NULL MODELS IN ECOLOGY

Paul H. Harvey
School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, Sussex, England

Robert K. Colwell
Department of Zoology, University of California, Berkeley, California 94720

Jonathan W. Silvertown
Department of Biology, Open University, Walton Hall, Milton Keynes MK7 6AA, England

Robert M. May
Department of Biology, Princeton University, Princeton, New Jersey 08544

INTRODUCTION

What determines the particular combinations of species found in biological communities, the number of those species, and their relative abundance? These questions are of perennial interest to ecologists, who often cite the influence of various factors—competition, predation, climate, nutrient availability, and chance dispersal events, to name a few. One way of testing hypotheses about the role of different processes is to search for patterns, i.e. similarities or differences between the compositions of specific communities. Process can be used to predict patterns, and pattern recognition can lead in turn to predictions about the consequences of experimentally manipulating natural communities. Such patterns typically involve: (a) some regularity in the relative abundance of species present in the community; (b) geographical correlates of the numbers of species; (c) the repeated occurrence or rarity of particular species or morphological combinations; and (d) differences in the behavior or morphology of a
given species that is dependent upon the presence of supposed competitors, mutualists, or natural enemies.

Recently, Simberloff, Strong, and others have raised serious questions about the extent to which apparent patterns in community structure bear witness to the biological processes that are invoked to explain them. In many cases, it can alternatively be argued that purely statistical models that make no assumptions about biological interactions among species provide an adequate explanation of the data. Such emphasis on the need to evaluate the significance of observed patterns against appropriately constructed "null models" (defined below) has helped set new standards of rigor for the discipline, as illustrated by the studies of Graves & Gotelli (45) and others. In practice, however, it is often very difficult to decide what constitutes a truly null, purely statistical, model: if some rearranged or reshuffled version of the data is used in constructing the null model, one cannot be sure that some biological interactions that may have shaped the data are not woven into the model. As reviewed below, in some instances clean and unambiguous null models can be constructed, and they sometimes do explain patterns in the data that earlier had been cited as evidence for competitive interactions. In other instances, however, the appropriateness and good sense of specific null models are disputed.

This review aims to survey the controversies surrounding the construction and use of null models in several areas of community ecology. In each case, we try to identify the biological questions at issue and the way null models can shed light on them, and then to focus on the technical problems involved in attempting to find an appropriate null model. The four authors of the review bring different perspectives to this task and the result may be frankly idiosyncratic. In places, space limitations have forced us to sweep aside the details.

We emphasize that we intend no evaluation of the overall importance of competition as a factor in community organization; our concern is the role of null models, as such. Indeed, our views are that community patterns may be influenced by temporal or spatial changes in the environment, by chance events, by competition, by mutualism, or by parasites or predators, and in general by the complicated (not to say chaotic) interplay of all of these factors.

**THE RELATIVE ABUNDANCE OF SPECIES**

When a community is sampled for individuals from a particular higher taxon, say birds or moths or diatoms, some species are typically more common than others. Efforts to make statistical generalizations about patterns of species abundance began at least 40 years ago (29). Until the 1970s almost all biological interpretations of the statistical patterns that emerged were vague and qualitative, without any rigorous theoretical basis (68, 69, 70, 77, 123, 124, 126). The principal "prediction" of these treatments was that in circumstances
that permit strong species interactions species abundance should be more equitable and the number of species greater. It occurred to Caswell (13) that one way to test this notion was to find out what relative abundance would look like for a community with no species interactions whatever. Drawing on mathematical results from the neutral gene controversy in genetics, Caswell developed models in which stochastic birth and death rates are equal and the number of new species entering the community is a Poisson stochastic parameter that is independent of the number of species already present. Precisely contrary to expectation, Caswell’s “neutral” distribution had relative abundance patterns that were more equitable than those from natural biotas that tradition held to have strong species interactions.

What is to be inferred from this result? As we will argue repeatedly in this review, all models in ecology have their biological assumptions, whether the models aim to incorporate certain species interactions or to eliminate them. In Caswell’s case, either the interactive model or the neutral one (or both of them) must be based on critically false assumptions—if the empirical data are accurate. To emphasize that hypotheses of any kind in ecology, including null hypotheses, are inevitably based on particular models and their assumptions and to avoid the more restrictive notion of neutrality, Colwell & Winkler (17) introduced the term null model for those models that aim to eliminate the effects of some particular ecological process.

As we will discuss below, Sugihara (118) has developed quantitative models of species abundance distributions for communities with strong interactions (competition), in part as a corrective for the vagueness of earlier models. On the other hand, Caswell’s models make assumptions that, if altered, might change the predictions of his models drastically. The assumption that the rate of immigration of new species is independent of the number of species already present implies an infinite species pool. Biomass per species invariably gives a more equitable distribution than examining the numbers per species in real communities (16). In addition, the birth rates of large organisms are generally lower than those of small organisms, yet models of relative abundance, including Caswell’s, simply count heads and assume the same birth rate parameters for all species.

A useful way of expressing the relative abundance of species is to plot the number of species with each level of abundance against the ranked abundance categories. Though there are many exceptions, when the abundance categories are logarithmically scaled, the resulting species-abundance curves often approximate a normal distribution (125, 126). If the curves were not constrained towards a particular family of lognormal distributions, a variety of hypotheses with different assumptions about the relative importance of biotic interactions in influencing community structure could accord with the data. However, natural communities do tend to aggregate towards one particular
form of lognormal distribution—the canonical lognormal (66, 85, 118). There is a specified, positive relationship between the variance of the distribution and the number of species in the community, so that one less parameter is needed to define the distribution (71). Furthermore, an inverse measure of the standard deviation of the distribution \( a = 0.71/\sigma \) is constrained so that it approximately equals 0.2 (56). Given reasonable sample sizes and a canonical lognormal species-abundance curve, this finding should not surprise us (71). But not all lognormal distributions are canonical and not all hypotheses proposed to account for the lognormal distribution of community structures predict a canonical form. In fact it seems that the only model presented so far that does so is Sugihara’s (118); he shows that a specific type of random, but sequential, division of multidimensional niche space among competing species leads to a distribution closely approximating the canonical lognormal. Indeed, Sugihara has been able to go further by examining the relative numbers of pairs and triplets of particular species in simple communities. The results lend additional support to his hypothesis.

