

7

CO-OCCURRENCE

COMMUNITY ASSEMBLY

A pervasive theme in community ecology is that the species composition of a community is governed by deterministic “assembly rules” (Cody and Diamond 1975; Case and Diamond 1986). These rules emphasize the importance of interspecific interactions in determining which species are found in a particular assemblage (Drake 1990). We have already examined two such “rules” in this book: (1) species that overlap “too much” in phenology, resource use, or other niche dimensions cannot coexist (Chapters 4 and 5); (2) species that do coexist must differ in body size or trophic morphology by a critical minimum that allows them to exploit different resources (Chapter 6).

In this chapter, we consider related assembly rules that predict the presence and absence of particular species rather than their sizes or patterns of resource utilization. The significance and even the reality of such assembly rules have been widely debated in community ecology. Proponents have argued for the importance of resource exploitation (Diamond 1975), competitive hierarchies (Gilpin et al. 1986), and priority effects (Drake 1991) in producing assembly rules. Critics have complained that many of the rules are trivial tautologies that lack predictive power (Connor and Simberloff 1979) and that the evidence for consistent patterns of community structure, much less for assembly rules, is hardly compelling (Wiens 1980; Wilson 1991a).

Laboratory studies have provided the strongest evidence for assembly rules. For example, Gilpin et al. (1986) thoroughly explored competitive interactions among 28 species of *Drosophila*. They varied the level of food resources, the length of the experiments, and, most importantly, the initial combinations and densities of species. Competitive interactions were strong and predictable, and few species persisted until the end of the experiments. For example, experiments that started with 10 species always ended with fewer than four. With 10 initial species, there were $2^{10} = 1,024$ different possible species combinations,

but fewer than a dozen of these persisted. These results contrast with nature, where dozens of *Drosophila* species may co-occur sympatrically.

Field experiments also provide some evidence for assembly rules. Abele (1984) experimentally “supersaturated” *Pocillopora* coral heads with component species of decapod crustaceans. Agonistic interactions and predation by voracious wrasses caused the fauna to “relax” to a predictable species number and composition for coral heads of a given size. Cole (1983) described assembly rules for five ant species on small mangrove islets in the Florida Keys. Two “primary species” (*Crematogaster ashmeidi* and *Xenomyrmex floridanus*) could never be introduced successfully on very small islands, perhaps because of frequent flooding. In contrast, the minimum island size occupied by two “secondary species” (*Pseudomyrmex elongatus* and *Zacryptocercus varians*) was set by the presence of primary species, which were superior competitors. The two primary species also did not coexist and formed a classic “checkerboard” pattern (Diamond 1975), in which only one of the two species, but not both, occurred on an island. Experimental transplants and behavioral arena experiments confirmed that aggressive interactions between workers prevented coexistence and that either primary species as an island resident could repulse an invader. Thus, observed species combinations could be predicted on the basis of island size and competitive relationships among colonizers.

However, the debate over assembly rules has not been concerned with such detailed experimental studies. Rather, the controversy has been over whether assembly rules can be inferred from nonexperimental data, specifically from combinations of coexisting species, usually on islands. Such data are conveniently summarized in a presence-absence matrix, which forms the fundamental unit of study in many analyses of community ecology and biogeography (McCoy and Heck 1987). Null models have been a useful tool for evaluating pattern in such matrices and for revealing the extent to which species combinations can be predicted on the basis of simple models of island colonization (Simberloff 1978a).

PRESENCE-ABSENCE MATRICES

Presence-absence matrices summarize data on the occurrence of a group of species at a set of sites (Table 7.1). The “sites” may range in size from 160-mm² vegetation quadrats (Watkins and Wilson 1992) to large oceanic islands (Diamond 1975) or entire continents (Smith 1983). Similarly, the set of “species” analyzed may be restricted to ecologically similar congeners (Graves and

Table 7.1
Presence-absence matrix for five fish taxa in 28 Australian springs

		Number of springs
Goby	x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x	28
Gudgeon	x x x x x x x x x x x x x x x x x x x x x x x	19
Catfish	x x x x x x x x x x x x x x x x	14
Hardyhead	x x x x x x x x x	9
Perch	x x x x x x x	7
Number of species	5 5 5 5 5 4 4 4 4 4 3 3 3 3 2 2 2 2 2 1 1 1 1 1 1 1 1 1	

Each row is a different taxon and each column is a different site (= spring). Each x represents the occurrence of a taxon in a particular spring. Springs are ordered by species richness. Goby = *Chlamydogobius* sp.; Gudgeon = *Mogurnda mogurnda*; Catfish = *Neosilurus* sp.; Hardyhead = *Craterocephalus* sp.; Perch = *Leiopotherapon unicolor*. Note the nearly perfect pattern of nestedness. From Kodric-Brown and Brown (1993).

Gotelli 1993) or to ecological or trophic guilds (Heatwole and Levins 1972), or it may encompass entire avifaunas (Connor and Simberloff 1979).

In an $r \times c$ matrix, each row of the matrix represents a different species and each column represents a different site (Simberloff and Connor 1979). Each cell in the matrix (a_{ij}) contains a 0 or a 1, denoting, respectively, the absence or presence of species i on site j . The row sum R_i represents the total number of occurrences of each species; it ranges from a minimum of 1 to a maximum of c , the number of sites censused. Similarly, the column sum C_j gives the number of species censused on site j and ranges from a minimum of 1 to a maximum of r , the number of species recorded. The grand matrix sum N represents the total number of site occurrences.

ASSUMPTIONS UNDERLYING THE ANALYSIS OF PRESENCE-ABSENCE MATRICES

Before discussing the analysis of such matrices, it is important to consider the assumptions implicit in the choice of the sampling universe and hence in the dimensions of the matrix. The decision of which species to analyze is critical. If ecologically diverse assemblages are analyzed together, species interactions may not be apparent because of the “dilution effect” (Diamond and Gilpin 1982), in which many species will be compared that are not interacting with one another. Or, as Grant and Abbott (1980) have more colorfully expressed it, the analysis is “in danger of throwing the baby out with the bathwater, or, more specifically, drowning the baby in a tub that is too deep.” On the other hand, if the analysis is restricted to very few species, the sample sizes may simply be too small to reveal any significant patterns, no matter what test is used (Biehl and Matthews 1984).

Equally important is the choice of which sites to include and which to exclude from the analysis. An implicit assumption in the analysis of presence-absence matrices is that individual islands serve as replicates to reveal repeated patterns of species co-occurrence. But if the islands differ in size or suitability for colonization, important adjustments to the null model are necessary. Indeed, if the islands vary too much in area, habitat, and isolation, it may be invalid to assume they share a common colonization history. Instead, it may be necessary to tailor source pools for each individual island (Graves and Gotelli 1983) or to control for island differences statistically (Schoener and Adler 1991). We suspect that many of the island archipelagoes that have been used for matrix analysis, such as the West Indies, probably should be analyzed on an island-by-island basis, rather than as “replicates” within an archipelago.

Published presence-absence matrices almost never include empty rows or columns, that is, sites with no species or species that occur on no sites. However, these “degenerate” arrangements may be very important in assessing nonrandomness of species on the sites they do occur on (Reddingius 1983). For example, suppose the source pool for a pair of islands contains many species that never occur on either island. In this case, null models that use the presence-absence matrix to construct the sampling universe will overestimate the expected number of shared species (Wright and Biehl 1982). We believe that the initial decisions about the dimensionality of the matrix may be more important in determining the results than the mechanics of the null model used.

There are other limitations of presence-absence matrices. By their very nature, these matrices do not include any information on absolute or relative abundances of species (Haila and Järvinen 1981). If abundances are known, more powerful null models of community assembly can be built (Gotelli et al. 1987; Graves and Gotelli 1993). The analysis of presence-absence matrices assumes that mechanisms controlling co-occurrence are reflected in the dominant or easily censused life-history stage (Pearson 1986). But in many assemblages, larval dispersal (Roughgarden et al. 1988) and interactions among early life-history stages (Wilbur 1988) are responsible for community patterns in the adult stage. Presence-absence matrices also require that residency criteria be unambiguous (Connor and Simberloff 1978). But it is sometimes unclear whether a species is “present” on an island. Breeding status may be difficult to establish, and nonbreeding species may have significant ecological effects on residents (Lynch and Johnson 1974; Simberloff 1976a). Finally, presence-absence matrices assume that all sites have received equal sampling effort, which may not be true even for well-studied archipelagoes (Connor and Simberloff 1978).

In spite of these limitations, the analysis of presence-absence matrices holds the promise of revealing general patterns of community structure (Pearson 1986). In some cases, prescriptions for conservation biology have been made on the basis of presence-absence matrices (Patterson 1987), so it is especially important that their analysis be on solid ground.

TWO MODES OF ANALYSIS

Using the terminology of numerical taxonomy (Sneath and Sokal 1973), there are two modes of analysis for a presence-absence matrix (Simberloff and Connor 1979). The Q-mode analysis assesses the similarity of different columns, revealing how similar sites are in the species they contain. In R-mode

analysis, we compare the rows of the matrix and ask how similar species are in the set of islands they occupy. Both modes of analysis are relevant to the issue of assembly rules, although the Q-mode analysis has its roots in biogeographic studies of faunal similarity (McCoy and Heck 1987).

Q-MODE ANALYSES IN BIOGEOGRAPHY

Given a list of species for two sites, how can we quantify the degree of similarity between the sites? The sites will obviously be most similar if they share identical species lists and most dissimilar if they share no species. In analyses of biogeography, a plethora of correlated indices has been used to quantify site similarity (Simpson 1960; Cheetham and Hazel 1969; Jackson et al. 1989). For example, Jaccard's (1908) index scales similarity to range from 0 to 1:

$$J = N_c / (N_1 + N_2 - N_c) \quad (7.1)$$

N_1 and N_2 are the number of species present on each site, and N_c is the number of species common to both sites. As with measures of species diversity (Chapter 2) and niche overlap (Chapter 4), these indices are sensitive to sample size—the value of the index depends as much on the total species richness as it does on the number of shared species. Although some indices have been derived from probability theory (Goodall 1966, 1974; Baroni-Urbani and Buser 1976), most lack an underlying statistical distribution. Consequently, it is difficult to say whether a particular value of an index is statistically significant.

If such indices are to be used at all, they must be combined with simulations to assess the degree to which the index is unusually large or small (Wolda 1981). Rice and Belland (1982) used this approach to examine the similarity of moss floras in five lithophysiographic regions of Bonne Bay, Newfoundland. They combined the species lists from all five regions to generate an aggregate species pool, and weighted each species by its number of occurrences (one to five). Without such weighting, the analysis would have assumed that all species have equivalent probabilities of dispersal and persistence. From the weighted species pool, they drew randomly the observed number of species found in each region, and then calculated Jaccard's index between all possible pairs of sites. The simulations were repeated 20 times to generate 95% confidence intervals for Jaccard's index. The result was that all the observed pairwise similarities were less than expected, and half of the combinations were significantly less than expected (Figure 7.1). This suggests that the floral regions are indeed distinct, although if the regions were originally delineated solely on the basis of these species lists, the test would be circular.

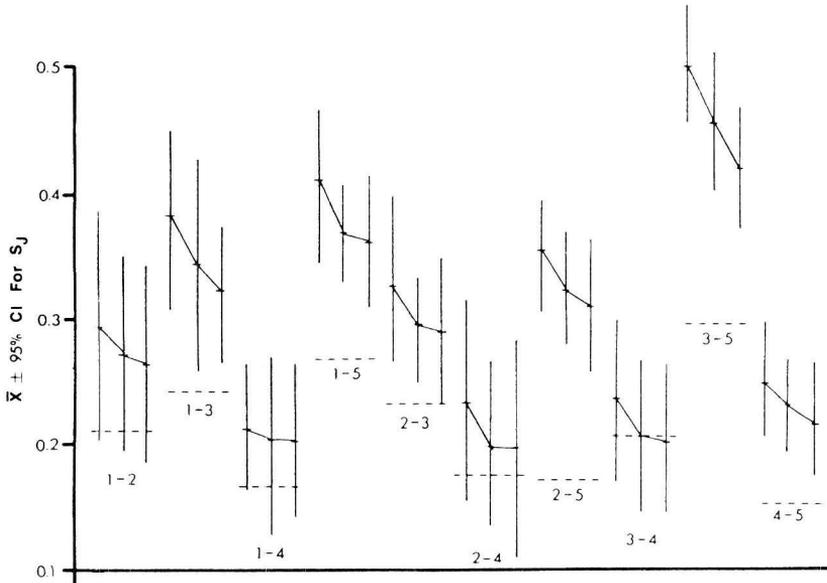


Figure 7.1. Mean and 95% confidence intervals for randomized values of Jaccard's coefficient of similarity between all possible pairs of five floristic provinces. The triad of vertical lines represents different source pool designations for mosses of Newfoundland. The dashed line is the observed similarity between pairs of provinces. From Rice and Belland (1982), with permission.

In a biogeographic context, faunal provinces are delineated by grouping many sites together on the basis of pairwise similarities (Pielou 1979a). A variety of cluster analyses are available for this purpose, which hierarchically group sites on the basis of similarity (Jackson et al. 1989). However, all of these methods assume that the resulting "groups" are biologically meaningful. Because groupings can be made for sets of random numbers, some method is needed for distinguishing groups that are more similar than expected by chance. Strauss (1982) described randomization techniques to ensure that, whatever the clustering algorithm, the clusters are actually more similar than expected by chance. In an evolutionary context, similar bootstrapping methods have been used to assess the significance of phylogenetic reconstructions (Felsenstein 1985).

