Ecology Letters, (2006) **9**: 683–693 doi: 10.1111/j.1461-0248.2006.00926.x

LETTER

Multivariate dispersion as a measure of beta diversity

Marti J. Anderson[1]*, Kari E. Ellingsen[2]† and Brian H. McArdle[3]

[1]Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

[2]Marine Biodiversity, Department of Biology, University of Oslo, PO Box 1066, Blindern, 0316 Oslo, Norway

[3]Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

**Correspondence: E-mail: mja@stat.auckland.ac.nz*

†Present address: Norwegian Institute for Nature Research (NINA), Polarmiljøsenteret, 9296 Tromsø, Norway.

Abstract

Beta diversity can be defined as the variability in species composition among sampling units for a given area. We propose that it can be measured as the average dissimilarity from individual observation units to their group centroid in multivariate space, using an appropriate dissimilarity measure. Differences in beta diversity among different areas or groups of samples can be tested using this approach. The choice of transformation and dissimilarity measure has important consequences for interpreting results. For kelp holdfast assemblages from New Zealand, variation in species composition was greater in smaller holdfasts, while variation in relative abundances was greater in larger holdasts. Variation in community structure of Norwegian continental shelf macrobenthic fauna increased with increases in environmental heterogeneity, regardless of the measure used. We propose a new dissimilarity measure which allows the relative weight placed on changes in composition vs. abundance to be specified explicitly.

Keywords

Beta diversity, community structure, dissimilarity measures, kelp holdfasts, macrobenthos, multivariate dispersion, species abundance, species composition, transformations, variability.

Ecology Letters (2006) 9: 683–693

INTRODUCTION

Whittaker (1960) originally proposed partitioning diversity into alpha, beta, and gamma components to characterize different aspects or levels of diversity. Alpha diversity (α) is commonly measured as the number of species in a single sampling unit, while gamma diversity (γ) is generally defined as the overall number of species within a defined geographical area. Beta diversity can be measured in many different ways (Koleff *et al.* 2003; Magurran 2004). Whittaker's (1960, 1972) original measure [$\beta_W = \gamma/\bar{\alpha}$ or $\beta_W = (\gamma/\bar{\alpha}) - 1$], the proportion by which a given area is richer than the average of samples within it, has been one of the most frequently used measures of beta diversity (Koleff *et al.* 2003).

We consider that beta diversity can be measured as the variability in species composition among sampling units for a given area at a given spatial scale. We propose that beta diversity for a group of units sampled from a given area can be measured as the average distance (or dissimilarity) from an individual unit to the group centroid, using an appropriate dissimilarity measure. The centroid must be defined in the principal coordinate space of the dissimilarity measure chosen. This concept of beta diversity is quite flexible, because it can be based on any chosen ecologically meaningful dissimilarity measure. It also has the added advantage over Whittaker's original measure in that it can be used to test for differences in beta diversity among areas or groups, through a multivariate test for homogeneity in dispersions (Anderson 2006). However, this approach does require the definition of what is meant by 'variability in community structure' (dictated by the choice of transformation and dissimilarity measure) to be carefully articulated.

The most widely used ecological measures of compositional dissimilarity include the classic measures described by Jaccard (1900) and Sørensen (1948) (e.g. Chao *et al.* 2005). Beta diversity has also been measured using dissimilarity measures which include relative abundance information, such as the Bray–Curtis measure (Bray & Curtis 1957) (e.g. Ellingsen & Gray 2002; see also Magurran 2004; Olszewski 2004). However, an important ecological issue is to understand how much of the dissimilarity is driven by compositional difference and how much is driven by differences in relative abundance, which is difficult to ascertain for the Bray–Curtis and related measures. We propose a new dissimilarity measure which weights an

order-of-magnitude change in abundance the same as a change in species composition.

Our purposes here are (i) to demonstrate how multivariate dispersion can be used as a measure of beta diversity; (ii) to demonstrate the test for differences in dispersion to investigate differences in beta diversity among areas or groups of sampling units; (iii) to outline the potential pitfalls in extending the concept of beta diversity to include abundances; and (iv) to propose a new dissimilarity measure.

We illustrate these concepts with two ecological examples. The first example examines the hypothesis of a relationship between beta diversity of invertebrate assemblages inhabiting holdfasts of the kelp *Ecklonia radiata* (C. Agardh) J. Agardh and increases in the size (volume) of the holdfast habitat (Anderson *et al.* 2005a). The second example examines the hypothesis that compositional heterogeneity (beta diversity) is related to environmental heterogeneity in soft-sediment benthic assemblages from the Norwegian continental shelf (Ellingsen & Gray 2002).

A TEST FOR HOMOGENEITY

Individual dissimilarities between pairs of sampling units do not alone constitute a measure of beta diversity over a large area. Although the average of all dissimilarities between pairs of samples within an area can be calculated (e.g. Ellingsen & Gray 2002), these individual values are not independent of one another, which prevents a direct statistical test. Whittaker's β_W measures beta diversity for a given area, but does not test for differences among different areas. We propose that a test of the null hypothesis of no difference in beta diversity among two or more areas can be obtained by implementing the test for homogeneity of multivariate dispersions (Anderson 2006) on the basis of an appropriate measure of dissimilarity. Note that what concerns us here is the structure within groups – the test says nothing about potential differences in location among groups in multivariate space, for which other tests are available (e.g. Clarke 1993; Anderson 2001; McArdle & Anderson 2001).

The test for homogeneity of multivariate dispersions (Anderson 2006) is a multivariate analogue to Levene's (1960) test and can be based on any dissimilarity measure of choice. In essence, one calculates an *F*-statistic to compare the average distance of observation units to their group centroid (or spatial median), defined in the space identified by the chosen dissimilarity measure. A *P*-value is then obtained by permuting appropriate residuals: either least-squares residuals (in the case of centroids) or least-absolute-deviation residuals (in the case of spatial medians). Here, we shall consider the test which uses distances to centroids and obtains *P*-values using permutation of least-squares residuals, which was found to be quite robust and powerful under various simulations (Anderson 2006).

An important complexity concerns the calculation of centroids for measures which are not Euclidean embeddable (Gower & Legendre 1986). Many commonly used dissimilarity measures (Bray–Curtis, Jaccard, etc.) are not embeddable in Euclidean space [i.e. they produce negative eigenvalues in a principal coordinate analysis (PCO), see pp. 432–438 in Legendre & Legendre 1998 for details]. For analyses based on Euclidean distance, the centroid is simply the arithmetic average for each variable. However, if the multi-species space is to be defined using a dissimilarity measure such as Jaccard's (1900), then the centroid cannot be calculated in this way and is not directly obtainable in the space of the original species variables. In this case, what is necessary before proceeding is to place the observations into a Euclidean space which preserves the original dissimilarities among them. This is achieved using PCO (Gower 1966). Specifically, the Euclidean distance between two points in the space defined by the principal coordinate axes is equivalent to the original dissimilarity between those two points using the chosen dissimilarity measure on the original variables (Gower 1966; Legendre & Legendre 1998). The result holds for non-Euclidean embeddable dissimilarities as well. For more details on the test and the underlying calculations, see Anderson (2006).