THE NUMBER OF SPECIES

How many species will be found in a particular area? The answer to this question depends to some extent upon the size of the area. If larger areas contain proportionately more organisms and if the species-abundance curve fits a canonical lognormal distribution, the number of species \( S \) represented in a sample would be expected to increase with the community area \( A \) in accordance with the equation:

\[
S = cA^z
\]

where \( z \) is about 0.26 and \( c \) is a constant related to the density of organisms (66, 71, 85). A relationship of this form graphs as a straight line after both variables are logarithmically transformed, thus

\[
\log(S) = \log(c) + z \log(A)
\]

For the past 60 years there has been a recurring controversy about whether this double logarithmic transformation gives a better fit to linearity than do other relationships (5, 18, 26, 33, 119). These include:

\[
S = c + z A
\]

2.

\[
\log(S) = c + z A
\]

3.

and
\[ S = c + z \log(A) \]

Note that for empirical tests, \( c \) and \( z \) are fitted constants and usually will not have the same values as in Equations 1a and 1b.

Connor & McCoy's recent compilation (18) of 100 case studies indicates that formula (1b) provides a reasonable fit to the data in 75 cases, compared with 47, 22, and 38 cases for formulas (2), (3), and (4) respectively. However, they argue that transformation of the data is statistically inappropriate when both (1b) and (2) provide equally good fits to the data and thus prefer formula (2). We do not understand their argument, since the transformation does not require incorporating additional constants. The preference seems even more unreasonable because no body of theory predicts a linear relationship between the untransformed variables. Connor & McCoy also claim that the tendency for \( z \) to be about 0.25 when estimated using formula (1b) is a consequence of the regression model used; if the standard deviation of \( \log(S) \) divided by that for \( \log(A) \) and the correlation coefficient between the two variables are both allowed to vary randomly between 0 and 1, then \( z \) will often lie in the region of 0.25. But this "artifact hypothesis" predicts the distribution as well as the mean of calculated \( z \) values and, using their own data, it can be rejected at the 0.99 level of probability (119) because the real data are too closely clustered around the expected value. This result further supports the general applicability of the canonical lognormal distribution model, from which the expected value of \( z \) can be derived.

As Williamson has pointed out, the log \( S \) against log \( A \) plot "is the commonest satisfactory and good fit, but there are plenty of data sets not linearized by it" (129, p. 64). Similarly, although most estimated \( z \) values lie in the 0.25 region, there are many that do not. Some of the exceptions to both of these ecological generalities describe real variation in patterns of community structure and may lead to testable hypotheses concerning the factors involved. For example, it has been suggested that values of \( z \) are lower than expected when comparing either areas of mainland of increasing size (66, 85) or islands in particularly isolated archipelagos (18, 46, 101). Although these patterns are weak at best (129), various hypotheses have been proposed to account for them (see 18, 129). Existing explanations make no firm quantitative predictions and may be premature.

The models we have discussed concern discrete communities occupying single habitats. Obviously, as we incorporate new communities into a sample by increasing the number of habitats covered, we expect new assemblages of species to be included. Under some circumstances, habitat heterogeneity may be a simple consequence of increases in area. For example, sand islands smaller than about 100 m in diameter are saturated with salt water, while a lens of fresh water accumulates beneath larger islands (122). As a consequence, sudden
increases in the number of plant species able to survive are expected at this threshold. There is also a critical minimum territory size for many bird species, and this may introduce a threshold into the relationship between habitat area and species number. Larger areas not only have more heterogeneity of habitats, but they are also usually less carefully sampled (129); although all of the insect species in a small plot may be recorded, new species remain to be found within the continent as a whole. In addition to the correlations between species richness and habitat diversity, positive correlations between species richness and potential productivity and negative ones between richness and latitude or elevation are well documented for many taxa. These patterns received little attention during the decade after Pianka's (80, 81) discussion of alternative hypotheses (see 93), but this important topic shows signs of heating up again (7, 16, 55, 98) and generating its own controversies over null models (3, 13, 72, 78, 120).

WHICH SPECIES?

Searching for Patterns

The islands of an archipelago constitute a series of alternative microcosms. Communities differ somewhat among islands and, within a particular guild or taxonomic group, the presence of some species may be associated with the presence or absence of others. Detecting such patterns provides a promising start toward analyzing structure in natural communities. The most commonly available data can be summarized in a table that records the presence or absence of each species on the different islands. Such tables have commonly been analyzed in three different ways (109): (a) Islands have been compared by testing whether species compositions are more similar or more distinct than those expected by chance. (b) Species have been compared by testing whether particular species combinations are found more or less often than expected by chance. (c) Missing species combinations have been counted and compared with the number expected by chance. Clearly, the third approach demands a cruder level of analysis than the second, and some relevant data are ignored.