Q-MODE ANALYSES IN ECOLOGY

In ecological studies, pairwise similarities have also been used to relate faunal similarities to other site characteristics. For example, Terborgh (1973a) re-

gressed pairwise similarities for West Indian avifaunas against interisland distances. He concluded that island size and position explained between 80 and 93% of the variation in avifaunal composition, and attributed the remainder to habitat differences and species interactions. Power (1975) regressed interisland similarities of Galápagos bird and plant species against physical characteristics of the islands. He concluded that bird similarity among islands was explained by plant similarity, but that plant similarity was explained by interisland distance, along with island area and elevation. For Galápagos finches, Abbott et al. (1977) also used similarity indices to evaluate the effects of floral composition on avifaunal composition. All of these analyses are problematic because the similarity indices are ad hoc and because the pairwise points in the regression analysis are not independent of one another.

A Simple Colonization Model (Null Hypothesis 0)

In an unpublished presentation, Johnson (1974; cited in Simberloff 1978a) pioneered a different approach. He used the number of shared species as a simple index of similarity between sites and then asked what the expected number of shared species is under the simplest colonization model (Null Hypothesis 0). If two islands with m and n species, respectively, are colonized randomly by a pool of P equiprobable species, the expected number of shared species (E_{ss}) is

$$E_{ss} = mn / P \quad (7.2)$$

with a variance (from Connor and Simberloff 1978) of

$$\sigma_{ss}^2 = \frac{\sum_{i=0}^m \binom{P-n}{i} \binom{n}{m-i} (m-i - E_{ss})^2}{\binom{P}{m}} \quad (7.3)$$

As a standardized index of site similarity, Connor and Simberloff (1978) recommended

$$(\text{Observed}_{ss} - E_{ss}) / \sqrt{\sigma_{ss}^2} \quad (7.4)$$

For the 29 Galápagos Islands, there are $\binom{29}{2} = 406$ island pairs. Of these, the observed number of shared species exceeded the expected in 369 cases. Connor and Simberloff (1978) obtained similar results at the generic level for Galápagos plants, and Simberloff (1978a) found that all pairs of nine mangrove

islands in the Florida Keys shared more insect species than expected if species had equal colonization probabilities. Note that this hypothesis does not assume that all islands are identical, because differences in species richness among islands are maintained in this model.

A Small-Island Limitation Model (Null Hypothesis I)

Johnson (1974) noted that habitat availability might be responsible for the fact that most sites shared more species than expected compared with Null Hypothesis 0. In particular, species may be missing from small islands if these islands lack critical habitat. A modified protocol also placed species randomly and equiprobably, but the species pool for each island was composed of only those species found on islands of that size or smaller. In this case, the expected number of shared species is mn/P_n where P_n is the number of species in the pool of the larger island ($m \leq n$). Incorporating this small-island limitation improved the fit of the Galápagos plant data to the expected values: 285 of the 406 pairs of islands shared more species than expected. If species number (rather than island size) characterized the lower minimum, 259 of the 406 pairs of islands shared more species than expected.

Limits at the upper size end are also possible. Diamond (1975) hypothesized that certain “supertramp” species were restricted to species-poor communities by diffuse competition. To incorporate this constraint into the null model, the species pool would consist of all species that occurred on islands of a particular size or larger. The extent to which the probability of occurrence of a species is related to community size, island area, or other site attributes is the “incidence function” of the species (Diamond 1975). Incidence functions may serve as realistic constraints in null models of species co-occurrence. They may also represent a type of “assembly rule” that can be tested against other colonization models, as we describe later in this chapter.

A Nonrandom Dispersal Model (Null Hypothesis II)

Null Hypothesis 0 is unrealistic, in part because it assumes species are identical. But even if colonization were stochastic, species would be expected to occur at different frequencies on islands because they differ in their abilities to disperse and persist. Dispersal and persistence abilities are probably a function of many species-level attributes, including population size (Terborgh and Winter 1978) and variability (Karr 1982a), body size (Pimm et al. 1988), and geographic range (Graves and Gotelli 1983; Jablonski 1986). When species occurrence probabilities can be measured independently, they can be a power-

ful tool in null model analyses. But for most archipelagoes, this information is lacking. One must either assume (unrealistically) that species occurrence probabilities are equal or somehow estimate those probabilities from the matrix itself.

The “occurrence distribution” (Connor and Simberloff 1978) has been used to weight species, so that the probability of occurrence of a species is proportional to its frequency of occurrence (Abele and Patton 1976). Many critics have objected that this procedure is circular—the incidence data are used to estimate probabilities of occurrence, which are then used to test the pattern of co-occurrence (Grant and Abbott 1980; Diamond and Gilpin 1982; Case and Sidell 1983). However, constraining marginal totals does not uniquely determine the pattern of co-occurrence in the data (Manly 1991). As in contingency table analysis (Fienberg 1980), some arrangements of the data may be highly improbable given a set of marginal constraints (Connor and Simberloff 1983), and this should be true whether or not the marginals themselves are influenced by competition (Simberloff 1978a). As we shall see, the consequences of marginal constraints on the resolution of co-occurrence patterns depend greatly on whether the constraints are absolute or probabilistic.

Analytic expressions for Null Hypothesis II are not known, so the expected number of shared species must be determined with a simulation. In such a simulation, species are drawn from an aggregate source pool for the archipelago, with probabilities weighted by species occurrences. As in Null Hypotheses 0 and I, the observed number of species on each island is maintained as an absolute constraint. For the Galápagos plants, 338 of the 406 island pairs had more shared species than expected under Null Hypothesis II (Connor and Simberloff 1978). Johnson (1974) performed a mixed-model simulation that simultaneously incorporated both the small-island limitation (Null Hypothesis I) and the weighted species pool (Null Hypothesis II). For the Galápagos plant data, this was the only analysis in which there were too many island pairs with fewer shared species than expected. However, only about 5% of the pairwise comparisons were statistically significant (Figure 7.2).

From this result, Simberloff (1978a) concluded that the mixed model “might not be far from an accurate description of colonization in this archipelago.” He also reasoned that an archipelago structured by competition should show a predominance of island pairs with fewer shared species than expected. However, he warned that shared species number is a weak statistic for detecting diffuse competition, because it says nothing about taxonomic or ecological relationships of the species that do co-occur.

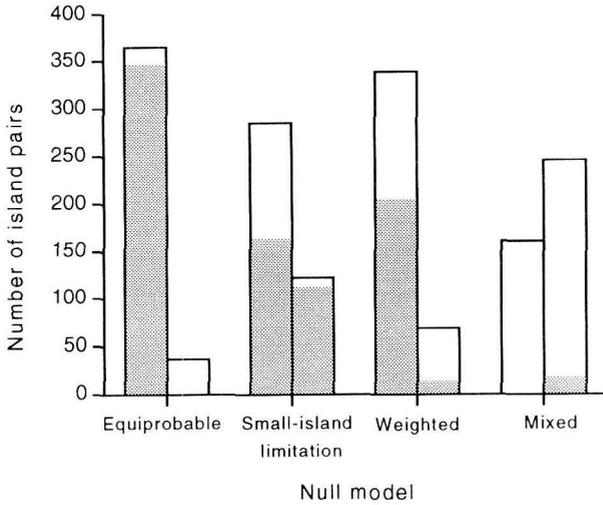


Figure 7.2. Four null models of shared plant species between pairs of Galápagos Islands. For each model, the left-hand bar represents the number of island pairs that shared more species than expected and the right-hand bar represents the number of island pairs that shared fewer species than expected. The hatched region represents the number of pairs that deviated at the 5% level of significance. “Equiprobable” is the model in which all species had equal colonization probabilities. “Small-island limitation” allowed for equiprobable colonization but restricted species occurrences to the minimum island size observed. “Weighted” assigned colonization probabilities for each species proportional to the number of island occurrences. “Mixed” model included both the small-island limitation and weighted colonization probabilities. If species occurrences were independent of one another, the height of the two bars would be approximately equal, and roughly 2.5% of each bar would be shaded. Note the predominance of island pairs that shared more species than expected for the first three null models. Data from Simberloff (1978a).

Criticisms of Ecological Q-Mode Analysis

Wright and Biehl (1982) argued that shared-island analyses may not reveal competitive effects, because many of the pairwise island comparisons are between islands that have the same species sets. Consequently, the shared-species analysis would fail to detect a highly significant “checkerboard distribution” in which a pair of species never occur together on the same island (Diamond 1975). The problem is not as severe if more than a single pair of species occurs in a checkerboard. With an assemblage composed of several competitive guilds, some shared species patterns may be improbable, although

it will still be difficult to detect competition from these sorts of data (Connor and Simberloff 1984).

Hamill and Wright (1988) pointed out two other complications with Q-mode analysis. First, the number of species pairs should be tallied above and below the simulated median, not the simulated mean (see also Hendrickson 1981). Second, because the pairs of islands are not independent, it is not appropriate to ask whether more than 5% of the pairs are significantly different from the expectation. They recommended examining shared species among randomly selected, independent pairs of islands, rather than among all possible pairs. These island pairs could be compared to randomizations for each, rather than to a single randomization of the entire presence-absence matrix. With these refinements, Hamill and Wright (1988) found a good agreement between observed and expected number of shared species for the Galápagos avifauna. In contrast, Connor and Simberloff's (1978) analyses suggested a significant excess of shared species among islands. Neither analysis implicated a role for interspecific competition in the distribution of the Galápagos avifauna, although these studies did not consider body size differences among coexisting species (see Chapter 6).

We think that using Q-mode analysis to detect competition is a case of the right answer to the wrong question. Q-mode analysis should be used as a biogeographic tool to group sites on the basis of similarity in species composition, rather than to infer species interactions on the basis of dissimilarity. The analysis is trivial for a single pair of species, but for larger assemblages, Equation 7.4 is a natural index for classifying pairs of sites on the basis of species similarity. Of course, such an analysis depends on the assumption of a common species pool and on whether small island constraints and species weighting are retained, but these assumptions are implicit in any similarity analysis. By generating an explicit null expectation and allowing for increasingly complex colonization models, Q-mode analysis provides a reliable framework for comparing species composition of sites (McCoy and Heck 1987). If the question is one of species interactions, then an R-mode analysis, which compares species in the sites they share, is more appropriate.

ASSEMBLY RULES

Missing Species Combinations

Most analyses of presence-absence matrices have examined the number of shared sites for a set of species. These R-mode analyses have emphasized the

detection of nonrandom patterns of species association. Two broad classes of interactions can lead to distinctive species combinations. (1) *Interactions of species and sites*. These include site characteristics, such as island area, isolation, and habitat availability, and species characteristics, such as dispersal, persistence, and habitat affinities. Even if species colonize sites independently of one another, interactions between species and sites may lead to characteristic species combinations. (2) *Interactions between species*. Competition, predation, and mutualism between particular species pairs and among larger groups of species will also lead to nonrandomness. Biotic interactions may lead to “forbidden combinations” of species (Diamond 1975) and also to combinations of species that are unusually stable and occur more frequently than expected.

The search for assembly rules of community organization is predicated on the assumption that biotic interactions are strong enough to produce discernible patterns. The null model approach has been to build colonization models that incorporate only species-site interactions (autecological factors) and to compare their predictions to the empirical data. The goal is to provide a baseline for recognizing nonrandom patterns caused by species interactions. Because a variety of forces can lead to either positive or negative species associations (Schluter 1984), it is a challenge to properly frame and interpret null models for assessing species interactions. Even if a pattern can be shown by a null model analysis to be nonrandom, it may not be possible to determine the underlying cause (Simberloff and Connor 1981).

Statistical Tests for Missing Species Combinations

The pioneering work of E. Chris Pielou (Pielou and Pielou 1968; Pielou 1972a) on the R-mode analysis of species associations was ignored and uncited for over a decade (Simberloff and Connor 1981). Pielou and Pielou (1968) noted that for presence-absence matrices with r species (= rows), there are 2^r possible species combinations (including the combination of none of the r species being present). With enough replicated sites, the observed and expected frequencies of these combinations can be compared with a chi-squared test. This comparison holds the row totals of the presence-absence matrix constant. The null hypothesis is that the observed frequency of different species combinations is no different than expected by chance, given that some species in the collection are common and others are rare.

Unfortunately, this test is rarely practical, because the number of species combinations is usually very large compared to the number of sites. Pielou and Pielou (1968) instead advocated a simulation procedure in which each species

is placed randomly and independently in sites. From these simulations, one can obtain the expected number of different species combinations.

The observed number of combinations may be less than this expectation either because (1) some sites are uniformly better than others for all species, or because (2) species associate nonrandomly with one another or with different sites. In either case, there will be fewer combinations of species than expected under the null model. By carefully comparing the number of occupied sites with the curve generated by the null model, it may be possible to distinguish between segregative and nonsegregative association. Pielou and Pielou (1968) did this for 13 collections of insect and spider fauna associated with *Polyperus* fungus brackets. The number of combinations did not differ from the expected in five collections. In the remaining eight, there were fewer associations than expected, although there was evidence for segregative association in only three of the samples.

Two restrictions are important in Pielou and Pielou's (1968) analysis. First, all sites are assumed equiprobable. If this assumption can be relaxed, it may be easier to reveal the effects of species interactions. Second, the analysis requires an estimate of the frequency of unoccupied sites. Unfortunately, these data are not usually collected, and ad hoc procedures for dealing with empty sites (e.g., Siegfried 1976) may not be valid (Simberloff and Connor 1981).

Whittam and Siegel-Causey (1981a) expanded on Pielou and Pielou's (1968) method and used a contingency table analysis to test for species associations in Alaskan seabird colonies. Rather than looking just at the number of species combinations, Whittam and Siegel-Causey (1981a) used a series of hierarchical log linear models to tease apart positive and negative species associations. Observed species combinations differed in frequency from those expected under a model of independent assortment. Models that best fit the data included mostly positive two-way associations between species (Figure 7.3). Higher-order interactions were uncommon, and positively associated pairs of species tended to overlap in diet. These sophisticated analyses derive their strength from very large sample sizes [20 species, 902 colonies (= sites), and 19 diet categories] and provide considerable insight into species interactions.