DESCRIPTION OF SOME DISSIMILARITY MEASURES

The Jaccard (1900) and Sørensen (1948) measures were originally described as similarities, but are given here in terms of their complement as dissimilarities. The Jaccard dissimilarity between two sampling units is

$$d_J = (b + c)/(a + b + c), \quad (1)$$

where a is the number of species shared, b the number of species in unit 1 that do not occur in unit 2, and c the number of species in unit 2 that do not occur in unit 1. Thus, it is the proportion of unshared species out of the total number of species recorded in the two units. The Sørensen measure, also described by Dice (1945), is

$$d_S = (b + c)/(2a + b + c). \quad (2)$$

It is monotonically related to the Jaccard measure (with $d_S < d_J$); however, the Sørensen measure gives double value to shared species. The biggest difference between (1) and (2) will occur when a is in the range from $(b + c)$ to $2(b + c)$. Moreover, when Whittaker's measure (β_W) is calculated between a single pair of units, then $\beta_W = (2–1/d_S)$ (Vellend 2001). Although there are many other measures based on presence/absence data (Legendre & Legendre 1998), d_J and d_S are favoured in ecology because they exclude joint absences and can be interpreted in a probabilistic framework (Chao *et al.* 2005).

Many dissimilarity measures also incorporate relative abundance information (Legendre & Legendre 1998). We shall focus on those considered to be extensions of either the Jaccard or Sørensen measures. A modified version of the Sørensen index (Magurran 2004), described by Bray & Curtis (1957) and also earlier by Steinhaus (Motyka 1947) and Odum (1950), is now one of the most commonly used measures in ecology (Clarke & Warwick 2001):

$$d_{\mathrm{BC}} = \frac{\sum_{k=1}^{p} |x_{1k} - x_{2k}|}{\sum_{k=1}^{p} (x_{1k} + x_{2k})}, \tag{3}$$

where x_{1k} is the abundance of species k in sampling unit 1, x_{2k} the abundance of species k in sampling unit 2, and p the total number of species recorded across both units. If the variables differ in their scales, then a transformation of abundances (e.g. to square-roots or fourth-roots) is often applied first (Clarke & Green 1988). The Bray–Curtis measure varies from 0 to 1 and if data are reduced to presence/absence, then $d_{\mathrm{BC}} = d_{\mathrm{S}}$.

Gower (1971, 1987) defined a flexible dissimilarity measure as follows:

$$d_{\mathrm{G}} = \frac{\sum_{k=1}^{p} w_k |x_{1k} - x_{2k}| / R_k}{\sum_{k=1}^{p} w_k}, \tag{4}$$

where R_k is the range of the kth species and w_k is an optional weight that can be given to each species (having a default value of $w_k = 1$). Dividing differences by the range (or SD) is optional; it is intended to eliminate differences in scale among variables. A version of Gower's measure which excludes joint absences can be obtained by setting $w_k = 0$, whenever $x_{1k} = x_{2k} = 0$ and $w_k = 1$ elsewhere. When calculated on presence/absence data, d_{G} excluding double zeros $= d_{\mathrm{J}}$.

The Jaccard measure is also interpretable as the probability that two species, one drawn at random from each sample, will not be shared. Taking this probabilistic approach, an abundance-based Jaccard measure was proposed by Chao *et al.* (2005), as

$$d_{\mathrm{ChJ}} = 1 - UV/(U + V - UV), \tag{5}$$

where U is the sum of the proportional abundances of species in sampling unit 1 that are shared with unit 2, and V is the sum of the proportional abundances of species in sampling unit 2 that are shared with unit 1. That is,

$$U = \sum_{k=1}^{p} w_k \frac{x_{1k}}{x_{1+}},$$

where x_{1+} is the sum of the abundances for all species in sampling unit 1 and $w_k = 1$ if species k occurs in both unit 1 and unit 2 (i.e. is shared) or else $w_k = 0$. The quantity V is defined similarly, but for sampling unit 2. The measure d_{ChJ} is interpretable as the probability that two *individuals*, one drawn at random from each sample, will not belong to a shared species. It takes values from 0 to 1 and $d_{\mathrm{ChJ}} = d_{\mathrm{J}}$ for presence/absence data. Chao *et al.*'s (2005) further contribution includes estimation of unseen shared species in order to reduce the sampling bias caused by not being able to census the entire community. This bias-corrected version (which we shall not describe here, see Chao *et al.* 2005) we shall refer to as d_{ChJB}.

POTENTIAL PITFALLS IN EXTENDING THE CONCEPT OF BETA DIVERSITY TO INCLUDE RELATIVE ABUNDANCE INFORMATION

Many of the dissimilarity measures that ecologists commonly use on abundance data include intrinsic standardizations by row sums or column sums (e.g. Bray–Curtis, Canberra, chi-squared, CY dissimilarity, Hellinger, Orlóci's chord, Kulczynski, etc.). Indeed, Legendre & Gallagher (2001) have demonstrated how many of these distance measures can be obtained directly by calculating Euclidean distances on variables that have been standardized or transformed in some way by row sums, column sums, or both. Gower's measure standardizes by the range for each variable and probabilistic measures, such as d_{ChJ}, also have intrinsic standardizations. Generally, such standardizations affect patterns of relative dispersion in ways that cannot easily be either predicted or interpreted by reference to the original variables.

Measures that do not have intrinsic standardizations, such as the Manhattan measure or the Euclidean distance measure, are more transparent in that they model raw abundance information directly. However, these measures: (i) do not exclude joint absences, and (ii) are very sensitive to differences in the scale of the variables (Legendre & Legendre 1998; Clarke & Warwick 2001). Furthermore, measuring variability in the counts of species abundances, even for univariate data, may not be best understood by a statistical measure of either variance or standard deviation on raw abundance values, *per se* (e.g. Gaston & McArdle 1994). In particular, counts of species abundances have intrinsic mean–variance relationships: increases in mean values are generally accompanied by increases in variance (dispersion) too. This is the basis of Taylor's power law (Taylor 1961; see also McArdle *et al.* 1990; Gaston & McArdle 1994). To examine heterogeneity in abundances, we should therefore ultimately seek to distinguish between a fundamental difference in variability (such as a shift in underlying mean–variance relationships) vs. a difference in variances caused by a simple change in mean values.