The form of analysis also depends upon the assumptions made about what constitutes chance, which are incorporated as constraints in the analysis. Three common constraints have been applied in various combinations. First, the incidence frequency of species has been incorporated into the model. Because some species occur more often than others, some combinations should also occur more often—pairs of very common species will be found on most islands, while pairs of very rare species may never exist on the same island. Second, the numbers of species on the different islands have been held constant so that larger islands, islands with more habitats, or islands near mainlands retain their high species numbers. These two constraints involve holding the
marginal totals in the original islands $x$ species table constant. The third
constraint arises because some species are more likely than others to be found
on islands of particular sizes—some species are restricted to small islands,
while others are found only on larger ones. [The same phenomenon occurs for
"habitat islands" (e.g. 76).] The third constraint amounts to incorporating what
have been termed incidence functions (23) into the analysis by adjusting the
probability of species occurrence for island size. Island size might be measured
directly or estimated by the number of species present. In either event, it is
important to distinguish this constraint from the second [contra the assertion of
Wright & Biehl (131)], because some species are only found on small islands.
A selection of studies using each of the three approaches with various combina-
tions of constraints is recorded in Table 1. Three studies not included in this
table incorporate the number of individuals present in the community as a
further constraint in null models of coral reef fish (97), snail communities (25),
and a guild of limestone plants (103).

Most analyses have sought to identify the role of competition in structuring
communities by identifying negatively correlated species distributions. Com-
petition may not be the correct explanation, however, of the presence of such
distributions. We shall return to the problems of interpretation below. For the
moment we assume that negatively correlated species distributions are at least
consistent with a competition hypothesis, and we discuss various criticisms of
the different analytical approaches.

Simberloff defended interisland comparisons by arguing that "competitive
interactions would tend to reduce the number of shared species below that
expected on the null hypothesis, because pairs or groups of competitors could
only exist on different islands" (106, p. 716). Most such comparisons have only
been made at the grossest level and have not identified the species responsible
for the patterns observed. Islands that differ in species composition may
contain different habitats (104). Searching for such correlates would be a
necessary second stage of any analysis; competition could only be inferred if
habitats were similar. Interspecific comparisons have been more widely used
and seem better suited to identifying the relevant patterns. Using only the
missing species combinations is a weak approach at best, because pertinent
data are inevitably discarded. For example, under a given null model we might
expect a particular species pair to coexist on most of the islands in an archi-
pelago while, in fact, they are only found on one. They would not be recorded as a
missing species combination, and therefore these species would not be identi-
fied as possible competitors. Null models of the distribution of Darwin’s
finches among the Galapagos Islands are much more firmly rejected when tests
take account of the species’ entire distributions than when data only on missing
species combinations are used [compare Connor & Simberloff (20) with
Alatalo (4) or Abbott et al. (2) with Grant & Schluter (44)].
Table 1  Various studies that have used species lists (rows) for a variety of communities (columns, usually islands) for statistical analyses*  

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Compare islands</th>
<th>Compare species</th>
<th>Missing species combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row (R)</td>
<td>Galapagos birds (4)</td>
<td>Galapagos birds (131)</td>
<td>Bracket fungi arthropods (82)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Galapagos Geospiza (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>West Indies bats (131)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alaskan seabird colonies (126a)</td>
<td></td>
</tr>
<tr>
<td>Column (C)</td>
<td>Galapagos birds (19)</td>
<td>Galapagos</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galapagos Geospiza (2,110,44)</td>
<td>Island rodents (38)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mangrove arthropods (105)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R x C</td>
<td>Galapagos birds (19)</td>
<td>Galapagos</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galapagos Geospiza (110)</td>
<td>Island rodents (110)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galapagos Geospiza (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>New Hebrides birds (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bismark Islands birds (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R x Incidence (I)</td>
<td>Bismark flycatchers (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bismark cuckoodoves (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C x I</td>
<td></td>
<td></td>
<td>Limestone plants (1,110)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sonoran Desert rodents (110)</td>
</tr>
<tr>
<td>R x C x I</td>
<td>New Hebrides birds (20,24)</td>
<td>New Hebrides birds (20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>West Indies birds (20)</td>
<td>West Indies birds (20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>West Indies bats (20)</td>
<td>West Indies bats (20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bismark Islands birds (24)</td>
<td>Bismark Islands birds (24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Comparing islands, comparing species distributions, and searching for missing species combinations are three of the methods used. Constraints have been used in various combinations, including keeping the relative incidence of species (rows), the number of species per island (columns), or the incidence functions (see text) constant. Some studies (italicized references) have resorted to simulation rather than using analytical methods to calculate the expected frequencies, in order to compare them with those observed in their empirical species x island table. When more than one cited paper has analyzed the same set of communities, they are referenced from left to right in order of improved analytical and statistical techniques and/or improved species lists.
Whichever of the three approaches is used, a major problem with incorporating the constraints is that relative species abundance and incidence functions may well have been shaped by competition (17, 24). Therefore, only additional effects of competition can possibly be revealed. The question of the general role of competition in structuring communities will not have been addressed. There have been recent exchanges in the literature about the relative merits of working within the different constraints (20, 21, 24, 32), and the choice obviously depends on the assumptions that the investigator wishes to make.

The analyses themselves pose some difficult problems. Analytical expectations become more difficult to estimate as more constraints are incorporated into the null model. For example, a recent study by Gilpin & Diamond (31) computed the expected "probabilities" for a species x island table using an inappropriate log-linear model; 235 of the 1568 cell "probabilities" were calculated to be greater than 1.0 (see 21). Their method then involved using an ad hoc procedure to redistribute the excess values. Because of the analytical problems involved with calculating expected values, several other studies have resorted to simulating the expected distributions of species occurrences (Table 1). With large data sets the simulations can become prohibitively expensive (20, 24, 131) and occasionally, with algorithms that involve reordering cell values in an observed matrix, too few iterations may have been performed (e.g. 20) so that the expected distribution barely diverges from the observed (24). The problem of lack of independence arises in another guise; the type II statistical error (failure to reject an incorrect null hypothesis) is extremely high in some instances (17). Observed data are used to calculate expected cell frequencies that, as the constraints are increased, may necessarily approximate the data from which they were derived (24, 41). As a consequence, some tests cannot detect that data sets comprised of pairs of species with mutually exclusive distributions are significantly nonrandom (24, 131). In addition, the expected species pool is derived from empirical species records used in the analysis. Therefore, it is perfectly possible that the real pool is larger but that some species do not successfully colonize any of the islands, although they commonly disperse from a nearby mainland (17, 131).

