Ecological Applications, 22(6), 2012, pp. 1762–1771

Biological surrogacy in tropical seabed assemblages fails

P. R. Sutcliffe,[1,4] C. R. Pitcher,[2] M. J. Caley,[3] and H. P. Possingham[1]

[1]*School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072 Australia*
[2]*CSIRO Marine and Atmospheric Research, EcoSciences Precinct, 41 Boggo Road, Dutton Park, Queensland 4102 Australia*
[3]*Australian Institute of Marine Science, PMB Number 3, Townsville MC, Townsville, Queensland 4810 Australia*

Abstract. Surrogate taxa are used widely to represent attributes of other taxa for which data are sparse or absent. Because surveying and monitoring marine biodiversity is resource intensive, our understanding and management of marine systems will need to rely on the availability of effective surrogates. The ability of any marine taxon to adequately represent another, however, is largely unknown because there are rarely sufficient data for multiple taxa in the same region(s). Here, we defined a taxonomic group to be a surrogate for another taxonomic group if they possessed similar assemblage patterns. We investigated effects on surrogate performance of (1) grouping species by taxon at various levels of resolution, (2) selective removal of rare species from analysis, and (3) the number of clusters used to define assemblages, using samples for 11 phyla distributed across 1189 sites sampled from the seabed of Australia's Great Barrier Reef. This spatially and taxonomically comprehensive data set provided an opportunity for extensive testing of surrogate performance in a tropical marine system using these three approaches for the first time, as resource and data constraints were previously limiting. We measured surrogate performance as to how similarly sampling sites were divided into assemblages between taxa. For each taxonomic group independently, we grouped sites into assemblages using Hellinger distances and medoid clustering. We then used a similarity index to quantify the concordance of assemblages between all pairs of taxonomic groups. Surrogates performed better when taxa were grouped at a phylum level, compared to taxa grouped at a finer taxonomic resolution, and were unaffected by the exclusion of spatially rare species. Mean surrogate performance increased as the number of clusters decreased. Moreover, no taxonomic group was a particularly good surrogate for any other, suggesting that the use of any one (or few) group(s) for mapping seabed biodiversity patterns is imprudent; sampling several taxonomic groups appears to be essential for understanding tropical/subtropical seabed communities. Consequently, where resource constraints do not allow complete surveying of biodiversity, it may be preferable to exclude rare species to allow investment in a broader range of taxonomic groups.

Key words: *assemblage patterns; Australia; biodiversity; cluster; congruence; cross-taxon surrogates; Great Barrier Reef; invertebrates; marine communities; rarity; taxonomic assemblages; tropical seabed.*

Introduction

Shallow tropical marine habitats host some of the most species-rich ecosystems on Earth. Because these systems are complex and difficult to observe directly, quantifying this biodiversity and understanding the processes responsible for its assembly and maintenance continues to be a significant challenge (Hughes et al. 2002, Hawkins and Agrawal 2005, Rex et al. 2005), particularly where effective conservation and management in the face of environmental change is a primary goal. This situation is exacerbated further because the collection and analysis of such data is expensive, time consuming, and demands sophisticated taxonomic and analytical capacity.

Where data are lacking for some sets of species for any of the above reasons, surrogate taxa are often used to represent the overall patterns of species assemblages, richness and abundance, or develop conservation plans (Mellin et al. 2011). The use of surrogate taxa that may be easier to count or monitor assumes that those surrogates are representative. The few studies that have tested directly the efficacy of biological surrogacy in tropical marine systems (e.g., Beger et al. 2003, 2007) illustrate both the potential utility and limitations (Mellin et al. 2011) of biological surrogates for representing patterns of marine biodiversity where data are sparse and/or expensive to collect and our current lack of knowledge about how such surrogates might perform.

To test the efficacy of biological surrogates, it is necessary to first identify how they will be used. The intended application will then dictate the most appropriate method applied to identify surrogates. Where surrogates are used for conservation planning, metrics that measure representation of a given feature are best (e.g., Magierowski and Johnson 2006, Beger et al. 2007,

Manuscript received 2 June 2011; revised and accepted 28 February 2012; final version received 23 March 2012. Corresponding Editor: S. S. Heppell.
[4] E-mail: p.sutcliffe@uq.edu.au

Grantham et al. 2010, Johnson and Hering 2010), whereas, where the goal is to increase understanding of a system, metrics that measure congruence between patterns are preferred (e.g., Wilsey et al. 2005, Lovell et al. 2007, Fattorini 2010, Heino 2010, Qian and Kissling 2010). Previous studies have examined the effectiveness of surrogates in relation to conservation planning (e.g., Vanderklift et al. 1998, Ward et al. 1999, Gladstone 2002, Beger et al. 2003, 2007, Smith 2005), impact assessment (e.g., Olsgard and Somerfield 2000), community structure (e.g., Karakassis et al. 2006, Magierowski and Johnson 2006, Hirst 2008), and the use of habitats as biodiversity indicators (e.g., Mumby et al. 2008). These studies, while advancing what we know about surrogates, also highlight what remains to be learned and current impediments to achieving this knowledge.

Difficulties arise in selecting and testing biological surrogates because different quantity and quality of data are available for different taxa in the same areas. Where data are limited, samples of different taxa are often added together—"pooled." On coral reefs, for example, fishes and hard corals are often monitored because they are conspicuous and readily identifiable. Also, these groups of species are of particular interest because of their social, economic, and ecological values. Even in particularly well-studied groups, however, the degree of taxonomic resolution can vary considerably, leading to different levels of pooling in analyses of these systems. Further, while data sets for these fishes and corals are often large, data available for other groups against which they can be compared are typically very limited. As a result of limited data availability or taxonomic expertise for many invertebrate groups, these taxa are often grouped broadly either as "benthic invertebrates" (e.g., Vanderklift et al. 1998, Ward et al. 1999, Mumby et al. 2008) or by phylum (e.g., Echinodermata and Mollusca; Magierowski and Johnson 2006). Alternatively, the dominant fauna of a study area is assessed and used as a surrogate for other species in the same place (e.g., polychaetes in Olsgard and Somerfield 2000). This inconsistent pooling of taxa, often driven by data availability rather than any consideration of biological relevance, does not allow comprehensive evaluation of surrogate performance.

Inconsistent pooling of taxa in this way can also affect predictions derived from surrogates. For instance, the collection and analysis of increasingly comprehensive data sets for gastropod and bivalve mollusks in the Pacific and eastern North Atlantic changed our understanding of the geographic distribution of these groups from being similar to very different (Rex et al. 2005), indicating that some taxa should not be grouped at phylum level. Conversely, grouping with a taxonomic resolution that is too fine can severely limit the amount of data available for any particular taxon. Therefore, understanding the consequences of grouping taxa at different levels will be important for maximizing their effective and efficient application as surrogates.

