SPECIES CO-OCCURRENCE: A META-ANALYSIS OF J. M. DIAMOND’S ASSEMBLY RULES MODEL

NICHOLAS J. GOTELLI1 AND DECLAN J. MCCAIBE2

Department of Biology, University of Vermont, Burlington, Vermont 05405 USA

Abstract. J. M. Diamond’s assembly rules model predicts that competitive interactions between species lead to nonrandom co-occurrence patterns. We conducted a meta-analysis of 96 published presence–absence matrices and used a realistic “null model” to generate patterns expected in the absence of species interactions. Published matrices were highly nonrandom and matched the predictions of Diamond’s model: there were fewer species combinations, more checkerboard species pairs, and less co-occurrence in real matrices than expected by chance. Moreover, nonrandom structure was greater in homeotherm vs. poikilotherm matrices. Although these analyses do not confirm the mechanisms of Diamond’s controversial assembly rules model, they do establish that observed co-occurrence in most natural communities is usually less than expected by chance. These results contrast with previous analyses of species co-occurrence patterns and bridge the apparent gap between experimental and correlative studies in community ecology.

Key words: community assembly rules; meta-analysis; null models; presence–absence matrix; species co-occurrence.

INTRODUCTION

A fundamental question in ecology is whether general assembly rules determine the structure of natural communities (Weiher and Keddy 1999). Many types of assembly rules have been described, including constant body-size ratios (Dayan and Simberloff 1994), favored states (Fox and Brown 1993), guild proportionality (Wilson 1989), species nestedness (Patterson and Atmar 1986), and trait–environment associations (Keddy and Weiher 1999). However, the most influential model remains Diamond’s (1975) original treatment of community assembly rules (Gotelli 1999).

Diamond (1975) argued that interspecific competition among bird species of the Bismark Archipelago generated a number of community assembly rules, including forbidden species combinations, checkerboard distributions, and incidence functions. Connor and Simberloff (1979) used a Monte Carlo null model analysis to demonstrate that many of the patterns attributed by Diamond (1975) to interspecific competition could also arise in communities that were assembled by random colonization and were competition-free. These exchanges touched off a debate in community ecology that has continued over the past 25 yr (reviews in Strong et al. 1984, Wiens 1989, Gotelli and Graves 1996, Weiher and Keddy 1999).

The initial debates focused on the statistical issues surrounding null models and potential flaws in the analysis of Connor and Simberloff (Diamond and Gilpin 1982, Gilpin and Diamond 1982, Connor and Simberloff 1983, 1984). More recent studies (Stone and Roberts 1990, Manly 1995, Sanderson et al. 1998) and the availability of software for null model analysis (Gotelli and Entsminger 1999, Patterson 1999, Colwell 1999) have clarified these issues. The statistical properties of many null model algorithms have been studied by evaluating their performance with random matrices and with structured matrices that include stochastic noise (e.g., Kelt et al. 1995, Gotelli et al. 1997, Shenk et al. 1998, Gotelli 2000).

Even with the availability of these improved statistical methods, the generality of Diamond’s (1975) assembly rules model remains unknown. Many authors have made broad assertions about the generality (or lack thereof) of assembly rules based on the analysis of a single presence–absence matrix. Much of the debate over assembly rules has rested on the reanalysis of a handful of presence–absence matrices for the birds of island archipelagoes (Gotelli and Graves 1996), such as the Vanuatu (formerly New Hebrides) Islands. In this paper, we use a null model analysis of 96 published presence–absence matrices to test the predictions of Diamond’s (1975) assembly rules model. A meta-analysis demonstrates the generality of these patterns and reveals taxon-specific differences in community organization.

METHODS

Data matrices

To test the predictions of Diamond’s (1975) model, we compiled 96 data sets from studies that report the distribution of species assemblages across a set of replicated sites. The sites range from small quadrats in old...
fields (0.25 m^2) to large islands in oceanic archipelagoes (2.3 × 10^6 m^2). The data from each study were organized as a presence–absence matrix in which each row represents a species or taxon and each column represents a site. The entries in the matrix indicate the presence (1) or absence (0) of a particular species in a particular site (Simberloff and Connor 1979).

We used a subset of modified data matrices from those provided by Patterson (1999). Data sets were assembled from original literature citations, according to the following criteria: only taxa described in the original citation that were extant at the time of the collection or census were included; nonnative species were included only when the data were included by the original authors; data sets that included both aquatic and terrestrial taxa were partitioned into two matrices; species occurrences labeled as questionable by the original authors were excluded; the most inclusive data set from overlapping studies was selected; data sets from artificial substrates were excluded; data sets that were partitioned on the basis of habitat types rather than sites were excluded; subspecies were lumped and treated as a single taxon. The original investigators in these studies were usually not addressing whether competition was an important force in structuring communities. Rather, they were seeking to construct species lists for a prescribed taxonomic group. Thus, the data probably do not suffer from a preselection bias towards communities that are structured by competition.