The problem of a high type II statistical error may be exacerbated by the use of statistics that have weak discriminatory ability. For example, one test (20) examines all pairwise species combinations and counts how many of them are found on various numbers of islands—33 species pairs on no islands, 14 species pairs on one island and so on. This distribution is then compared with a simulated distribution using a chi-squared test. Although the two distributions may not differ significantly, individual cases in the observed distribution might be highly unlikely. On the Bismark archipelago, 1690 pairs of bird species share no islands compared with an expected 1388 pairs (24). Although the difference is actually significant, "this test greatly understates the case [because] buried in this total are pairs of ecologically close species with exclusive
distributions despite each species occupying many Bismark islands, such that the probability of attaining such a result by chance is as low as $6 \times 10^{-9}$ (24, p. 71). Indeed, when observed numbers are of the same magnitude as expected, then some null models have been considered suitable approximations. For example, it has been claimed that "for the West Indies birds and bats . . . there are as many mutually exclusive species pairs as would have been expected" (20). There are 12,757 mutually exclusive species pairs of birds compared with an expectation of 12,448.1, and for the bats the numbers are 996 compared with 941.7. However, Connor & Simberloff (20) show that the observed numbers for birds and bats lie 3.9 and 4.7 standard deviations above the expected value, respectively; there are, in fact, far more mutually exclusive species pairs than expected.

The final important area of criticism combines ecological and statistical considerations. Competition may occur over access to resources, and only a few of the species of an island flora, fauna, or even avifauna will overlap in resource utilization. These groups of species may be defined as belonging to separate guilds (95). In order to prevent significant results from being diluted by a morass of irrelevant comparisons, analysis should be restricted to the members of a guild. This procedure has not been common practice and has led to pointed criticisms of analyses that incorporate, for example, owls with hummingbirds or ducks with warblers (see 4, 24, 131). In a study of body size in desert rodent assemblages, Bowers & Brown (6) showed that body size is extremely nonrandom within the granivore guild, but that the pattern is lost when members of other guilds are included in the analysis. It is important to emphasize that taxonomic groups and guilds are not synonymous. Members of a guild may contain representatives from several high-level taxa, while members of a taxon may be distributed among different guilds—e.g. the arboreal fruit-eater guild in New Guinea contains members of 14 bird families (24), while the New Guinea parrots may be nectarivores, frugivores, bark feeders or seedeaters (24). There are substantial difficulties, however, in defining guild membership (21). Useless data may also be contained in a data set containing guild members only; if some species are present on all islands there is no point in incorporating them into analyses that aim to spotlight positive and negative species associations, yet at least one recent study does exactly this (131).

The final outcome of a satisfactory analysis reveals pairs, trios, and higher-level combinations of species that show positively or negatively associated distribution patterns across islands. Notwithstanding the statistical shortcomings mentioned above, which are reviewed by Connor & Simberloff (21), the most complete analysis reported so far is Gilpin & Diamond's (31) study on the birds of the Bismark Islands. They identify pairs of species that are found together more or less often than three standard deviations from the expected value of a null model that does not employ incidence function constraints.
Then, using unpublished data on the habitats occupied, diet types, foraging techniques, morphologies, incidence functions, and geographical distributions within the archipelago, they suggest causes for the associated distributions. Although other factors might be important, such as unequal distributions of habitats across the islands (see above), in the Bismark case Gilpin & Diamond conclude that the highly significant negative species associations are probably due to competitive exclusion, differing distributional strategies (i.e. incidence functions), and varying geographical origins. On the other hand, Gilpin & Diamond considered positively associated species distributions (which are far more numerous than negative ones) as the result of shared habitats, endemism (i.e. speciation tends to occur on large islands where extinction rates are low), shared geographical origins, or shared distributional strategies. Without the supporting evidence cited by Gilpin & Diamond (31), the data in the species x island occurrence table is insufficient to enable one to distinguish among the possible explanations. Of course, the additional data only provide circumstantial evidence that make particular explanations more or less likely to be correct. At the very least, the consequences of introduction and removal experiments can now be predicted with more certainty.

**Predicting Patterns**

**SPECIES PER GENUS RATIOS** Ecological theories predict that particular species combinations should be more or less common than the levels expected under carefully formulated null models. In an early example that set a baseline for studies in plant geography, Raunkaier (90) showed that particular plant life forms were overrepresented in regional floras when compared with a pool of 1,000 representative species drawn from the global flora. Regional floras tend to be dominated by only a few of the life-forms found among vascular plants as a whole. Analysis on this geographical scale emphasizes the similarity of co-occurring species. On a finer scale—within habitats or parts of habitats—competition theory predicts a minimum dissimilarity between co-occurring species. Close competitors are unlikely to coexist in the same community. If competitive exclusion is an important factor in determining the species composition of communities, then the size of guilds should be limited. We have already mentioned that taxonomic affinity does not correlate perfectly with guild membership, but congeners usually do occupy similar ecological niches. In the third chapter of his *Origin of Species*, Darwin wrote, “As species of the same genus have usually, though by no means invariably, some similarity in habits and constitution, and always in structure, the struggle will generally be more severe between species of the same genus, when they come into competition with each other, than between species of distinct genera” (22, Chap. 3). We might therefore predict that the average number of species per genus ($S/G$ ratio) in a community would be smaller than would be expected from a random
sample of a regional species pool. However, it is extremely important to formulate the correct null hypothesis to use in testing the available data.

