Ecological Indicators 88 (2018) 214–218

Contents lists available at ScienceDirect

Ecological Indicators

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

# Biological surrogates: A word of caution

Gustavo Filete de Morais[a], Luiz Guilherme dos Santos Ribas[b], Jean Carlo Gonçalves Ortega[c], Jani Heino[d], Luis Mauricio Bini[b,*]

[a] Departamento de Ecologia, Universidade Federal de Goiás, 74001-970 Goiânia, GO, Brazil
[b] Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, 74001-970 Goiânia, GO, Brazil
[c] Programa de Pós-Graduação em Ecologia de Ambientes Aquáticos Continentais, Universidade Estadual de Maringá, 87020-900 Maringá, PR, Brazil
[d] Finnish Environment Institute, Biodiversity, PO Box 413, FIN-90014, University of Oulu, Finland

ARTICLE INFO

*Keywords:*
Biodiversity assessment
Biomonitoring
Concordance among biological groups
Surrogate group
Meta-analysis
Systematic review

ABSTRACT

The value of biological surrogates has been tested for many ecosystems and biological groups. Biological surrogates are biological groups whose biodiversity patterns (e.g. abundance, species richness or assemblage composition) correlate strongly with those of other biological groups. They should thus be cost-effective proxies for overall diversity variation in biodiversity assessment and biomonitoring projects. We assessed whether the available evidence support the use of surrogate groups in biodiversity assessment and biomonitoring studies considering aquatic and terrestrial ecosystems. To achieve this goal, we carried out a meta-analysis of studies testing the strength of different surrogacy approaches (relationship between species richness, ordination patterns and compositional (dis)similarity matrices of different biological groups). The strengths of relationships between species richness of biological groups were higher for plants and microorganisms than for animals, were similar for terrestrial and aquatic ecosystems and for different types of data. The variation in the strength of relationships between compositional dissimilarity matrices was not explained by the explanatory variables 'taxa', 'realms' or 'types of data'. However, as main results, we found that the weighted effect sizes, measuring the value of surrogates, were low, highly variable and mostly unpredictable (at least considering our explanatory variables). Therefore, the available evidence suggests caution in the use of surrogate groups and that biodiversity assessment and biomonitoring programs should be based on multiple taxonomic groups, whenever possible.

## 1. Introduction

Implicitly or explicitly, applied ecological studies often rely on surrogate groups (Rodrigues and Brooks, 2007; Lindenmayer et al., 2015; Hunter et al., 2016). This is a necessity due to our insufficient knowledge of species identities (i.e. the field of taxonomy) and distributions (i.e. the field of biogeography), commonly referred to as Linnean and Wallacean shortfalls, respectively (Brown and Lomolino, 1998). In systematic conservation planning, for example, surrogacy power can be quantified by assessing the biodiversity representation (sensu Margules and Pressey, 2000) of a particular biological group (e.g. reptile species) in a network of protected areas that was originally selected considering the spatial patterns of a second biological group (e.g. birds, collected in the same sampling units). In biomonitoring studies, the efficiency of surrogate groups can be quantified by analyzing concordance between biological assemblages (also known as community concordance, cross-taxon congruence, cross-taxon correlation, and variations of these terms). There is concordance when ordination or classification patterns generated independently by two biological groups (e.g. diatoms and macroinvertebrates collected in the same sampling units) are significantly similar (e.g. Spitale et al., 2012). An analysis of concordance can also be done for species richness (or another univariate attribute, such as abundance or biomass) and, in this case, a concordant pattern emerges when the species richness of a given group is significantly correlated with that of a second biological group (e.g. Hofmeister et al., 2014). In this context, it is important to emphasize that the efficiency of surrogate groups should be quantified and not only assumed (Vieira et al., 2015). Although widely used in biodiversity assessment, biomonitoring programs and systematic conservation planning, the surrogates approach would be justified only if there is a strong relationship between biological groups (Heino, 2010).

The relationships between pairs of biological taxa have been tested using different types of data (species richness and community composition) and a variety of statistical approaches (e.g. correlation tests, Mantel test, Procrustes Analysis; Gioria et al., 2010). In general, results from these tests support different conclusions regarding the validity of

* Corresponding author.
*E-mail addresses:* lmbini@gmail.com, bini@ufg.br (L.M. Bini).

https://doi.org/10.1016/j.ecolind.2018.01.027
Received 9 October 2017; Received in revised form 12 January 2018; Accepted 15 January 2018

surrogates. For example, in aquatic ecosystems, there are evidences for (e.g. Bilton et al., 2006; Bini et al., 2008; Gioria et al., 2010; Johnson and Hering, 2009; 2010; Johnson et al., 2007) and against the use of surrogates (Bini et al., 2007; Dolph et al., 2011; Heino, 2010; Heino et al., 2005; Larsen et al., 2012; Padial et al., 2012; Vieira et al., 2015). However, it must be borne in mind that the interpretation of evidence 'for' or 'against' the use of surrogate groups depends on what the researchers have considered to be 'strong surrogacy'. Some studies have considered surrogacy to be strong when the among-groups correlations have been merely significant, whereas others have focused more on effect sizes (e.g. correlation coefficients).

Given these uncertainties on the validity of the surrogate approach, we carried out a meta-analysis to assess whether there is evidence supporting the use of surrogate groups in bioassessment and biomonitoring studies. Based on the results of previous studies (Bae et al., 2014; Bini et al., 2007, 2008; Dolph et al., 2011; Juen et al., 2013; Traversetti et al., 2013), we predicted that, in general, levels of concordance between assemblages would be statistically significant. However, considering that these same assemblages also tend to respond differently to biotic interactions, specific environmental gradients and are subject to different stochastic processes, we predicted that the levels of concordance, albeit significant, would not be high enough (e.g. $> 0.7$; following Heino (2010) and references therein) to justify the use of surrogate groups in biomonitoring studies.