Pielou (1972a) developed a second type of R-mode analysis. She analyzed the variance of the row sums in the presence-absence matrix, that is, the variance in the frequency of occurrence of each species. If the species are distributed independently of one another, then the row sums will behave as a series of random variables, with a covariance of zero. As described in Chapter 5, the ratio of variance in total species number in sites to the sum of the variances of individual species provides a simple test for positive or negative covariance in a presence-absence matrix. The variance ratio was derived inde-

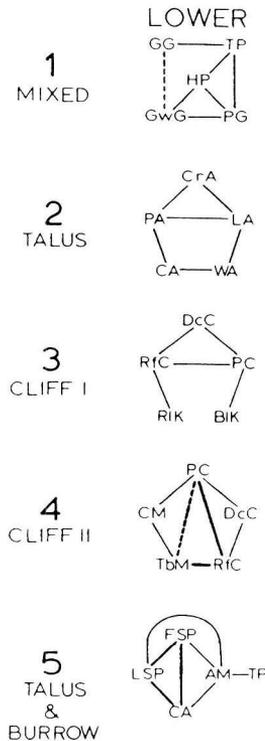


Figure 7.3. Complex species associations in Alaskan seabird colonies, as detected by contingency table analysis. For each of five habitat types, the diagram summarizes statistically significant interactions. Solid line = positive pairwise interaction; dashed line = negative pairwise interaction; heavy line = negative three-way interaction. GG = Glaucous Gull (*Larus hyperboreus*); TP = Tufted Puffin (*Fratercula cirrhata*); HP = Horned Puffin (*Fratercula corniculata*); GwG = Glaucous-winged Gull (*Larus glaucescens*); PG = Pigeon Guillemot (*Cepphus columba*); CrA = Crested Auklet (*Aethia cristatella*); PA = Parakeet Auklet (*Cyclorhynchus psittacula*); LA = Least Auklet (*Aethia pusilla*); CA = Cassin's Auklet (*Ptychoramphus aleuticus*); WA = Whiskered Auklet (*Aethia pygmaea*); DcC = Double-crested Cormorant (*Phalacrocorax auritus*); RfC = Red-faced Cormorant (*Phalacrocorax urile*); PC = Pelagic Cormorant (*Phalacrocorax pelagicus*); RfK = Red-legged Kittiwake (*Rissa brevirostris*); BfK = Black-legged Kittiwake (*Rissa tridactyla*); CM = Common Murre (*Uria aalge*); TbM = Thick-billed Murre (*Uria lomvia*); FSP = Fork-tailed Storm Petrel (*Oceanodroma furcata*); LSP = Leach's Storm Petrel (*Oceanodroma leucorhoa*); AM = Ancient Murrelet (*Synthliboramphus antiquus*). From Whittam and Siegel-Causey (1981a), with permission.

pendently by Järvinen (1979) in the analysis of temporal fluctuations and by Schluter (1982) in the analysis of spatial distributions.

The variance ratio and related tests do not focus on assembly rules per se or missing species combinations, but on the average amount of association between each species and all others in the matrix. Schluter (1984) analyzed several published presence-absence matrices with the ratio test and found that positive covariation, rather than neutral or negative co-occurrence, was the most common pattern. However, the variance ratio test assumes that species number per site can vary freely, which may account for the predominance of positive aggregations.

Simberloff and Connor (1981) thoroughly reviewed the literature on missing species combinations. They used Maxwell-Boltzmann statistics to calculate the probability of observing a particular number of missing species combinations in an archipelago. The Maxwell-Boltzmann analysis assumes that each combination of species is an equiprobable group of colonists. An alternative analysis assumes that the probability of a particular combination is proportional to the product of the colonization probabilities of the component species. Thus, combinations that include common (= widespread) species are more probable than combinations with rare species. This proportional weighting always raises the observed tail probability. If some combinations are more likely than others, the number of combinations expected under the null model decreases.

Simberloff and Connor (1981) investigated claims that missing species combinations were nonrandom for assemblages of plants (Abbott 1977), birds (Siegfried 1976; Abbott et al. 1977), and mammals (Grant 1972b; M'Closkey 1978; King and Moors 1979). Few significant patterns emerged from these analyses—in most cases, the observed number of missing species combinations was about what one would expect, given a random sample of all species in the archipelago.

One notable exception was the Galápagos finches, which had fewer multi-species combinations than expected. With six species of *Geospiza*, there are $\binom{6}{3} = 20$ possible three-species combinations. But only two of the 20 possible combinations were found on the islands ($p = 0.000002$; weighted $p = 0.001$). On the other hand, the observed number of one- and two-species combinations did not differ from the null model, and the significance of the multispecies patterns hinged upon which species and islands were included in the source pool. Nevertheless, the presence of very few species combinations of Galápagos finches is consistent with the evidence of nonrandom body size patterns in this assemblage (see Chapter 6).

Diamond's Assembly Rules

Diamond (1975) popularized the study of community assembly rules in a detailed treatise on the distribution of 141 land-bird species on New Guinea and its satellite islands in the Bismarck Archipelago. Diamond (1975) summarized many years of his own field studies on species-area relationships, incidence functions, species combinations, and resource use patterns in this archipelago. Although Diamond (1975) discussed the importance of factors such as dispersal, habitat availability, and chance colonization, he emphasized that interspecific competition within groups of related species (ecological guilds) was the most important determinant of observed species combinations.

Diamond (1975) codified and generalized his findings in a list of seven "rules" of community assembly:

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."*
3. *"A combination that is stable on a large or species-rich island may be unstable on a small or species-poor island."*
4. *"On a small or species-poor island a combination may resist invaders that would be incorporated on a larger or more species-rich island."*
5. *"Some pairs of species never coexist, either by themselves or as part of a larger combination."*
6. *"Some pairs of species that form an unstable combination by themselves may form part of a stable larger combination."*
7. *"Conversely, some combinations that are composed entirely of stable sub-combinations are themselves unstable."*

The publication of Diamond's (1975) rules touched off an acrimonious debate over null models that has spanned more than 20 years in the ecological literature [see Wiens (1989) for an even-handed review of the controversy]. In this section, we review the null models that have been used to test Diamond's (1975) rules, and the various claims and counterclaims that have followed. We think that all of the null models proposed to date contain one or more serious flaws. Later in this chapter, we propose a hybrid model that combines the best features of each and seems to overcome the most serious deficiencies.

First, we need to consider the nature of the assembly rules. Wiens (1989) noted that Diamond's (1975) rules are statements of pattern and do not explicitly describe mechanisms, although they do require that communities be in an equilibrium state. Nevertheless, the use of terms such as "forbidden combinations" clearly implicates the role of competitive interactions and implies that the patterns are nonrandom and would not be expected in the absence of species interactions.

It has been very difficult to make these rules operational. On the one hand, the rules as stated are so broad that they would be difficult to apply to real data, much less test them (Haefner 1988b). Null model analyses have contributed considerably to the study of assembly rules by forcing the issue of precisely what patterns must be established in order to document a competitively structured community. On the other hand, Diamond's (1975) own interpretation of these rules for the Bismarck Archipelago seems so detailed, anecdotal, and nonstatistical that it resembles more an historical narrative than a test of community assembly. We agree with Simberloff (1978a) that "one is left with the uneasy feeling that the rules lack predictive power, in that all the data are required before any prediction can be made."

The Connor and Simberloff Procedure

Connor and Simberloff (1979) launched an aggressive attack on Diamond's (1975) assembly rules. They argued that Rules 2, 3, 4, 6, and 7 were either tautologies or restatements of other rules. Connor and Simberloff (1979) did test Rules 1 and 5 with a null model analysis. Rules 1 and 5 are identical, except that Rule 1 mentioned "related species" and Rule 5 was restricted to species pairs. Rule 5 described a checkerboard pattern of species occurrences, which is perhaps the simplest and most clear-cut of Diamond's (1975) assembly rules (Graves and Gotelli 1993). The requirement of a complete distributional checkerboard is especially stringent: two species might have exclusive distributions on 99 islands, but if they occurred together on a single island, the pair would not be scored as a true checkerboard. Diamond (1975) presented seven examples of pairs of closely related, ecologically similar species that were distributed as checkerboards in the Bismarck Archipelago.

The checkerboard patterns (Figure 7.4) were striking and provided the strongest evidence for Diamond's (1975) hypothesis. For two of the examples (*Macropygia* cuckoo-doves and *Pachycephala* flycatchers), the probability that the species within each genus formed a distributional checkerboard [the "shared islands" analysis of Wright and Biehl (1982)] was unusually small (Connor and Simberloff 1979).

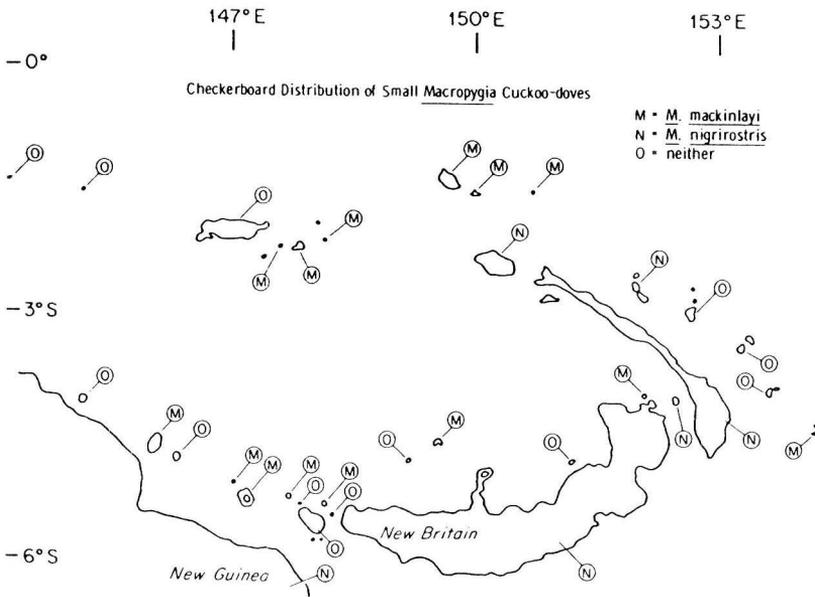


Figure 7.4. Checkerboard distribution of two *Macropygia* cuckoo-dove species in the Bismarck Archipelago. Each island supports only one of the two species. M = *Macropygia mackinlayi*; N = *Macropygia nigrirostris*; O = neither species present. From Diamond (1975). Reprinted by permission of the publisher from *Ecology and Evolution of Communities*. M. L. Cody and J. M. Diamond (eds). Cambridge, Mass.: The Belknap Press of Harvard University Press. Copyright © 1975 by the President and Fellows of Harvard College.

But even these examples may not be so clear-cut. *Macropygia nigrirostris* occurred mostly on large, species-rich islands, whereas its putative competitor (*M. mackinlayi*) occurred mostly on small and medium-sized islands with few species. If these incidence constraints were not a direct result of competitive interactions, then the resulting checkerboard may be due to independent colonization by each species of different-sized islands (Wiens 1989). Moreover, Diamond's (1975) Figure 20 of this example (and also Diamond's (1975) Figure 22 for the genus *Ptilinopus*) did not show all of the "empty islands" that supported neither species. When these were incorporated into the calculation, the *p* value was marginally nonsignificant (Connor and Simberloff 1983).

However, Connor and Simberloff (1979) were not satisfied with the demonstration that some pairs of species showed unusually exclusive distributions. They argued that, with $\binom{141}{2} = 9,870$ possible species pairs, it was not surpris-

ing that seven pairs showed exclusive distributions. Malanson (1982) made a similar point about checkerboard distributions of plant species in the hanging gardens of Zion National Park.

Because Diamond (1975) did not provide data on all guild designations for this assemblage, we cannot say how improbable it is that all five pairs would be ecologically similar congeners. In order to assert that species checkerboards are a manifestation of community assembly rules, Connor and Simberloff (1979) required that the observed number of pairs, trios, etc. of species that exhibit checkerboards be significantly greater than the expected number generated by an appropriate null model.

Diamond's (1975) Bismarck data were never published, so Connor and Simberloff (1979) instead used presence-absence data for West Indian birds and bats, and New Hebridean birds. Connor and Simberloff (1979) used a null model that randomized an observed presence-absence matrix subject to the following three constraints:

1. The row totals of the randomized matrix were maintained.
2. The column totals of the randomized matrix were maintained.
3. For each row, species occurrences were restricted to those islands for which total species richness fell within the range occupied by the species.

Constraint (1) maintained differences between species in their frequency of occurrence. Constraint (2) maintained differences among islands in the number of species they contained. Constraint (3) maintained the observed incidence function for each species (it could not occur in assemblages larger or smaller than those observed). The simulation first ordered the rows of the matrix from most common to least common species, and then randomly placed species on islands until all three constraints were satisfied.

When the matrix was sparsely filled, this algorithm was satisfactory. But if many of the species were widespread (= large row totals), it was sometimes impossible to place species late in the simulation and still maintain the constraints. In these cases, Connor and Simberloff (1979) repeatedly interchanged submatrices in the matrix so that row and column sums were not altered. After the interchanges, new matrices that were nonequivalent to the observed matrix were retained to estimate the null distribution. More recent advances in randomization algorithms (Stone and Roberts 1990; Daniel Simberloff, personal communication) have overcome this problem.

