Acta Oecologica 37 (2011) 187–194

Contents lists available at ScienceDirect

Acta Oecologica

journal homepage: www.elsevier.com/locate/actoec

Original article

Evaluating and interpreting cross-taxon congruence: Potential pitfalls and solutions

Margherita Gioria [a,*], Giovanni Bacaro [b], John Feehan [a]

[a] School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland
[b] BIOCONNET, Biodiversity and Conservation Network, Department of Environmental Science "G. Sarfatti", University of Siena, Via P.A. Mattioli 4, 53100 Siena, Italy

ARTICLE INFO

Article history:
Received 4 October 2010
Accepted 2 February 2011
Available online 24 February 2011

Keywords:
Biodiversity
Co-correspondence analysis
Conservation planning
Procrustes analysis
Surrogate taxon

ABSTRACT

Characterizing the relationship between different taxonomic groups is critical to identify potential surrogates for biodiversity. Previous studies have shown that cross-taxa relationships are generally weak and/or inconsistent. The difficulties in finding predictive patterns have often been attributed to the spatial and temporal scales of these studies and on the differences in the measure used to evaluate such relationships (species richness *versus* composition). However, the choice of the analytical approach used to evaluate cross-taxon congruence inevitably represents a major source of variation. Here, we described the use of a range of methods that can be used to comprehensively assess cross-taxa relationships. To do so, we used data for two taxonomic groups, wetland plants and water beetles, collected from 54 farmland ponds in Ireland. Specifically, we used the Pearson correlation and rarefaction curves to analyse patterns in species richness, while Mantel tests, Procrustes analysis, and co-correspondence analysis were used to evaluate congruence in species composition. We compared the results of these analyses and we described some of the potential pitfalls associated with the use of each of these statistical approaches. Cross-taxon congruence was moderate to strong, depending on the choice of the analytical approach, on the nature of the response variable, and on local and environmental conditions. Our findings indicate that multiple approaches and measures of community structure are required for a comprehensive assessment of cross-taxa relationships. In particular, we showed that selection of surrogate taxa in conservation planning should not be based on a single statistic expressing the degree of correlation in species richness or composition. Potential solutions to the analytical issues associated with the assessment of cross-taxon congruence are provided and the implications of our findings in the selection of surrogates for biodiversity are discussed.

1. Introduction

Promoting conservation efforts and mitigating the effects of human-induced changes, such as habitat destruction and degradation, climate change, and the spread of alien species, requires a solid understanding of the distribution and main drivers of patterns of biodiversity in space and time (Gaston, 1996, 2000; Howard et al., 1998; Gioria et al., 2010a). The collection of extensive information on biodiversity is, however, hampered by the complexity of spatio-temporal dimensions of this variable, and by financial constraints and by often limited taxonomic expertise (e.g., Gaston, 1996; Howard et al., 1998; Sauberer et al., 2004). To address these issues, there has been an increasing interest in the use of correlates for biodiversity (Noss, 1990; Prendergast et al., 1993; Margules and Pressey, 2000), with evidence showing that the use of surrogate taxa in conservation planning is substantially more effective than that of surrogates based on environmental data only (see Rodrigues and Brooks, 2007 for a review).

The effectiveness of the use of one taxon to predict community patterns for other taxonomic groups ultimately depends upon the strength of the relationship among such groups and on its underlying mechanisms (e.g., Prendergast et al., 1993; Paszkowski and Tonn, 2000; Sætersdal et al., 2003; Warman et al., 2004). The factors that affect cross-taxa relationships include 1) a similar but independent response of two taxonomic groups to the same set of environmental conditions (Jackson and Harvey, 1993; Paavola et al., 2003), 2) trophic interactions or functional interdependence (Jackson and Harvey, 1993), 3) a shared bio-geographical and evolutionary history at a large/global scale (e.g., Howard et al.,

---

Abbreviations: CCA, canonical correspondence analysis; Co-CA, co-correspondence analysis; NMDS, non-metric multidimensional scaling; PCA, principal components analysis.

* Corresponding author. Tel.: +353 86 067 5182.
E-mail address: margheritagioria@yahoo.com (M. Gioria).

1146-609X/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.actao.2011.02.001

1998), and 4) species–energy relationships (e.g., Gaston, 2000; see Gaston and Williams, 1996 and Heino, 2010 for a summary).

Cross-taxon congruence analysis can be expressed as the correlation in patterns of species richness and/or diversity (Pearson and Carroll, 1999) or, in a multi-species context, as community concordance, i.e., the relationship among compositional patterns of multiple taxonomic groups across sites (Jackson and Harvey, 1993; ter Braak and Schaffers, 2004; Virtanen et al., 2009). To date, cross-taxa relationships have been characterized using groups belonging to the same taxonomic rank, such between different invertebrate groups (e.g., Briers and Biggs, 2003; Bilton et al., 2006) or taxonomic or functional ranks, often evaluating whether vascular plants can be used as surrogates for other groups such as invertebrates, birds, bryophytes, etc. (e.g., Prendergast et al., 1993; Sætersdal et al., 2003; Su et al., 2004; Wolters et al., 2006). A quantification of congruence in species richness or composition between different taxa has been proved to be useful to identify hotspots of biodiversity (Prendergast et al., 1993; Howard et al., 1998), although evidence of consistent patterns that can justify the use of one taxon to effectively predict patterns for other taxonomic groups is rather weak (see Rodrigues and Brooks, 2007; Heino, 2010 for a review).