A NEW PROPOSED DISSIMILARITY MEASURE

The following modification to Gower's (1971, 1987) dissimilarity measure can be used which explicitly weights

an order-of-magnitude change in abundance the same as a change in species composition (from 0 to 1). First, values of x are transformed to a new multiplicative scale which makes a simple allowance for zeros, as follows: let $x' = \log_{10}(x) + 1$, unless $x = 0$, in which case $x' = 0$. Thus, for $x = \{0, 1, 10, 100, 1000\}$, $x' = \{0, 1, 2, 3, 4\}$. Then, a modified Gower dissimilarity measure is

$$d_{\mathrm{MG}} = \frac{\sum_{k=1}^{p} w_k |x'_{1k} - x'_{2k}|}{\sum_{k=1}^{p} w_k}, \tag{6}$$

where the weights w_k are used to provide the desired exclusion of joint absences by setting $w_k = 0$, whenever $x_{1k} = x_{2k} = 0$ and $w_k = 1$ elsewhere. Note that d_{MG} will generally be non-Euclidean and also does not include any intrinsic standardization by the range, as was specified for d_{G}.

Although not a unitless measure, this dissimilarity measure has the advantage of being directly interpretable as the average change in orders of magnitude per species between two sampling units. Note that the above transformation is *not* the same as $\log(x + 1)$. The base of the logarithm used for the transformation can also be altered. For example, using $\log_2$ would weight a compositional change equal to a doubling in abundance (i.e. for $x = \{0, 1, 2, 4, 8\}$, transformed values would be $x' = \{0, 1, 2, 3, 4\}$). The emphasis placed on compositional change vs. changes in abundance is specified directly by the choice in the base of the logarithm. For presence/absence data, $d_{\mathrm{MG}} = d_{\mathrm{J}}$.

Unlike d_{BC} or d_{ChJ}, d_{MG} does not have an upper bound. Although Clarke & Warwick (2001) suggested that an ecological dissimilarity measure should reach a maximum when there are no species in common, a common criticism of d_{BC} is its lack of discrimination near its upper bound (Cao *et al.* 1997) and its erratic behaviour for sparse data (Clarke *et al.* 2006). However, d_{MG} does not suffer from these problems.

The proposed measure can easily be partitioned into a component driven by compositional differences (d_{J}) and a component driven by order-of-magnitude changes in abundance ($d_{\mathrm{MG}} - d_{\mathrm{J}}$). Furthermore, if raw values of x are used instead of the transformed values x', then the measure would simply be the Manhattan (city-block) measure, but with weights to exclude double zeros:

$$d_{\mathrm{Manx}} = \frac{\sum_{k=1}^{p} w_k |x_{1k} - x_{2k}|}{\sum_{k=1}^{p} w_k}. \tag{7}$$

Legendre & Legendre (1998) described this measure as a modification of Czekanowski's (1909) mean character difference to exclude double zeros. Weights can also be introduced in this manner to exclude double zeros for the Euclidean distance measure (d_{Eucx}):

$$d_{\mathrm{Eucx}} = \frac{\sqrt{\sum_{k=1}^{p} w_k (x_{1k} - x_{2k})^2}}{\sum_{k=1}^{p} w_k}. \tag{8}$$

Thus, we can articulate a series of distance measures that provide a continuum in emphasis from pure species composition through to (virtually) pure relative abundance information, as follows: d_{J}, d_{MG}, d_{MG2}, d_{Manx}, d_{Eucx}, where d_{MG} uses $\log_{10}$ and d_{MG2} uses $\log_2$ for the transformation (and clearly other choices for the base of the log can also be used).

ECOLOGICAL EXAMPLES

Beta diversity in New Zealand kelp holdfast assemblages

Eighty holdfasts of the kelp, *E. radiata*, were collected from subtidal forest stands (*c.* 11–15 m depth) along the north-eastern coast of New Zealand in 2002 (Anderson *et al.* 2005a). Invertebrates inhabiting each holdfast (either attached to haptera or retained on a 0.5 mm sieve) were identified and enumerated, resulting in a total of 351 taxa. The volume of each holdfast was also measured (in mL) using water displacement.

The structure of holdfast assemblages has a significant relationship with the volume of the holdfast (e.g. Smith *et al.* 1996). As the plant grows, the number of species and the total number of organisms inhabiting the holdfast increase (Anderson *et al.* 2005a). Naturally, the size of the sampling unit is also expected to affect assemblage structure (Cao *et al.* 2002; Mac Nally *et al.* 2004). Explicit models of changes in community structure with holdfast volume, including species accumulation curves, are given elsewhere (Anderson *et al.* 2005a,b). Here, we wished to examine patterns of variability (dispersion) in assemblage structure for holdfasts of different volume. Thus, the 80 holdfasts were allocated into four groups of 20 on the basis of their rank volume (size). The resulting groups were (1) 36–76 mL, (2) 82–110 mL, (3) 114–150 mL, and (4) 150–285 mL.

Tests of homogeneity in multivariate dispersions were done based on d_{J}, d_{MG}, d_{MG2}, d_{Manx} and d_{Eucx}. These measures represent a spectrum in the amount of emphasis placed on species composition vs. relative abundance, and none of them include any extra intrinsic standardizations by row sums, column sums or ranges. For comparison, we also did analyses based on d_{ChJ}, d_{ChJB} and on d_{BC} after first applying various transformations to the data: none, square-root or presence/absence, with the latter being equal to d_{S}. Bray–Curtis, when used with different transformations, is also considered to be a method of weighting the relative importance of rare vs. common (or abundant) species in community analysis (e.g. Clarke & Green 1988).

Tests done using either d_{J} or d_{S}, emphasizing compositional differences, indicated that there was significantly

greater beta diversity (dispersion) among small holdfasts than among larger ones (Table 1, Figs 1*a,b* and 2*a,e*). Average dissimilarities to centroids based on either the Jaccard or Sørensen dissimilarity measures have a clear correspondence to Whitakker's β_W (Table 2). In contrast, analyses emphasizing differences in relative abundances (d_{Manx}), indicated the opposite trend: significantly greater variation among large holdfasts compared with smaller ones (Table 1, Figs 1c and 2d). Thus, variation in species composition is greater in smaller holdfasts, while variation in relative abundances is greater in larger holdfasts. Although the analysis based on d_{Eucx} yielded a similar pattern to that seen for d_{Manx} (Figs 1c and 3a), the test did not detect statistically significant heterogeneity in dispersions (Table 1).

The analysis based on the modified Gower measure, which explicitly weights a compositional change equal to an order-of-magnitude change (d_{MG}) or a doubling (d_{MG2}) in abundance, indicated that there were no significant differences in dispersion among the four groups (Table 1, Figs 1b and 2b,c).

A second-stage non-metric multi-dimensional scaling (MDS) plot (Somerfield & Clarke 1995) was used to compare the various dissimilarity measures. This technique allows different analytical approaches to be compared, by calculating the Mantel correlations between every pair of dissimilarity matrices. This matrix of correlations is then input as a new matrix of similarities into the MDS algorithm. Here, we used Pearson correlations among the distance matrices (rather than Spearman rank correlations), because we wished to consider two measures to be more similar if they were linearly related, as opposed to simply related in their rank order. The pattern obtained shows the gradient in emphasis on changes in species composition through to changes in relative abundance (from left to right) with the use of different dissimilarity measures (Fig. 3a).