Most species are rare (Gaston 1994, Lennon et al. 2004), and these rare species can also present difficulties for the analysis of assemblage data. Such difficulties are often circumvented by their exclusion from analyses (Clarke and Warwick 2001), but the exclusion of rare species from assemblage analyses can lead to an underestimation of the differences between assemblages (Cao et al. 1998). The exclusion of rare species when constructing surrogates may also compromise their performance by indicating greater similarity between surrogate and target taxa than actually exists. Moreover, the removal of rare species can be problematic, as distribution patterns of rare species can differ substantially between taxa, despite very similar patterns of total species richness (Grenyer et al. 2006). Therefore, it is important to quantify the effects of removing rare species on the performance of biological surrogates with respect to cross-taxon congruency of assemblage patterns in order to better understand how to construct surrogates that perform best.

In order to derive an estimate of a surrogate assemblage structure and apply it to a target assemblage, decisions must also be made about what constitutes an assemblage within the set of assemblages that comprise the ecosystem of interest. Commonly, clustering techniques are used to delineate assemblages (e.g., Proches 2005, Heikinheimo et al. 2007, Bandelj et al. 2009, Rueda et al. 2010). However, the best way to choose the number of clusters (corresponding to assemblage resolution) to describe an assemblage remains contentious. Methods proposed include the Calinski and Harabasz index (Calinski and Harabasz 1974), Krzanowski and Lai's index (Krzanowski and Lai 1988), and a range of approaches discussed in Milligan and Cooper (1985). More contemporary methods include the gap statistic (Tibshirani et al. 2001), "jump" method (Sugar and James 2003), Random Simulation Test (Guidi et al. 2009), cross validation techniques (Wang 2010), and a range of other methods referred to within these articles. Irrespective of the method used to define assemblages based on clustering, the relationship between how finely assemblages are resolved and the performance of surrogates has not been explored.

While of considerable importance to surrogate performance, a general lack of taxonomically and spatially comprehensive data sets for a large number of taxa in the same geographic region has largely precluded the study of surrogate performance and how it might be affected by such things as taxonomic resolution, the inclusion or exclusion of rare species, and how assemblages are operationally defined. A data set that provides an unprecedented opportunity for testing such effects was generated by the Great Barrier Reef Seabed Biodiversity Project, which surveyed almost 1400 sites on the Great Barrier Reef seabed between 2003 and 2006, collecting and identifying more than 5300 species from all major phyla (Pitcher et al. 2007).

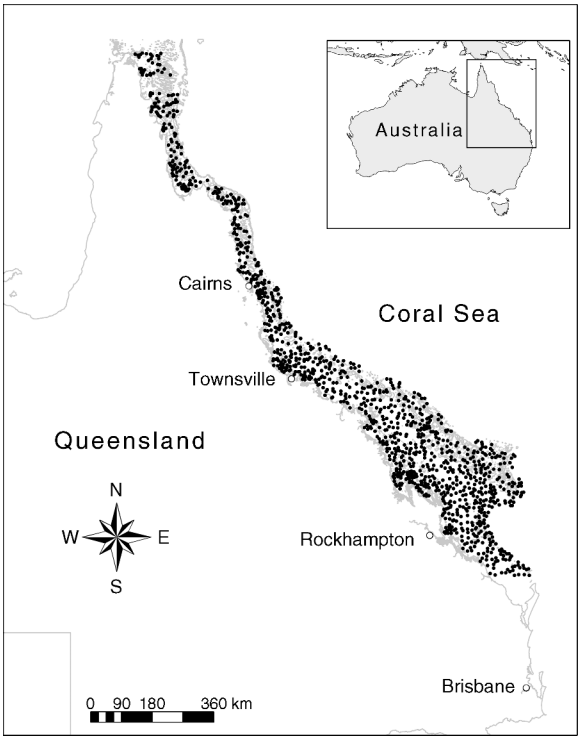

FIG. 1. The 1189 sites sampled by the epibenthic sled (solid dots) on the continental shelf seabed of the Great Barrier Reef, Australia.

Using these data, we explored three factors (taxonomic grouping, rare species removal, number of clusters) with the potential to influence the performance of cross-taxon surrogates for congruency of assemblage patterns. First, we assessed the effect of grouping species at different taxonomic levels to see which level of grouping produces the most effective cross-taxon surrogates. We did this by comparing surrogate performance where species were grouped by phylum, with surrogate performance where species were grouped at a more refined taxonomic level (class or order). Second, we quantified the effect of excluding rare species on surrogate performance by comparing results obtained using the whole assemblage against a range of results obtained by removing different numbers of rare species. Third, we quantified the effect of the scaling of assemblage resolution on surrogate performance by varying the number of clusters used to define assemblages on surrogate performance. Finally, we provide recommendations for the effective use of surrogates in tropical marine seabed ecosystems.

METHODS

Study area and sampling design

The Great Barrier Reef (GBR) World Heritage Area spans almost 350 000 km^2, of which ~7% is reef area, and 61% continental shelf seabed. The remainder is composed of islands, continental slope, or abyss (Great Barrier Reef Marine Park Authority 2009). The GBR shelf seabed was comprehensively sampled for the first time during the Great Barrier Reef Seabed Biodiversity Project from 2003 to 2006 (Pitcher et al. 2007). This project surveyed almost 1400 sites on the Great Barrier Reef seabed, using multiple sampling devices, collecting and identifying more than 5300 species from 15 phyla.

Sampling was predominantly limited to depths shallower than 80 m (except across the Capricorn Trough, which was sampled to ~105 m), deeper than 7 m near the coast, and deeper than 12 m over shoals. The sampling design was stratified based on analysis of 21 environmental variables on a 0.01 decimal degree grid weighted by biological importance (Pitcher et al. 2002). Sites (geographic coordinates) were chosen to optimally represent different strata, minimize spatial autocorrelation, and to be environmentally and spatially representative of the entire GBR seabed region. In this paper, we used data for specimens collected with an epibenthic sled at 1189 sites (Fig. 1). The sled was 1 m long, 1.5 m wide, and 0.5 m high, constructed of 20-mm square steel mesh and fitted with a 25-mm stretched net at the rear of the frame. Each tow was deployed at the preselected site coordinates and towed for 200 m at a speed of 2 knots (1 m/s). The sled data set comprised 70 860 site-by-species records of 4723 nominal species (i.e., operational taxonomic units, OTU) representing 49 classes from 15 phyla, which were identified, weighed, and recorded in the laboratory. Full details of the sampling design, field sampling, and laboratory processing are available in Pitcher et al. (2007).