## 2. Material and methods

### 2.1. Systematic review and meta-analysis

We conducted a systematic review searching for relevant articles, published between 1994 and November 30$^{th}$ 2017, on the ISI *Web of Science* database. We used the following search terms in the "Topic" field: Communit* concordance OR Communit* congruence OR Assemblage* concordance OR Assemblage* congruence OR Cross-tax* congruence OR Cross-tax* concordance OR Cross-tax* correlatio* OR Cross-tax* relationship OR Concordance between communit* OR Concordance among communit* OR Concordance between assemblage* OR Concordance among assemblage*.

From the results of this search (1926 articles), we selected those published in the following research areas: environmental sciences, ecology, marine freshwater biology, biodiversity conservation, evolutionary biology and plant science. After applying these filters, we recorded 580 articles. These articles were read in full to verify the availability of relevant information about biological surrogates (Supplementary Material – Appendix A). We excluded 494 studies because they did not test the relationship between biological groups and, therefore, were out of the scope of this review. For example, we excluded 357 articles that tested the relationship between biological groups and environmental or spatial gradients and 58 articles that tested the relationship between genetic or molecular information from two groups of species. We included a study if it contained correlations between compositional (dis)similarities matrices (Mantel test), ordination scores (Procrustes analysis) or univariate correlations (e.g. Pearson or Spearman correlations between species richness or total abundance for two biological groups). Thus, we excluded 17 studies because they did not present sufficient information for the analysis and two studies that presented correlations based on co-inertia analysis. Eighty-six articles met all our selection criteria and were used in our meta-analysis (Fig. 1). Many papers included concordance measurements between several pairs of biological groups, resulting in 2939 effect sizes, which were analyzed in this study.

From each study, we retrieved different measures of concordance between biological groups (Mantel or Procrustes correlations for compositional data or coefficient of determination, Spearman and Pearson correlation for richness data). We also classified the studies according to the type of ecosystem (Aquatic or Terrestrial), taxonomic group [Animal, Vascular plants, Microorganisms and Mixed (when more than one of the previous categories were evaluated)] and to the type of attribute used in the analysis of concordance (relationship between total abundance or species richness of two biological groups, for simple correlations). For Mantel's test, we also classified the studies according to the numerical resolution of their data (presence-absence and abundance).

### 2.2. Data analysis

We took the square root of coefficients of determination ($R^2$) from studies that used simple linear regressions to rescale them to the Pearson's correlation coefficients. When a negative relationship was detected we multiplied this correlation by $-1$. This multiplication is needed because the inclusion of negative values would artificially reduce the average effect size. We converted Spearman rank correlation coefficients (for both univariate Spearman correlation and Spearman-based Mantel correlation) to Pearson's correlation coefficients following equations from Lajeunesse (2013).

We calculated Fisher's Z (and its variance) as a measure of effect size for studies that provided simple correlations (Borenstein et al., 2009). Some studies reported a correlation coefficient (r) of 1.0. In these cases, we converted these values to 0.99999 because an $r = 1.0$ corresponds to Fisher's Z = infinite. We used Mantel's $r_M$ and Procrustes $r_P$ as measures of effect size and sample size as a measure of precision for studies that provided multivariate measures of association (Rosenberg et al., 2013).

We calculated cumulative (average) effect sizes to quantify the magnitude of concordance between biological groups using random effect-models (Borenstein et al., 2009). A random-effects model assumes that the true effect is not the same across all studies (Borenstein et al., 2009; Nakagawa and Santos, 2012). These random-effects models consisted of multilevel meta-analysis to control for within-study dependence between effect sizes (Nakagawa and Santos, 2012). We modeled the dependence between effect sizes with a within-study random-effect term (see details in Nakagawa and Santos, 2012). For univariate data, effects sizes were weighted by the inverse of their variances, while the effect sizes of Mantel and Procrustes-based studies were weighted by studies' sample size. As in any meta-analysis, the goal of this procedure is to give more weight to more precise estimates (Borenstein et al., 2009; Koricheva and Gurevitch, 2013). We carried out separate analyses for univariate, Mantel and Procrustes data (Rosenberg et al., 2013).

We reported the between-study variance ($T^2$) as a measure of heterogeneity for univariate data (Borenstein et al., 2009; Senior et al., 2016). We assessed the effect of broad taxonomic categories, ecosystem type and numerical resolution (moderator variables or explanatory variables) on Fisher's Z variation (response variable) with a subgroup analysis. We also used the same subgroup model to assess variation in effect sizes for concordance measured with Mantel statistic (response variable). We did not explore variation in concordance as measured by the Procrustes statistic because few studies used this method. All analyses were conducted in R (R Core Team, 2015) with the *metafor* package (Viechtbauer, 2010).

## 3. Results

Among the studies included in our meta-analysis, 51 reported simple correlations, 31 used Mantel test and 20 utilized Procrustes analysis to measure the level of concordance between biological groups. Fisher's Z values varied from zero to 6.10 (mean $\pm$ SD: 0.46 $\pm$ 0.47), Mantel's $r_M$ varied from zero to 1.0 (0.34 $\pm$ 0.23) and Procrustes' $r_P$ varied from 0.04 to 0.98 (0.45 $\pm$ 0.20). Cumulative effect sizes were low, but significant for simple correlation [cumulative Fisher's Z $\pm$ 95% confidence interval ($CI_{95}$) = 0.50 $\pm$ 0.09, Z = 10.78, $P < .01$; Fig. 2] and multivariate analyses (cumulative