Results obtained using d_{BC} on untransformed or square-root transformed data indicated that there was significantly greater variation in smaller holdfasts compared to larger ones, although the size of the effect was not as large as for presence/absence data (Table 1, Fig. 1a and 2f,g). The analysis using untransformed data may emphasize abundant taxa more than rare taxa; however, the relative contribution of species composition vs. relative abundance towards this measure is unclear, because d_{BC} includes an intrinsic standardization by the sum of the abundances in the two

Table 1 Results of tests for homogeneity of multivariate dispersions for each example data set on the basis of each of several dissimilarity measures

	Holdfast data		Macrobenthos data	
Distance measure	*F*	*P*-value	*F*	*P*-value
Sørensen (Bray–Curtis, presence/absence)	15.392	0.0001	53.944	0.0001
	1 2 3 4		3 5 4 2 1	
Bray–Curtis, square-root transformation	11.613	0.0001	56.764	0.0001
	1 2 3 4		3 5 4 2 1	
Bray–Curtis, no transformation	4.479	0.0107	38.447	0.0001
	1 2 3 4		3 5 2 4 1	
Jaccard	15.630	0.0001	49.778	0.0001
	1 2 3 4		3 5 4 2 1	
Modified Gower (base 10)	1.512	0.2449	73.037	0.0001
			3 2 5 4 1	
Modified Gower (base 2)	2.416	0.0850	87.506	0.0001
			3 2 5 4 1	
Manhattan, excluding joint absences	5.626	0.0036	26.077	0.0001
	4 3 2 1		2 3 1 5 4	
Euclidean, excluding joint absences	2.416	0.2146	6.752	0.0011
			2 1 3 5 4	
Chao's abundance-based Jaccard	13.644	0.0001	55.102	0.0001
	1 2 3 4		3 5 4 2 1	
Chao's abundance-based Jaccard	6.279	0.0010	47.851	0.0001
	1 2 3 4		3 5 4 2 1	

Where there was a statistically significant overall *F*-ratio comparing groups ($P < 0.05$, 9999 permutations), pairwise comparisons were done. Numbers correspond to areas 1–5 for the macrobenthos data and to groups 1–4 for the holdfast data, as described in the text.
Numbers are given in decreasing order of average dispersion, and underlining bars indicate groups that were not statistically significantly different ($P > 0.05$).

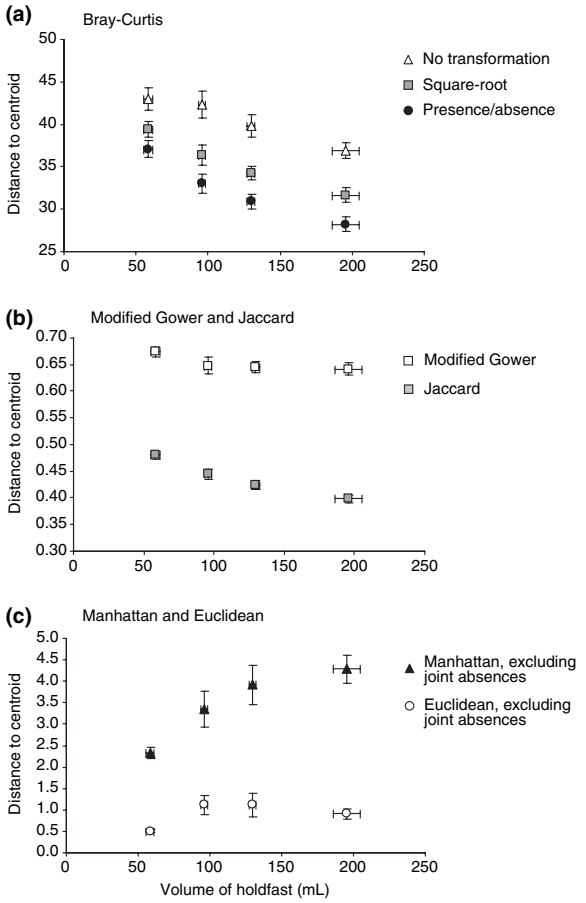

Figure 1 Average distance to centroid (± 1 SE, $n = 20$) of New Zealand holdfast assemblages classified into each of four groups on the basis of holdfast volume, using each of several dissimilarity measures, as indicated. The modified Gower measure in panel *b* was based on a $\log_{10}$ transformation.

Table 2 Average richness ($\bar{\alpha}$), gamma diversity (γ), beta diversity [$\beta_W = (\gamma/\bar{\alpha}) - 1$] and average distance from centroid on the basis of the Sørensen ($\bar{z}_S$) or Jaccard ($\bar{z}_J$) measures for (a) New Zealand holdfast data and (b) Norwegian continental shelf data

	$\bar{\alpha}$	γ	β_W	$\bar{z}_S$	$\bar{z}_J$
(a) New Zealand holdfasts					
Volume 1	67	252	2.8	0.371	0.480
Volume 2	82	271	2.3	0.330	0.444
Volume 3	91	274	2.0	0.309	0.425
Volume 4	105	290	1.8	0.282	0.397
(b) Norwegian continental shelf					
Area 1	66	177	1.7	0.361	0.246
Area 2	101	307	2.0	0.425	0.312
Area 3	83	477	4.7	0.564	0.487
Area 4	72	297	3.1	0.463	0.351
Area 5	106	405	2.8	0.470	0.359

sampling units (3). Chao's abundance-based Jaccard measure (either in raw or bias-corrected form) also clearly emphasized compositional structure (Fig. 3a, Table 1), although high stress (> 0.20) precludes close interpretation of the MDS plot (Fig. 2h).

Relating beta diversity to environmental heterogeneity on the Norwegian continental shelf

Samples of soft-sediment macrobenthic organisms were obtained from 101 sites occurring in five large areas along a transect of 15° of latitude, collected in 1996, 1997 and 1998 (Ellingsen & Gray 2002). The number of sites in each area was 16, 21, 25, 19, and 20, respectively, in a sequence from the most southern area (area 1) to the northernmost (area 5). At each site, five replicates were taken with a 0.1 m^2 van Veen grab, and these were sieved on a 1 mm sieve to provide a sample of the macrobenthos. There were 809 taxa overall, and analyses were done on abundance data pooled over the five grabs from each site. The upper 5 cm of one additional grab at each site was also sampled to measure environmental variables, which included: median, SD, skewness and kurtosis of sediment grain size, percentage silt and clay (fraction < 0.063 mm), total organic matter (%) and water depth (m) (Ellingsen & Gray 2002).

First, we tested the null hypothesis of homogeneity in the multivariate dispersions among the five areas in terms of the environmental variables, based on Euclidean distances to centroids for normalized data. We also tested the null hypothesis of homogeneity in the multivariate dispersions among the five areas on the basis of the Sørensen dissimilarity measure for the biotic variables. We then related the environmental to the biological measures of distances to centroids directly.

