Syst. Biol. 52(5):717–723, 2003

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150390235584

Biotic Element Analysis in Biogeography

BERNHARD HAUSDORF[1] AND CHRISTIAN HENNIG[2]

[1]*Zoologisches Museum der Universität Hamburg, Martin-Luther-King-Platz 3, 20146 Hamburg, Germany; E-mail: hausdorf@zoologie.uni-hamburg.de*
[2]*ETH Zürich (LEO), Seminar für Statistik, 8092 Zürich, Switzerland; E-mail: hennig@stat.math.ethz.ch, and Universität Hamburg, Fachbereich Mathematik–SPST, Bundesstrasse 55, 20146 Hamburg, Germany*

Abstract.— Biotic element analysis is an alternative to the areas-of-endemism approach for recognizing the presence or absence of vicariance events in a given region. If an ancestral biota was fragmented by vicariance events, biotic elements or clusters of distribution areas should emerge. We propose a statistical test for clustering of distribution areas based on a Monte Carlo simulation with a null model that considers the spatial autocorrelation in the data. The hypothesis tested is that the observed degree of clustering of ranges can be explained by the range size distribution, the varying number of taxa per cell, and the spatial autocorrelation of the occurrences of a taxon alone. A method for the delimitation of biotic elements which uses model-based Gaussian clustering is introduced. We demonstrate our methods and show the importance of grid size by means of a case study, an analysis of the distribution patterns of southern African species of the weevil genus *Scobius*. The example highlights the difficulties in delimiting areas of endemism if dispersal has occurred and illustrates the advantages of the biotic element approach. [Area of endemism; biogeography; biotic elements; null model; *Scobius*; South Africa; vicariance.]

According to the vicariance model (Croizat et al., 1974; Rosen, 1976, 1978; Platnick and Nelson, 1978; Nelson and Platnick, 1981; Wiley, 1988; Humphries and Parenti, 1999), an ancestral biota was fragmented by the appearance of a barrier. The barrier interrupted the gene flow between the populations separated by the barrier and, consequently, this vicariance event resulted in allopatric speciation of many of the species formerly constituting the ancestral biota. In this way, two new biotas originate, separated by the barrier. By repetitions of this process, areas of endemism with distinct biotas, i.e., with many species restricted to the individual areas, emerge. On average, the ranges of the species that have originated in the same area of endemism will be more similar to each other than to ranges of species that have originated in other areas of endemism. Thus, the vicariance model predicts a nonrandom congruence of species ranges (Morrone, 1994; Hausdorf, 2002). The existence of such a distribution pattern has not been tested in most biogeographic analysis. Only recently, Mast and Nyffeler (2003) introduced a test for the nonrandom co-occurrence of pairs of taxa into the vicariance biogeography framework.

It must be emphasized that the pattern for which we can examine distribution data is nonrandom congruence, i.e., clustering of ranges. Areas of endemism as such cannot be found in distribution data. They can only be inferred from clusters of species ranges. This distinction is obscurred in the operational methods for identifying and delimiting areas of endemism (Morrone, 1994; Linder, 2001; Szumik et al., 2002; Mast and Nyffeler, 2003).

The delimitation of areas of endemism is not problematic as long as species do not disperse across the barriers separating the areas of endemism. Under these conditions, the borders of the areas of endemism can be drawn between the range clusters. However, usually there is stochastic dispersal of species across barriers with time. Moreover, many barriers weaken or disappear with time. If dispersal across the barriers that separated the areas of endemism resulted in an overlap of ranges of species which originated in different areas of endemism, biogeographical data alone are insufficient for delimiting areas of endemism (Hausdorf, 2002).

However, biotic elements, i.e., groups of taxa whose ranges are more similar to each other than to those of taxa of other such groups (Hausdorf, 2002), can be recognized, even if some of the species that originated by vicariance dispersed across the barriers that separated the areas of endemism. Previous computational approaches for the determination of biotic elements did not provide a decision about how many biotic elements should be recognized and which species cannot be adequately assigned to any biotic element (e.g., Holloway and Jardine, 1968; Jardine, 1972; Dennis et al., 1998).

We propose a statistical test for clustering of distribution areas based on a Monte Carlo simulation and a method for the delimitation of biotic elements which uses model-based Gaussian clustering.

MATERIAL AND METHODS

Distribution Data

As a case study, we analyzed the ranges of the 47 southern African weevil species of the genus *Scobius* Schoenherr, because this example was previously used to introduce three other methods for the analysis of distribution patterns (Morrone, 1994; Szumik et al., 2002; Mast and Nyffeler, 2003).