Connor and Simberloff (1979) used 10 such randomizations to construct a histogram of the mean and standard deviation of the number of pairs (or

Table 7.2

Null model analyses of co-occurrence in island faunas

Taxon	Number of possible species pairs	Observed number of checkerboard species pairs	Expected number of checkerboard species pairs	<i>p</i>
New Hebrides birds	99	1	0.90	>0.99
West Indies birds	1,029	621	437.0	<10 ⁻⁸
West Indies bats	499	325	208.6	<10 ⁻⁸

For each fauna, the table lists the total number of confamilial pairs, the expected number of checkerboard pairs, and the observed number of checkerboards. The chi-squared significance test is for the overall fit of the distribution. Note the excess of checkerboard pairs for the West Indian faunas. From Connor and Simberloff (1979) and Diamond and Gilpin (1982).

triplets) of species that shared 0, 1, 2 . . . *n* islands. The first bar of this histogram is the zero class and represents the expected number of species pairs that share no islands, i.e., have a checkerboard distribution. Connor and Simberloff (1979) lumped adjacent size classes that had small expected values and then compared the entire distribution of observed and expected frequencies with a chi-squared test, using *n*-1 degrees of freedom for an archipelago with *n* islands. They tested the fit of the observed and expected number of shared islands for pairs and trios of species at the level of the entire assemblage and within taxonomic families.

What was the outcome of these tests? For the New Hebrides bird data, the fit between observed and expected was suspiciously good (*p* > 0.99). This matrix had to be randomized by the interchange of submatrices, and Connor and Simberloff (1979) noted that the matrices randomized in this way were very similar to the observed matrix. They pointed out that there were very few arrangements of this matrix that maintained row, column, and incidence constraints. As a consequence, the evidence for competition was weak, unless one were to argue that competition was responsible for the marginal totals.

Patterns for West Indian birds and bats were quite different. In all cases, the null hypothesis was strongly rejected, and in all analyses, there were more exclusively distributed species pairs (checkerboards) than expected by chance (Alatalo 1982). Connor and Simberloff (1979) did not emphasize these highly nonrandom results (Table 7.2). Instead, they presented graphical comparisons of observed and expected species combinations in each frequency class (Figure 7.5). From these, they argued that the overall fit of the observed and

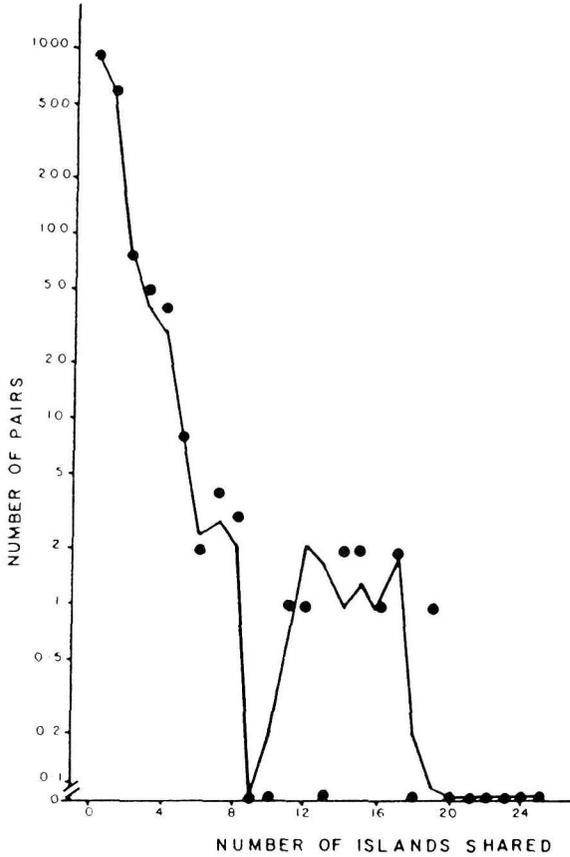


Figure 7.5. Observed and expected numbers of species pairs versus number of islands shared for West Indies birds. The solid line is the expected value generated by a null model that held row and column totals constant and maintained species area restrictions. Dots are the observed values. Although the logarithmic scale suggests a good fit between the observed and expected values, the difference is highly significant ($p < 0.01$) and in the direction predicted by the competition hypothesis (more species pairs that share zero islands than expected by chance). See also Table 7.2. From Connor and Simberloff (1979), with permission.

expected data was good, and that therefore Diamond's (1975) seven examples of checkerboards were not a compelling demonstration of Assembly Rules 1 and 5. For birds of the Bismarck Archipelago, and for the West Indian bird and bat faunas, Connor and Simberloff (1979) suggested that allopatric speciation and limited dispersal of single-island endemics were alternative hypotheses

that did not invoke interspecific competition but might account for checkerboard distributions.

The Connor and Simberloff (1979) model has been used in two additional tests. Matthews (1982) analyzed the occurrence of 13 minnow species distributed in six streams of the Ozark watershed. Although some species pairs that never co-occurred in watersheds were morphologically and ecologically similar, the observed number of checkerboard pairs matched the predictions of the null model. However, Matthews's (1982) analysis was based on a binomial distribution, sampling with replacement, whereas a more appropriate analysis uses the hypergeometric distribution, sampling without replacement (Biehl and Matthews 1984). Jackson et al. (1992) used the Connor and Simberloff (1979) model to analyze five presence-absence matrices for Ontario lake fish and also found no evidence of nonrandomness.

Criticisms of the Connor and Simberloff Procedure

Connor and Simberloff's (1979) analysis provoked several critiques (Alatalo 1982; Diamond and Gilpin 1982; Gilpin and Diamond 1982, 1984; Gilpin et al. 1984) and subsequent rebuttals (Connor and Simberloff 1983, 1984; Gilpin et al. 1984; Simberloff and Connor 1984). Here, we summarize the most important of the criticisms:

1. The dilution effect. Because Connor and Simberloff (1979) analyzed confamilial groups or entire avifaunas, competitive effects were not apparent. Diamond's (1975) choice of examples suggested that the ecological guild was the correct unit of analysis for revealing competitive effects. However, guilds must be established a priori by criteria that are independent of the co-occurrence data being tested (Connor and Simberloff 1983). Delineating guilds is not an easy task (Jaksić and Medel 1990; Simberloff and Dayan 1991), and guild designations clearly affect the outcome of null model tests.

For example, Graves and Gotelli (1993) tested the significance of checkerboard distributions in mixed-species flocks of Amazonian forest birds. There was no evidence of unusual patterns for the entire assemblage of flocking species, or for species grouped into ecological foraging guilds. Only when the analysis was restricted to congeneric species within feeding guilds was there evidence of unusual checkerboard distributions. Even at this level, the patterns were statistically significant only for null models based on abundance and population structure, rather than presence-absence data (Table 7.3). Vuilleumier and Simberloff (1980) also found that co-occurrence patterns of Andean birds were affected by the designation of ecological and taxonomic guilds in

Table 7.3

Observed and expected numbers of perfect checkerboard distributions among pairs of Amazonian bird species in mixed species flocks

Level	<i>n</i>	Obs.	Null model					
			SPEC		ABUN		DEMO	
			Exp.	<i>p</i>	Exp.	<i>p</i>	Exp.	<i>p</i>
Flock	71	1,012	1,018.91	0.63	927.58	<0.01	992.47	0.25
Guild								
1	27	111	122.11	0.912	94.04	0.027	114.73	0.706
2	14	36	38.12	0.758	40.47	0.876	38.34	0.783
3	8	10	7.29	0.137	6.37	0.046	7.28	0.123
4	3	3	2.63	0.667	2.46	0.526	2.65	0.687
5	10	14	11.04	0.142	10.14	0.095	9.84	0.046
6	7	15	14.70	0.576	14.81	0.602	14.62	0.556
7	2	1	0.89	0.887	0.88	0.879	0.89	0.893
Fisher's combined probabilities test (df = 14)			$\chi^2 = 10.77$ <i>p</i> > 0.50		$\chi^2 = 20.91$ <i>p</i> > 0.10		$\chi^2 = 13.68$ <i>p</i> > 0.50	

Genera								
<i>Monasa</i>	2	0	0.03	1.000	0.13	1.000	0.09	1.000
<i>Xiphorhynchus</i>	4	2	0.35	0.041	0.24	0.024	0.32	0.036
<i>Philydor</i>	2	0	0.00	1.000	0.00	1.000	0.00	1.000
<i>Automolus</i>	3	3	1.28	0.045	0.09	0.014	1.30	0.052
<i>Xenops</i>	2	1	0.05	0.054	0.07	0.067	0.06	0.062
<i>Thamnomanes</i>	2	0	0.00	1.000	0.00	1.000	0.00	1.000
<i>Myrmotherula</i>	7	3	1.30	0.119	0.89	0.048	0.95	0.051
<i>Hylophilus</i>	2	1	0.19	0.189	0.02	0.020	0.14	0.142
Fisher's combined probabilities test (df = 16)			$\chi^2 = 24.76$ $p < 0.10$		$\chi^2 = 33.43$ $p < 0.01$		$\chi^2 = 27.86$ $p < 0.05$	

Three null models were tested: (1) SPEC (randomization of presence-absence matrices), (2) ABUN (randomization of abundance data), and (3) DEMO (randomization of abundance data with intraspecific demographic constraints). These analyses were carried out at three hierarchical levels: (1) entire flocks, (2) foraging guilds, and (3) congeneric groups within feeding guilds. Guild designations: 1 = arboreal gleaning insectivores; 2 = arboreal sallying insectivores; 3 = arboreal dead-leaf-searching insectivores; 4 = bark interior insectivores; 5 = superficial bark insectivores; 6 = arboreal omnivores; 7 = arboreal frugivores. Note that significant checkerboards were consistently revealed only for congeners within the same feeding guild. From Graves and Gotelli (1993).

the source pool. Although the designation of guilds is a crucial step in the analysis of co-occurrence matrices, it is a procedure that is quite distinct from the null model randomizations themselves.

2. *Effects of randomization constraints.* The three simultaneous constraints imposed by Connor and Simberloff (1979) were severe and made it less likely that the null hypothesis would be rejected. For example, relaxing the “incidence constraint” prevented the simulation for the New Hebrides matrix from hanging up and revealed significant negative associations between species (Wilson 1987). For a similar model with only row and column sum constraints, Wilson et al. (1992) detected no evidence of nonrandomness in the distribution of rock-pool algae, whereas both positive and negative associations were revealed in the flora of islands in Lake Manapouri, New Zealand (Wilson 1988).

However, if row and column totals are constrained, some, but not all, checkerboard distributions cannot be detected by the null model (Diamond and Gilpin 1982; Connor and Simberloff 1984). In an empirical comparison of several R-mode analyses, the Connor and Simberloff (1979) procedure was the only one that could not detect nonrandomness in presence-absence matrices for fish assemblages (Jackson et al. 1992). More formally, Roberts and Stone (1990) showed that the average number of shared islands among all species pairs cannot change in a simulation if row and column totals are constrained. Somewhat paradoxically, matrices with mutually exclusive species pairs will always contain some species pairs that co-occur more than expected (Stone and Roberts 1992).

The sample variance of the number of shared islands reflects this pattern. For the New Hebrides bird data, this variance was significantly larger than expected under the null model that held row and column totals constant, suggesting that some species pairs shared too many islands and others shared too few (Roberts and Stone 1990). Indices of “checkerboardedness” (Stone and Roberts 1990) and “togetherness” (Stone and Roberts 1992) also confirmed that there were too many exclusive and aggregated pairs of species within con-familial subsets of the New Hebridean avifauna.

Some authors (Grant and Abbott 1980; Colwell and Winkler 1984) have claimed that it is circular to constrain marginal totals, because the marginals also reflect interspecific competition. This claim has never been validated, but if it is true, then the Connor and Simberloff (1979) model will reveal only competitive effects above and beyond those expressed in the row and column sums. The hypothesis that competition sets the total number of island occurrences for a particular species (or the total number of species on a particular

island) is quite distinct from the hypothesis of forbidden species combinations, and deserves to be tested in its own right. There are many reasons besides competition that species occur on many or few islands and that species number varies among islands. We agree with Connor and Simberloff (1983) that these factors need to be incorporated into null models, although imposing absolute marginal constraints may be too severe a restriction.

Finally, the debate over marginal constraints also reflects the way in which the null models are interpreted. As we noted in Chapter 1, one interpretation is that the null model is simply a statistical randomization. This interpretation is consistent with the use of absolute marginal constraints, because the question is whether the co-occurrence patterns are nonrandom, given the observed “sample” of species and islands. On the other hand, if the randomization is viewed as a model of community colonization in the absence of competition, the marginal constraints may not be appropriate. In a group of randomly colonized archipelagoes, we would not expect each replicate of an island to have exactly the same number of species, nor would we expect each species to occur on exactly the same number of islands in the different archipelagoes. We develop these ideas later in this chapter.

3. *Significance tests.* Connor and Simberloff (1979) compared the observed and expected distributions with a chi-squared test. However, this may be inappropriate because of the nonlinear dependence imposed by the marginal constraints. Even if the chi-squared test were appropriate, the number of degrees of freedom used by Connor and Simberloff (1979) was roughly double the “best fit” value found empirically (Roberts and Stone 1990). The same criticism applies to Gilpin and Diamond’s (1982) null model, which also compared observed and expected distributions of species pairs with a chi-squared test.

The Wright and Biehl Procedure

Wright and Biehl (1982) suggested a “shared-island” test for detecting unusual species co-occurrences. For each species pair, they calculated the tail probability of finding the observed number of co-occurrences. The calculation is identical to Connor and Simberloff’s (1978) “shared species” analysis, but with the rows and columns of the presence-absence matrix transposed. Wright and Biehl (1982) argued that an assemblage exhibited nonrandomness if more than 5% of the species pairs shared more islands than expected (at the 5% significance level). By this test, the New Hebridean avifauna exhibited aggregation, because 8% of the species pairs shared significantly more islands than expected by chance.