There were significant differences in environmental heterogeneity among the five areas ($F = 24.67$, $P < 0.001$). Area 3 had significantly greater environmental variation, while area 1 had significantly lower variation than any of the other areas (Fig. 4). There were statistically significant differences in environmental heterogeneity between every pair of areas ($P < 0.001$), except for area 2, which did not differ significantly from either area 4 or area 5 (pairwise comparisons, $P > 0.13$). These results coincide with patterns in values of coefficients of variation for each environmental variable (Ellingsen & Gray 2002).

The analysis of the biological data effectively mirrored the results obtained using the environmental data. There was general agreement in the rank order of measures of beta diversity, using β_W (see Ellingsen & Gray 2002) or using distances from centroids on the basis of either the Sørensen ($\bar{z}_S$) or the Jaccard ($\bar{z}_J$) measures (Table 2). The test revealed significant differences in beta diversity among the five areas, using either Sørensen or Jaccard: area 3 demonstrated the

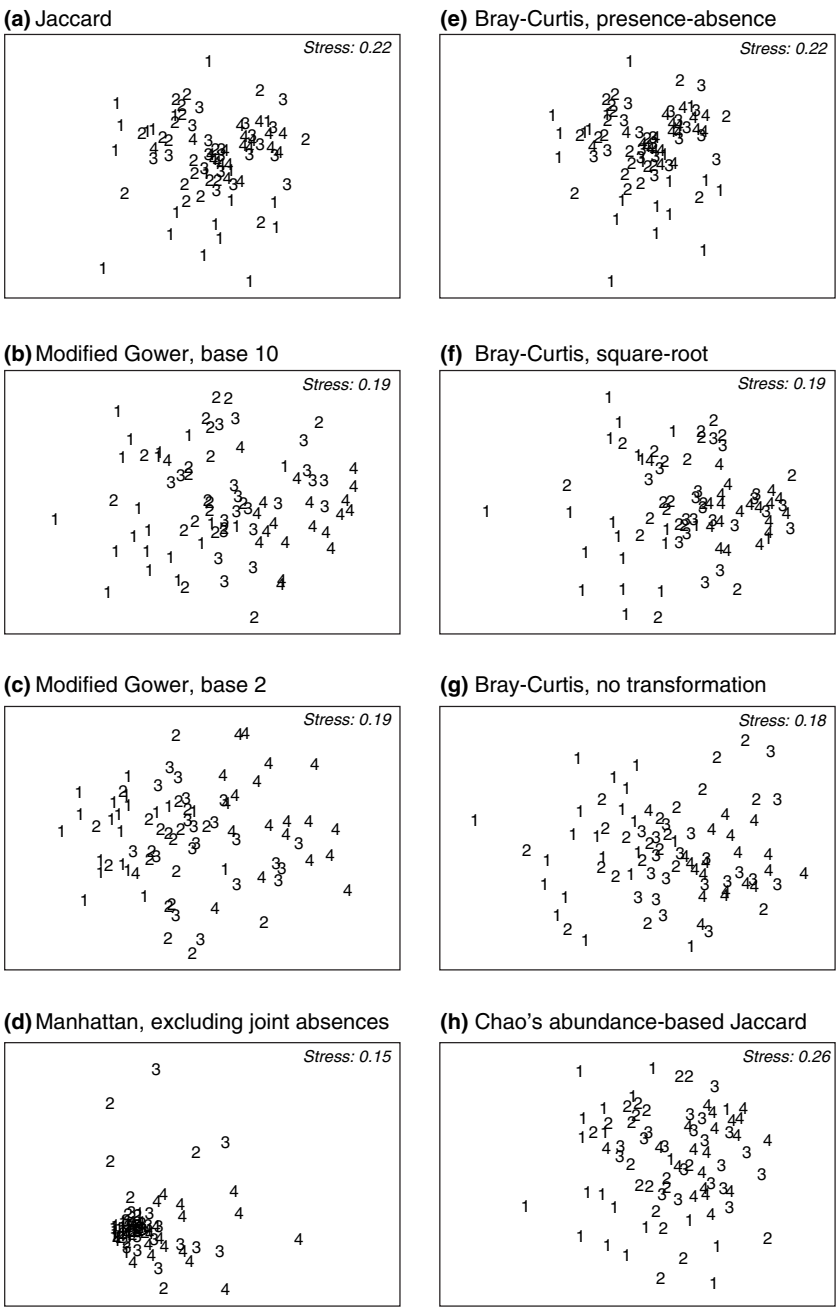

Figure 2 Non-metric MDS plots of New Zealand holdfast assemblages on the basis of each of several choices of transformation and dissimilarity measure, as indicated. Numbers indicate the volume class of each holdfast: 1, 36–76 mL; 2, 82–110 mL; 3, 114–150 mL; 4, 150–285 mL.

greatest beta diversity, followed by areas 5, 4 and 2, which did not differ significantly, followed by area 1 (Table 1, Fig. 4). Furthermore, there was a clear direct relationship between the distances to centroids based on the biological compositional data vs. the distances to centroids based on the environmental data (Fig. 4). This is in accordance with the results obtained by Ellingsen & Gray (2002), using the average of all dissimilarities (square-root transformed Bray–Curtis) between pairs of samples within each area, although in their study no direct statistical test on the values was performed. Some sites in area 3 were characterized by big distances from the area centroid for the environmental data, but not for the biological data (i.e. observations larger than three on the *x*-axis in the top of Fig. 4). Closer inspection revealed that all of these sites except one occurred in deep water (> 300 m), and also contained a high percentage of silt-clay and total organic matter. However, these sites did not correspondingly result in an unusually high biological deviation from the 'typical' assemblage. Despite such individual anomalies, the overall pattern of relationship between beta diversity and environmental heterogeneity was clear (Fig. 4).