We used Morrone's (1994) presence/absence matrix for 2° latitude × 2° longitude grid cells as corrected by Mast and Nyffeler (2003) as well as the presence/absence matrix for 1° latitude × 1° longitude grid cells generated by Mast and Nyffeler (2003).

Test for Clustering of Distribution Areas

As first step in the analysis, we investigated whether distribution areas are clustered. We test the hypothesis that the observed degree of clustering of ranges can be

explained by the varying number of taxa per cell and the spatial autocorrelation of the occurrences of a taxon alone.

This test is based on distances between the ranges of the examined taxa. Three specifications must be made: a distance measure, a test statistic, and a null model for the generation of sets of ranges. The distribution of the test statistics under the null hypothesis is approximated by Monte Carlo simulation, which depends on some parameters which are estimated from the data. Details of the test procedure are described by Hennig and Hausdorf (in press).

The tests as well as the subsequently described method for the delimitation of biotic elements is implemented in the program package PRABCLUS which is an add-on package for the statistical software R. These programs are available at http://cran.r-project.org.

Distance measure.—$R = \{1, \ldots, c\}$ is a set of geographic units (e.g., $c = 25$ cells). The range of a taxon A is the subset of R where the taxon is present. Let $|A|$ denote the number of elements of a set A, e.g., the number of cells occupied by the taxon. If $A_1, A_2 \subseteq R$ are the ranges of two taxa, $|A_1 \cap A_2|$ is the number of cells where both taxa are present.

Often, the Jaccard distance d_J is used in biogeographic analyses (e.g., Jardine, 1972; Dennis et al., 1998):

$$d_J(A_1, A_2) = 1 - \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|}. \quad (1)$$

We found that this distance measure is not appropriate for our analyses because it is approximately equal to 1 if one range is much smaller than the other, even if the smaller range is a subset of the larger one. However, it might well be that one species occupies only a small part of an area of endemism (e.g., because it is restricted to a rare habitat), whereas another species that originated in the same area of endemism has a much larger range because it occurs in several habitats. In such a case, the distance between the ranges should not approach 1, because we would like to recognize that they belong to the same biotic element. Therefore the Jaccard distance is misleading because it relates the common distribution area only to the size of the more widely distributed species (which equals the union if the smaller area is a subset of the larger one).

We used the Kulczynski distance d_K (1 − "Kulczynski unnamed 2"; Shi, 1993):

$$d_K(A_1, A_2) = 1 - \frac{1}{2}\left(\frac{|A_1 \cap A_2|}{|A_1|} + \frac{|A_1 \cap A_2|}{|A_2|}\right), \quad (2)$$

which relates the common area to both ranges of the individual species in a properly balanced way. This distance approximates 0.5 when one range is small subset of the other, as opposed to 1 for species whose ranges do not overlap. Thus, it is better suited for our analysis.

Test statistics.—A significant clustering of ranges means that the distances are small between ranges of the same cluster, while the distances between ranges of distinct clusters are large. The variation of distances of a homogeneously distributed set of ranges is expected to be lower, since there is no clear distinction between ranges that belong together and ranges that should be separated.

Based on these considerations, we propose a new test statistic. If n is the number of ranges, then there are $n(n-1)/2$ distances between ranges. Let $d_{1:n(n-1)/2} \leq d_{2:n(n-1)/2} \leq \cdots \leq d_{n(n-1)/2:n(n-1)/2}$ denote the ordered Kulczynski distances. Let $0 < \pi < 0.5$ be a proportion. Then the test statistic

$$T := \frac{\sum_{i \leq \pi n(n-1)/2} d_{i:n(n-1)/2}}{\sum_{i \geq (1-\pi)n(n-1)/2} d_{i:n(n-1)/2}}, \quad (3)$$

i.e., the ratio between the π smallest and the π largest distances, can be expected to be small for clustered data compared to homogeneous data. The proportion of distances that occur inside clusters is not known, so it is not immediately clear how to choose π. However, in our setup the result does not turn out to be very sensitive to the choice of π. We use $\pi = 0.25$, i.e., half of the distances are used for the test. See also Hennig and Hausdorf, in press, for a comparison of test statistics for clustering including the present proposal.

Null model.—The null model should simulate the case in which all inhomogeneities of the data can be attributed to varying range sizes, to varying numbers of taxa per geographic unit, and the spatial autocorrelation of the occurrences of a taxon. We developed a null model in which the nonoccurrence of range clusters is modeled so that all ranges are generated independently according to the same probabilistic routine. This routine yields ranges such that their cell number distribution approximates the actual distribution of the number of cells per range, the richness distribution of the cells approximates the actual richness distribution of the cells, and the tendency to form disjunct areas is governed by a parameter which is estimated from the real data set. Computational details have been described elsewhere (Hausdorf and Hennig, 2003; Hennig and Hausdorf, in press).