A series of studies over the last 60 years have pointed out that communities with few species have lower S/G ratios than those characterizing the species pools from which they were drawn (27, 37, 59, 73, 74). Such findings have usually been interpreted as support for the competitive exclusion principle, in spite of the fact that small, randomly constructed communities also have reduced S/G ratios. More appropriate analyses compare S/G ratios in real communities with those expected from a random sample of the species pool. These expected values can be estimated either analytically or by simulation. Analytical solutions require assumptions about how species are distributed among genera in the species pool, while simulations obtain results from the actual species pool (45, 104, 127). When these more appropriate procedures are used it is almost invariably the case that S/G ratios are higher than expected—more congers and fewer genera than expected are found in impoverished communities (35, 48, 67, 79, 83, 104, 127, 128). Simberloff presents the most comprehensive set of analyses to date, and he attributes the high S/G ratios to “similarity of congeneric species [either] in ecological requirements [or] in dispersal capabilities” (104, p. 46). When source pools are carefully standardized for the geographic range of their constituent species (as an indication of dispersal ability) and habitat match within recipient islands, the avifaunas of land-bridge islands closely match source pools in taxonomic composition (45). Jarvinen (62) gives a detailed account of the earlier controversies but, in the light of Darwin’s (22) statement that presaged all subsequent efforts, it is amusing to note that one of the few well-established cases of lowered S/G ratios is found among the birds of the Galapagos Islands (19, 41, 107).

**MORPHOLOGICAL SIMILARITY** S/G ratios provide a crude tool at best and, when they deviate from the expected value, a variety of competing hypotheses are usually available (107). Furthermore, this gross statistic gives no indication of which species within a genus are more or less likely to coexist. One indicator of similarity in the use of resources may be morphological similarity (57). For example, if bird populations are regulated by their food supplies and if beak size is a reasonable indicator of food size, then we might expect that members of the same feeding guild with very similar beak sizes are less likely to coexist than those with more distinct beak sizes. For the moment we assume that the morphological trait being measured is a species-specific characteristic and varies little among populations.

Islands from an archipelago that is adjacent to a mainland with a known species pool provide a suitable series of natural samples. Using the null hypothesis that the different islands are randomly populated by members of a
guild, we can compare the morphological divergence among members of the
guild on each island with that expected according to the null hypothesis. Strong
et al (117) adopted this approach for birds of the Tres Marias and California
Channel Islands. These workers compared observed culmen and wing length
ratios with the expected values, calculating the latter by drawing artificial
communities at random for the various islands. The artificial communities
contained the same number of species in each bird family as had been observed
on each island. Each community was simulated one hundred times. Analyses
were performed within families—e.g. if three species from a family were
present on an island, then two ratios were calculated that compared the size of
the smallest with the intermediate species and the intermediate species with the
largest. Only about one half of the ratios of both culmen and wing length on the
two sets of islands were greater than expected under the null model—the other
half were smaller than expected. Strong et al (117) concluded that there was no
evidence of nonrandom species assortment on the basis of the measured
characters.

Their study has been criticized on a number of grounds. Family membership
is not a good indicator of guild membership, the species source pools were too
large, differential dispersal abilities among species were not taken into account,
mistakes in data compilation and analysis were made, and the statistics used
were either inappropriate or invalid (41, 50, 102). Using a modified species
pool and an improved battery of statistical tests, Hendrickson (50) reanalyzed
the data from the Tres Marias Islands and confirmed the lack of evidence for
larger than expected ratios of culmen length among sympatric confamilials. He
did, however, find larger than expected differences in wing length. The
significance of this finding is not clear, since wing length is probably well
correlated with body size and the relationship between body size and diet in
many taxa may be complex (115, 130). An analysis performed among species
within two genera found in the Tres Marias Islands reveals no evidence that
coexisting congenerics are unusually divergent in either bill or wing length
(107).

Differential dispersal ability and competitive exclusion of potential colonists
should have opposite effects on the morphological similarity of coexisting
species. Colwell & Winkler (17) demonstrated this with a null model based on
stochastically generated phylogenies, morphologies, and biogeographic pat-
terns. They drew random samples from stochastically evolved source “faunas”
to colonize imaginary islands, then subjected each sample to a specified level of
competitive reduction and removed vulnerable species using an algorithm that
maximizes the separation among surviving species in morphospace. When
every species in the source fauna is equally likely to be drawn for colonization,
the effects of competition on the island faunas are easily detected by comparing
them with the interspecific distribution in morphospace of the source fauna. If,
instead, the state of one of the randomly evolved characters of source species
determines the probability of each species’ inclusion among colonists, the
effects of subsequent competition may be completely nullified and thus statisti-
cally undetectable. Because evolution is a branching process with extinction as
one possibility, even stochastic phylogenies produce interspecific correlations
among randomly chosen, independently evolving, unselected characters (17,
91); certain chance character combinations are approximately “copied” through
speciation, while other combinations perish. Thus two species alike in one
character (e.g. colonizing ability) are often alike in other characters as well
(e.g. bill size), with the result that precompetitive island samples are actually
more similar morphologically than expected from the source pool, even for
characters that have no bearing on colonizing ability.

In the real world, the effect of natural selection on functionally correlated
characters will only amplify this effect. Thus any effects of competition on the
morphology of island species will be systematically underestimated by any null
model that ignores the differential colonizing ability of the species in the source
pool. Although they did not treat morphology as such, Graves & Gotelli’s (45;
see above) approach provides a promising antidote for this problem. Moulton
& Pimm (75) took the ideal approach in comparing morphological spacing
among coexisting introduced bird species in the Hawaiian Islands by randomiz-
ing the actual list of historic introductions.