Biological data

For our purposes, a subset of the sled data was selected that included only taxa with reliable species or OTU level identifications. Each taxonomic group was required to have been collected from at least 292 sites. This was the minimum number of sites needed to run our finest assemblage resolution cluster analysis (see *Methods: Number of clusters and surrogate performance*). The main taxonomic groups excluded by the application of these criteria were anemones, ascidians, cephalopods, crinoids, hard corals, hydroids, polychaetes, and zoanthids. The benthic taxa retained for our analysis (see Appendix A: Table A1) were effectively sampled by the benthic sled and were representative for assessing biological surrogacy for seabed environments.

Analysis

Site-by-species matrices were constructed for each taxonomic group. Biomass, rather than abundance, was used because counts of marine plants and colonial animals such as sponges and corals were not possible. Also abundance does not necessarily represent the relative importance of these species in assemblages where single colonies can cover considerable area.

All analyses were done using the R statistical computing environment (R Development Core Team

2009). For each taxonomic group considered, we used the Hellinger transformation with Euclidean distance (referred to as Hellinger distance; Legendre and Gallagher 2001) to calculate a matrix of dissimilarity values between sites. These were computed using functions `decostand` and `vegdist` in the R package vegan (Oksanen et al. 2007). All sites were then clustered based upon these dissimilarity values, using the partitioning-around-medoids clustering method (function `pam` in R package `cluster`; Maechler et al. 2005). To examine the effects of taxonomic grouping and rare species removal, the number of clusters was set to 16. This number of clusters was chosen as it was the number of clusters used to delineate assemblages of the Great Barrier Reef seabed (Pitcher et al. 2007). We did not force an equal number of sites into each cluster or assemblage. All sites with no data for any given taxonomic group were preassigned to a separate "zero" cluster.

We estimated surrogate performance by calculating the similarity of clusterings between pairs of taxonomic groups using the functions `confusion.matrix` and `similarity.index` in the R package `clv` (Nieweglowski 2009). Greater similarity between clusters of a pair of taxonomic groups was taken to indicate better surrogate performance. The confusion matrix enumerated the number of sites assigned to each pair of clusters between clusterings of any two taxa. From this confusion matrix, the similarity index calculated the maximum number of sites that were commonly allocated to clusters between pairwise sets of taxonomic groups, relative to the total number of sites. The labeling of clusters was arbitrary and without order, but is unimportant in calculating the similarity index and does not affect the result. Sites without data for both taxonomic groups in a pairwise comparison were excluded from the calculation of the similarity index.

The index returns a value of similarity between zero and one, with one being a perfect match between the assignment of sites to clusters for the two taxa being compared. In practice, however, the minimum value that can be obtained is greater than zero and depends on the number of clusters and the number of sites. Thus, to provide a "null" expectation against which clustering performance could be assessed, the similarity indices of a series of random assignments of sites to the same number of clusters, was also calculated. We then subtracted the null value from the similarity index for each pairwise comparison and rescaled all pairwise similarity values between 0 and 1. This scaling procedure provided an estimate of surrogate performance relative to a random surrogate. Therefore, any value of similarity greater than zero is better than random. A very good surrogate, however, would have a similarity index approaching one. Any value <0.5 is considered here to be a relatively poor surrogate.

The similarity index is best explained with a simple example containing five sites where fish and corals are present (Table 1). Based on a dissimilarity matrix calculated for fish, all five sites were allocated to cluster A or cluster B. Similarly for corals, sites were allocated to cluster X or Y. In this example, site 3 was placed in cluster B for fish and cluster X for corals, and the other four sites were similarly partitioned. The similarity index, S, computed as $S(P,P') = (A(P,P') - 1)/(N - 1)$, where $A(P,P') =$ maximum number of sites from taxa P (e.g., fish) and P' (e.g., corals), which are similarly partitioned, and $N =$ total number of sites, returns a value of 0.75 for the similarity of assemblage patterns between fish and corals. In this example, the mean null value from many randomized allocations of sites to clusters was 0.585, so the standardized index is $\sim$0.4.

TABLE 1. Hypothetical example of a simple cluster assignment of five sites for two taxa with a similarity index of 0.75 from the Great Barrier Reef, Australia.

Site	Fish	Corals
1	A	X
2	A	X
3	B	X
4	B	Y
5	B	Y

Notes: The five sites are allocated to cluster A or B based on cluster analysis of a dissimilarity matrix calculated from fish species composition data. Similarly, the same five sites are allocated to cluster X or Y by clustering a dissimilarity matrix from coral species composition. The values of cluster labels are arbitrary and irrelevant for computation of similarity. In this example, four of the five sites are similarly partitioned between the two taxa.

Taxonomic grouping and surrogate performance

To quantify the effects of grouping species at different taxonomic levels on surrogate performance, we tested the similarity between assemblage patterns of pairs of taxonomic groups for species aggregated at two levels.

First, we compared species grouped by phylum. Eleven phyla were analyzed, including Chordata (fishes and sharks), Arthropoda (crustaceans), Bryozoa (lace corals), Chlorophyta (green algae), Cnidaria (hard corals, soft corals, anemones, black corals, zoanthids, sea pens), Echinodermata (sea stars, sea cucumbers, brittle stars, urchins), Magnoliophyta (sea grasses), Mollusca (gastropods, bivalves, octopus, cuttlefish, squid), Phaeophyta (brown algae), Porifera (sponges), and Rhodophyta (red algae). A full list of phyla analyzed, with the number of species and sites represented, can be found in Appendix A: Table A2. Sites were clustered using data for each phylum, and cluster allocations between phyla were compared using the similarity index as a measure of surrogate performance. Boxplots were used to summarize surrogate performance for species grouped by phylum.

Second, we grouped species at class level, with three exceptions. This grouping at class level is hereafter referred to as "refined" taxonomic grouping. For a surrogate taxon to be practically applied, it must be

readily collectable. Therefore, the exceptions not aggregated at class level were species from the class Anthozoa, phylum Bryozoa, and kingdom Plantae. First, the order Alcyonacea (soft corals) was extracted from the class Anthozoa, which also included the orders Scleractinia (hard corals), Actinaria (anemones), and Zoantharia (zoanthids). These orders are easily separable in the field, making them usable as surrogate taxa, and contain species with fundamental ecological differences. For the specimens from the seabed, only the order Alcyonacea contained sufficient species for this analysis. Therefore the orders Scleractinia, Actinaria, and Zoantharia were excluded from our analyses. Second, the phylum Bryozoa was not split beyond phylum level, as identification to class level requires microscopic analysis. Third, the phyla Magnoliophyta (seagrass), Chlorophyta (green algae), Phaeophyta (brown algae), and Rhodophyta (red algae) were grouped as plants because class and phylum level identification of these phyla requires significant expertise.