One advantage of the Wright and Biehl (1982) method is that it directly pinpoints particular species pairs that show aggregated or segregated distributions. However, assessing the statistical significance of all possible non-independent species pairs is problematic. Even if the null hypothesis is rejected for more than 5% of the pairs, this pattern can be caused by nonrandomness in the distribution of only a few species (Connor and Simberloff 1983). The same criticism applies to the shared species analysis of Connor and Simberloff (1978) and to Gilpin and Diamond's (1982) null model for R-mode analyses. An excess of species pairs that are significant at the 5% level does not mean that the patterns are biologically meaningful for all of these pairs.

A more serious problem is that Wright and Biehl's (1982) shared island model assumes that all sites are equivalent. Consequently, it confounds species-site associations with the effects of species interactions. The slight excess of New Hebrides species pairs that shared more islands than expected may simply reflect differences in island suitability.

The Gilpin and Diamond Procedure

Gilpin and Diamond (1982) developed their own R-mode analysis as an alternative to the Connor and Simberloff (1979) approach. Gilpin and Diamond's (1982) null model was based on the principles of contingency table analysis. For species i on island j , they calculated the probability of occurrence as

$$P_{ij} = R_i C_j / N \quad (7.5)$$

where R_i is the row total for species i , C_j is the column total for island j , and N is the grand total for the presence-absence matrix. Next, they calculated the expected overlap for each species pair by summing the product of these probabilities across all islands. Observed and expected overlaps for each species pair were standardized and then compared with a chi-squared test to a standard normal distribution. If the null hypothesis of independent placement were true, the histogram of normalized deviates would follow a normal distribution. Species pairs that showed unusual aggregation would appear in the right-hand tail of the distribution, and species pairs that showed unusual segregation would appear in the left-hand tail.

For the New Hebridean birds, the observed distribution of deviates was significantly different from a standard normal, and weakly skewed towards the right. The Gilpin and Diamond (1982) analysis did not reveal any significant negative associations, in contrast to the results of Wilson (1987) and of Stone and Roberts (1992). Returning to Diamond's (1975) original Bismarck data, Gilpin and Diamond (1982) found a strong excess of positive associations and

a weak excess of negative associations. They examined the extreme species pairs in detail and attributed positive associations to shared habitat requirements, clumping of single-island endemics on large islands, shared geographic origins, and shared distributional strategies. Negative associations were attributed to competitive exclusion, differing distributional strategies, and differing geographic origins. Compared to Diamond's (1975) original conclusions, the null model interpretations of Gilpin and Diamond (1982) placed considerably less emphasis on competitive interactions in producing community assembly rules for the birds of the Bismarck Archipelago.

Using the Gilpin and Diamond (1982) model, Jackson et al. (1992) also found mostly positive associations between species pairs of fishes in lakes. These positive associations were usually between pairs of cold-water species with similar habitat requirements. Negative associations were usually between piscivorous fish species and their prey. Finally, McFarlane (1989) found an excess of positive associations for the Antillean bat fauna and attributed this pattern to a large number of single-island endemics in the northern Antilles.

Gilpin and Diamond's (1982) approach was important because it introduced the idea that the marginal totals may represent *expected values* rather than *absolute constraints*. In different runs of a stochastic model, we would not expect each island to support precisely the observed number of species, or each species to always occur with its observed frequency. In fact, putting a strict "cap" on the number of species that can occur on an island, as Connor and Simberloff (1979) did, could be interpreted as a competitive limit to local species richness. In contrast, if islands behave as "targets" that are colonized independently by different species (Coleman et al. 1982), we would expect some variance about the expected species number in our null model. Empirically, a "target" model is consistent with the finding that island area typically explains only 50% of the variation in species number (Boecklen and Gotelli 1984). Although Gilpin and Diamond (1982) did not emphasize this point, their approach was an important first step toward incorporating this sort of variability into null models.

Unfortunately, there are several problems with the Gilpin and Diamond (1982) model. First, the formula for cell probabilities ($R_i C_j / N$) is actually the cell expectation (Connor and Simberloff 1983), and consequently some of the expected "probabilities" in a matrix can take on values > 1.0 . Even when a corrected formula is used, the contingency table model is not strictly correct, because it allows for multiple occurrences in each cell of the matrix, whereas a presence-absence matrix only contains zeros and ones. This may not be a problem in practice, because Monte Carlo simulations of cell probabilities gave qualitatively similar results (appendix to Gilpin and Diamond 1982).

Second, the Gilpin and Diamond (1982) model implicitly allows for “degenerate” distributions: empty islands and missing species, which correspond to matrix columns or rows with all zero entries (Connor and Simberloff 1983). In the Gilpin and Diamond (1982) model, occasionally a given island would never be “hit” by any of the species. Unless the observed presence-absence matrix contained the full sampling universe of missing species and empty islands, the expected values from this model would be biased. On the other hand, Haefner (1988b) found that degenerate distributions introduced only a slight bias toward accepting the null hypothesis.

Finally, the Gilpin and Diamond (1982) model is susceptible to Type I error. Wilson (1987) constructed a random presence-absence matrix with a probability of occurrence of 0.33 for every cell. This matrix was significantly non-random by the Gilpin and Diamond (1982) model ($p < 0.001$), with a slight excess of positive species associations. Gilpin and Diamond (1987) objected that such randomly constructed matrices may contain biological structure. Nevertheless, Wilson’s (1987) analysis suggests that the Gilpin and Diamond (1982) model is probably not appropriate for the analysis of co-occurrence data.

Summary of the Controversy and a New Approach

We think the controversy over R-mode analysis boils down to four issues:

1. *Which species and which islands should be analyzed?* It is surprising how little attention has been given to this point. The same data sets have been analyzed over and over with little discussion of source pools, colonization potential, and habitat availability of islands (Terborgh 1981; Graves and Gotelli 1983). Some of these factors have been discussed post hoc, but they must be considered in a systematic fashion *before* any analysis is attempted. Careful a priori selection of sites and species by explicit criteria is at least as important as the particular null model used for analysis.

2. *Which metric should be used?* In other words, how do we properly quantify nonrandomness and species associations in a presence-absence matrix? There are many different kinds of structure in a presence-absence matrix, but for the purposes of recognizing species associations, we think the following five metrics are most informative:

a. *The number of species combinations.* This is perhaps the most basic measure of community organization (Pielou and Pielou 1968). If assembly

rules are operative, there should be fewer species combinations observed than expected under an appropriate null model.

b. The number of checkerboard distributions. The checkerboard distribution is the simplest and most clear-cut of Diamond's (1975) assembly rules. It represents the strongest possible pattern of species repulsion.

c. The "checkerboardedness" index of Stone and Roberts (1990). This statistic measures the overall tendency for species pairs in a matrix to co-occur. It may reveal species pairs that associate negatively but do not occur in a perfect checkerboard.

d. The "togetherness" index of Stone and Roberts (1992). Both positive and negative associations are possible in the same matrix, and this index measures the tendency for species to co-occur.

e. Schluter's (1984) variance ratio. Negative co-occurrence as measured by the variance ratio is not always equivalent to that measured by the "checkerboardedness" index. This difference arises because the variance ratio test does not constrain column totals. Different patterns of negative covariation may be revealed by comparing the variance ratio to null model predictions.

3. Which simulation procedure should be used? We accept the logic of Connor and Simberloff (1979) that neither islands nor species are equiprobable and that this should be reflected in the null model. However, their simultaneous constraint of row and column totals was simply too severe. We prefer an approach similar to Gilpin and Diamond's (1982), which treats the row and column totals as expectations; the actual row and column totals should be allowed to vary from one simulation to the next. Although both the Connor and Simberloff (1979) and the Gilpin and Diamond (1982) approaches are flawed, we think the following two simulation procedures are acceptable alternatives:

a. Row totals fixed, column totals probabilistic. This model takes the observed frequency of occurrence of each species as a constraint but treats each site as a "target"; the probability of occurrence of each species at each site is proportional to the total number of species at that site. Thus, species number in each site will vary somewhat from one simulation to the next, although the relative rankings of sites in their species richness will be maintained on average. Simberloff and Gotelli (1984) used this procedure to examine incidence functions, which will be described later. This protocol is the inverse of Connor and Simberloff's (1978) Null Hypothesis II and Patterson and Atmar's (1986) RANDOM1 algorithm, in which column totals were fixed and row totals were probabilistic.

b. Row and column totals probabilistic. A less constrained null model would allow both the row and column totals to vary randomly. In this case, what is simulated is not the placement of each individual species across the set of sites, but rather, the placement of the N species-occurrences across the entire matrix. For cell a_{ij} in the matrix, the initial probability of a “hit” in the simulation is the joint probability of selecting the species (R_i/N) and selecting the site (C_j/N). This cell probability is thus $(R_i C_j / N^2)$, which corrects the error in Gilpin and Diamond (1982). In this simulation model, the most likely occurrence will be of the most common species on the most species-rich site, and the least likely occurrence will be of the rarest species on the most depauperate site.

One complication is that these estimated probabilities will apply only to the very first occurrence that is placed in the matrix. After the initial placement, the distribution of less probable combinations will be more “even” than predicted by these probabilities. Because sites can only be occupied once per species, the less probable sites are consecutively hit as the matrix fills up. If the marginal probabilities are squared and then rescaled, the resulting marginal distributions in the simulated matrix will more closely resemble the original column and row totals (Bruce D. Patterson, personal communication).

A second complication is that both protocols, and particularly protocol (*b*), could occasionally lead to degenerate distributions (= empty rows or columns) in simulated matrices. As we noted earlier, degenerate matrices should not be compared to observed presence-absence matrices, unless the full sampling universe of sites and species is known. Simulated degenerate matrices should be either discarded from the analyses or have a single species occurrence randomly repositioned to fill the empty row or column. Because all occurrences can be easily placed by either algorithm (*a*) or (*b*), difficulties in filling the matrix and simultaneously maintaining row and column totals (Connor and Simberloff 1979) are not present.

This simulation protocol may be expanded so that marginal probabilities do not depend on row and column totals but on independently measured attributes of species and islands. For example, site probabilities could be set proportional to island area (Coleman et al. 1982; Simberloff and Gotelli 1984) rather than total species richness. Species probabilities could be scaled proportional to density (Haila and Järvinen 1981) or geographic range size (Graves and Gotelli 1983) rather than to the total number of occurrences for each species. This sort of analysis requires a good deal of biological insight and information about sites and species, but it avoids the circularity (and convenience) of using marginal sums to estimate cell probabilities.

4. *How should the observed and simulated distributions be compared?* Both Connor and Simberloff (1979) and Gilpin and Diamond (1982) used a chi-squared test to decide statistical significance. This is problematic, because none of the frequency classes in their analyses were independent of one another. We recommend summarizing matrix patterns in a single metric (or five metrics!). In this way, statistical significance can be directly estimated by classical randomization procedures (Edgington 1987)—the metric for the observed matrix can be directly compared to the distribution of values from a large number of simulated matrices (usually ≥ 1000) to estimate the probability value. This is much more straightforward than trying to compare a simulated with an observed *distribution* when the frequency classes of that distribution are not independent of one another.

This approach also highlights an important difference in the strategy of analyzing presence-absence matrices. Connor and Simberloff (1978), Wright and Biehl (1982), and Gilpin and Diamond (1982) all advocated examining deviations of all possible pairs of islands or species and then attributing significance to those pairs that showed extreme values. We think there are statistical and conceptual difficulties in this. The statistical difficulty is that none of the pairs are independent of one another, so it is unclear which pairs are statistically and biologically significant. The conceptual difficulty is that this procedure comes perilously close to “data-dredging” (Selvin and Stewart 1966). There may be a temptation to *infer the hypothesis from emergent patterns, rather than explicitly stating an a priori hypothesis and testing for it with the null model*. Rather than examining the individual pairs to reveal potential interaction, we prefer to carefully select species and sites for this purpose ahead of time, and then test whether there are nonrandom patterns in the matrix. In this way, the patterns and the mechanisms that produce them are kept conceptually distinct from one another. We cannot overemphasize that the selection criteria must *not* include the distributional patterns being tested!

Other Tests of Assembly Rules

Diamond’s (1975) approach to assembly rules emphasized the particular combination of species *within* a guild that coexist and the intermeshing resource utilizations that allow these species to persist. Other studies have expanded on Diamond’s (1975) model and used null models to test for patterns of community assembly. For example, Pulliam (1975) used estimates of seed availability, bill sizes, and diets of wintering sparrow species to construct a “coexistence matrix” that predicted the bill sizes of species combinations that could coexist. The initial application of the model was successful, but additional data for three

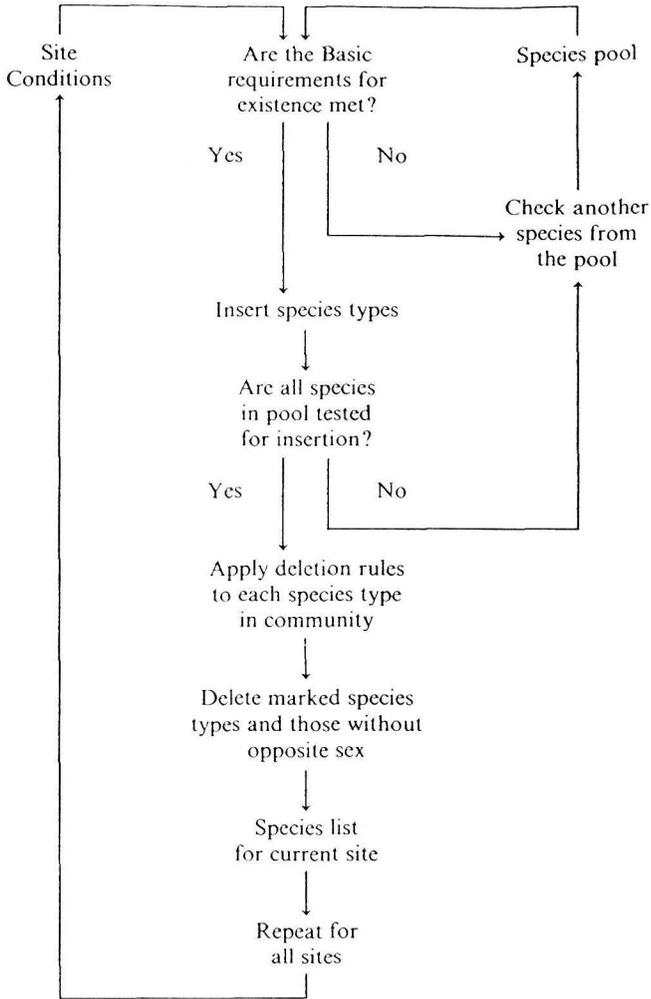


Figure 7.6. Flow chart for null models used to test assembly rules of Greater Antillean *Anolis* lizards. From Haefner, J. W. 1988b. Assembly rules for Greater Antillean *Anolis* lizards. Competition and random models compared. *Oecologia* 74:551–565, Figure 1. Copyright © 1988 by Springer-Verlag GmbH & Co. KG.

subsequent years fit the model very poorly (Pulliam 1983). A null model that incorporated habitat preferences of the species but did not include species interactions fit the data about as well as the assembly rules, although some patterns were still nonrandom with respect to this model.