Determination of Biotic Elements

For the determination of biotic elements, we used model-based Gaussian clustering (MBGC) as implemented in the software MCLUST (Fraley and Raftery, 1998), because this method provides a decision about the number of meaningful clusters and about ranges that cannot be assigned adequately to any biotic elements. MBGC operates on a data set where the cases are defined by variables of metric scale. Therefore, we performed nonmetric multidimensional scaling (MDS; Kruskal, 1964) on the matrix of Kulczynski distances.

The principle of MBGC is as follows. For a given number of clusters k and a number of variables p, MBGC fits

a probability mixture model of the form

$$\varepsilon_0 U_C + \sum_{i=1}^{k} \varepsilon_i N_p(\mu_i, \Sigma_i), \quad \sum_{i=1}^{k} \varepsilon_i = 0, \quad \varepsilon_i \geq 0, \quad i = 0, \ldots, k \qquad (4)$$

to the data by means of maximum likelihood estimation. $N_p(\mu, \Sigma)$ denotes a p-dimensional normal distribution with mean vector μ and covariance matrix Σ, and U_C denotes the uniform distribution on a convex set C containing all data points. This means that the k clusters are modeled by distinct p-dimensional normal distributions (ε_i gives the proportion of the ith cluster). All data points that do not fit well into any normal cluster are assigned to the mixture component defined by U_C ("noise component"). A data point x is assigned to the mixture component j that maximizes $\varepsilon_j \ f_j(x)$ over $j = 0, \ldots, k$, where $f_j(x)$ denotes its estimated probability density. Such a fit is calculated for various values of k, the number of clusters. An optimal k can be found by means of the Bayesian information criterion. The same criterion is used to choose an optimal model for the relation between the covariance matrices of the different clusters (see Fraley and Raftery, 1998, for further details).

MCLUST needs an initial estimate of noise, which was performed by the software NNCLEAN (Byers and Raftery, 1998) as suggested by Fraley and Raftery (1998). We used as NNCLEAN constant k = number of species/40 and four MDS dimensions.

RESULTS

Test for Clustering of Distribution Areas

The test statistic T, i.e., the ratio of the sum of the 25% smallest d_K distances to the sum of the 25% largest distances, is 0.272 for the southern African *Scobius* species data set generated with a 2° grid. It is significantly smaller ($P = 0.01$) than should be expected under our null model (for 1000 artificial populations, T varied between 0.214 and 0.446, mean = 0.335). The test indicates that the distribution areas of the southern African *Scobius* species are significantly clustered. In contrast, the test with the data set generated with a 1° grid indicates that there is no significant clustering of distribution areas ($T = 0.415$; T varied between 0.373 and 0.709 for 1,000 artificial populations, mean = 0.478; $P = 0.09$).

Determination of Biotic Elements

A nonmetric MDS representation of the partition of the 47 southern African *Scobius* species to biotic elements is shown in Figure 1. If 2° grid squares are used, as in the analysis of Morrone (1994), most species ranges were assigned to four biotic elements. Only four ranges (8.5%) are classified in the noise component. The geographic distribution of the four biotic elements is shown in Figure 2. Biotic element 1 (seven species) is restricted to northeastern South Africa. The highest species richness can be found in grid squares B, E, and F (nomenclature of the grid cells after Morrone, 1994; Mast and Nyffeler, 2003). The geographic center of biotic element 2 (18 species) is in Natal, in the grid squares I, J, L, and M. The highest species richness of biotic element 3 (10 species) can be found in grid squares N and O in the Eastern Cape Province. Most species of biotic element 4 (eight species) are restricted to grid square P in the Eastern Cape Province. Only one species is more widespread.

If 1° grid squares are used, there is greater scatter in the occupancy of grid cells. Biotic elements 1 and 3 found in the analysis based on the 2° grid are no longer recognized as clusters due to the greater scatter. The species that were