The difficulties in finding predictive patterns of congruence between different taxonomic groups have often been attributed to the spatial scale at which these relationships have been investigated (Pearson and Carroll, 1999; Lamoreux et al., 2006; Paavola et al., 2006). While strong cross-taxa relationships have been identified, for certain groups, at the regional scale (Prendergast et al., 1993; Howard et al., 1998; Wolters et al., 2006; see Gaston, 2000 for a review), contrasting results have been identified at the local/community level (e.g., Heino et al., 2009a; Warman et al., 2004; Oertli et al., 2005; Bilton et al., 2006). These differences may also be due to the choice of the variable(s) used to evaluate such relationships (e.g., species richness *versus* abundance or species occurrence data), with patterns in species composition often stronger than those in species richness, both in terrestrial (e.g., Su et al., 2004; Oertli et al., 2005; Schaffers et al., 2008) and aquatic ecosystems (e.g., Bilton et al., 2006; Heino et al., 2009b).

Although there is no doubt that the spatio-temporal scales of a study and the choice of the variable used to summarize biodiversity patterns may substantially affect the degree of association between two or more taxa, the characteristics of the statistical approach(es) used to analyse community data could also act as a major source of variation and of inconsistent patterns. To date, a multiplicity of methods has been used to assess cross-taxon congruence in species richness and/or composition. In a univariate context, the correlation between two or more taxonomic groups has been typically quantified using simple regression analysis applied to either species richness, diversity indices, or taxonomic distinctness indices (e.g., Prendergast et al., 1993; Howard et al., 1998; Su et al., 2004; Warman et al., 2004; Oertli et al., 2005; Wolters et al., 2006), using either the Pearson or the Spearman rank correlation, depending on the error distribution for the data to analyse. Rarefaction (Sanders, 1968; Gotelli and Colwell, 2001) has also been used to compare species richness among multiple taxonomic groups across different habitats (e.g., Chiarucci et al., 2008). As for patterns in species composition, a recent review by Heino (2010) showed that Mantel tests (Mantel, 1967; Manly, 1997) and Procrustes analysis (Gower, 1971) were the methods that had been most frequently applied to evaluate cross-taxon congruence in freshwater ecosystems.

The Mantel approach has been widely used by ecologists to test hypotheses concerning the degree of association between two or more data matrices (see Legendre and Legendre, 1998). This is a robust and flexible method that calculates the concordance between two datasets by correlating dissimilarity or distance matrices calculated using any resemblance measure of choice. However, the use of resemblance matrices rather than the original ones as the starting point of the correlation calculations has been much criticized (Dutilleul et al., 2000; Peres-Neto and Jackson, 2001; Legendre and Fortin, 2010). To address this issue, Jackson and Harvey (1993) suggested the use of Procrustes analysis (Gower, 1971) as an alternative approach to calculate the degree of association among ecological data matrices. This is a family of multivariate ordination techniques that aims to calculate the correlation among data matrices using a rotational-fit algorithm. This algorithm attempts to minimize the total sum-of-squared residuals between data values in two matrices using a series of steps, including centering, scaling, reflection, rotation, and dilation (see Jackson, 1995; Legendre and Legendre, 1998). Randomization (Procrustean Randomization tests or PROTEST; Jackson, 1995) or bootstrapping tests (Husson et al., 2007) have then been used to test the significance of the Procrustean fit, i.e., the measure of matrix concordance. One of the advantages of this method over the Mantel approach is the possibility of using the results of multivariate ordination analyses as the input data matrices, thus allowing a reduction in the dimensionality of the original data (Peres-Neto and Jackson, 2001). The majority of applications of this approach to the analysis of cross–taxa relationships have used the resulting axes of gradient analyses, including Principal Components Analysis (PCA; e.g., Bini et al., 2008), Correspondence Analysis (CA; e.g., Paszkowski and Tonn, 2000), and Detrended Correspondence Analysis (DCA; e.g., Bini et al., 2007), or those of non-parametric ordination methods such as non-metric multidimensional scaling (NMDS; e.g., Paavola et al., 2006; Virtanen et al., 2009) as the input data matrices.

More recently, ter Braak and Schaffers (2004) developed an ordination method specifically to predict one biological community from another (see also Schaffers et al., 2008), i.e., co-correspondence analysis (Co-CA). This method directly relates two community data matrices by maximizing the weighted covariance between weighted averages species scores of one community and weighted averages species scores of the other community. Two versions of Co-CA are available, a symmetric explanatory version, and a predictive one, based on weighted averaging and partial least squares regression (see ter Braak and Schaffers, 2004 and references therein). The predictive version of Co-CA has been used to assess the value of vascular plants in predicting invertebrate community patterns in terrestrial (Schaffers et al., 2008) and aquatic systems (Gioria et al., 2010a), as well as among aquatic invertebrate groups (e.g., Bitušík et al., 2009).

A number of indirect approaches have also been used to analyse cross–taxa relationships. For instance, congruence in composition has been evaluated by calculating pair-wise correlations of the resulting ordination axes of indirect gradient analyses or reciprocal averaging (e.g., Chiarucci et al., 2007; Maccherini et al., 2009 and references therein). Other approaches have included the calculation of the correlation between the results of DCA performed on community data and a set of environmental variables (Sætersdal et al., 2003).

In this study, we wished to comprehensively evaluate cross-taxon congruence and describe the use of a range of statistical approaches that are available to ecologists to quantify the degree of association between two or more community datasets. Specifically, we aimed to 1) characterize the potential differences arising from the use of alternative approaches to calculate cross-taxa relationships; 2) to describe the issues associated with the use of such methods; and 3) to provide potential solutions in order to overcome such issues when attempting to identify surrogate taxa for biodiversity. To do so, we used abundance data for plant and water beetle (Coleoptera) assemblages obtained from a study on farmland