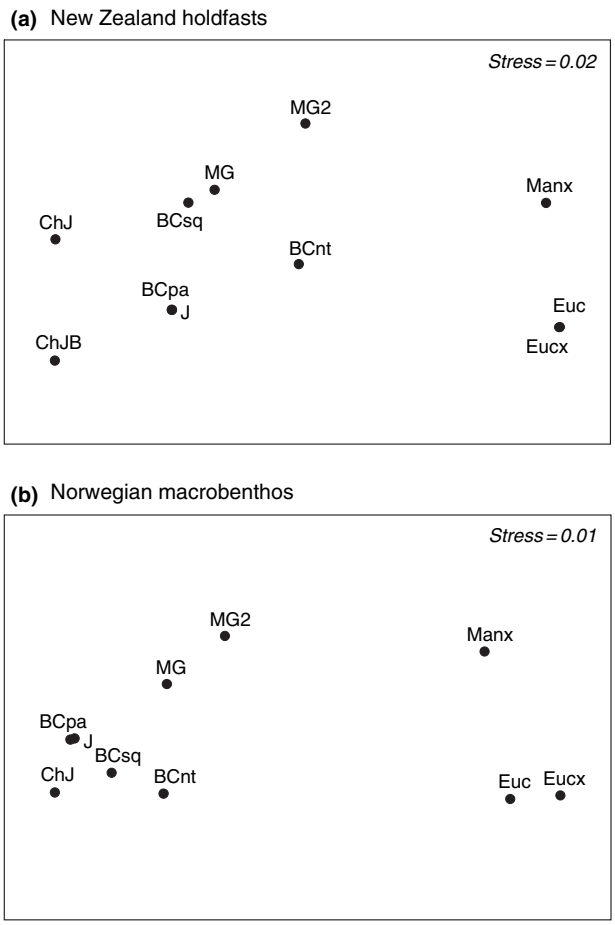

Figure 3 Second-stage MDS plots drawn on the basis of Pearson product-moment correlations among several dissimilarity matrices calculated on New Zealand holdfast data and on Norwegian continental shelf data. ChJ, Chao's abundance-based Jaccard; ChJB, Chao's abundance-based Jaccard with bias correction; J, Jaccard; BCpa, Bray–Curtis on presence/absence data; BCsq, Bray–Curtis on square-root transformed data; BCnt, Bray–Curtis on untransformed data; MG, modified Gower measure using $\log_{10}$; MG2, modified Gower using $\log_2$; Manx, Manhattan excluding double zeros; Eucx, Euclidean excluding double zeros; and Euc, Euclidean distance. *Note*: ChJB was not included in plot *b* as it caused the MDS to collapse into two points, due to its lack of correlation (near zero) with Eucx, Euc and Manx. However, the correlation of ChJB to ChJ for these data was $r = 0.96$.

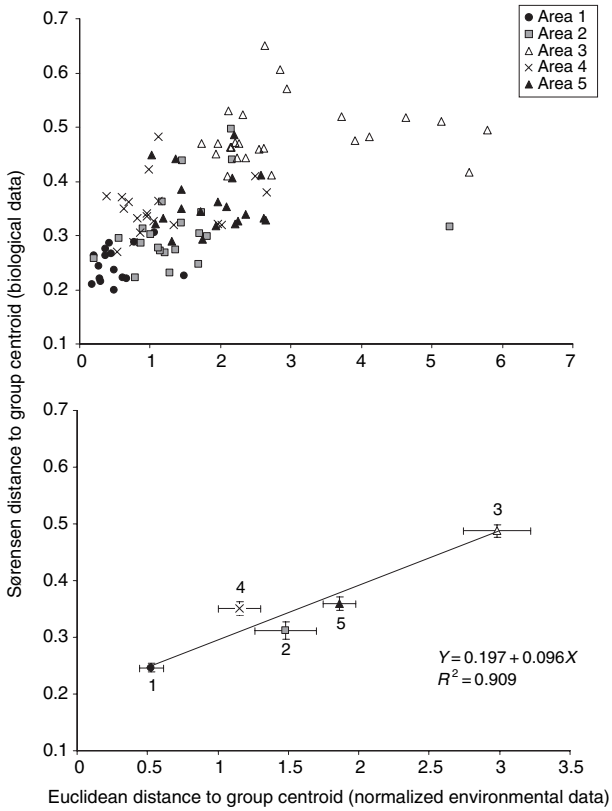

Figure 4 Sørensen distance to group centroids on biological data vs. Euclidean distances to group centroids on normalized environmental data for individual samples (top) and on average (± 1 SE, bottom) for each of five areas from the Norwegian continental shelf. Regression line was drawn using reduced major axis (model II) regression.

We next analysed the Norwegian data set using d_{MG}, d_{MG2}, d_{Manx}, d_{Eucx}, d_{Euc}, d_{ChJ}, d_{ChJB} and d_{BC} with different transformations, as for the holdfast data. For these data, the choice of transformation and dissimilarity measure had very little effect on the general outcome of the analysis (Table 1). In essence, there was generally significantly larger dispersion in area 3 and significantly smaller dispersion in area 1 compared with the other three areas. Thus, regardless of whether one considers variation in species composition or in abundance, the areas with greater environmental heterogeneity also had the greatest multivariate dispersion. Differences in sample size and in the area sampled could have played a role in determining results (Cao *et al.* 2002; Mac Nally *et al.* 2004). However, there were little or no differences between results obtained using d_{ChJB} (which substantially reduces sample-size bias) and d_{ChJ}. In fact, the linear correlation between these two dissimilarity matrices was $r = 0.96$ for the Norwegian data and $r = 0.85$ for the holdfast data, suggesting that sample-size issues were not the primary drivers behind the patterns observed.

Interestingly, the analysis placing a change in species composition on an equal footing with a doubling in abundance (d_{MG2}) resulted in the strongest measured effects (the greatest value for the *F*-ratio, Table 1). This contrasts with the results obtained for the holdfast data, where there were effects in opposite directions for compositional vs. relative abundance-based dispersion. These may have effectively 'cancelled each other out' when the modified Gower measure was used for the holdfast data (Table 1).

Despite the differences inherent in these two data sets, the second-stage MDS plot in both cases shows a gradient in the change of emphasis from compositional to abundance-based information from left to right on the plot (Fig. 3). In either case, it is interesting to note that d_{BC} on untransformed data does not necessarily go very far along this path and d_{ChJ} (or d_{ChJB}) does not fit easily into this continuum, but appears even to go in a different direction entirely.

DISCUSSION

As a single value calculated for a given area, Whittaker's (1960, 1972) β_W does not allow statistical comparisons of beta diversity between two or more areas. We propose that a test for homogeneity in multivariate dispersions (Anderson 2006), when based on a suitable dissimilarity measure, provides such a test. There is a clear correspondence between the average distances to centroids based on compositional dissimilarity (such as d_S or d_J) and Whittaker's β_W. This correspondence does not necessarily occur, however, when multivariate dispersion is measured using dissimilarities that include abundance information.

Beta diversity was positively related to environmental heterogeneity for macrobenthic assemblages on the Norwegian continental shelf. In fact, multivariate dispersion was positively related to environmental heterogeneity regardless of the dissimilarity measure or transformation used as the basis for the analysis. In contrast, the dissimilarity measure used had a strong effect on patterns of multivariate dispersion for kelp holdfast fauna. There was high variation in species composition during early stages of succession (small holdfasts), while relative abundances of species were more variable in larger holdfasts. This could have been a simple consequence of Taylor's power law (variance increasing with the mean) or it could have been driven by recruitment pulses or increasingly variable immigration or extinction as the holdfast ages. The consequences of intrinsic mean–variance relationships inherent in counts of species abundances for measuring multivariate dispersion on the basis of different dissimilarity measures is clearly an area that warrants further research.

