczynski, and Relativized Manhattan (i.e., each species standardized to percent abundance within a sample and normalized to its maximum abundance in any sample) coefficients produce the closest rank-order and linear correlation between compositional dissimilarity and ecological distance along the gradient under most conditions (Faith et al. 1987). Although some coefficients perform better than others with simulated ecological data, there is no objective guide to choosing a distance metric (see Faith et al. (1987) for a detailed review of several distance measures). The Bray-Curtis coefficient was designed for interpretation of ecological data (Bray and Curtis 1957) and performs well on simulated datasets (Faith et al. 1987). Because it enjoys widespread use and is intuitive – simply calculate the absolute value of the difference in abundance of a species (X_i–X_j) between two samples divided by the total abundance of that species in the two samples (X_i+X_j), and sum the results for all species – it is a good choice for use in NMDS. It is by no means the only possibility, as other metrics perform well and yield robust solutions, but Bray-Curtis similarity will be used in the NMDS ordination in this chapter. Other distance metrics, such as the Kulczynski coefficient (Faith et al. 1987), yield very similar solutions.

In contrast to the eigenanalysis methods, NMDS simply arranges the points in *n*-dimensional space (where *n* is a number of dimensions, usually two or three, chosen by the user) rather than extracting multiple orthogonal axes that explain decreasing amounts of variance (Clarke 1993). This has two major implications for viewing and interpreting NMDS ordinations. First, axes in an NMDS ordination have no specific meaning and the data plot can be rotated, translated, or scaled without altering the results, as long as the relative position of the points remains unchanged (Clarke 1993). For example, the ordination is unchanged after rotating points so that the direction of maximum variability is horizontal to facilitate visual comparison with DCA or PCA plots (e.g., Tomasovych and Siblik 2007). Second, as the NMDS algorithm specifically arranges the samples to fit in *n*-dimensional space, the best solution for two dimensions will not necessarily be a two-dimensional projection of the best solution for three dimensions (Clarke 1993). This contrasts with eigenanalysis methods, where many eigenvectors are calculated and a plot of axis 1 vs. axis 2 will simply be a two-dimensional projection of the multivariate data cloud. Although it is rare to require more than three dimensions, how should the number of dimensions be chosen? The iteration will yield a better solution (i.e., lower stress) as additional dimensions are included because the original dataset is multivariate, until the number of dimensions is one less than the number of samples. However, a plot of stress against the dimensionality (called a "scree plot," available in some software packages) will often exhibit a distinct break in slope, providing a rough guideline for the minimum number of dimensions that should be analyzed (Kruskal 1964). It is also reasonable to interpret one extra dimension if the ordination results in additional structure (Kruskal 1964). One caveat: the resulting ordination plot may be distorted when the specified dimensionality is greater than the actual number of dimensions in the original environmental gradient (Austin 1976).

NMDS was adopted slowly due to its intensive computational requirements but the calculations are now routine and it is used widely in paleoecological analysis (Olszewski and West 1997; Bonuso et al. 2002; Clapham et al. 2003; Dominici and Kowalke 2007; Tomasovych and Siblik 2007; Zuschin et al. 2007). Many studies have shown that NMDS produces more accurate ordinations than CA/DCA, for simulated ecological data at least (Fasham 1977; Kenkel and Orloci 1986; Minchin 1987), although each technique has individual strengths and weaknesses. In particular, NMDS is less susceptible to distortion when beta diversity is high and when there are multiple underlying gradients with different beta diversities (Fasham 1977). The non-metric technique may be less adept than the eigenanalysis techniques at delineating clusters within the data, but may perform better at optimizing relative distance between samples even when there are clusters (Kenkel and Orloci 1986). As discussed more below, NMDS may be of reduced use when sample size is small;

in analyses with fewer than 30 samples, use of rank-order may result in significant loss of information (Anderson 1971). Overall, NMDS is a robust ordination technique that performs excellently with ecological data.