A total of 12 groups was analyzed at this refined taxonomic resolution, including Class Actinopterygii (fishes), Order Alcyonacea (soft corals), Class Asteroidea (sea stars), Class Bivalvia (bivalves), Phylum Bryozoa (bryozoans), Class Crustacea (crustaceans), Class Demospongiae (sponges), Class Echinoidea (urchins), Class Gastropoda (gastropods), Class Holothuroidea (sea cucumbers), Class Ophiuroidea (brittle stars), and Kingdom Plantae (plants) (Appendix A: Table A1).

Sites were clustered based on dissimilarity values, using data for each taxonomic group separately. Similarity between clusterings of each pair of taxonomic groups was calculated. Boxplots were used to provide a summary of overall surrogate performance for species grouped by phylum and refined taxonomic structure, respectively.

Rare species and surrogate performance

To test the effect of excluding spatially rare species on surrogate performance, we removed species sampled at less than 1%, 2%, 4%, and 6% of sites from each of the refined taxonomic groups. This refined grouping was preferred over grouping at the phylum level, given the low similarity between taxonomic groups within phyla (see *Results*). These values of rare species removal were chosen because they cover the range of thresholds at which rare species have been suggested for removal in analyses of assemblages (e.g., Clarke and Warwick 2001). Sites where there were no representatives from either taxonomic group present (joint absences), could not be added to a cluster based on biological data and were removed from the analysis. Surrogate performance was summarized for each of the truncated assemblages using boxplots derived from pairwise comparisons of assemblage patterns between all taxonomic groups. We compared surrogate performance obtained using these truncated assemblages against surrogate performance obtained using the complete data set. A full list of taxa with corresponding number of species and sites represented for the complete data set and each of the truncated assemblages is provided in Appendix A: Table A1.

We repeated this analysis without excluding sites where no data were present for both taxonomic groups in the pairwise comparison. This was done to quantify the degree to which similarity between taxa could result from joint absences of groups among sites.

Number of clusters and surrogate performance

Surrogate performance was assessed using a range of numbers of clusters (c) to resolve assemblages corresponding to multiples $m = 2, 3, 4, 5, \ldots n$ of the average inter-site geographic distance (d), computed as $c = [A/(\mathrm{md})^2]$, where A is the area (km^2) of the GBR shelf.

Thus, the square root of the average area represented by each cluster type, for each number of clusters, corresponded to approximately equal steps in the average inter-site distance. The maximum number of clusters occurred at $m = 2$, and the maximum value of m occurred at $c = 2$ as two clusters was the minimum number of clusters for any of these analyses. This provided a range of 17 different scales of assemblage resolution for consideration, from fine through to coarse (including 292, 130, 73, 47, 32, 24, 18, 14, 12, 10, 8, 7, 6, 5, 4, 3, and 2 clusters). These multiple levels of assemblages subsuming other smaller assemblages do not represent different spatial scales in a strict nested sense, but rather are associated with a range of absolute spatial scales (e.g., see Kotliar and Wiens 1990) as sites being grouped were not necessarily contiguous. Sixteen clusters was also considered, as this was the number of clusters used to define the Great Barrier Reef seabed assemblage when species from all taxonomic groups were included in a previous analysis (Pitcher et al. 2007). Refined taxonomic grouping was used for this analysis due to low similarity between taxonomic groups from the same phylum, and all species were included. Pairwise similarity between all taxonomic groupings was calculated at each number of clusters.

Results

Taxonomic grouping and surrogate performance

Surrogates performed slightly better between phyla than between the refined taxonomic groups (Fig. 2), with number of clusters set at 16. At phylum level, the standardized pairwise similarity index ranged from 0.06 to 0.35 with a mean of 0.14. Within the refined taxonomic groups, the standardized pairwise similarity index among taxa ranged from 0.04 to 0.29, with a mean of 0.13. Similarities between taxonomic groups within phyla were also low. For example, the similarity between classes Bivalvia and Gastropoda, in the phylum Mollusca, was 0.08, highlighting substantial differences among assemblage patterns of different classes in the same phylum (highlighted in Fig. 2). This low similarity was true for all other classes within phyla (Appendix A: Table A3). No

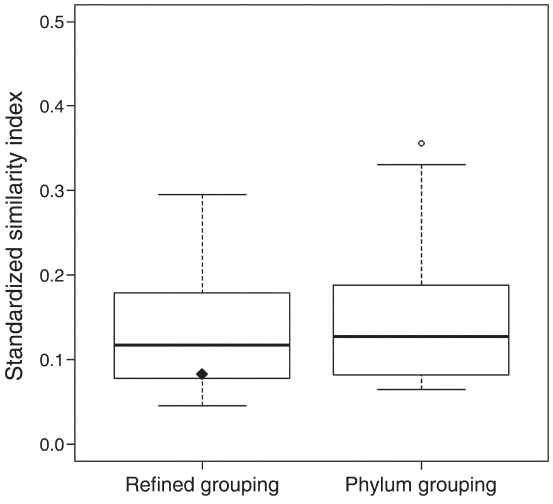

FIG. 2. Boxplots representing the range of standardized similarity index values for pairwise comparisons at two taxonomic resolutions: refined grouping (12 taxonomic groups) and phylum grouping (11 taxonomic groups), with number of clusters set at 16. The diamond represents the pairwise similarity between Mollusk classes Gastropoda and Bivalvia; the open symbol indicates an outlier. The boxes represent the lower quartile, median, and upper quartile. The whiskers span data points that fall 1.5 times the interquartile range from the box.

taxon was a good surrogate for any other taxon, regardless of how they were grouped taxonomically.

Rare species and surrogate performance

There was no significant change in surrogate performance when rare species were removed from the analysis and joint absences were excluded (Fig. 3A). Joint absences are sites where there are no representatives from either taxonomic group present; therefore they cannot be added to a cluster based on biological data. With all data included in the analysis, pairwise similarity in assemblage patterns was low, with the similarity index of all pairs of taxa <0.3 and the mean similarity of all pairwise comparisons 0.13. This did not improve significantly at any level of rare species removal. With maximum truncation, where species occurring at $<6\%$ of sites were removed from the analysis, similarity ranged from 0.06 to 0.27 with a mean of 0.13.

Conversely, when joint absences were included (i.e., sites with no data for either taxon) in the similarity calculation, the removal of rare species did increase surrogate performance (Fig. 3B). The exclusion of species occurring at $<1\%$ of sites more than doubled the mean similarity between taxa to 0.31. When we increased exclusion to 6%, the mean similarity between taxa increased to 0.65, with a maximum similarity of 0.85. It was therefore the increasing number of sites with no data and their overlap, rather than an agreement of assemblage patterns between taxa where data remained, that increased similarity when the threshold was increased. Removing rare species did not affect overall surrogate performance.