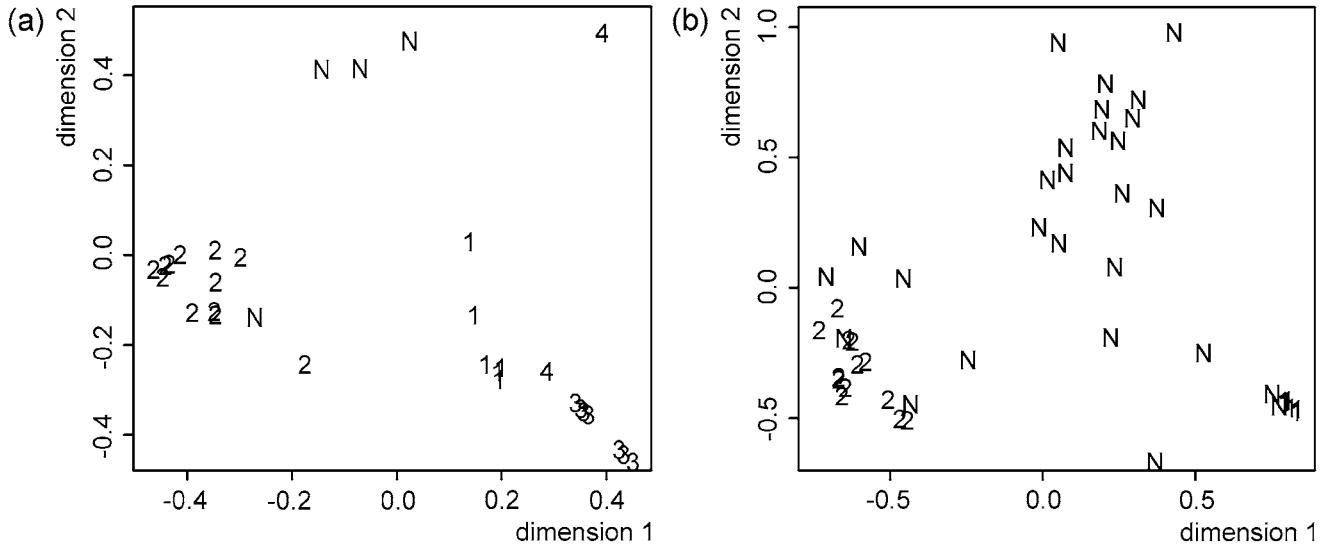

FIGURE 1. First two dimensions of the nonmetric MDS of the range data of the southern African *Scobius* species. 1–4-Biotic elements found by PRABCLUS; N = noise component. (a) Result if the distribution areas are mapped on a 2° grid. (b) Result if the distribution areas are mapped on a 1° grid.