Figure 1.5 shows NMDS ordination plots for Mistaken Point data using the Bray-Curtis similarity coefficient, based on 2D and 3D dimensionality. As discussed above, the Bray-Curtis coefficient was designed for comparison of ecological samples and produces good results with simulated ecological data (Faith et al. 1987). Three different low-stress two-dimensional solutions (arbitrarily chosen from many) yield virtually the same arrangement of points, differing slightly in the position of the three E surface subsamples (Fig. 1.5a–c). In contrast, the three randomly-selected 3D solutions are highly erratic (the substantial difference in arrangement does not result from only viewing two dimensions) (Fig. 1.5d–f). The stress value is lower, but recall that stress is always reduced when extra dimensions are added (Kruskal 1964). The unpredictable performance of the 3D solution may indicate that the original environmental gradient was two-dimensional (Austin 1976), but could instead result from the small number of samples and low species richness within each sample. Replicate random analyses did not converge upon a repeated solution of lowest stress, suggesting that there was no single configuration of points that best fit the criteria, likely because there were few samples with few shared taxa.

1.5 Comparison and Interpretation of Results

These results raise an obvious question – which ordination yields the “correct” ecological representation of Mistaken Point communities? As with all natural datasets, there is no objective way of testing the accuracy of the ordination results because the positions of samples along the underlying environmental gradients are unknown. Analyses of simulated ecological data suggest that DCA and NMDS should yield the most accurate and robust results (Fasham 1977; Kenkel and Orloci 1986; Minchin 1987) and, with the exception of the placement of the PC surface, they produce very similar ordination of the data (Fig. 1.6). Note that the NMDS ordination has been rotated from the plots in Fig. 1.5 to highlight the similarity with the DCA plot, and recall that rotations such as this do not alter the results of the ordination. Two primary gradients can be observed: one from samples with abundant *Fractofusus* to samples with no *Fractofusus* and also containing more *Charniodiscus* and *Bradgatia* (vertical axis) and one from samples with abundant *Pectinifrons* to samples with abundant *Charnia* and “*Charnia*” B (horizontal axis). Placement of the PC surface is erratic with different methods, almost certainly because that assemblage only contains three taxa (*Thectardis*, *Ivesheadia*, and *Charnia*). DCA ordination places it close to the SH surface because of their shared abundance of *Ivesheadia*, but its position in the NMDS plot closer to the frond-rich LMP community (with *Charnia*), E surface (with *Thectardis*), and G surface seems intuitively more reasonable.