Schoener’s (102) work on bird-eating hawks is the most thorough study to
date of the morphology of coexisting species on a geographic scale. He
collected data on the worldwide distribution of all 47 species in the genus Accipiter and on 13 other primarily bird-eating hawks with similar habitat
requirements. Two sets of analyses were performed: one on the genus Accipiter
and another on the modified data set composed of the bird-eating hawk guild.
Careful criteria were used to establish areas of sympathy among species, and
observed ratios were calculated for each size-adjacent species pair. When more
than two ratios were available (e.g. there were four ratios with five coexisting
species), they were analyzed separately according to their position in the
species size ranking for that area. Appropriate expected distributions were
computed from all combinations of species pairs, trios, quartets, and so on.
Each observed distribution was compared with the corresponding expected
distribution using a Kolmogorov-Smirnov cumulative-frequency test. Variants
on Schoener’s basic test included correcting the data for the relative common-
ness of the different species. The study’s primary conclusion was that the most
similar sympatric species “are 20–35% more different in size than expected by
chance,” and that the differences between observed and expected distributions
are significant when sample sizes are above about 20. However, even though
Accipiter species are often highly sexually dimorphic with associated dietary
differences between the sexes, Schoener used the mean of both the male and
female sizes for his species values. Whether this averaging influenced his
results is unclear, although it is not obvious how it could create more uniform distributions than expected under his null models.

Bowers & Brown (6) present a convincing case for the existence of nonrandom patterns of body size in desert granivorous-rodent guilds. On the other hand, morphological studies of other mainland faunas have confirmed the fit of null models to empirical patterns (e.g. 47, 61, 87, 94).

SIZE RATIOS In an influential paper, Hutchinson (57) proposed that when size-adjacent close competitors are found in sympatry, “the ratio of the larger to the smaller form varies from 1.1 to 1.4, the mean being 1.28 or roughly 1.3. This latter figure may tentatively be used as an indication of the kind of difference necessary to permit two species to co-occur in different niches but at the same level of a food web” (57, p. 152). Hutchinson’s ratio (which he applied to linear dimensions—for weight the ratio would be about 2.0) has been widely accepted as an important ecological rule, but “as much as it has been asserted, the empirical truth of the rule has rarely been measured” (96, p. 396). This lack of testing may, in part, result from the ratio being variously interpreted as a modal, mean, optimum, or minimum figure and even as “a value that some proper summation of differences along all n axes of niche space should attain” (see 96, p. 396). There has, therefore, been a clear need for a statistical investigation of the available data.

In their recent review of 31 studies reporting on 105 data sets containing between 3 and 52 sympatric species, Simberloff & Boecklen (108) test for constancy of size ratios and for the existence of a larger than expected minimum size ratio. They conclude that, while most of the studies claim constant and minimum size ratios, there is surprisingly little evidence for either pattern when the data are subjected to statistical analysis.

For their null model, Simberloff & Boecklen fix the sizes of the smallest and the largest species; then they compare the observed size ratios with those expected on the assumption that other species in the sample are equally likely to be of any intermediate size on a logarithmically scaled axis. However, source pools will usually be composed of species with sizes drawn from approximately lognormal, rather than loguniform, distributions (11, 49, 102). Any unimodal distribution makes the majority of expected ratios smaller than those derived from the Simberloff & Boecklen null model. Therefore, their model predicts a higher proportion of large ratios than we might reasonably expect (17). This criticism also applies to Roth’s (96) model, which examines the same problem.

Simberloff & Boecklen’s test for constancy of size ratios entails comparing four size ratios: \(a\) the smallest, \(b\) the largest, \(c\) the second largest, and \(d\) the second smallest. Several data sets contain only 3 species. For these, they test the ratio of the smaller to the larger size differences. For larger samples, they ask whether ratios \(a\) and \(b\), \(a\) and \(c\), and \(b\) and \(d\) are more
similar than expected under their null model. Even in the case where there are 52 species with 51 ratios, only four ratios are used—the other 47 ratios between size-adjacent species are ignored. With such weak statistical tests, the Type II statistical error will be very large. Similarly, their test for whether there is a minimum size ratio searches for the smallest value and calculates the probability that no ratios will be any smaller. The rest of the distribution of ratios is ignored. As Schoener (102) has pointed out, the magnitude of the average ratios in a particular data set is not considered. For example, given three species and two ratios between size-adjacent species pairs, the pair of ratios 1.01 and 1.01 provide more evidence for a minimal difference \( p = 0.06 \) than the pair of ratios 1.75 and 2.50 \( p = 0.15 \)!

The more carefully constructed null model that Schoener uses in his analysis of size differences between bird-eating hawks (see above) demonstrated that, although sympatric species have larger ratios than expected by chance, there is no tendency for ratios to be particularly constant (102). Similarly, on a group of West Indian Islands, differences in body size among coexisting bird species from the same guilds are larger and more uniformly distributed than expected by chance invasion from the Colombian mainland, although again there is little tendency for ratios to be particularly constant (11). It seems that an adequate statistical analysis of the remaining size ratio data has yet to be performed.

**SIMILARITY IN THE USE OF RESOURCES** Behavioral similarity may often provide a more direct indicator than morphological likeness of niche overlap between two species, although the problems involved in pattern interpretation are similar whichever features are used. For example, Case (9) reports a study of lizard communities on 37 islands in the Sea of Cortez and on the adjacent California Baja mainland. The food size and type, microhabitat, and time-temperature of activity of all 18 lizard species present were assessed. The island communities contained species with lower behavioral overlaps than expected on the basis of a random invasion of colonists from the mainland (10). Competitive exclusion may have caused this pattern but, as Case stresses, differential dispersal abilities and the nonrandom distribution of niche space on the islands as compared with the mainland might also be involved.