For each matrix, we measured from survey maps, gazetteers, navigational charts, and atlases the following variables: the latitude and longitude of the geographical center of the set of sites, the average area of the censused sites, and the minimum north-south and east-west spans that included all of the sites. Studies were classified according to the biogeographical regions designated by Pielou (1979). Average site area was calculated from data provided in the original papers or from encyclopedias and atlases. Presence–absence matrices and raw data are available from Patterson (1999).

**Co-occurrence indices**

A major stumbling block to testing Diamond’s (1975) model has been the difficulty in quantifying patterns in a presence–absence matrix and relating these patterns to assembly rules. Although Connor and Simberloff (1979) noted that some of Diamond’s (1975) rules are tautologies, others can be made operational by using an appropriate co-occurrence index, specifying a priori the pattern expected with Diamond’s (1975) model, and comparing the pattern to that generated by a well-tempered null model.

Pielou and Pielou (1968) first introduced the number of species combinations as an index of community structure. This index is directly related to Diamond’s (1975:344) first and second assembly rules: “1. If one considers all the combinations that can be formed from a group of related species, only certain ones of these combinations exist in nature. 2. These permissible combinations resist invaders that would transform them into a forbidden combination.” If assembly rules 1 and 2 are met, a set of islands or sites should harbor significantly fewer species combinations than expected by chance.

A second useful index is the number of species pairs that never co-occur, forming “checkerboard” distributions. This index describes Diamond’s (1975:344) fifth assembly rule: “5. Some pairs of species never coexist, either by themselves or as part of a larger combination.” If this assembly rule is in operation, there should be significantly more species pairs in a matrix forming perfect checkerboards than expected by chance. A related index is Stone and Robert’s (1990) C score. This index also measures the degree to which species co-occur, but it is not as stringent as the checkerboard measurement because it does not require perfect segregation between species. For a community structured by species interactions, the C score should be significantly larger than expected by chance. Diamond’s (1975) other assembly rules (numbers 3, 4, 6, and 7) are more difficult to test with simple null models because they involve complex comparisons of patterns in species-rich and species-poor communities.

Thus, if Diamond’s (1975) assembly rules are in operation, real communities should contain fewer species combinations, more checkerboard pairs, and a larger C score than randomly assembled communities that are not structured by species interactions.

The number of checkerboard pairs is calculated by counting the number of unique pairs of species that never co-occur. Stone and Robert’s (1990) C score is calculated for each pair of species as \((R_i - S)(R_j - S)\) where \(R_i\) and \(R_j\) are the matrix row totals for species \(i\) and \(j\), and \(S\) is the number of sites in which both species occur. This \(C\) score is then averaged over all possible pairs of species in the matrix. The number of species combinations is calculated by counting the number of distinct arrangements represented by the columns of the matrix. Gotelli (2000) summarizes the statistical properties of these indices.

**Randomization algorithm**

We used a Monte Carlo “null model” simulation to randomize each matrix in the data set. The random matrices retained the row and column totals in the matrix, so that each random matrix had the same number of species per site and the same number of sites per species as did the real matrix (Connor and Simberloff 1979). Random matrices were created with a “swap” algorithm (Gotelli and Entsminger 2001), in which the cells of randomly chosen \(2 \times 2\) submatrices are exchanged so as to maintain row and column totals (Manly 1995). For each original matrix, 1000 initial transpositions were used to randomize the original pattern, and then the next 1000 consecutive transpositions were
retained to create 1000 null matrices. The co-occurrence index was calculated for each null matrix, and the statistical significance of the observed matrix was calculated as the frequency of simulated matrices that had indices that were equal to or more extreme than the index of the observed matrix (one-tailed test).

Sanderson et al. (1998) have criticized some aspects of this swapping algorithm when used with Roberts and Stone’s (1990) $S^2$ metric. However, the performance of our algorithm with the C score, number of checkerboards, and number of species combinations has been validated by extensive comparisons with random matrices and with structured matrices that have simulated “noise” added (Gotelli 2000, Gotelli and Entsminger 2001). In particular, the C score (Stone and Roberts 1990) used with the swap algorithm has good Type I error properties and does not reject the null hypothesis too frequently when tested with random matrices (Gotelli 2000). Thus, the highly nonrandom patterns observed in real matrices (Fig. 1) cannot be attributed to a Type I error caused by an artifact of the simulation procedure. We generated similar results using an unbiased version of Sanderson et al.’s (1998) “Knight’s Tour” algorithm, in which an empty matrix is filled one cell at a time (Gotelli and Entsminger 2001; see also A. Zaman and D. Simberloff, unpublished manuscript). Null model analyses were conducted with EcoSim 4.0 simulation software (Gotelli and Entsminger 1999).