The test for homogeneity of dispersions (Anderson 2006) can be calculated on the basis of any dissimilarity measure of choice. In addition, a canonical partitioning of the variability in any dissimilarity matrix also can be done (McArdle & Anderson 2001; Legendre *et al.* 2005) and the sizes of pseudo multivariate variance components can be calculated from these (e.g. Anderson *et al.* 2005a), in a manner analogous to the estimation of univariate variance components (Searle *et al.* 1992). In addition, differences in the sizes of such multivariate variance components can be tested using bootstrapping procedures (A. Terlizzi, M.J. Anderson, S. Fraschetti, L. Benedetti-Cecchi, unpublished data). While all of these methodologies are useful in a general framework for analysing multivariate variability, they do not necessarily constitute analyses of beta diversity, *per se*.

Kiflawi & Spencer (2004) have shown how β_W can be expressed as an odds, and therefore how comparisons between two areas can be done by building a confidence interval for the odds ratio. Multiple pairwise tests using this approach and also a measure of beta-diversity based on rarefaction curves suggested by Olszewski (2004) were recently used by Davis (2005) to compare and test for differences in beta diversity across multiple areas in time and space. These approaches are specific to particular measures, however, and do not allow for a simultaneous test among several areas (or times).

The test proposed here has a bit more flexibility and allows ecologists to rigorously test hypotheses concerning differences in multivariate dispersions (variability in community structure) or beta diversity (for species composition data) among groups of multivariate samples. The potential power of this method for investigating new and interesting hypotheses in ecology comes, however, with a serious dose of responsibility. It is very important to be clear about what is meant by 'variability in community structure' in a given analysis. Do we mean variability in abundances? Variability in the kinds of species present (beta diversity)? Variability in their proportions? The analysis can be based on any dissimilarity measure, but ecological interpretations can be made only with care. Transformations also have strong impacts on patterns and results.

It is generally recommended, for the multivariate analysis of ecological species abundance data, that the data be transformed (e.g. to square-roots or logs) in order to reduce the influence of very abundant species (e.g. Clarke & Green 1988). Such transformations are often used in univariate analysis in order to fulfil the assumption of homogeneity of variances (e.g. Underwood 1981; McArdle & Anderson 2004). Such transformations are called 'non-affine' in that they do not preserve the collinearity (hence straightness or parallelism) in the system. The effects of non-affine transformations occur for multivariate data, just as they do for univariate data. Therefore, comparisons of multivariate dispersions after transforming the data must be interpreted with this in mind, particularly if there is any evidence for differences in means (locations) among groups.

Although the effects of transformations are relatively well understood, at least for univariate analysis, the effect of using different dissimilarity measures on patterns of multivariate dispersions has so far been completely overlooked. Most of the literature on distance measures concentrates on whether particular distance or dissimilarity functions fulfil various mathematical properties such as metricity (e.g. Gower & Legendre 1986), whether they

correspond to some particular notion of simulated 'ecological distance' (e.g. Faith *et al.* 1987), whether they are related to hypothetical environmental gradients (e.g. Legendre & Gallagher 2001), whether they are sensitive to specific kinds of changes in assemblages (e.g. Hajdu 1981; Cao *et al.* 1997), whether or not they include or exclude information on joint absences or whether they are influenced by differences in the scale of variables (Legendre & Legendre 1998).

Most of the measures considered suitable for ecological species abundance data (e.g. Bray–Curtis, Canberra, chi-squared, CY dissimilarity, Hellinger, Orlóci's chord, Kulczynski, etc.) have intrinsic non-affine transformations. They relativise absolute differences in abundance to capture specific notions of 'community structure' (e.g. if two communities have the same *proportions* of individuals in different species, then they may be considered similar). However, differences in *absolute* abundances can also be important ecologically, as they may correspond to differences in an ecosystem's productivity (e.g. Roughgarden *et al.* 1991; Menge *et al.* 1997) or responses to a pollutant or other impact. The intrinsic non-affine transformations inherent in these measures can make relative multivariate dispersions based on them difficult to interpret by reference to the original species variables.

It is clear that much thought is needed concerning the choice of transformation and dissimilarity measure for the analysis of multivariate dispersions. Interpretations of results are generally clearer for those measures where any transformations used are explicit. We have provided here an abundance-based dissimilarity measure that is directly interpretable by reference to original species abundances on a log scale. The measure is flexible as the relative value of the change from absence to presence can be specified. It reduces to the Jaccard measure for presence/absence data and can be partitioned directly according to the contribution of species composition vs. relative abundance information.

Comparisons of dispersions among groups can reveal dramatically different stories, depending on which aspect of assemblage structure is emphasized. Given the difficulties already encountered in obtaining a consensus concerning appropriate measures of variability for a single population (e.g. Williamson 1984; McArdle & Gaston 1992; Leps 1993; Gaston & McArdle 1994; McArdle & Gaston 1995), it is going to be even more difficult to obtain a consensus concerning how to measure variability in whole assemblages (Underwood 1986). Therefore, we recommend using a range of dissimilarity measures which cover the spectrum from emphasizing compositional change to changes in abundances. Armed with this information, the nature of heterogeneity in species composition (beta diversity) or in dispersions based on other measures of assemblage structure can be articulated with greater specificity and confidence.

ACKNOWLEDGEMENTS

OLF (the Norwegian Oil Industry Association) permitted us to use the macrobenthic data, and Det Norske Veritas and Akvaplan-niva prepared them. KEE acknowledges the support of the Research Council of Norway. Kelp holdfast data were obtained by MJA, in collaboration with C. Diebel and W. Blom at the Auckland War Memorial Museum, funded by the Royal Society of New Zealand's Marsden fund. The manuscript was improved by comments provided by E.B. Davis, T.D. Olszewski, P. Legendre and one anonymous referee. A computer program for calculating measures of dispersion and for performing tests is freely available from http://www.stat.auckland.ac.nz/~mja/Programs.htm.

REFERENCES

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Aust. Ecol.*, 26, 32–46.

Anderson, M.J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62, 245–253.

Anderson, M.J., Diebel, C.E., Blom, W.M. & Landers, T.J. (2005a). Consistency and variation in kelp holdfast assemblages: spatial patterns of biodiversity for the major phyla at different taxonomic resolutions. *J. Exp. Mar. Biol. Ecol.*, 320, 35–56.

Anderson, M.J., Millar, R.B., Blom, W.M. & Diebel, C.E. (2005b). Nonlinear multivariate models of successional change in community structure using the von Bertalanffy curve. *Oecologia*, 146, 279–286.

Bray, J.R. & Curtis, J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.*, 27, 325–349.

Cao, Y., Williams, W.P. & Bark, A.W. (1997). Similarity measure bias in river benthic Aufwuchs community analysis.. *Water Env. Res.*, 69, 95–106.

Cao, Y., Larsen, D.P., Hughes, R.M., Angermeier, P.L. & Patton, T.M. (2002). Sampling effort affects multivariate comparisons of stream assemblages. *J. Nth. Amer. Benth. Soc.*, 21, 701–714.

Chao, A., Chazdon, R.L., Colwell, R.K. & Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.*, 8, 148–159.

Clarke, K.R. (1993). Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.*, 18, 117–143.