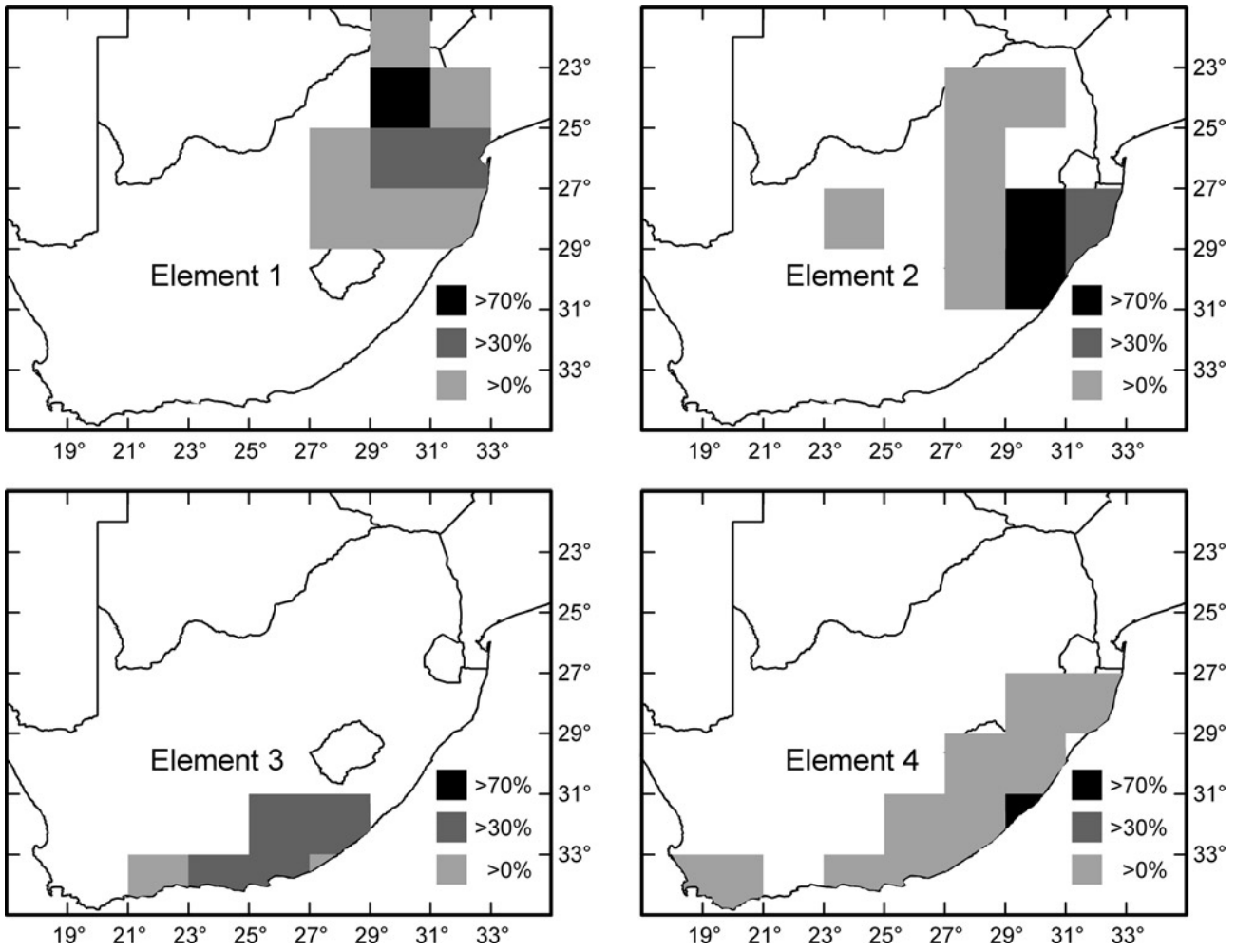

FIGURE 2. Distribution maps of four biotic elements found if the distribution areas of the southern African *Scobius* species are mapped on a 2° grid (some grid squares in Namibia are not shown). The different shadings indicate the areas where >70%, >30%, and >0% of the species of an element are present.

assigned to these biotic elements in the previous analysis are now classified in the noise component, which includes more than half of the species ranges (57.4%). Two biotic elements are recognized (Figs. 1, 3). Biotic element 1 (seven species) is restricted to grid square J15. It corresponds to biotic element 4 of the previous analysis with the exception that the single more widespread species is excluded. The geographic center of biotic element 2 (13 species) is in Natal. It corresponds to biotic element 2 but includes fewer species.

DISCUSSION

Test for Clustering of Distribution Areas

Most methods for the analysis of areas of endemism or biotic elements proposed so far take for granted that there is nonrandom structure in the data without testing this hypothesis.

We propose a test with which it can be examined whether the analyzed distribution areas are clustered. Only if this is the case, it is meaningful to determine biotic elements, which would be a consequence of a subdivision of an ancestral biota by vicariance events (Hausdorf, 2002).

Our test for clustering of distribution areas should not be confounded with the test of Mast and Nyffeler (2003) for the significance of pairwise co-occurences of taxa. The objective of our test is to check whether the entire set of ranges shows a degree of clustering that differs significantly from the degree of clustering observed in a set of similar ranges positioned independently of each other in the same area. Even if the null hypothesis of independent distribution of ranges is not rejected by our test, there may be some pairs of taxa whose ranges are more similar than should be expected by chance. On the other hand, no conclusion about the degree of clustering of the entire set of ranges is possible by a rejection of the null hypothesis of independent distribution for some pairs of taxa by the test of Mast and Nyffeler (2003).

The tests proposed by Mast and Nyffeler (2003) and by us are not directly comparable, but the null models to generate ranges are. The major differences between

the null model of Mast and Nyffeler (2003) and our null model are that our null model can consider the richness distribution of the cells and can produce ranges with a similar amount of disjunctions as in the real range set, whereas the null model of Mast and Nyffeler (2003) does not consider the richness distribution of the cells and cannot consider disjunctions of ranges (or of grid cells). Our null model was already applied in a test for nestedness of distribution areas (Hausdorf and Hennig, 2003).

Scobius *Example*

The *Scobius* example demonstrates our test for clustering of distribution areas and our method for the delimitation of biotic elements.

If 2° grid squares are used, the test indicates a significant clustering of distribution areas. Four biotic elements are found with our method. We can investigate how far these biotic elements correspond to areas of endemism as identified and delimited in previous studies.

Our biotic element 1 corresponds more or less to sets 2, 10, and 11 of Szumik et al. (2002), biotic element 2 corresponds to area 1 of Morrone (1994) and Mast and Nyffeler (2003) and sets 1, 4, 5, 6, 7, 8, and 9 of Szumik et al. (2002), biotic element 3 corresponds to area 2 (or 2′) of Morrone (1994) and Mast and Nyffeler (2003) and sets 3, 12, 13, and 14 of Szumik et al. (2002), and biotic element 4 corresponds to area 3 of Morrone (1994) and Mast and Nyffeler (2003).