Before their introduction to biogeography, null models were commonly used in ecology to test the observed patterns of resource use against randomized patterns within communities (or among related communities) (30, 52, 58, 88). The principal pitfall in the randomization of species/resource state matrices and related analytic methods is that the relative availability of appropriate resources must be known independently in order to permit an unequivocal inference of nonrandom resource partitioning. This problem is equivalent to the need for information on habitat differences among islands in an archipelago. Cognizant of the problem of resource availability, Rathcke (89) cites mixed results for
temporal patterning of flowering among coexisting entomophilous plants. Patterns of food use by diurnal and nocturnal raptors (60) and lizards (53) are not significantly distinct, so the question of equivalent resource availability does not arise. Likewise, differences among coexisting stem borers are no greater than expected at random (88), but the variance in escape behavior similarity is lower than expected for coexisting lizard species (99).

EVOLUTIONARY CHANGE IN RESPONSE TO COMPETITORS

When two potentially competing species with similar behavior or morphology become sympatric, natural selection may produce character divergence between them [for the conditions under which this is likely to occur, see Slatkin (111)]. This phenomenon has been called character displacement (8), while its complement, evolution towards the character states of a newly removed competitor, is termed character release (39). Contemporary evidence from natural populations might reveal increased morphological or behavioral differences between two species in sympatry as compared to allopatry; this finding would be compatible with the occurrence of either character displacement or release. After reviewing several cases that allegedly demonstrated this pattern, Grant (39, 40) found that the evidence in each instance was either weak or nonexistent. By 1979, Strong et al were able to write: "In only two instances do known morphological differences really conform to the pattern, one between two congeneric skinks [51, 54] and another between two species of snails [28]. The role of competition in these differences is not well established" (117, p. 897). A recent experimental study (65) casts some doubt on the snail example.

One of the most commonly quoted cases of supposed character displacement concerns the finches on the Galapagos Islands (64). A recent study (117) has returned to this example in searching for patterns that accord with character displacement or release at a community-wide level. Populations of the same species often differ among islands in beak size (64), and Strong et al (117) reasoned that character displacement might be recognized by size-adjacent species that were in sympathy and had unusually divergent beak sizes. One hundred "null communities" containing the correct number of species were constructed for each island. First, a species and then an island population of that species were selected at random; this procedure was repeated until the appropriate number of species-populations had been obtained. These null communities were used to calculate the average expected beak size differences between all pairs of size-adjacent species on each island. Finches were classified as either ground living or tree living and two measures of beak size (length and depth) were used, but in each case about one half of the expected values exceeded the observed. Hendrickson (50) pointed out that the binomial test used to compare
observed and expected values was inappropriate because the expected value was the average of a skewed distribution; in this case the test was biased against rejection of the null hypothesis. He reanalyzed the data using improved statistical procedures and correcting various data compilation, transcription, and computational errors and found evidence for "community wide character displacement" in beak length among the ground finches. However, since the null communities did not contain the actual combinations of species found in the real communities, the patterns revealed could have resulted from either competitive exclusion or character displacement. A more recent analysis by Case & Sidell (12) distinguished between the two processes and found patterns only in accord with a competitive exclusion process.

One other comparison made by Strong et al (117) in their search for character displacement involved comparing beak shapes as opposed to length and width. The test was appropriate for detecting character displacement because, in this case, the null communities for each island were composed of the correct combinations of species, with beak shapes drawn from a randomly chosen island population of each appropriate species. No evidence for character displacement was detected. Since about 99% of the differences between populations in the shape measure used are directly attributable to beak length (50), this finding is similar to Case & Sidell's (12; see above).

Strong et al (117) designed one test to search for character convergence that did yield significant findings. If an island has a population of one finch species composed of either large or small individuals, does this mean that finches of other species will deviate from their median species size in the same direction? To order answer this question, Strong et al ranked the populations of each species by size across all islands and tested whether the average species ranking for each island was larger or smaller than expected. An excess of islands with much higher and much lower rankings than expected under a random-colonization-with-no-divergence null model would suggest that character convergence had occurred. There was indeed an excess of such islands. Hendrickson (50) used a similar but improved statistical procedure and, contrary to his assertion, confirmed the original finding for bill length. [His Figure 1 clearly demonstrates the phenomenon; see also Strong & Simberloff (115).]

However, in tests of this sort, whether they are for character convergence or for character divergence, positive results can often be explained either by an evolutionary response or by competitive exclusion. For example, Stiles (113) reported the flowering times of 11 hummingbird-pollinated plants in a Costa Rican rain forest. After some controversy (84, 114), it now seems reasonably well established that flowering times of the different species are more evenly spaced than a random flowering-time model would predict (15, 34). This finding raises the intriguing possibility that competition for pollinators or interspecific hybrid inviability has resulted in the evolution of staggered
flowering times. Nonetheless, it is also possible that the same two processes have merely prevented other hummingbird-pollinated plants from invading the community. Scenarios involving both phenomena to varying extents are also plausible. Information about flowering times in other populations is obviously crucial if we wish to distinguish between them.

CONCLUDING REMARKS

During the past twenty years, the data base for analyses of the sort described in this review has increased enormously. Given the ease with which such data can now be manipulated, it is reasonable to search for patterns revealed by the data and attempt to interpret them. Although some authors (notably Simberloff and his coworkers) have stressed that pattern detection and pattern interpretation should be ideally independent exercises, it is clear that the two cannot always easily be separated. For example, the choice among alternative null models and the choice of constraints employed in them may rest upon a prior assumption about whether competition, mutualism, or predation has an important influence on the composition of a species pool or on some aspect of the colonization process in archipelagos. While detecting such patterns can often lead to our dismissing some hypotheses about factors that influence species distributions, an unknown number of competing hypotheses invariably remain (86). Other data are necessary before we can refine them or distinguish among them. Sometimes, as in Gilpin & Diamond's (31) Bismark Islands bird survey, many of the additional data may already have been collected. In other cases, as in the recent work of Grant & colleagues (36, 42, 43, 44, 100) on Galapagos finches, predictions made from interpreting distribution patterns require the collection of additional fine-grain distributional, morphological, or behavioral data. Another category of cases entails manipulating population densities, usually by species addition or removal experiments, such as in Stewart & Aldrich's (112) study of interspecific territoriality. As much progress as we may make, some patterns will have been caused by historical effects that, because they were not recorded at the time, are likely to remain as imponderable possibilities. Colwell & Winkler's (17) simulation study in which they produced phylogenies, dispersed species among the islands of an archipelago, subjected populations to exclusion by competition, and then attempted to analyze the distribution patterns so produced, revealed the possible importance of such effects.