Number of clusters and surrogate performance

Surrogate performance gradually increased as the number of clusters decreased (Fig. 4), corresponding to coarser resolution of assemblages and larger absolute scales. However, pairwise similarity between any pair of taxa fluctuated as the number of clusters changed, demonstrating that the number of clusters defining an

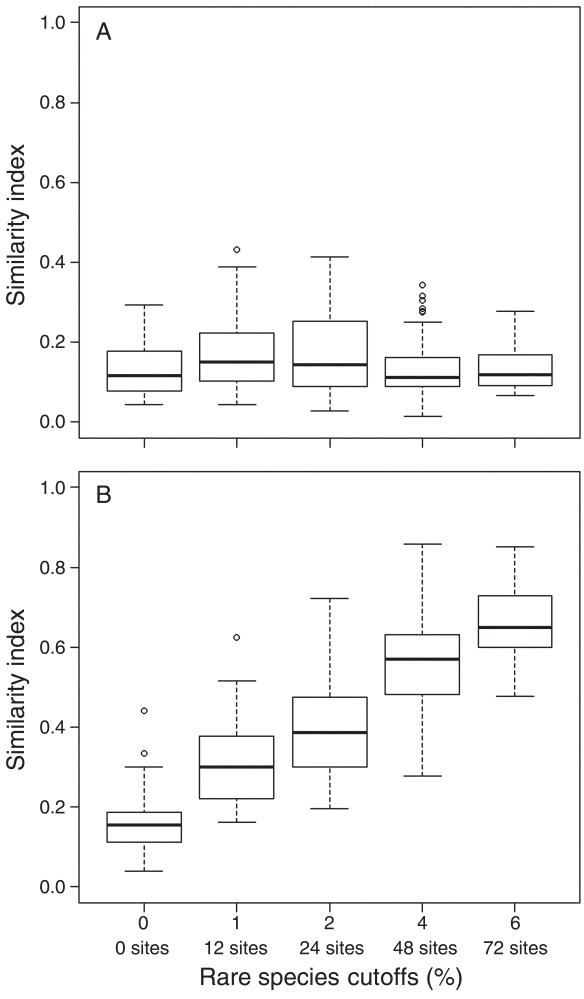

FIG. 3. Each boxplot represents the range of standardized similarity index values between pairs of refined taxonomic groups (12 groups) at a different threshold for rare species removal: 0% (3806 species; complete community), 1% (845 species; community truncated to exclude those species found at $<1\%$, or 12, of the sites), 2% (494 species; excludes species at $<2\%$, or 24, of the sites), 4% (237 species; excludes species at $<4\%$, or 48, of the sites), 6% (148 species; excludes species at $<6\%$, or 72, of the sites) where (A) joint absences are excluded from the analysis and (B) joint absences are included in the analysis. Joint absences are sites where no representatives from either taxonomic group are present. Open symbols indicate outliers. The boxes represent the lower quartile, median, and upper quartile. The whiskers span data points that fall 1.5 times the interquartile range from the box.

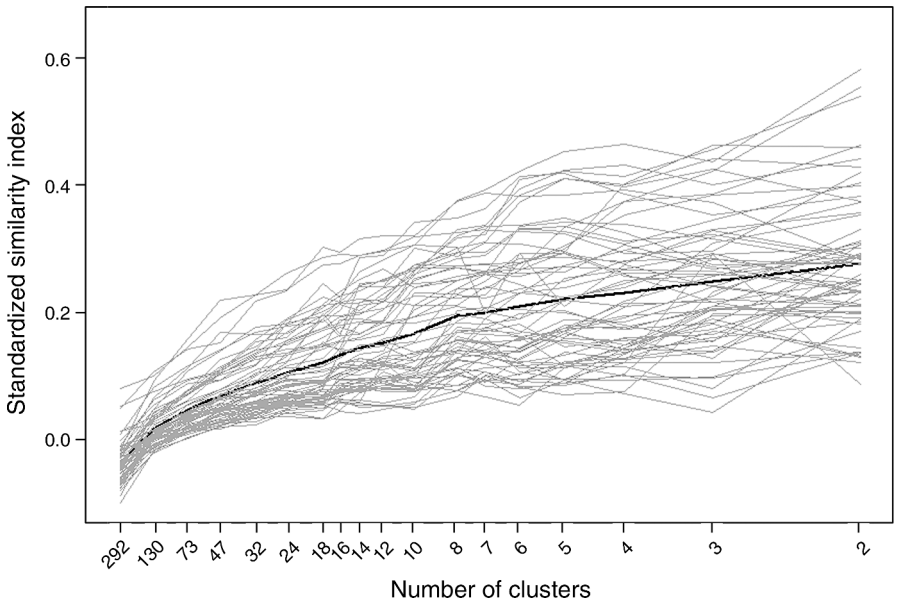

FIG. 4. Standardized similarity index values between pairs of refined taxonomic groups (gray), and average standardized similarity (bold), at a range of numbers of clusters. As the number of clusters decreases, the resolution of assemblages and absolute scale increases. Two clusters is a very large-scale community. That is why the scale goes from large to small numbers of clusters. The scale corresponds to multiples of the average inter-site geographic distance (see *Methods*).

assemblage affected surrogate performance. Surrogate performance was better than random for only five pairs of taxa at the maximum number of 292 clusters. Maximum standardized pairwise similarity was 0.58 when assemblages were divided into two clusters.

Fishes, which are a relatively commonly studied taxon, did not perform well as surrogates (Appendix A: Table A3), with an average similarity of 0.13 at 16 clusters. Average surrogate performance for fishes was worse overall than five other taxonomic groups (Table 2). The taxa with the most dissimilar patterns to all other taxa were Bryozoa (bryozoans), Asteroidea (sea stars), and Echinoidea (urchins). Demospongiae (sponges) and Holothurians (sea cucumbers) had the greatest pairwise similarity (0.29; Appendix A: Table A3) at 16 clusters. Average similarity at 16 clusters, calculated for each taxonomic group as a measure of overall surrogate performance (Table 2), showed that Holothurians (sea cucumbers) had the highest average pairwise similarity (0.2) of all taxonomic groups. This value, however, indicated low surrogate performance. Overall, surrogate performance increased as the number of clusters decreased; however, taxa that performed best as surrogates changed as the number of clusters changed. Importantly, the most readily available data (e.g., fish) were not good surrogates for any other taxon.