Haefner’s (1988b) analysis of *Anolis* lizard coexistence in the Greater Antilles represents the most ambitious attempt to test Diamond’s (1975)

assembly rules against an appropriate set of null models. Haefner (1988b) analyzed the occurrence of eight species of *Anolis* at 11 Puerto Rican sample sites and asked what percentage of the observed species-site occurrences could be successfully predicted by a suite of 20 (!) different assembly models (Figure 7.6). “Random insertion and deletion models” placed species on sites independently of one another, but with varying degrees of habitat affinity and niche requirements. These models constituted null hypotheses of varying complexity with respect to competitive interactions. In contrast, a set of “simple deletion models” removed species according to rules of body size and niche overlap. “Complex deletion models” removed species according to rules derived from Williams’s (1972, 1983) ecomorph model.

Some of Haefner’s (1988b) null models provided a good fit to the data, especially those that constrained row or column sums and incorporated habitat affinities. As expected, the best-fitting models were the complex deletion models, which were calibrated to maximize the fit to the Puerto Rican data set. Simple deletion models did not fit the data very well, suggesting that if competition structured this community, competitive asymmetries between species were probably common. When these same models were applied to co-occurrence of Jamaican *Anolis*, the data were best fit by a null model that maintained column sums, or that maintained row sums and habitat affinities of species. As with the Puerto Rican data, the fit of the complex deletion models was good and that of the simple deletion models was poor.

Although Haefner’s (1988b) results enhance our understanding of community assembly, the interpretations are by no means clear-cut. On the one hand, the good fit of the data to several of the simple null models suggests that competition is unimportant in community assembly. On the other hand, the ability of the calibrated complex models to successfully predict community structure of Jamaican *Anolis* supports Williams’s (1972, 1983) ecomorph model and implicates competitive interactions.

OTHER ASSEMBLY RULES

Although most of the controversy surrounding assembly rules has been over forbidden combinations of species and competitive interactions, there are other types of assembly rules that may dictate the organization of communities. These other assembly rules have also been addressed productively with null models, and here we review these approaches.

Incidence Functions

Diamond (1975) introduced the *incidence function* to describe the probability of occurrence of a species with respect to ordered site characteristics, such as species number. To calculate the incidence function, Diamond (1975) classified the islands of an archipelago into discrete size classes, according to species richness. He then plotted the proportion of islands occupied ($0.0 \leq f \leq 1.0$) in each size class, ordered from species-poor to species-rich classes. The resulting incidence curve has been viewed as a characteristic “fingerprint” of the distributional ecology of a species.

For most species, the incidence curve increases monotonically, though not linearly (Figure 7.7). Diamond interpreted different incidence functions for birds of the Bismarck Archipelago as reflecting different distributional strategies. “High-S” species occurred mostly on large, species-rich islands, whereas the much less common “supertramp” species showed the opposite pattern and were found only on species-poor islands. Diamond (1975) speculated that high-S species competitively excluded supertramps from large, species-rich

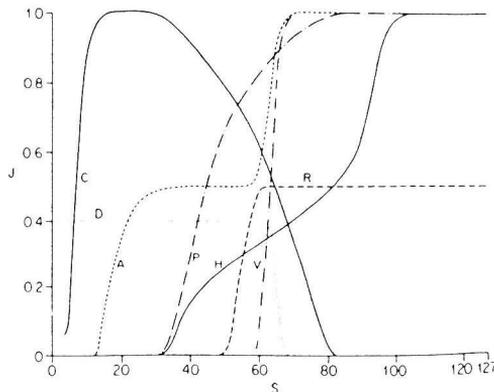


Figure 7.7. Incidence functions for gleaning flycatchers of the Bismarck Archipelago. The x axis is the number of species on an island, and the y axis is the proportion of islands in a given size class that were occupied by the species. C = *Monarcha cinerascens*; D = *Pachycephala melanura dahli*; A = *Myiagra alecto*; P = *Pachycephala pectoralis*; H = *Myiagra hebetior*; V = *Monarcha verticalis*; R = *Monarcha chryso-mela*. From Diamond (1975). Reprinted by permission of the publisher from *Ecology and Evolution of Communities*. M. L. Cody and J. M. Diamond (eds). Cambridge, Mass.: The Belknap Press of Harvard University Press. Copyright © 1975 by the President and Fellows of Harvard College.

islands, and that shifts in incidence functions in different archipelagoes reflected different species pools and subsequent levels of competition.

In contrast, Gilpin and Diamond (1981) explored the connection between the incidence function and the equilibrium theory of island biogeography. They viewed an assemblage as having no competitive interactions, with each species at a dynamic equilibrium in which colonizations and extinctions were balanced. In this scenario, the incidence function represents the fraction of *time* that a species occupies islands of a particular size class. From these incidence functions, Gilpin and Diamond (1981) inferred species-specific colonization and extinction rates, and derived the community-level extinction and colonization curves of the MacArthur and Wilson (1967) equilibrium theory (see Chapter 8). Hanski (1992) successfully applied this model to the insular distribution of shrews in eastern Finland, where field data have corroborated the assumption of dynamic turnover (Peltonen and Hanski 1991).

Other interpretations of incidence functions are possible. The incidence function may reflect nothing more than the distribution of habitat types among islands. Thus, high-S species may simply be habitat specialists, such as birds that are restricted to lakes, marshes, and high-elevation forest found only on large islands (Wiens 1989). Even in the absence of habitat specialization, the incidence function will reflect sampling variability, both in terms of the number of islands censused and the number of islands “sampled” by the species (Taylor 1991). Common and widespread species will have their incidence functions shifted to the left because they will occur on a variety of different island sizes (Haila et al. 1983). Whatever the cause of the incidence function, environmental stochasticity should lead to a shallower curve as occurrences become more unpredictable with respect to ordered site characteristics (Schoener 1986b).

Null models clarify some of these differing interpretations by asking how an incidence function would appear in the absence of any structuring force. Whittam and Siegel-Causey (1981b) made the first attempt to analyze incidence functions statistically, for seabird colonies of Alaska. For an archipelago of islands that have been sorted into size classes, they constructed an $S \times C$ table, where each species is a row and each column is a size class (*not* an individual island). The entries in the table were the number of islands of a particular size class occupied by a species. Contingency table analysis was used to assess statistical significance, and standardized residuals pinpointed particular cells in the table that contributed positive or negative deviations.

For Alaskan seabird colonies, the distributions of 20 species were highly nonrandom with respect to colony richness. Each species showed one or more significant deviations, and the sign of the deviations showed only one change in direction, running from small to large colonies (Figure 7.8). The analysis

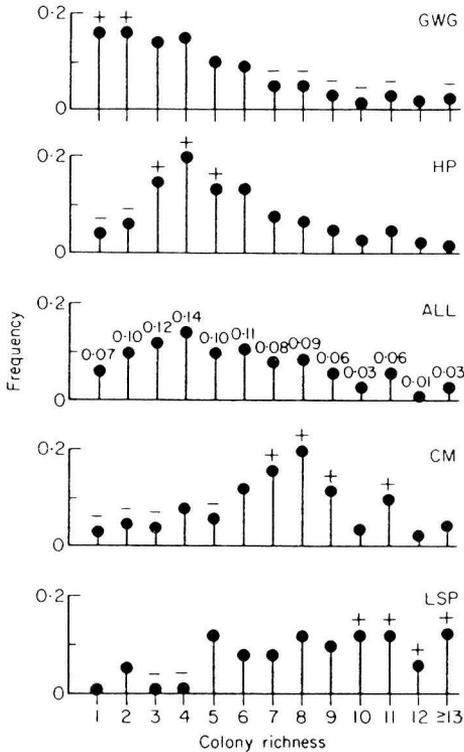


Figure 7.8. Statistical analysis of incidence functions for species in Alaskan seabird colonies. The x axis is the number of species per colony, and the y axis is the fraction of species occurrences in colonies of a given size. Plus and minus signs indicate significant deviations for a particular size class. Numbers indicate overall frequency of occurrence of species in colonies of different size. ALL = all species studied. Other species abbreviations as in Figure 7.3. From Whittam and Siegel-Causey (1981b), with permission.

identified supertramp species, such as the Glaucous-winged Gull (*Larus glaucescens*) and the Tufted Puffin (*Lunda cirrhata*), which occurred less often than expected in large colonies and more often than expected in small colonies. The more typical deviation was the high-S pattern, typified by several species of murres and auklets that were unusually present in large colonies and unusually absent from small colonies. Finally, some species showed no significant pattern of deviation, suggesting that occurrences were random with respect to colony size.

There are two points to note about Whittam and Siegel-Causey's (1981b) analysis of incidence functions. First, they defined incidence differently from Diamond (1975). Whereas Diamond (1975) defined incidence as the proportion of occupied islands in a size class, Whittam and Siegel-Causey (1981b) used the proportion of the total occurrences of a species that fell in a particular size class. Second, unusual deviations in the Whittam and Siegel-Causey (1981b) analysis were measured relative to the distribution of all other species in the assemblage. Thus, high-S species were those which occurred relatively more frequently on species-rich islands than did all other species. If all species showed an identical supertramp distribution, nonrandomness would not be revealed.

Incidence functions may have important implications for conservation biology. In particular, an incidence function analysis can be used to identify unusual minimum area requirements for particular species. Simply scanning a list of occupied sites may not be sufficient for this purpose, because even if a species were distributed randomly, it might be missing from some small sites by chance. Simberloff and Gotelli (1984) tested for unusually small minimum area occurrences of plant species in five archipelagoes of remnant prairie and forest patches in the North American prairie-forest ecotone. For each species, they ordered the patches on the basis of site area and then calculated the tail probability of finding the observed number of smaller, unoccupied sites. For three of the five archipelagoes, there were more species showing nonrandom patterns than expected by chance.

However, it might not be correct to conclude from this result that species have *unusual* minimum area requirements. In particular, this null model assumes that all sites are equivalent. But even if species have no special minimum area requirements, we expect large sites to be disproportionately occupied. A simulation model placed each species randomly in sites, with the probability of occurrence being proportional to the area of the patch. The simulation was repeated 10 times, and then the cumulative distribution functions for the observed and expected site occupancy were compared. The result was that far fewer of the species showed unusual deviations. In other words, nonrandomness in the occupancy of sites could be effectively accounted for by differences in the areas of the sites. Nevertheless, the distributions of seven plant species in one of the assemblages were significantly nonrandom by this test. Interestingly, none of these species showed an unusually large minimum area requirement. Instead, they followed the supertramp pattern and occurred too frequently in small prairie and forest patches (Simberloff and Gotelli 1984). Wilson (1988) tested for the occurrence of plant species with respect to island area and also found only a few species that exhibited supertramp distributions.

Finally, Schoener and Schoener (1983) expanded Diamond's (1975) idea of incidence functions beyond considerations of island area or species richness. They pointed out that sites can be ordered on any number of criteria, and then the occurrence of species tested against this ordering. "Haphazard" sequences of presence and absence would suggest that a variable does not play a direct role in determining the occurrence of a species. But if the sequence is highly ordered, the occurrence of the species is predictable with respect to that factor (Figure 7.9). The Mann-Whitney U test is a simple nonparametric analysis of the degree of this ordering for any particular variable. Within an archipelago, each species displays the same total number of presences and absences, so the

islands could be predicted by vegetation and habitat structure, although the two taxa were ordered on slightly different vegetation characteristics.

Within each group, distributional complementarities were more clear-cut when habitat variables were controlled for with a logistic regression (Schoener and Adler 1991). Thus, some checkerboards will be detected only when habitat differences among sites are measured and incorporated into the analysis. The more orderly the distributions of species are with respect to site characteristics, the less the patterns will conform to a simple checkerboard in which two species never co-occur within a set of similar islands. Such ordering at the community level may lead to “nested” species distributions, which we discuss in the next section.

Assembly rules, as originally described by Diamond (1975), have been very difficult to discern in nature. The extent to which the occurrence of a species can be predicted by the distribution of other species is still open to debate. In contrast, the evidence seems quite compelling that site characteristics can be used to successfully predict the occurrence of most, if not all, species in an assemblage. We do not deny the potential importance of species interactions, but we do suggest that assembly rules might best be developed in an autecological framework. Such rules would emphasize the ecological requirements of each species and the characteristics of individual sites, rather than the presence of forbidden combinations of species (Graves and Gotelli 1983).