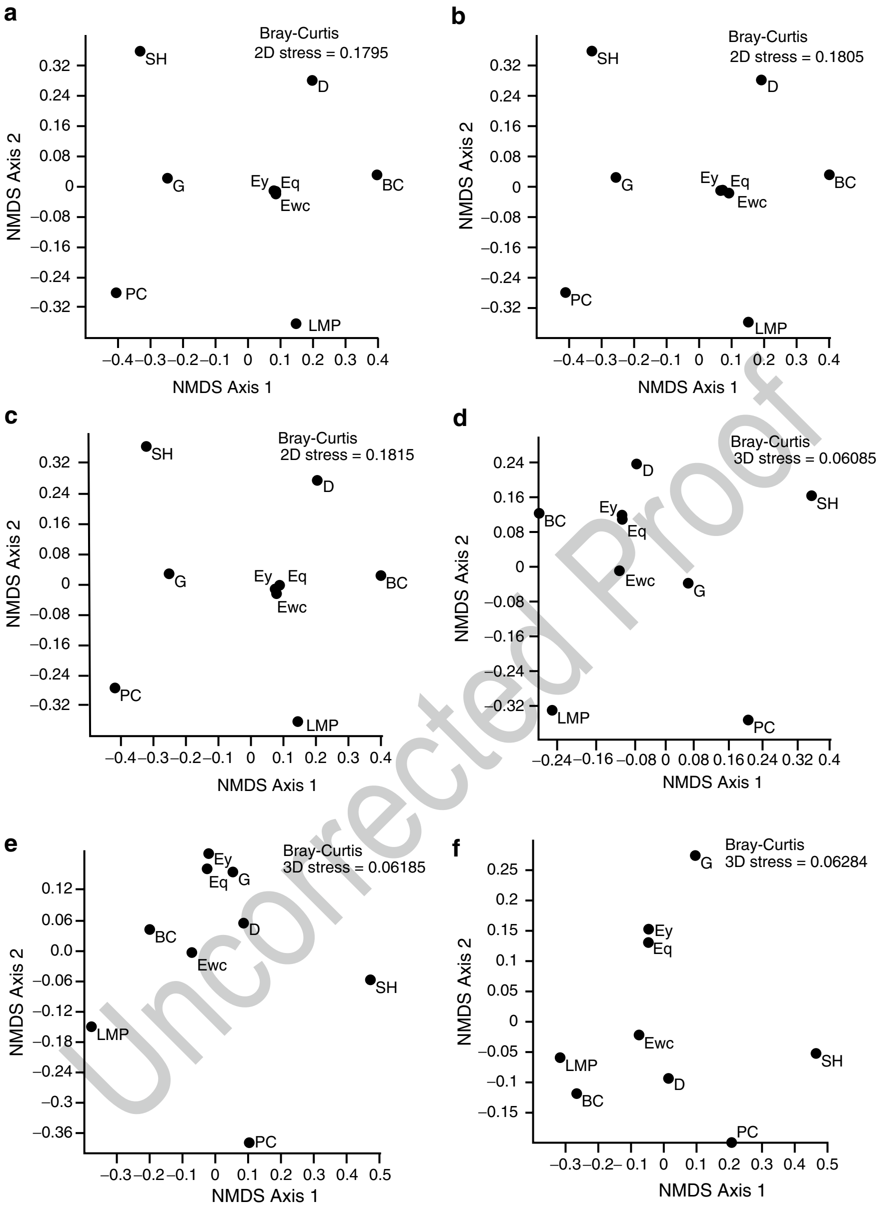

Fig. 1.5 Non-Metric Multidimensional Scaling ordinations of Mistaken Point data. (**a**–**c**) Three results of a two-dimensional ordination (arbitrarily chosen) using the Bray-Curtis coefficient. (**d**–**f**) Three random results of a three-dimensional ordination using the Bray-Curtis coefficient. Note the instability of the three-dimensional solution compared to the stable two-dimensional results. As discussed in the text, this may indicate that the underlying ecological gradient is two-dimensional or may be a limitation of the small dataset

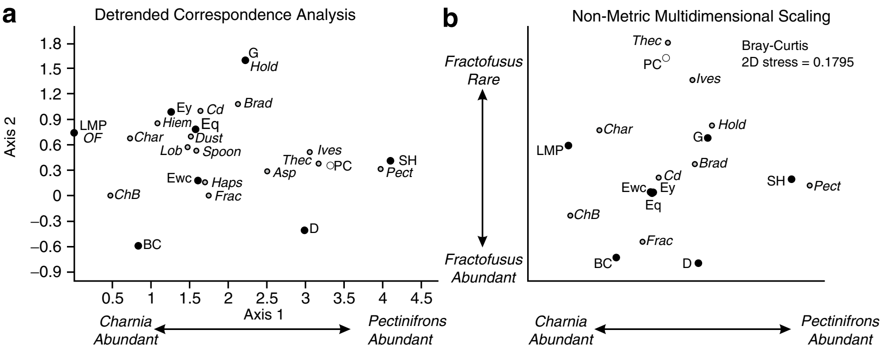

Fig. 1.6 Comparison of DCA and NMDS ordinations of Mistaken Point data, showing interpreted taxonomic gradients responsible for the position of samples. (**a**) Detrended Correspondence Analysis, with samples shown by *black filled circles* (except for the PC surface, which is shown by an *open circle*) and species by *gray filled circles*. (**b**) Non-Metric Multidimensional Scaling (two dimensions, Bray-Curtis coefficient) rotated to align with the DCA results. Samples shown by *black filled circles* except for the PC surface (*open circle*). The positions of selected abundant species are shown by *gray filled circles*; PAST does not ordinate species so their positions were obtained from NMDS ordination in R

The next step in an indirect gradient analysis such as DCA or NMDS is inferring the environmental, evolutionary, ecological, or other gradients that may have been responsible for generating the taxon gradients identified above. This is most often accomplished qualitatively by comparing other lines of evidence, for example from sedimentological interpretation, to the results of the ordination plot, although it is possible to quantitatively include independent lines of evidence using linear regression (e.g., Clapham et al. 2003). Although the interpretation is less important than the methodology for the purposes of this chapter, it may be instructive to briefly consider possible interpretations. Because Mistaken Point assemblages span >10 Myr of time, there is the possibility that evolutionary changes are an important control; thus, ordination position may be related to stratigraphic position. Stratigraphic position is also an indicator of paleoenvironment at Mistaken Point, shallowing upward from basin plain in the lowest samples (PC, BC) to toe-of-slope deposits in the Mistaken Point Formation (LMP, D, E, G) to lower slope in the uppermost sample (SH) (Wood et al. 2003). There is a broad but weak correspondence between axis 1 position (DCA) and stratigraphic position (Fig. 1.6), excluding the PC surface, likely because the BC and LMP samples predate the first appearance of *Pectinifrons* whereas the SH sample postdates the last appearance of *Fractofusus*. This implies that evolutionary changes in the Mistaken Point biota were one of the probable controls on community composition, as expected for samples spanning such a long time interval. Clapham et al. (2003) argued that each fossil surface may also reflect a snapshot of a different stage within an ecological succession trend from *Pectinifrons* and *Fractofusus*-dominated early successional communities to frond-dominated late successional communities. Mistaken Point communities have the temporal resolution to preserve a near-instantaneous snapshot of the community at a point in time; although the snapshots occur randomly throughout the stratigraphic section. Ordination techniques are ideally suited to take a jumbled series of snapshots and organize them into a gradient, whether that gradient is related to ecological succession or some other factor. The taxonomic gradient inferred for ecological succession can be observed along the vertical axis 2 in the DCA and NMDS plots, although it is modified by evolutionary changes in the biota along axis 1 (Fig. 1.6). Other ecological characteristics are at least consistent with expectations of a succession model (Clapham et al. 2003) but more surfaces may be needed to strengthen or falsify the link to ecological succession.