DISCUSSION

Understanding surrogate performance is important for identifying knowledge gaps, reserve system design, and designing survey programs that have comprehensive biological representation in marine or terrestrial systems. The purpose of our study was to assess surrogate performance in a tropical seabed system and to quantify the effect of varying factors that may affect surrogate performance. We showed that surrogate taxa do not reflect assemblage patterns of any other tropical seabed taxon regardless of taxonomic grouping or whether rare species are included. This suggests that taxonomically comprehensive studies that exclude rare species would provide a better understanding of seabed assemblage patterns than studies that focus on fewer taxonomic groups and sample rare species. This is important for future studies where a cost/benefit trade-off decision must be made.

TABLE 2. Summary of the surrogate performance of each taxonomic group in the "refined" analysis, quantified by the average standardized similarity index (SI) across all pairwise comparisons.

Taxon	Average SI, 16 clusters	Average SI, 2 clusters
Sea cucumbers (Holothuroidea)	0.20	0.35
Soft corals (Alcyonacea)	0.16	0.33
Crustaceans (Crustacea)	0.16	0.42
Sponges (Demospongiae)	0.15	0.45
Bivalves (Bivalvia)	0.14	0.35
Fish (Actinopterygii)	0.13	0.39
Brittle stars (Ophiuroidea)	0.13	0.41
Plants (Plantae)	0.12	0.33
Gastropods (Gastropoda)	0.09	0.31
Urchins (Echinoidea)	0.09	0.27
Sea stars (Asteroidea)	0.08	0.31
Lace corals (Bryozoa)	0.07	0.31

Note: These data were compiled for 16 clusters and two clusters, respectively (for complete results of pairwise similarity for all taxa, see Appendix A: Table A3).

Our results demonstrate that assemblage patterns of the seabed fauna on the GBR differed between phyla and between refined taxonomic groups, thus limiting their utility to act as surrogates for one another. These results therefore indicate that attempts to simplify the assessment of such assemblages by grouping taxa broadly by phylum (e.g., Beger et al. 2003), or groups of phyla (e.g., benthic invertebrates; Mumby et al. 2008) may compromise surrogate effectiveness and confound assemblage patterns. In contrast to our results demonstrating poor performance of cross-taxa surrogacy, two studies of Indo-Pacific coral reefs (Beger et al. 2003, 2007) found that fishes and corals were adequate surrogates for corals and mollusks, respectively. In the Caribbean, Mumby et al. (2008) found fish species were good surrogates for benthic species; however, benthic species were not good surrogates for fishes. Our study does not support the use of fish assemblage patterns as a surrogate for assemblage patterns of any other taxonomic group in the GBR seabed ecosystems. The contrast between our results and findings of these previous studies is potentially due to the large number of species represented in our study. Also, this study of the tropical seabed represents a different biome to tropical coral reefs, where the relationships and dependences between groups differ from those of coral reef organisms. This study reflects the differences in complex assemblage patterns between taxonomic groups on the seabed.

Our results are also in contrast to findings in temperate marine regions. Mollusks were found to be good surrogates for overall species richness (Smith 2005) and in selecting representative areas for conservation (Gladstone 2002) in temperate rocky intertidal regions; however they performed poorly as surrogates for assemblage patterns in our tropical seabed system. Fish were good surrogates in developing comprehensive reserve systems in temperate shallow-water regions (Ward et al. 1999), which was also not supported by our study. These differing results highlight that surrogate performance cannot be extrapolated into different habitats or different trophic levels than where the study was undertaken, and that caution must be taken when defining surrogates in new regions.

In terrestrial systems, large data sets are available and extensive testing of surrogates has found some support for the use of biological (e.g., Lund and Rahbek 2002) and environmental (e.g., Ferrier 2002, Sarkar et al. 2005) surrogates. In tropical systems, it is thought that taxa with fine-scale distribution and high richness are good surrogates for other less diverse, widely distributed species (Moritz et al. 2001). We found no support for this in tropical marine seabed systems, where sponges and crustaceans were the taxa with the highest diversity; however, sea cucumbers and soft corals with relatively low diversity had the highest overall similarity of assemblage patterns (Table 2). Low surrogate performance in tropical marine seabed assemblages is likely due to a combination of the higher diversity in tropical systems compared to temperate systems, as well as the use of complex assemblage patterns rather than hotspots or reserve design, to test surrogate performance.

The lack of congruence demonstrated here, between assemblage patterns of different taxonomic groups, highlights the need for taxonomically broad-based data collection. Taxonomic groups that are difficult to identify or expensive to census will not necessarily be well represented by another taxonomic group. Our understanding of assemblage patterns within an ecosystem will remain incomplete where this is true and where obstacles to direct estimation of these taxa cannot be overcome.

While rare species can confuse the interpretation of assemblage patterns (Clarke and Warwick 2001), the exclusion of rare species here made little difference to surrogate performance with respect to assemblage patterns (Fig. 3A). The exclusion of rare species, however, limited the number of sites at which particular taxonomic groups were represented. When joint absences of pairs of taxa from sites were included in calculation of the similarity indices between taxa, the exclusion of rare species did affect surrogate performance and increased similarity between taxonomic groups (Fig. 3B). This should not be incorrectly interpreted as indicating that we would achieve more similar assemblage patterns and increased surrogate performance as the rare species threshold was raised. The assemblage patterns of the more common species were dissimilar, and the assemblage patterns displayed when all species were included were dissimilar; however, there was increasing agreement of sites that contained no data for any pair of taxonomic groups, as rare species were removed. This agreement of sites with no data for any given pair of taxa does not increase surrogate performance with respect to assemblage patterns; similarity of assemblage patterns between taxonomic groups remains low at all levels of rare species removal. This suggests that a survey design that includes a comprehensive set of taxonomic groups may not necessarily be compromised by excluding the rarest species in the assemblages being assessed. Indeed it suggests that rather than comprehensively survey a small number of phyla, we may be better off sampling more phyla less well. The exclusion of rare species, which are difficult to detect and thus expensive to survey effectively, will substantially increase the efficiency of studies where assessing assemblage patterns is the primary goal.

The number of clusters used to define the resolution and absolute scale of assemblages affected surrogate performance. Surrogate performance, however, remained low even when very few clusters were used. Consequently, no useful surrogate taxa could be identified. The fluctuation of pairwise similarities as the number of clusters changed (Fig. 4) indicates that biological surrogates are unstable with respect to changes in the resolution and scale used to define

assemblages. While no taxonomic group was a good surrogate for any other, irrespective of the number of clusters used, similarity was greatest when taxonomic groups were divided into two clusters. When the number of sites in each of the two clusters was examined, we found that sponges (Demospongiae), fishes (Actinopterygii), brittle stars (Ophiuroidea), and crustaceans (Crustacea) clustered most sites into a single cluster (Appendix B) with very few sites in the second cluster. In contrast, urchins (Echinoidea), gastropods (Gastropoda), sea stars (Asteroidea), and bryozoans (Bryozoa) divided sites more evenly divided between the two clusters (Appendix C).