Clarke, K.R. & Green, R.H. (1988). Statistical design and analysis for a biological effects study. *Mar. Ecol. Prog. Ser.*, 46, 213–226.

Clarke, K.R. & Warwick, R.M. (2001). *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*, 2nd edn. Primer-e Ltd, Plymouth Marine Laboratory, Plymouth, UK.

Clarke, K.R., Somerfield, P.J. & Chapman, M.G. (2006). On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray–Curtis coefficient for denuded assemblages. *J. Exp. Mar. Biol. Ecol.*, 330, 55–80.

Czekanowski, J. (1909). Zur Differentialdiagnose der Neandertalgruppe. Korrespondenz-Blatt deutsch. *Ges. Anthropol. Ethnol. Urgesch.*, 40, 44–47.

Davis, E.B. (2005). Mammalian beta diversity in the Great Basin, western USA: palaeontological data suggest deep origin of modern macroecological structure. *Global Ecol. Biogeogr.*, 14, 479–490.

Dice, L.R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26, 297–302.

Ellingsen, K.E. & Gray, J.S. (2002). Spatial patterns of benthic diversity: is there a latitudinal gradient along the Norwegian continental shelf? *J. Anim. Ecol.*, 71, 373–389.

Faith, D.P., Minchin, P.R. & Belbin, L. (1987). Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, 69, 57–68.

Gaston, K.J. & McArdle, B.H. (1994). The temporal variability of animal abundances: measures, methods and patterns. *Phil. Trans. R. Soc. Lond. B*, 345, 335–358.

Gower, J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53, 325–338.

Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27, 857–871.

Gower, J.C. (1987). Introduction to ordination techniques. In: *Developments in Numerical Ecology* (eds Legendre, P. & Legendre, L.). Springer-Verlag, Berlin, pp. 3–64.

Gower, J.C. & Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *J. Classif.*, 3, 5–48.

Hajdu, L.J. (1981). Graphical comparison of resemblance measures in phytosociology. *Vegetatio*, 48, 47–59.

Jaccard, P. (1900). Contribution au problème de l'immigration post-glaciare de la flore alpine. *Bull. Soc. Vaudoise Sci. Nat.*, 36, 87–130.

Kiflawi, M. & Spencer, M. (2004). Confidence intervals and hypothesis testing for beta diversity. *Ecology*, 85, 2895–2900.

Koleff, P., Gaston, K.J. & Lennon, J.J. (2003). Measuring beta diversity for presence–absence data. *J. Anim. Ecol.*, 72, 367–382.

Legendre, P. & Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129, 271–280.

Legendre, P. & Legendre, L. (1998). *Numerical Ecology*, 2nd English edition. Elsevier Science BV, Amsterdam, The Netherlands.

Legendre, P., Borcard, D. and Peres-Neto, P.R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecol. Monogr.*, 75, 435–450.

Leps, J. (1993). Taylor's power law and the measurement of variation in the size of populations in space and time. *Oikos*, 68, 349–356.

Levene, H. (1960). Robust tests for equality of variances. In: *Contributions to Probability and Statistics* (eds Olkin, I., Ghurye, S.G., Hoeffding, W., Madow, W.G. & Mann, H.B.). Stanford University Press, Stanford, CA, pp. 278–292.

Mac Nally, R., Fleishman, E., Bulluck, L.P. & Betrus, C.J. (2004). Coparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *J. Biogeogr.*, 31, 917–929.

Magurran, A.E. (2004). *Measuring Biological Diversity*. Blackwells, Oxford, UK.

McArdle, B.H. & Anderson, M.J. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*, 82, 290–297.

McArdle, B.H. & Anderson, M.J. (2004). Variance heterogeneity, transformations and models of species abundance: a cautionary tale. *Can. J. Fish. Aquat. Sci.*, 61, 1294–1302.

McArdle, B.H. & Gaston, K.J. (1992). Comparing population variabilities. *Oikos*, 64, 610–612.

McArdle, B.H. & Gaston, K.J. (1995). The temporal variability of densities: back to basics. *Oikos*, 74, 165–171.

McArdle, B.H., Gaston, K.J. & Lawton, J.H. (1990). Variation in the size of animal populations: patterns, problems and artifacts. *J. Anim. Ecol.*, 59, 439–454.

Menge, B.A., Daley, B., Wheeler, P.A. & Strub, P.T. (1997). Rocky intertidal oceanography: an association between community structure and nearshore phytoplankton concentration. *Limnol. Oceanogr.*, 42, 57–66.

Motyka, J. (1947). *O zadaniach I metodach badan geobotanicznych. Sur les buts et les méthods des recherches géobotaniques*. Annales Universitas Mariae Curie-Sklodowska, Lublin, Polonia, Sectio C, Supplementum I viii + 168pp.

Odum, E.P. (1950). Bird populations of the Highlands (North Carolina) plateau in relation to plant succession and avian invasion. *Ecology*, 31, 587–605.

Olszewski, T.D. (2004). A unified mathematical framework for the measurement of richness and evenness within and among multiple communities. *Oikos*, 104, 377–387.

Roughgarden, J., Pennington, J.T., Stoner, D., Alexander, S. & Miller, K. (1991). Collisions of upwelling fronts with the intertidal zone: the cause of recruitment pulses in barnacle populations of central California. *Act. Oecol.*, 12, 35–51.

Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application in analysis of the vegetation on Danish commons. *Biol. Skr.*, 5, 1–34.

Searle, S.R., Casella, G. & McCulloch, C.E. (1992). *Variance Components*. John Wiley & Sons, New York, USA.

Smith, S.D.A., Simpson, R.D. & Cairns, S.C. (1996). The macrofaunal community of *Ecklonia radiata* holdfasts: description of the faunal assemblage and variation associated with differences in holdfast volume. *Aust. J. Ecol.*, 21, 81–95.

Somerfield, P.J. & Clarke, K.R. (1995). Taxonomic levels in marine community studies, revisited. *Mar. Ecol. Prog. Ser.*, 127, 113–119.

Taylor, L.R. (1961). Aggregation, variance and the mean. *Nature*, 189, 732–735.

Underwood, A.J. (1981). Techniques in analysis of variance in experimental marine biology and ecology. *Oceanogr. Mar. Biol. Ann. Rev.*, 19, 513–605.

Underwood, A.J. (1986). What is a community? In: *Patterns and Processes in the History of Life* (eds Raup, D.M., Jablonski, D.). Springer-Verlag, Berlin, Germany, pp. 351–367.

Vellend, M. (2001). Do commonly used indices of b-diversity measure species turnover? *J. Veg. Sci.*, 12, 545–552.

Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.*, 22, 1–44.

Whittaker, R.H. (1972). Evolution and measurement of species diversity. *Taxon*, 21, 213–251.

Williamson, M. (1984). The measurement of population variability. *Ecol. Entomol.*, 9, 239–241.

Editor, Aaron Ellison
Manuscript received 21 December 2005
First decision made 23 January 2006
Manuscript accepted 17 February 2006