**Meta-analysis**

To compare results across studies, we calculated a standardized effect size for each matrix. The standardized effect size measures the number of standard deviations that the observed index is above or below the mean index of the simulated communities. The null hypothesis is that the average standardized effect size, measured for the set of 96 presence–absence matrices, is zero.

In meta-analysis, an “effect size” is calculated by standardizing the difference between “control” and “treatment” groups (Gurevitch et al. 1992). In our analysis, the observed index ($I_{obs}$) corresponds to the “treatment” group. The mean of the 1000 indices from the simulated communities ($I_{sim}$) corresponds to the “control” group, because it reflects the pattern expected in the absence of species interactions. We used the standard deviation of the 1000 indices from the simulated communities ($\sigma_{sim}$) to calculate the standardized effect size (SES) as

$$\text{SES} = \frac{(I_{obs} - I_{sim})}{\sigma_{sim}}.$$

We analyzed the variation in SES of the C score that could be attributed to factors measured for each of the data matrices. We used the C score for this analysis because it has the greatest statistical power of the three indices (C score, number of checkerboards, number of species combinations) for detecting nonrandomness (Gotelli 2000). The continuous predictor variables were the size of the matrix (number of rows × number of columns), the percentage of the matrix that was filled with ones, the latitude, longitude, average site area, and the geographic area encompassed by the sites (east–west span × north–south span). The discrete predictor variables were the biogeographic province (Australasian, Ethiopian, Nearctic, Neotropical, Oceanian, Oriental, Palearctic), community type (continental vs. island), and taxon (bat, bird, mammal, reptile/amphibian, fish, invertebrate, ant, plant).

We used simple correlations and one-way ANOVAs to initially screen the set of predictor variables and retained the subset of statistically significant predictor variables ($P < 0.05$). We also reanalyzed the data after...
deleting outliers and influential points to ensure the
stability of the patterns. Finally, we used nonparametric
tests to ensure that the patterns were robust to the non-
ormality of the data in Fig. 1. Matrix size was a sig-
nificant predictor of effect size \( (P = 0.048) \), because
large matrices enhanced the statistical power of the
analysis. However, when matrix size was used as a
covariate, differences among taxonomic groups were
still statistically significant \( (P = 0.035) \).

RESULTS

For all three co-occurrence indices, the standardized
effect size differed significantly from zero: across the
96 studies, there were fewer species combinations,
more species pairs exhibiting perfect checkerboard dis-
btributions, and larger \( C \) scores than expected by chance
(Fig. 1). All of these patterns are in the direction pre-
dicted by Diamond’s (1975) assembly rules model.

We next analyzed in detail the variation in the \( C \)
score. Standardized effect size was not correlated with
the latitude \( (P = 0.544) \) or longitude \( (P = 0.904) \) of
each study site, and there were no differences in effect
size when matrices were classified according to bio-
geographic province \( (P = 0.794) \) or according to con-
tinental-island status \( (P = 0.536) \). The standardized
effect size also appears to be scale-invariant: neither
the average area of the sites \( (P = 0.550) \) nor the geo-
graphic area encompassed by the set of sites \( (P = 0.240) \)
were correlated with the standardized effect
size.

However, the standardized effect size differed sig-
nificantly among taxonomic groups (Fig. 2). Effect size-
nes were significantly larger for matrices of homeo-

Discussion

Our findings settle a long-standing controversy over
community patterns: in most natural communities of
plants and nonparasitic animals, there is less species
c-co-occurrence than expected by chance, in accordance
with the predictions of Diamond’s (1975) assembly
rules model. These results contrast with the conclusions
of two earlier studies that examined co-occurrence pat-
terns for a broad set of plant and animal matrices (Sim-
berloff 1970, Schluter 1984), although these studies did
not attempt to test the specific predictions of Diamond’s
(1975) model.

Simberloff (1970) compiled and analyzed patterns
in the species/genus (S/G) ratio of local communities
and found that observed S/G ratios were usually greater
than expected, in contrast to the predictions of com-
petition theory (Gotelli and Colwell 2001). However,
the S/G ratio may be a weak indicator of competitive
effects (Hairston 1964), and the null model that was
used assumed that colonization potential of all species
was equiprobable.