The results we report here form a spatially and taxonomically comprehensive assessment of the performance of biological surrogates in a tropical marine system that is unprecedented in scale, which can help guide the design of future sampling programs with the aim of assessing assemblage patterns using biological surrogates. Species should not be combined at phylum level or higher to assess assemblage patterns. In addition, large expenditure to achieve robust estimation of rare species would be misguided given that their exclusion had no discernable effect on surrogate performance. We recommend surveys of tropical systems to include sampling of all taxonomic groups in inter-reef seabed areas, even if this has to be at the expense of rare species, to gain a more complete understanding of these complex and biodiverse ecosystems, and to investigate the inclusion of fewer sites in a survey rather than fewer taxa where resources do not allow for comprehensive sampling of taxa and space. The generality of these recommendations, however, will need to be tested beyond the GBR seabed ecosystem before they are routinely applied. Our understanding of surrogate performance would be enhanced if the congruence between richness and abundance patterns among a range of taxa were tested, as this may reveal different results compared to assemblage pattern surrogates. Additionally, the environmental drivers of community patterns, and the similarity between drivers for different taxonomic groups is another avenue recommended for further research.

Acknowledgments

We thank the many scientists who participated in the Great Barrier Reef Seabed Biodiversity Project, a collaboration between the Australian Institute of Marine Science (AIMS), the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Queensland Department of Primary Industries and Fisheries (QDPIF), and the Queensland Museum (QM); funded by the CRC Reef Research Centre, the Fisheries Research and Development Corporation, and the National Oceans Office. The project was led by R. Pitcher (Principal Investigator, CSIRO), P. Doherty (AIMS), J. Hooper (QM), and N. Gribble (QDPIF). We also thank the field team and crew of the *RV Lady Basten* (AIMS) that collected the sled samples. This work was supported by the Marine Biodiversity Research Hub, a collaborative partnership between the University of Tasmania, CSIRO Wealth from Oceans Flagship, Geoscience Australia, Australian Institute of Marine Science, and Museum Victoria, and funded by the Commonwealth Environmental Research Facilities (CERF) programme, an Australian Government initiative. We thank members of the Hub for their contributions during workshops, especially W. Venables and N. Ellis. We are grateful for the comments provided by two anonymous reviewers, which enhanced this paper considerably.

Literature Cited

Bandelj, V., D. Curiel, S. Lek, A. Rismondo, and C. Solidoro. 2009. Modelling spatial distribution of hard bottom benthic assemblages and their functional response to environmental parameters. Ecological Modelling 220:2838–2850.

Beger, M., G. P. Jones, and P. L. Munday. 2003. Conservation of coral reef biodiversity: a comparison of reserve selection procedures for corals and fishes. Biological Conservation 111:53–62.

Beger, M., S. A. McKenna, and H. P. Possingham. 2007. Effectiveness of surrogate taxa in the design of coral reef reserve systems in the Indo-Pacific. Conservation Biology 21:1584–1593.

Calinski, R. B., and J. Harabasz. 1974. A dendrite method for cluster analysis. Communications in Statistics 3:1–27.

Cao, Y., D. D. Williams, and N. E. Williams. 1998. How important are rare species in aquatic community ecology and bioassessment? Limnology and Oceanography 43:1403–1409.

Clarke, K. R., and R. M. Warwick. 2001. Changes in marine communities: an approach to statistical analysis and interpretation. Second edition. PRIMER-E, Plymouth, UK.

Fattorini, S. 2010. Biotope prioritisation in the central Apennines (Italy): species rarity and cross-taxon congruence. Biodiversity and Conservation 19:3413–3429.

Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? Systematic Biology 51:331–363.

Gaston, K. J. 1994. Rarity. Chapman and Hall, London, UK.

Gladstone, W. 2002. The potential value of indicator groups in the selection of marine reserves. Biological Conservation 104:211–220.

Grantham, H. S., R. L. Pressey, J. A. Wells, and A. J. Beattie. 2010. Effectiveness of biodiversity surrogates for conservation planning: different measures of effectiveness generate a kaleidoscope of variation. PLoS ONE 5(7):e11430.

Great Barrier Reef Marine Park Authority. 2009. Great Barrier Reef outlook report 2009. Great Barrier Reef Marine Park Authority, Townsville, Australia.

Grenyer, R., et al. 2006. Global distribution and conservation of rare and threatened vertebrates. Nature 444:93–96.

Guidi, L., F. Ibanez, V. Calcagno, and G. Beaugrand. 2009. A new procedure to optimize the selection of groups in a classification tree: applications for ecological data. Ecological Modelling 220:451–461.

Hawkins, B. A., and A. A. Agrawal. 2005. Latitudinal gradients. Ecology 86:2261–2262.

Heikinheimo, H., M. Fortelius, J. Eronen, and H. Mannila. 2007. Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. Journal of Biogeography 34:1053–1064.

Heino, J. 2010. Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? Ecological Indicators 10:112–117.

Hirst, A. J. 2008. Surrogate measures for assessing cryptic faunal biodiversity on macroalgal-dominated subtidal reefs. Biological Conservation 141:211–220.

Hughes, T. P., D. R. Bellwood, and S. R. Connolly. 2002. Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. Ecology Letters 5:775–784.

Johnson, R. K., and D. Hering. 2010. Spatial congruency of benthic diatom, invertebrate, macrophyte, and fish assemblages in European streams. Ecological Applications 20:978–992.

Karakassis, I., A. Machias, P. Pitta, K. N. Papadopoulou, C. J. Smith, E. T. Apostolaki, M. Giannoulaki, D. Koutsoubas, and S. Somarakis. 2006. Cross-community congruence of patterns in a marine ecosystem: Do the parts reflect the whole? Marine Ecology Progress Series 310:47–54.

Kotliar, N. B., and J. A. Wiens. 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59:253–260.

Krzanowski, W. J., and Y. T. Lai. 1988. A criterion for determining the number of groups in a data set using sum-of-squares clustering. Biometrics 44:23–34.

Legendre, P., and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

Lennon, J. J., P. Koleff, J. J. D. Greenwood, and K. J. Gaston. 2004. Contribution of rarity and commonness to patterns of species richness. Ecological Letters 7:81–87.

Lovell, S., M. Hamer, R. Slotow, and D. Herbert. 2007. Assessment of congruency across invertebrate taxa and taxonomic levels to identify potential surrogates. Biological Conservation 139:113–125.

Lund, M. P., and C. Rahbek. 2002. Cross-taxon congruence in complementarity and conservation of temperate biodiversity. Animal Conservation 5:163–171.