For the 1° grid data set, our test resulted in a P value of 0.09, which means that a clustering of distribution areas cannot significantly be established. However, the value of the test statistic is still clearly below the average of the values for the simulated populations, and therefore it may be interesting to consider the result of the cluster analysis also for this grid. We found two biotic elements. All species belonging to biotic element 1 belong also to biotic element 4 of the 2° grid data set, which corresponds to area iv of Mast and Nyffeler (2003). All species belonging to biotic element 2 belong also to biotic element 2 of the 2° grid data set, which corresponds to areas i and ii of Mast and Nyffeler (2003). No biotic elements were found in the 1° grid data set that would correspond to the biotic elements 1 and 3 of the 2° grid data set. The distribution areas of the species that form these biotic elements are not similar enough at the finer scale to form distinct clusters.

The differences between the results of the analysis of the 1° grid and those of the 2° grid data set show the importance of grid size (Jardine, 1972; Mast and Nyffeler, 2003). Distinct biotic elements may be amalgamated if the grid used is too coarse. In the *Scobius* example, no elements are found in the analysis with the 1° grid data set that are not also found with the 2° grid data set. Thus, there is no indication that the 2° grid is too coarse. If the grid used is too fine and the distribution data are not interpolated, insufficient sampling may introduce artificial noise in the data set. This can cause a failure to recognize existing biotic elements. This seems to be the case in the analysis of the 1° grid *Scobius* data set, where the clustering is not significant at the 5% level, the cluster analysis does not find two biotic elements that are recognized with the 2° grid data set, and the MDS representation (Fig. 1b) indicates a much weaker clustering pattern.

The distribution maps of the biotic elements (Figs. 2, 3) show that the elements often overlap geographically and that the species belonging to one element are usually not homogeneously distributed. In several cases, a center can be discerned, where most species belonging to an element occur. With increasing distance from the center less and less of the species belonging to the element are present. One process that can create such a pattern is diversification by vicariance and subsequent chance dispersal of the species that originated in an area of endemism. The problems in delimitating areas of endemism are due to this pattern.

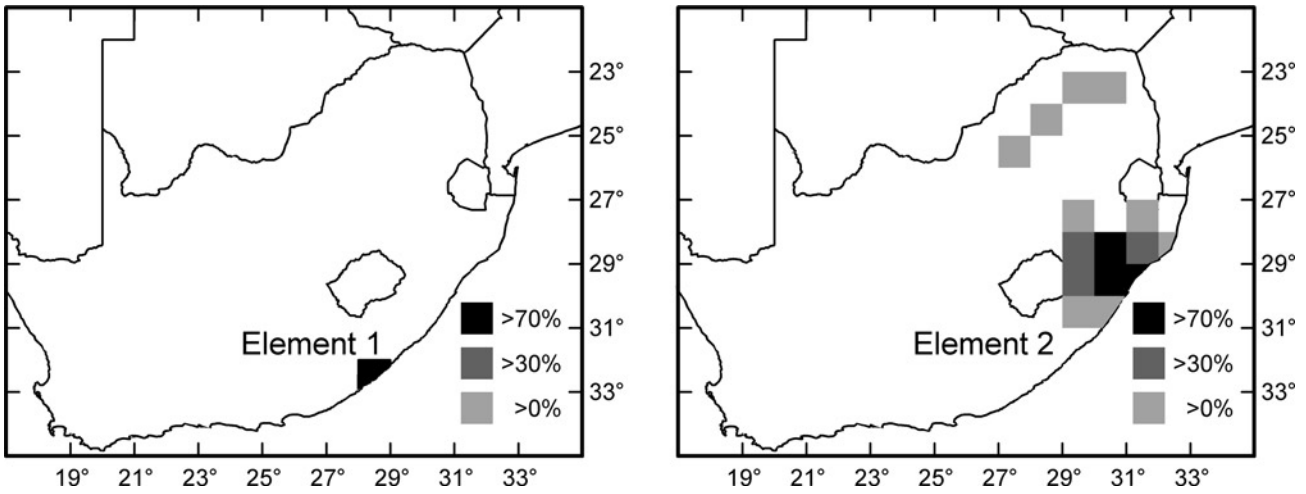

FIGURE 3. Distribution maps of two biotic elements found if the distribution areas of the southern African *Scobius* species are mapped on a 1° grid (some grid squares in Namibia are not shown). The different shadings indicate the areas where >70%, >30%, and >0% of the species of an element are present.

Biotic Elements versus Areas of Endemism

Both biotic element analysis and the areas of endemism approach try to identify similarities in distribution areas, which might be the result of a common biogeographic history (primary biogeographic homology, sensu Morrone, 2001).