Our results also differ from those of Schluter (1984),
who used the variance ratio as an index of co-occur-
cence and found that patterns in published matrices
tended to be random or slightly positive. However, the
variance ratio measures variability in the total number
of species in an assemblage; it does not directly mea-
sure the pattern of species co-occurrence as we have
done here. Moreover, Schluter’s (1984) null model
implicitly assumed that all sites are equiprobable. There-
fore, negative co-occurrence patterns may have been
masked by heterogeneity among sites.

Although our results confirm the basic predictions
of Diamond’s (1975) model, they should not be con-
strued as a definitive test of the model, because some
important alternative hypotheses can also produce non-
randomness in the directions we observed. First, some
species may exhibit “habitat checkerboards” and seg-
regate because of affinities for nonoverlapping habitats,
not because of competitive interactions. If data are
available on habitat availability and affinities, it may
be possible to statistically distinguish these effects
(Schoener and Adler 1991). Second, some species may
exhibit “historical checkerboards” and co-occur infre-
quently because of allopatric speciation and other
events that reflect biogeographic and evolutionary his-
tory (Vuilleumier and Simberloff 1980, Cracraft 1988).
Reliable phylogenies and details of biogeographic his-
tory are necessary to sort out such historical effects
(Losos 1990, 1995). Moreover, historical, habitat, and
ecological checkerboards may not represent mutually
exclusive hypotheses if evolutionary change reflects
species interactions (the “ghost of competition past”; Connell 1980, Ricklefs and Schluter 1993). Recent studies in macroecology have emphasized that the way species partition energy is a fundamental constraint on community structure (Brown 1995, Blackburn and Gaston 1998, Maurer 1999). Our results confirm basic differences in co-occurrence patterns for poikilotherm and homeotherm groups. The finding that plant matrices are highly structured is consistent with widespread evidence of plant competition for nutrients, light, and water (Tilman 1982, Keddy 1989). The higher organization of ant matrices relative to other poikilotherm groups may be a reflection of eusociality and complex social structure in ants and the widespread belief that ant communities are largely structured by interspecific competition (Hölldobler and Wilson 1990).

However, many other explanations for these taxonomic differences are possible. For example, the random structure of the non-ant invertebrate matrices could reflect the fact that “functional groups” of stream assemblages may include heterogeneous, ecologically mixed taxa that might not be expected to show non-random structure (P. Giller, personal communication). To the extent that the assemblages represented in the matrices do not correspond to true ecological guilds (Simberloff and Dayan 1991), all of these analyses potentially suffer from the so-called “dilution effect” (Diamond and Gilpin 1982), which might cause our tests for segregation patterns to be conservative. Alternatively, the highly nonrandom structure of the homeotherm matrices might reflect the large body size of individuals and the higher energy requirements of free-living taxa, leading to greater niche saturation (Rohde 1991, Gotelli and Rohde 2002). Finally, differences among taxa in systematics and criteria used for delineating species, collecting effort, and guild and site designations can also affect the results of null model tests (Gotelli and Graves 1996).

Such issues cannot be resolved by statistical analyses that are based solely on presence–absence matrices. Simple null models are best viewed as statistical tools for recognizing nonrandom patterns (Gotelli 2001) rather than as a critical “litmus test” for competitive effects (Roughgarden 1983). Nevertheless, the history of community assembly rules has been so controversial that documenting nonrandom patterns in nature is an important first step in establishing the generality of such models.

In summary, our results confirm the basic predictions of Diamond’s (1975) assembly rules model and reveal intriguing differences among taxa that hint at the importance of physiological constraints in producing community patterns. A widespread misperception in ecology is that the results of small-scale experimental studies of species interactions (Resetarits and Bernardo 1998) are not concordant with the results of large-scale nonexperimental analyses of species co-occurrence (Peters 1991, Cornell and Lawton 1992, Maurer 1999). Our analyses demonstrate that species co-occurrence, measured for a variety of taxa at many different spatial scales, is usually less than expected by chance. Additional research is needed to elucidate the detailed operation of assembly rules in particular communities and to determine the extent to which abiotic and historical factors reinforce or weaken species co-occurrence patterns (Dunson and Travis 1991).

ACKNOWLEDGMENTS
We thank S. Heard, R. Paine, D. Simberloff, the University of Vermont Ecology Discussion Group, and the NCEAS working group in Biogeography for comments on the analysis, and G. Ensminger and C. Bliss for computer programming. EcoSim software development supported by NSF grants DBI 97–25930 and DEB 01–07405.

LITERATURE CITED