In short, the recent emphasis on the need to evaluate perceived patterns in community structure against null hypotheses is—as seen with hindsight—overdue. But legitimate enthusiasm for sound methodology must go hand in hand with the realization that null hypotheses in ecology, as elsewhere, depend on null models, and that all models make assumptions. If these assumptions are not appropriate, or create systematic biases, no amount of mathematical and
statistical precision will produce biologically valid answers. In Tukey’s words “Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise” (121).

ACKNOWLEDGMENTS

We thank all those authors who replied so rapidly to our many requests for material that is still in press. D. Strong, D. Simberloff, M. Williamson, and J. Maynard Smith made useful comments on earlier drafts of the manuscript.

Literature Cited


43. Grant, P. R., Grant, B. R. 1980. The breeding and feeding characteristics of Darwin’s finches on Isla Daphne Major, Galápagos. Ecol. Monogr. 50:381–410


60. Jaksic, F. M. 1982. Inadequacy of activ-
ity time as a niche difference: the case of
diurnal and nocturnal raptores. Oecologia
52:171–75
61. James, F. C., Boecklen, W. 1983. Inter-
specific morphological relationships and
the densities of birds. See Ref. 116. In
press
ratios in biogeography: a historical note.
J. Biogeogr. 9:363–70
63. Johnson, M. P. 1974. Species number,
endemism and equilibrium in the Galapa-
Sci. Galapagos Symp., San Francisco
64. Lack, D. 1947. Darwin’s Finches. Cam-
bridge: Cambridge Univ. Press
65. Leviton, J. S. 1982. The body size-prey
size hypothesis: the adequacy of body
size as a vehicle for character displace-
ment. Ecology 63:869–72
The Theory of Island Biogeography.
Princeton Univ. Press
67. Mailléter, A. 1929. Le coefficient génér-
ique de P. Jaccard et sa signification.
68. Margalef, R. 1963. Succession in marine
2:137–88
69. Margalef, R. 1963. On certain unifying
principles in ecology. Am. Nat. 97:357–
74
70. Margalef, R. 1968. Perspective in
Ecological Theory. Chicago Univ. Press
71. May, R. M. 1975. Patterns of species
abundance and diversity. See Ref. 14,
pp. 81–120
72. McCoy, E. D., Connor, E. F. 1980. Lati-
tudinal gradients in the species diversity
of North American mammals. Evolution
34:193–203
73. Moreau, R. E. 1948. Ecological isolation
17:113–26
74. Moreau, R. E. 1966. The Bird Faunas
of Africa and Its Islands. NY: Mouton
experimental test of community wide
character displacement. Evolution. In press
of birds on small islands: an invalid
77. Odum, E. P. 1969. The strategy of
ecosystem development. Science 164:
262–70
Patterns of species diversity: Fact or arti-
fact? Paleobiology 4:41–54
79. Palmgren, A. 1925. Die Artenzahl als
pflanzengeographischer Charakter so-
wie der Zufall und die sekuläre Land-
hebung als pflanzengeographische Fak-
toren. Ein pflanzengeographischer Ent-
wurf, basiert auf Material aus dem aländi-
1:1–143
in species diversity: a review of concepts.
Am. Nat. 100:33–46
diversity: North American flatland des-
82. Pielou, D. P., Pielou, E. C. 1968. Asso-
ciation among species of infrequent
occurrence: the insect and spider fauna of
Polyporus betulinus (Bulliard) Fries. J.
Theor. Biol. 21:202–16
83. Polyá, G. 1930. Eine Wahrscheinlich-
keitsaufgabe in der Pflanzensociologie.
Zürich. 75:211–19
84. Poole, R. W., Rathcke, B. J. 1979. Regu-
larity, randomness, and aggregation in
flowering phenologies. Science 203:
470–71
85. Preston, F. W. 1962. The canonical
distribution of commonness and rarity of
32
86. Quinn, J. F., Dunham, A. E. 1983. On
hypothesis testing in ecology and evolu-
ion. Am. Nat. In press
87. Ranta, E. 1982. Species structure of
North European bumblebee communi-
ties. Oikos 38:202–9
88. Rathcke, B. J. 1976. Competition and
coeXistence within a guild of herbivorous
insects. Ecology 57:76–87
89. Rathcke, B. J. 1983. Patterns of flower-
ing phenologies: testability and causal in-
fERENCE using a random model. See Ref.
116. In press
90. Raunkiaer, C. 1934. The Life Forms of
Plants and Statistical Plant Geography.
Oxford: Clarendon
Stochastic simulation and evolution of
morphology—towards a nomothetic
palaeontology. Syst. Zool. 23:305–32
92. Rice, J., Belland, R. J. 1982. A simul-
ation study of moss floras using Jaccard’s
coefficient of similarity. J. Biogeogr.
9:411–19
Chiron. 2nd ed.
94. Ricklefs, R. E., Cochran, D., Pianka,
E. R. 1981. A morphological analysis of the
structure of communities of lizards in
95. Root, R. 1967. The niche exploitation
pattern of the blue-grey gnat-catcher.
Ecol. Monogr. 37:317–50
96. Roth, V. L. 1981. Constancy in the size
118:394–404