Nestedness

Suppose that each species in an assemblage showed a perfectly ordered occurrence sequence with respect to species richness. In this case, the incidence function would be a steep J-shaped curve, with an incidence of zero below a critical species richness class and 1.0 above that class. At the community level, this occurrence sequence would correspond to a “nested” distribution of species (Patterson and Atmar 1986). Like a collection of puzzle boxes that fit perfectly within one another, the species composition of each small assemblage would be a perfectly nested subset of all larger assemblages.

Many presence-absence matrices conform to this pattern (Patterson 1990; Wright and Reeves 1992). For example, the distribution of five fish taxa in 28 South Australian springs (Kodric-Brown and Brown 1993) showed only a single species occurrence that deviated from a pattern of perfect nestedness (Table 7.1). Patterns of nestedness have been shown for assemblages of mammals (Patterson and Brown 1991), birds (Bolger et al. 1991), insects (Patterson 1990), plants (Wright and Reeves 1992), and parasites (Dobson and Pacala 1992), for insular and oceanic islands (Patterson and Atmar 1986); for habitat

fragments (Bolger et al. 1991); and for fossil assemblages and colonizing insect faunas (Patterson 1990). Nestedness has been noted before in the ecological (May 1978) and biogeographic (Darlington 1957) literature, but it has only received widespread attention since the 1980s.

How can nestedness be quantified, and what level of nestedness is expected in a randomly constructed biota? Indices of nestedness can be derived for entire communities or for individual species. Given that most assemblages exhibit a significant pattern of nestedness, the latter may be more informative (Simberloff and Martin 1991). Deviations from nestedness occur when species are found where they are not expected (“outliers”) or are absent where they are expected (“holes”). For example, both montane mammals and birds of the Great Basin exhibit significant nestedness, but the bird distribution is outlier-rich, whereas the mammalian pattern is hole-rich (Cutler 1991). Atmar and Patterson (1993) measured “unexpectedness” as a diagonal deviation from a perfectly nested matrix, and Wright and Reeves (1992) introduced a scaled index that can be used to compare matrices of different dimension. Finally, Ryti and Gilpin (1987) used parameters from a logistic regression model to quantify the degree of nestedness in a matrix.

All of these metrics can be compared to values expected from randomly assembled matrices. Patterson and Atmar (1986) introduced two null models for randomizing matrices and comparing them with nestedness statistics. In the first model (RANDOM0), site richness was retained and species were drawn equiprobably from the source pool. In the second model (RANDOM1), species were drawn in proportion to their occurrences. These models are identical to Null Hypotheses 0 and 2, proposed for Q-mode analysis (Connor and Simberloff 1978). Because the observed number of species occurrences was not maintained in these simulations, species that occurred in every site artificially inflated the apparent degree of nestedness and should have been omitted from the analysis (Simberloff and Martin 1991). But even without such species, the pattern of nestedness in most archipelagoes is so strong that it is unlikely to have been caused by a simulation artifact.

What explanations have been offered for the nestedness pattern? The dominant interpretation, first offered by Patterson and Atmar (1986) for montane mammals, is that the nestedness reflects a contraction of insular area, followed by an orderly sequence of extinctions, so that the same subset of widespread species survives in small habitats. If this “selective extinction” hypothesis is true, it suggests that faunal collapse is a highly deterministic process. The implication for conservation biology is that the biotas of small habitat fragments converge by loss of species that are uncommon in the landscape (Wright and Reeves 1992). This interpretation of nestedness is also a restatement of the

Table 7.4

Relative nestedness of avifaunas of land-bridge and oceanic islands

Fauna	Observed N	Expected N	p
Oceanic	53	53.32	~0.50
Land-bridge	158	363.73	<0.00001

The nestedness index, N , indicates the degree to which a presence-absence matrix departs from perfect nestedness. The simulation procedure (RANDOM1) holds species number constant on islands; the probability of occurrence of each species is weighted by its frequency. Note that the faunas of land-bridge islands are more highly nested than the faunas of oceanic islands. From Patterson (1987).

SLOSS (single-large-or-several-small) debate in conservation biology (Simberloff and Martin 1991).

But this is by no means the only explanation for nestedness. Darlington (1957) first proposed that the pattern could be explained by “selective immigration”—nestedness arises because of differences among species in their ability to immigrate and successfully colonize islands. Finally, a nested pattern of species occurrences might be related to neither immigration nor extinction, but simply reflect a nesting of habitats on islands of different sizes. In the West Indian avifauna, for example, the proportion of single-island endemics is much higher on the four large Greater Antilles than on other smaller islands (Terborgh 1973a). The Greater Antilles also support a greater diversity of habitat types, and the nestedness of the avifauna may simply reflect this pattern of habitat diversity.

How can these mechanisms be distinguished? Support for the “selective extinction” hypothesis comes from the observation that the degree of nestedness is somewhat greater for faunas of land-bridge islands than oceanic islands (Table 7.4). However, land-bridge and oceanic islands may differ in many factors besides their previous mainland connection (Gotelli and Graves 1990), and the evidence for “faunal collapse” of land-bridge island faunas is not generally compelling (Faeth and Connor 1979; Boecklen and Simberloff 1986). Moreover, highly nested faunas have been discovered in archipelagoes that are dominated by stochastic colonization and extinction, such as defaunated mangrove islands and vacant city lots (Patterson 1990). These findings suggest that nestedness need not imply orderly extinction, and that the pattern of nestedness may be a general property of many communities, rather than a specific attribute of islands undergoing biotic relaxation (Patterson 1990).

Lomolino (in press) has shed light on the problem by measuring the degree of nestedness when islands are ordered first by area and then by the degree of isolation. If selective extinction is more important, the ordering by area should show the strongest pattern of nestedness. If immigration is more important, the pattern should be more nested when sites are ranked by isolation. Significance was assessed by randomly reordering the matrices and then examining the weighted deviations from perfect nestedness.

Lomolino (in press) applied this test to distributions of nonvolant mammals in three archipelagoes of the Great Lakes and the St. Lawrence River. For all three archipelagoes, faunas were nested with respect to both area and isolation. But for two of the three, patterns of nestedness were more orderly with respect to isolation than to area. This finding suggests that immigration dynamics may play an important role in determining the pattern of nestedness, even though isolation itself was not always significantly correlated with species richness. We need more tests of this sort on additional archipelagoes to evaluate the relative contributions of immigration and extinction to nestedness.

Finally, it is worth noting that nestedness will cause faunas of similar size to be similar in composition, whereas checkerboard distributions and other forbidden combinations of species will have the opposite effect and increase diversity of similar sites (Graves and Gotelli 1993). Consequently, the degree of nestedness in an archipelago is correlated with the average similarity in species composition among sites (Wright and Reeves 1992). Perhaps the frequent occurrence of nestedness is one reason that checkerboard patterns of distribution seem to be the exception, rather than the rule, for most assemblages.

Niche Limitation

Wilson et al. (1987) suggested that if niche limitation were generally important, then the number of species that coexist in a guild would be constrained by competition (Colwell 1979). This approach does not emphasize the identities of particular species combinations, but instead draws on the older literature of niche limitation initiated by studies of species/genus (S/G) ratios (see Chapter 1).

Wilson et al. (1987) measured niche limitation as the variance in species number among a set of similar sites. If competition limited the number of species that could coexist locally, then the *variance* in species number would be unusually small compared to the variance in an unstructured community. On the other hand, the variance in S will also be increased if sites differ in their suitability, or if the community is not in an ecological or evolutionary equilib-

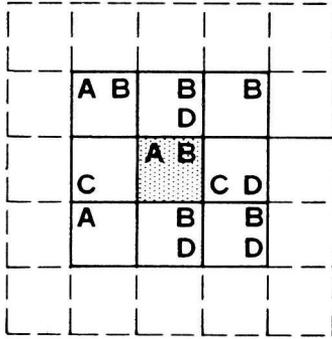


Figure 7.10. A patch model for small-scale niche limitation. The local source pool for colonization of a quadrat is the set of species occurring in the centered 3×3 grid. In this example, there are three local occurrences of species A, so it is placed with a simulation frequency of $3/9 = 0.333$ for the central quadrat. Reprinted with permission from A. J. Watkins and J. B. Wilson. 1992. Fine-scale community structure of lawns. *Journal of Ecology* 80:15–24.

rium. The expected variance can be calculated analytically (Barton and David 1959) or generated by a null model simulation in which species occurrences are randomly allocated to quadrats. The test is very similar to the variance ratio (Schluter 1984), because strong negative associations between species pairs will reduce the variance in total S .

Wilson et al. (1987) applied their test to a number of floristic data sets. To control for differences in suitability of sites, they searched for niche limitation of coexisting plant species at small spatial scales ($5 \text{ m} \times 5 \text{ m}$ contiguous quadrats) that were presumed to be environmentally uniform. For 164 sites in a homogeneous agricultural field, the observed variance in S was significantly greater than expected. Results were similar when the quadrats were subdivided floristically or environmentally. A large variance in S was also found for the vegetation of a uniform dune slope, but not for Patrick's (1968) data on diatom colonization of glass microscope slides. Similar results were obtained by Wilson and Sykes (1988). Variance in species richness was reduced for plants of 12 "lawns" in New Zealand and Fiji, but only for tiny quadrat subdivisions (360 and 160 mm^2 ; Watkins and Wilson 1992). Palmer (1987) found reduced variance in plant species number of Minnesota old fields, although his sites were unlikely to have been in equilibrium. Collectively, these results provide little evidence for strong niche limitation in floristic assemblages.

The Wilson et al. (1987) test assumes that all sites are colonized from a common species pool. An alternative assumption is that subtle microhabitat

variation generates spatial variation in the pool of species available to colonize plots. Watkins and Wilson (1992) tested this “patch” model by randomly selecting species for a quadrat from a set of adjacent quadrats (Figure 7.10). This null model is conservative, because the number of species that can be sampled from a few adjacent quadrats is usually much smaller than the total number from all quadrats. Nevertheless, there was still a variance deficit in species richness for most sites. The results suggest that niche limitation for plants may occur at an extremely small spatial scale, perhaps reflecting spacing constraints of individual plants in small quadrats.

Guild Structure

Whereas Wilson et al. (1987) emphasized the number of coexisting species within a particular guild, other studies have examined the relative frequency of different guilds. The null hypothesis, often implicit, is that the relative frequency of guilds in the assemblage represents a random sample of species from the colonizing source pool.

Two deviations from this null model are possible: the difference in guild frequencies between the source pool and the assemblage might be unusually large or unusually small. Biological significance has been attributed to both patterns of deviation. When the deviations are large, certain guilds are over- or underrepresented in local assemblages. In particular, many authors have noted the apparent “disharmony” of small island faunas compared to large mainland assemblages (Carlquist 1974). For example, MacArthur et al. (1972) attributed the consistent absence of some bird families from Neotropical land-bridge islands to interspecific competition. D. L. Lack (1976) also described examples of the limited coexistence of confamilial bird species in the West Indies and the role of reduced habitat and resource availability in producing these patterns. However, as in the analysis of species/genus (S/G) ratios (Chapter 1), we expect some deviations from the expected number of species in a guild because of sampling error. Drawing species randomly from an appropriate source pool is the correct procedure for evaluating the deviations of guild frequencies in small assemblages.

As described in Chapter 2, Gotelli and Abele (1982) used the hypergeometric distribution to test for deviations in species richness of West Indian land-bird families. For each island, the observed number of species in each family was compared to the expected number if species were drawn equiprobably from the archipelago list. Based on this test, the number of coexisting species of parrots (Psittacidae) was less than expected per island, whereas the number of pigeons and doves (Columbidae) and mockingbirds (Mimidae) was greater

than expected. For a set of Neotropical land-bridge islands, the Columbidae were also overrepresented, as were species in all families with large geographic ranges (Figure 2.5) or large body sizes (Gotelli and Graves 1990). Neither study provided evidence that competition within avian families limited insular coexistence in the Neotropics, or that certain families were extinction-prone on islands (cf. Terborgh and Winter 1978; Faaborg 1979).

Trophic Ratios

In some assemblages, guild frequencies show unusually small deviations, and this pattern has also been accepted as evidence for community assembly rules. The explanation is that biotic interactions, such as competition between different functional groups or predation between different food web components, constrain community structure (see Chapter 10). For example, the ratio of prey species number to predator species number was relatively constant in a sample of 100 community food webs (Cohen 1978), suggesting a trophic constraint on community assembly. Assemblages that did not contain the “correct” proportions of different groups were presumably unstable and did not persist. The hypothesis here is that the observed assemblage shows an unusually *small* deviation in guild frequencies from a colonizing source pool. This “constraints” hypothesis says nothing about what shapes guild frequencies in the source pool, only that, whatever these forces are, they dictate the stability of observed communities. The correct null hypothesis is that small deviations between observed and expected guild frequencies represent sampling error. This is the same null hypothesis for related tests of body-size matching (Schluter 1990) and convergence of morphology between communities (see Chapter 6).

Heatwole and Levins (1972) searched for trophic constraints in Simberloff’s (1969) mangrove insect recolonization data. They constructed a source pool list for six islands that were experimentally defaunated, assigned species to eight trophic classes (predators, parasites, detritivores, etc.), and calculated the expected frequencies in each trophic class. Using recolonization data, they calculated the summed deviation between observed and expected trophic class frequencies. This index decreased significantly with time since defaunation. In other words, as colonization proceeded, the observed fauna converged on the trophic structure of the source pool.