The *Scobius* example illustrates two major problems with the areas-of-endemism approach. First, a single set of nonoverlapping areas of endemism has to be identified. But the application of an optimality criterion, namely the number of the species that can be considered endemic, shows that there are many combinations of grid squares with similar scores of endemicity (Szumik et al., 2002). Optimality criteria can be formulated to select a set of nonoverlapping areas of endemism. However, in this way the information about overlapping distribution patterns gets lost. In the next step of the usual vicarance biogeography protocol, an area cladogram is constructed based on the previously selected set of areas of endemism. In this step it is tested whether the taxa that are supposed to have been originated in a given set of areas of endemism actually had a common biogeographic history (secondary biogeographic homology, sensu Morrone, 2001). This step depends crucially on the correct identification and delimitation of areas of endemism. Different delimitations of areas of endemism can result in different outcomes concerning area relationships (Henderson, 1991). Furthermore, the comparison of area cladograms is a possibility to check which candidate areas of endemism might actually be historical units that were separated by vicariance events (Harold and Mooi, 1994). However, such a test is not very effective, if many candidate areas of endemism have been excluded already during the selection of a set of nonoverlapping areas of endemism based on other criteria.

The second major problem of the areas-of-endemism approach is that it is hardly possible to reconstruct the borders between areas of endemism if dispersal resulted in overlapping distribution patterns even if we knew the true number and the approximate position of these areas (Hausdorf, 2002). This difficulty is illustrated in the *Scobius* example by the different delimitations of areas of endemism (Morrone, 1994; Szumik et al., 2002; Mast and Nyffeler, 2003), which are a consequence of the overlap of the biotic elements (Fig. 2). Another well-known example which highlights that problem are the numerous lines which have been proposed to delimit the Oriental and Australian regions (Mayr, 1944; Holloway and Jardine, 1968; Simpson, 1977; Vane-Wright, 1991). It is possible to formulate optimality criteria to choose among different delimitations of areas (Szumik et al., 2002). But the resulting borders vary with the choice of the criteria and there is no rationale for how to find the true borders. Again, this might create problems in the reconstruction of an area cladogram.

Our approach based on biotic elements presents solutions to both problems. First, there is no necessity to select a single set of nonoverlapping biotic elements. Overlap between the areas occupied by different biotic elements is the rule rather than the exception (Fig. 2) and indicates that processes other than vicariance (e.g., dispersal across barriers) were frequently involved in the origination of these distribution patterns. It can be investigated which biotic elements represent historical units that originated by vicariance, even if they overlap geographically. For that purpose, we proposed to convert a taxon cladogram into an element cladogram by substituting the names of the taxa with the respective biotic elements and looking for recurrent patterns in the resulting element cladograms (Hausdorf, 2002). Thus, overlapping distribution patterns are not ignored, as in the areas-of-endemism approach. Further elaboration of this topic is beyond the topic of this paper.

Second, in contrast to the area-of-endemism approach, where the correct delimitation of the borders of the areas of endemism is crucial for the conversion of a taxon cladogram into an area cladogram, it is not even necessary to reconstruct the borders between areas of endemism in the biotic element approach. Biotic element cladograms can be constructed and searched for patterns without fixing borders.

Other Methods for the Analysis of Biotic Elements

A well-known problem in cluster analysis is that most cluster analysis methods assign all units to clusters, i.e., they always yield an exhaustive clustering, if this is adequate for the data at hand or not, and no automatic decision about the number of clusters is possible. It is not clear from the results of previous analyses of biotic elements, which employed hierarchical distance based cluster analyses such as single or complete linkage (e.g., Holloway and Jardine, 1968; Jardine, 1972; Dennis et al., 1998), whether all biotic elements which were delimited are meaningful. In contrast to such methods, MBGC provides a decision about the number of meaningful biotic elements and about ranges that cannot be assigned adequately to any biotic element.

There are attempts to decide about the meaningfulness of biotic elements by the use of pairwise significance tests of co-occurrence. Márquez et al. (1997) combined a hierarchical average linkage (UPGMA) clustering with a sequence of tests concerning the number of significantly similar species in different branches of the dendrogram. Mast and Nyffeler (2003) used an informal clustering approach equivalent to single linkage for the discussed examples, where the P values of the pairwise comparisons are used as dissimilarities between species and the resulting tree is cutted at a level of 0.05. The appropriateness of P values as a measure of similarity between ranges is doubtful because P values do not only depend on the similarity between ranges, but also on the power of the test, which changes, e.g., with the scales of the grids. Moreover, the statistical validity of both approaches is questionable because they lead to multiple testing procedures, where no adequate adjustment of the P values is possible (see Mast and Nyffeler, 2003, for arguments

why the Bonferroni adjustment is too conservative for this task).