1.6 Conclusion

Quantitative ordination techniques are ideally suited for analyzing multivariate community ecology data. Principal Components Analysis (PCA) uses Euclidean distance to relate species dissimilarity to distance along the gradient, requiring the inappropriate assumption of a linear species response. Because that assumption is almost always invalid for community data, PCA ordinations are distorted, often severely, into a horseshoe shape or worse, usually making them uninterpretable. Although PCA is not suited for ecological ordination, it is still quite appropriate for data in morphometric or paleobiogeographic analyses, for example, where the dissimilarity-distance response is closer to linear (Chaps. 2 and 3). Detrended Correspondence Analysis (DCA) and Non-Metric Multidimensional Scaling (NMDS) can both yield accurate ecological ordinations. NMDS performs slightly better on simulated ecological data but DCA may be better at extracting the primary gradient. Detrending often results in a more accurate ordination of simulated data than in unaltered Correspondence Analysis (CA), but has also been criticized for its ad hoc methodology and assumptions. Thus, the best approach for analyzing your own paleocommunity data depends on the ultimate goal of the study. If you wish to identify the primary environmental gradient and reconstruct sample and species placement along that gradient, DCA is an excellent approach. If you are instead interested in the overall relationship among samples, NMDS may be the best option. Since DCA and NMDS ordinations have very complementary strengths, the most robust approach is to apply both methods; the overall result should be similar and gradients or clusters resolved by both DCA and NMDS are likely to be a good reflection of the real underlying ecological structure.

Acknowledgments Rodrigo Sala and Marc Laflamme assisted with data collection. Fieldwork in 2000 and 2001 in the Mistaken Point Ecological Reserve was carried out under Scientific Research Permits granted by the Parks and Natural Areas Division, Department of Tourism, Culture, and Recreation, Government of Newfoundland and Labrador. The original data collection was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant (to Guy M. Narbonne) and by an NSERC postgraduate scholarship (to Clapham). Thoughtful reviews from John Huntley, Richard Krause, and Amelinda Webb helped improve many aspects of this contribution.

References

Anderson AJB (1971) Ordination methods in ecology. J Ecol 59:713–726

Austin MP (1976) Performance of four ordination techniques assuming three different non-linear species response models. Vegetatio 33:43–49

Beals EW (1984) Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. Adv Ecol Res 14:1–55

Bonuso N, Newton CR, Brower JC et al (2002) Statistical testing of community patterns: uppermost Hamilton Group, Middle Devonian (New York State: USA). Palaeogeogr Palaeoclimatol Palaeoecol 185:1–24

Botquelen A, Gourvennec R, Loi A et al (2006) Replacements of benthic associations in a sequence stratigraphic framework, examples from the Upper Ordovician of Sardinia and Lower Devonian of the Massif Armoricain. Palaeogeogr Palaeoclimatol Palaeoecol 239:286–310

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27:325–349

Chardy P, Glemarec M, Laurec A (1976) Application of inertia methods to benthic marine ecology: practical implications of the basic options. Estuar Coast Mar Sci 4:179–205

Cisne JL, Rabe BD (1978) Coenocorrelation: gradient analysis of fossil communities and its applications in stratigraphy. Lethaia 11:341–364

Clapham ME, James NP (2008) Paleoecology of Early-Middle Permian marine communities in eastern Australia: response to global climate change in the aftermath of the late Paleozoic ice age. Palaios 23:738–750