Maechler, M., P. Rousseeuw, A. Struyf, and M. Hubert. 2005. Cluster analysis basics and extensions. http://cran.r-project.org/

Magierowski, R. H., and C. R. Johnson. 2006. Robustness of surrogates of biodiversity in marine benthic communities. Ecological Applications 16:2264–2275.

Mellin, C., S. Delean, J. Caley, G. Edgar, M. Meekan, R. Pitcher, R. Przeslawski, A. Williams, and C. Bradshaw. 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. PLoS ONE 6(6):e20141.

Milligan, G. W., and M. C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–179.

Moritz, C., K. S. Richardson, S. Ferrier, G. B. Monteith, J. Stanisic, S. E. Williams, and T. Whiffin. 2001. Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota. Proceedings of the Royal Society B 268:1875–1881.

Mumby, P. J., K. Broad, D. R. Brumbaugh, C. P. Dahlgren, A. R. Harborn, A. Hastings, K. E. Holmes, C. V. Kappel, F. Micheli, and J. N. Sanchirico. 2008. Coral reef habitats as surrogates of species, ecological functions, and ecosystem services. Conservation Biology 22:941–951.

Nieweglowski, L. 2009. Clv: cluster validation techniques. http://cran.r-project.org/

Oksanen, J., R. Kindt, P. Legendre, and R. B. O'Hara. 2007. Vegan: community ecology package. http://cran.r-project.org/

Olsgard, F., and P. J. Somerfield. 2000. Surrogates in marine benthic investigations—which taxonomic unit to target? Journal of Aquatic Ecosystem Stress and Recovery 7:25–42.

Pitcher, C. R., et al. 2007. Seabed biodiversity on the continental shelf of the Great Barrier Reef World Heritage Area. AIMS/CSIRO/QM/QDPI CRC Reef Research Task Final Report 320. CSIRO, Brisbane, Australia.

Pitcher, C. R., W. Venables, N. Ellis, I. McLeod, F. Pantus, M. Austin, M. Cappo, P. Doherty, and N. Gribble. 2002. GBR seabed biodiversity mapping project: phase 1 report to CRC-Reef. CSIRO/AIMS/QDPI Report. http://www.reef.crc.org.au/resprogram/programC/seabed/Seabedphase1rpt.htm

Proches, S. 2005. The world's biogeographical regions: cluster analyses based on bat distributions. Journal of Biogeography 32:607–614.

Qian, H., and W. D. Kissling. 2010. Spatial scale and cross-taxon congruence of terrestrial vertebrate and vascular plant species richness in China. Ecology 91:1172–1183.

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rex, M. A., J. A. Crame, C. T. Stuart, and A. Clarke. 2005. Large-scale biogeographic patterns in marine mollusks: a confluence of history and productivity? Ecology 86:2288–2297.

Rueda, M., M. A. Rodriguez, and B. A. Hawkins. 2010. Towards a biogeographic regionalization of the European biota. Journal of Biogeography 37:2067–2076.

Sarkar, S., J. Justus, T. Fuller, C. Kelley, J. Garson, and M. Mayfield. 2005. Effectiveness of environmental surrogates for the selection of conservation area networks. Conservation Biology 19:815–825.

Smith, S. D. A. 2005. Rapid assessment of invertebrate biodiversity on rocky shores: where there's a whelk there's a way. Biodiversity and Conservation 14:3565–3576.

Sugar, C. A., and G. M. James. 2003. Finding the number of clusters in a dataset: an information-theoretic approach. Journal of the American Statistical Association 98:750–763.

Tibshirani, R., G. Walther, and T. Hastie. 2001. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society B 63:411–423.

Vanderklift, M. A., T. J. Ward, and J. C. Phillips. 1998. Use of assemblages derived from different taxonomic levels to select areas for conserving marine biodiversity. Biological Conservation 86:307–315.

Wang, J. H. 2010. Consistent selection of the number of clusters via crossvalidation. Biometrika 97:893–904.

Ward, T. J., M. A. Vanderklift, A. O. Nicholls, and R. A. Kenchington. 1999. Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. Ecological Applications 9:691–698.

Wilsey, B. J., D. R. Chalcraft, C. M. Bowles, and M. R. Willig. 2005. Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. Ecology 86:1178–1184.

Supplemental Material

Appendix A

Tables outlining the number of species and sites included in each of the phylum and refined taxonomic groupings, and similarity indices for pairwise comparisons of refined taxonomic groups at the resolution of two and 16 clusters (*Ecological Archives* A022-094-A1).

Appendix B

Maps of the cluster allocation at each site when number of clusters was set at 16, for each of the 12 refined taxonomic groups (*Ecological Archives* A022-094-A2).

Appendix C

Maps of the cluster allocation at each site when number of clusters was set at two, for each of the 12 refined taxonomic groups (*Ecological Archives* A022-094-A3).

Ecological Applications

Title Details | Table of Contents

Basic Description

Title	Ecological Applications
ISSN	1051-0761
Publisher	Ecological Society of America
Country	United States
Status	Active
Start Year	1991
Frequency	8 times a year
Language of Text	Text in: English
Refereed	Yes
Abstracted / Indexed	Yes
Serial Type	Journal
Content Type	Academic / Scholarly
Format	Print
Website	http://www.esajournals.org/loi/ecap
Description	Publishes articles on the ecological basis for decision making and policy in the areas of global change, biogeochemistry, conservation biology, ecotoxicology, pollution ecology, fisheries and wildlife ecology, forestry, agroecosystems, range management, soils, hydrology and groundwater, landscape ecology, and epidemiology.

Subject Classifications

Additional Title Details

Publisher & Ordering Details

Commercial Publisher General

Ecological Society of America
Address: 1990 M St, NW, Ste 700, Washington, DC 20036 United States
Phone: 202-833-8773
Fax: 202-833-8775
Website: http://www.esa.org
Email: esahq@esa.org

Corporate Author General

Ecological Society of America
Address: 1990 M St, NW, Ste 700, Washington, DC 20036 United States
Phone: 202-833-8773
Fax: 202-833-8775
Website: http://www.esa.org
Email: esahq@esa.org

Price Data

▶ **Online Availability**

▶ **Abstracting & Indexing**

▶ **Other Availability**

▼ **Demographics**

Audience	academic
	special adult
Ad Contact	Eric Gordon, Ecological Society of America

▼ **Reviews**

Ecological Applications ,an Ecological Society of America journal, stands apart from most other ecology journals since articles "explicitly discuss the applications or implications of the work in regard to policy, management, or the analysis and solution to major environmental problems." The journal contains articles describing original and significant research and communications of short length for urgent application and scientific debate. There are also invited features, forums, and letters to the editor. The journal is a core title for university and special libraries with scholars and practitioners in the ecological and environmental science fields. It is available online and in print. URL: http://esapubs.org (LaBonte, Kristen)
ML 01-12-2012