However, the Heatwole and Levins (1972) index was sensitive to sample size. Because species richness increased during colonization, it may not be surprising that the deviation index decreased with time. Simberloff (1976b, 1979b) showed with a null model simulation that the observed deviations from the source pool were no different than expected, given the number of species

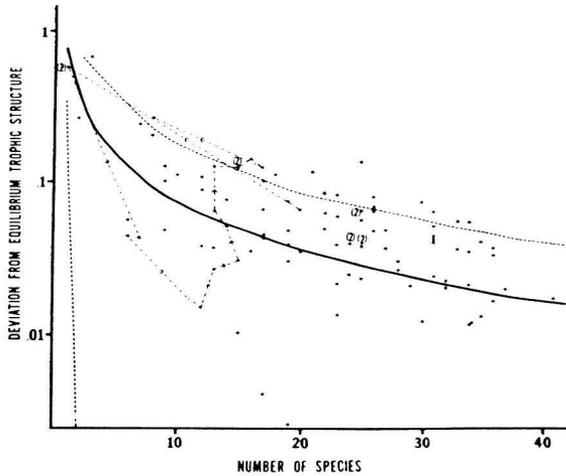


Figure 7.11. Sample-size dependency of the trophic deviation index. The solid line gives the expected deviation index for trophic classes of insects recolonizing six defaunated mangrove islands. The smooth dotted lines enclose two standard deviations. Each point represents a different recolonization sample, and the jagged line connects the observations for a single island. Note that although the trophic deviation index decreases with increasing species richness, most observations fall within the predictions of the null model. From Simberloff (1976b), with permission.

observed at that point in the recolonization sequence (Figure 7.11). Whether or not the trophic ratios were nonrandom in the source pool fauna, the mangrove insect assemblage did not exhibit unusual trophic constraints during recolonization.

Evans and Murdoch (1968) claimed that the ratio of herbivorous to entomophagous species number in a grassland insect community was unusually constant. In a reanalysis of Evans and Murdoch's (1968) data, Cole (1980) compared the observed and expected number of herbivorous species for each of 14 samples with the hypergeometric model. Cole (1980) found that all but one of the 14 samples were within two standard deviations of the expectation, and concluded that there was no evidence for unusual trophic structuring.

Finally, Wilson (1989) tested the guild proportionality of plants in a New Zealand rain forest by assigning species to one of eight stratification classes. He then measured the variance in the proportion of species in each stratum across 80 quadrats (each 100 m^2), and assigned species randomly to quadrats to generate the null expectation for this variance. Most strata did not deviate significantly, although the proportion of lianas was more variable than ex-

pected and inversely proportional to the number of epiphyte species in the quadrat. The observed proportion of canopy trees was unusually constant, perhaps reflecting space constraints in the number of trees per quadrat.

Functional Groups

Fox (1987) introduced a related type of assembly rule that also predicted pattern in the proportional representation of different guilds in an assemblage. He argued that the species pool for small-mammal communities could be divided into a number of distinct “functional groups,” and that these functional groups showed an unusual propensity to be equally represented in any local assemblage. The assembly rule specified that species are added from different groups until all groups are represented in the assemblage, and then the rule repeats. In other words, Fox (1987) hypothesized that the distribution of species among functional groups was not simply constant (as in the guild limitation tests) but was unusually *uniform*.

Functional groups do not reflect forbidden combinations of species per se, but only “favored” or “unfavored” states. Favored states have equal (or nearly equal) representation by each functional group and are predicted to occur more frequently than expected by chance. For example, if a species pool is composed of three functional groups, then a local three-species assemblage with all groups represented (1,1,1) would be a favored state. A four-species assemblage with (2,1,1) would also represent a favored state, because the mix of species is maximally even. In contrast, a three-species assemblage with a (0,1,2) mix would represent an unfavored state, because there are no species from the first functional group but two species from the third group.

As in Diamond’s (1975) analysis, competition for food resources is thought to be responsible for the uniform frequencies of different functional groups. However, Fox’s (1987) hypothesis is less restrictive than Diamond’s (1975) original assembly rules. Whereas Diamond (1975) emphasized the identities of coexisting species, Fox (1987) considered only functional group identity. Diamond’s (1975) analysis was restricted to taxonomic guilds of closely related species, but functional groups may be more phylogenetically diverse and could include unrelated species that are united by common morphology and foraging strategy. For example, Fox and Brown (1993) recognized one functional group of “quadrupedal non-heteromyids” that included four species of cricetids, three species of deer mice, one sciurid, and one ground squirrel. Other groups (e.g., “bipedal heteromyids”) included only closely related congeners and were more similar to Diamond’s (1975) taxonomic guilds. The species in different functional groups may differ in body size (Figure 7.12), which has been used as a

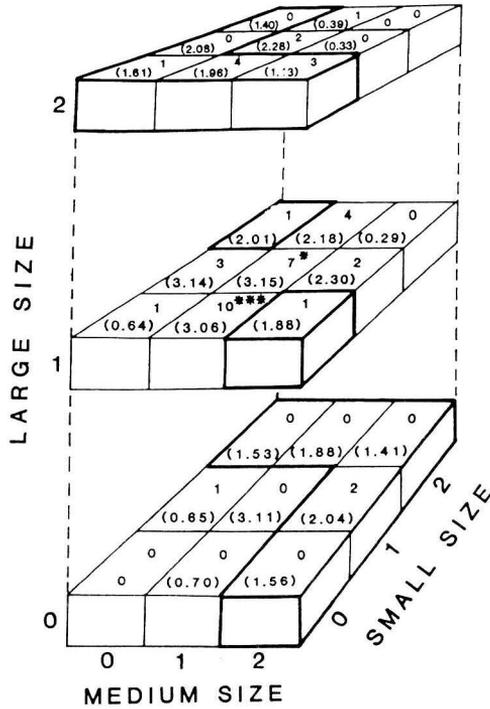


Figure 7.12. A three-dimensional model illustrating functional groups of sorcid assemblages from mesic forests in the northeastern United States. Three functional groups based on body size were recognized. Each cell represents the number of occurrences of a different combination of zero to two species of small, medium, or large sorcids. Expected values from the null model are indicated in parentheses, and unfavorable cells are shown with heavy outline. Significant departures from randomness are indicated by asterisks (* $p < 0.05$; *** $p < 0.001$). From Fox and Kirkland (1992), courtesy of the American Society of Mammalogists.

convenient character to define functional groups (Fox and Kirkland 1992). Whatever criteria are used, the functional groups must be defined a priori and independently of the co-occurrence data (Fox and Brown 1993).

Once the functional groups are defined, observed assemblages are classified as favorable or unfavorable states. For example, suppose there are three functional groups in the source pool, each with two species. In any local assemblage, each functional group will be represented by zero, one, or two species. Thus, there are $3^3 = 27$ different combinations of species from the three functional groups. In this hypothetical example, there are 15 favored combinations (including the 0,0,0 state) in which the functional groups are uniformly

represented, and 12 unfavored combinations. As in Pielou and Pielou's (1968) analysis of 2^k contingency tables, the different states are not equiprobable, and instead are conditional on the frequencies with which each species is represented in the source pool.

The expected number of favored and unfavored states was determined by a simulation in which species were drawn randomly and equiprobably from the source pool, without regard to their functional group. For a set of 52 mammal assemblages in temperate Australian heathlands, Fox (1989) divided the fauna into three functional groups (insectivore, herbivore, granivore-omnivore) of six, six, and three species each. Because no site contained more than two species in any functional group, the simulations were also constrained in this way to collapse the number of possible combinations. Fox (1989) found significantly more favored combinations and significantly fewer unfavored combinations than expected by chance. The same pattern held for mammals of Australian eucalypt forests (Fox 1987) and for North American soricid communities (Fox and Kirkland 1992).

However, these initial analyses used an algorithm of sampling with replacement that was not appropriate for small assemblages (Biehl and Matthews 1984). Fox and Brown (1993) constructed null models that sampled without replacement and extended the study of functional groups to North American desert rodent assemblages. For rodent assemblages in a small region of Nevada and at sites dispersed across the southwestern United States, there were again more favorable states represented than expected by chance (Figure 7.13).

Fox and Brown (1993) next examined the individual cells of the contingency table to pinpoint particular combinations that led to nonrandomness. Although favorable combinations were more probable than unfavorable ones, not all favorable combinations were equally likely. For example, the strongest deviation for the Nevada data set was a favored state of two quadrupedal heteromyid species, one bipedal heteromyid, and one generalist granivore. In contrast, the strongest deviations for the dispersed sites were two unfavored cells that had more assemblages than expected, even though the net result was that favored combinations were too frequent.

These results have been interpreted primarily in the context of interspecific competition: the different functional groups exploit food resources in different ways that are presumed to promote coexistence. The patterns reinforce experimental evidence for resource-based competition in both North American (Munger and Brown 1981) and Australian (Dickman 1986) rodent assemblages.

Two puzzles remain. First, why should the functional group rule repeat itself as more species are added to the system? In other words, if resource limitation is severe, how can two or more species from the same functional group coexist?

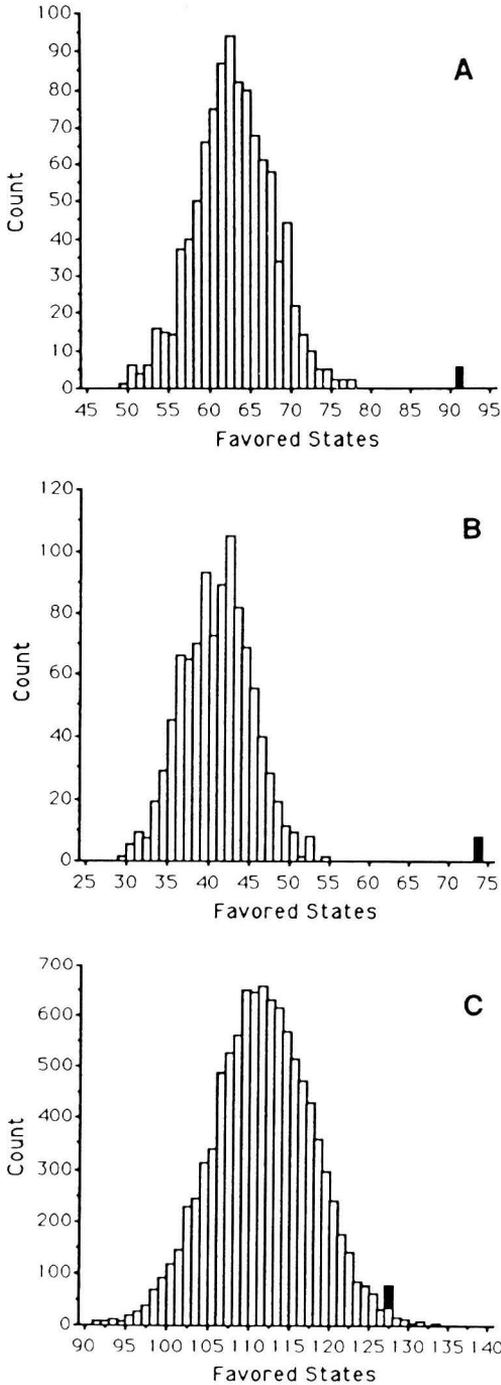


Figure 7.13. Observed and expected frequencies of favored states for North American desert rodent assemblages. The frequency histograms show the number of favored states when species are sampled equiprobably from the source pool. The observed number of favored states is shown with a solid bar. (A) 11 granivore species in three functional groups from 115 sites in Nevada; (B) 14 rodent species in five functional groups from 115 sites in Nevada; (C) 28 granivore species in three functional groups from 202 sites in southwestern deserts. From Fox and Brown (1993), with permission.

One might expect a superior competitor to exclude all other species within a functional group. Competition is not the only force leading to favorable states. Schluter (1990) showed that the same pattern will arise if source pools are small and if there is variation in colonization probabilities within a functional group. Wilson (1995) reanalyzed the Nevada data of Fox and Brown (1993) and showed that the statistical evidence for the functional group rule disappears if the assumption of equiprobable colonization is relaxed.

Second, what role does phylogeny play in producing functional group assembly rules? Because many of the functional groups are composed of closely related species, historical processes such as allopatric speciation may be responsible for the orderly composition of local assemblages. Whatever the role of interspecific interaction, phylogeny and history determine the mix of local species available for colonization (Cornell and Lawton 1992). Fox and Brown (1993; see also Brown 1988) argued that these phylogenetic effects are only present at a regional scale and that even at this scale, geographic range boundaries (and hence source pool composition) may be shaped primarily by ecological forces (Bowers and Brown 1982). However, phylogenetic correlates of geographic range boundaries may be strong, even within groups of closely related species (Taylor and Gotelli 1994; see Chapter 9). Unraveling these historical and ecological correlations is a major goal of the comparative method (Harvey and Pagel 1991), and the functional group assembly rules could be profitably analyzed from this perspective.

RECOMMENDATIONS

For Q-mode analyses of shared species, we suggest Simberloff's (1978a) weighted colonization models. For R-mode analyses of species combinations, there are several useful tests. The variance ratio (Schluter 1984) is a simple test for patterns of covariation that does not require Monte Carlo simulations. For detecting checkerboard distributions and unusual species combinations, we recommend a simulation in which row and/or column totals act as probabilistic constraints (Graves and Gotelli 1993). Other assembly rules and community patterns can also be addressed profitably with null model simulations, including incidence functions (Whittam and Siegel-Causey 1981b; Simberloff and Gotelli 1984), nestedness (Simberloff and Martin 1991; Atmar and Patterson 1993), niche limitation (Wilson et al. 1987), guild structure (Cole 1980; Wilson 1989), trophic ratios (Simberloff 1976b), and functional groups (Fox and Brown 1993). In all cases, analyses will be greatly strengthened by independent estimates of colonization potential, resource availability, source pool composition, and phylogenetic effects.