Clapham ME, Narbonne GM, Gehling JG (2003) Paleoecology of the oldest-known animal communities: Ediacaran assemblages at Mistaken Point, Newfoundland. Paleobiology 29:527–544

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Dominici S, Kowalke T (2007) Depositional dynamics and the record of ecosystem stability: early Eocene faunal gradients in the Pyrenean foreland, Spain. Palaios 22:268–284

Faith DP, Minchin PR, Belbin L (1987) Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69:57–68

Fasham MJR (1977) A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. Ecology 58:551–561

Gauch HG, Whittaker RH (1972) Comparison of ordination techniques. Ecology 53:868–875

Gauch HG, Whittaker RH, Singer SB (1981) A comparative study of nonmetric ordinations. J Ecol 69:135–152

Goodall DW (1954) Objective methods for the classification of vegetation. Aust J Bot 2:304–324

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol Electronica 4:1–9

Hill MO, Gauch HG (1980) Detrended correspondence analysis: an improved ordination technique. Vegetatio 42:47–58

Holland SM (2005) The signatures of patches and gradients in ecological ordinations. Palaios 20:573–580

Holland SM, Miller AI, Meyer DL et al (2001) The detection and importance of subtle biofacies within a single lithofacies: the Upper Ordovician Kope Formation of the Cincinnati, Ohio region. Palaios 16:205–217

Jackson DA, Somers KM (1991) Putting things in order: the ups and downs of detrended correspondence analysis. Am Nat 137:704–712

James FC, McCulloch CE (1990) Multivariate analysis in ecology and systematics: panacea or Pandora's box. Annu Rev Ecol Syst 21:129–166

Jongman RHG, ter Braak CJF, Van Tongeren OFR (1995) Data analysis in community and landscape ecology. Cambridge University Press, Cambridge

Kenkel NC, Orloci L (1986) Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology 67:919–928

Knox RG (1989) Effects of detrending and rescaling on correspondence analysis: solution stability and accuracy. Vegetatio 83:129–136

Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29:1–27

McKinney FK, Hageman SJ (2006) Paleozoic to modern marine ecological shift displayed in the northern Adriatic Sea. Geology 34:881–884

Minchin PR (1987) An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69:89–107

Oksanen J, Minchin PR (2002) Continuum theory revisited: what shape are species responses along ecological gradients? Ecol Model 157:119–129

Olszewski TD, West RR (1997) Influence of transportation and time-averaging in fossil assemblages from the Pennsylvanian of Oklahoma. Lethaia 30:315–329

Palmer MR (1993) Putting things in even better order: the advantages of canonical correspondence analysis. Ecology 74:2215–2230

Peet RK, Knox RG, Case JS et al (1988) Putting things in order: the advantages of detrended correspondence analysis. Am Nat 131:924–934

Podani J, Miklós I (2002) Resemblance coefficients and the horseshoe effect in principal coordinates analysis. Ecology 83:3331–3343

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge

Rodriguez J (2004) Stability in Pleistocene Mediterranean mammalian communities. Palaeogeogr Palaeoclimatol Palaeoecol 207:1–22

Scarponi D, Kowalewski M (2004) Stratigraphic paleoecology: bathymetric signatures and sequence overprint of mollusk associations from upper Quaternary sequences of the Po Plain, Italy. Geology 32:989–992

Shi GR (1993) Multivariate data analysis in palaeoecology and palaeobiogeography – a review. Palaeogeogr Palaeoclimatol Palaeoecol 105:199–234

Tomasovych A, Siblik M (2007) Evaluating compositional turnover of brachiopod communities during the end-Triassic mass extinction (Northern Calcareous Alps): removal of dominant groups, recovery and community reassembly. Palaeogeogr Palaeoclimatol Palaeoecol 244:170–200

Wartenberg D, Ferson S, Rohlf FJ (1987) Putting things in order: a critique of detrended correspondence analysis. Am Nat 129:434–448

Wood DA, Dalrymple RW, Narbonne GM et al (2003) Paleoenvironmental analysis of the late Neoproterozoic Mistaken Point and Trepassey formations, southeastern Newfoundland. Can J Earth Sci 40:1375–1391

Zuschin M, Harzhauser M, Mandic O (2007) The stratigraphic and sedimentologic framework of fine-scale faunal replacements in the middle Miocene of the Vienna basin (Austria). Palaios 22:285–295