ACKNOWLEDGMENT

We are grateful to Austin Mast for communicating details of an unpublished manuscript.

REFERENCES

BYERS, S., AND A. E. RAFTERY. 1998. Nearest neighbor clutter removal for estimating features in spatial point processes. J. Am. Stat. Assoc. 95:781–794.

CROIZAT, L., G. NELSON, AND D. E. ROSEN. 1974. Centers of origin and related concepts. Syst. Zool. 23:265–287.

DENNIS, R. L. H., W. R. WILLIAMS, AND T. G. SHREEVE. 1998. Faunal structures among European butterflies: Evolutionary implications of bias for geography, endemism and taxonomic affilation. Ecography 21:181–203.

FRALEY, C., AND A. E. RAFTERY. 1998. How many clusters? Which clustering method? Answers via model based cluster analysis. Comput. J. 41:578–588.

HAROLD, A. S., AND R. D. MOOI. 1994. Areas of endemism: Definition and recognition criteria. Syst. Biol. 43:261–266.

HAUSDORF, B. 2002. Units in biogeography. Syst. Biol. 51:648–652.

HAUSDORF, B., AND C. HENNIG. 2003. Nestedness of north-west European land snail ranges as a consequence of differential immigration from Pleistocene glacial refuges. Oecologia 135:102–109.

HENDERSON, I. M. 1991. Biogeography without area? Aust. Syst. Bot. 4:59–71.

HENNIG, C., AND B. HAUSDORF. In press. Distance-based parametric bootstrap tests for clustering of species ranges. Comput. Stat. Data Anal.

HOLLOWAY, J. D., AND N. JARDINE. 1968. Two approaches to zoogeography: A study based on the distributions of butterflies, birds and bats in the Indo-Australian area. Proc. Linn. Soc. Lond. 179:153–188.

HUMPHRIES, C. J., AND L. R. PARENTI. 1999. Cladistic biogeography. Oxford Univ. Press, Oxford, UK.

JARDINE, N. 1972. Computational methods in the study of plant distributions. Pages 381–393 *in* Taxonomy, phytogeography and evolution (D. H. Valentine, ed.). Academic Press, London.

KRUSKAL, J. B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29:1–27.

LINDER, H. P. 2001. On areas of endemism, with an example from the African Restionaceae. Syst. Biol. 50:892–912.

MÁRQUEZ, A. L., R. REAL, J. M. VARGAS, AND A. E. SALVO. 1997. On identifying common distribution patterns and their causal factors: A probabilistic method applied to pteridophytes in the Iberian Peninsula. J. Biogeogr. 24:613–631.

MAST, A. R., AND R. NYFFELER. 2003. Using a null model to recognize significant co-occurrence prior to identifying candidate areas of endemism. Syst. Biol. 52:271–280.

MAYR, E. 1944. Wallace's line in the light of recent zoogeographic studies. Q. Rev. Biol. 19:1–14.

MORRONE, J. J. 1994. On the identification of areas of endemism. Syst. Biol. 43:438–441.

MORRONE, J. J. 2001. Homology, biogeography and areas of endemism. Divers. Distrib. 7:297–300.

NELSON, G., AND N. PLATNICK. 1981. Systematics and biogeography: Cladistics and vicariance. Columbia Univ. Press, New York.

PLATNICK, N. I., AND G. NELSON. 1978. A method of analysis for historical biogeography. Syst. Zool. 27:1–16.

ROSEN, D. E. 1976. A vicariance model of Caribbean biogeography. Syst. Zool. 24:431–464.

ROSEN, D. E. 1978. Vicariant patterns and historical explanations in biogeography. Syst. Zool. 27:159–188.

SHI, G. R. 1993. Multivariate data analysis in palaeoecology and palaeobiogeography—A review. Palaeogeogr. Palaeoclimatol. Palaeoecol. 105:199–234.

SIMPSON, G. G. 1977. Too many lines; the limits of the Oriental and Australian zoogeographic regions. Proc. Am. Philos. Soc. 121:107–120.

SZUMIK, C. A., F. CUEZZO, P. A. GOLOBOFF, AND A. E. CHALUP. 2002. An optimality criterion to determine areas of endemism. Syst. Biol. 51:806–816.

VANE-WRIGHT, R. I. 1991. Transcending the Wallace line: Do the western edges of the Australian region and the Australian Plate coincide? Aust. Syst. Bot. 4:183–197.

WILEY, E. O. 1988. Vicariance biogeography. Annu. Rev. Ecol. Syst. 19:513–542.

First submitted 26 February 2003; reviews returned 23 May 2003; final acceptance 19 June 2003
Associate Editor: Roderic Page