The Distribution of Abundance in Neutral Communities

Graham Bell

The American Naturalist, Vol. 155, No. 5. (May, 2000), pp. 606-617.

Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28200005%29155%3A5%3C606%3ATDOAIN%3E2.0.CO%3B2-1

The American Naturalist is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Mon Feb 25 13:17:41 2008

VOL. 155, NO. 5 THE AMERICAN NATURALIST MAY 2000

The Distribution of Abundance in Neutral Communities

Graham Bell[*]

Redpath Museum and Biology Department, McGill University,
Montreal, Quebec H3A 2K6, Canada

Submitted January 6, 1999; Accepted December 22, 1999

ABSTRACT: The patterns of abundance generated by a simple stochastic birth-death-immigration model are described in order to characterize the diversity of neutral communities of ecologically equivalent species. Diversity is described by species number S and the variance of frequency or log abundance $q{\sim}$. The frequency distribution of abundance is very generally lognormal, skewed to the left by immigration and resembling descriptions of natural communities. Increased immigration and community size always cause S to increase. Their effect on $q{\sim}$ is more complicated, but given biologically reasonable assumptions, S and $q{\sim}$ will be positively correlated in most circumstances. Larger samples contain more species; the graph of log S on log individuals, equivalent to a species-area curve, is generally convex upward but becomes linear with a slope of about +0.25 when immigration is low and births exceed deaths. When individuals invade a new, vacant environment, both S and $q{\sim}$ increase through time. Thus, a positive correlation between S and $q{\sim}$ will usually be generated when sites of differing size or age are surveyed. At equilibrium, communities maintain roughly constant levels of S and $q{\sim}$ but change in composition through time; composition may remain similar, however, for many generations. Many prominent patterns observed in natural communities can therefore be generated by a strictly neutral model. This does not show that community structure is determined exclusively by demographic stochasticity, but rather demonstrates the necessity for an appropriate null model when functional hypotheses are being tested.

Keywords: abundance, species diversity, biodiversity, species-area relationship, neutral model.

The distribution of abundance among species with similar ways of life has been the subject of several classical papers in ecology, and it raises issues that have not yet been resolved. The literature of the subject is formidably complex, but a useful basic distinction can be drawn between phenomenological and mechanistic interpretations of abundances. A purely phenomenological approach identifies a particular frequency distribution that seems capable of generating the patterns of abundance reported from natural communities, so that the properties of these communities can be summarized as the parameters of the distribution. The two distributions that have received the most attention are the log series (Fisher et al. 1943) and the lognormal (Preston 1948). The alternative theoretical approach first identifies an ecological mechanism thought to be responsible for community structure, and then infers the frequency distribution of abundance to which it gives rise. The mechanisms concerned involve the partitioning of resources in different ways (MacArthur 1957; Sugihara 1980; Hughes 1986). The elegant account of the mathematical basis of these models provided by May (1975) remains useful, while an extensive review of the literature has been published by Tokeshi (1993).

There has been little effort devoted to constructing a H_0 of community structure, by studying the properties of neutral models in which species are on average identical and diverge through stochastic effects alone. In population genetics, the neutral theory of allele frequencies has reached a high degree of sophistication (see Kimura 1983), but no similar theory has been elaborated in community ecology, perhaps because ecologists have been reluctant to concede that the visible differences among species, often striking, have no effect on abundance. The major exception is the paper in which Caswell (1976) applied the neutral theory of allele frequencies to species abundances, showing that it gave rise to distributions resembling those generated by phenomenological or mechanistic models. In particular, his adaptation of infinite-allele models generated log-series distributions of species abundance similar to those observed in many samples of natural communities. A different model, involving speciation within communities of fixed size, has subsequently been elaborated as a theory of "community drift" by Hubbell (1979, 1995, 1997, 1998) and Hubbell and Foster (1986). It seems worthwhile to investigate the behavior of neutral models more closely. Their value is chiefly to identify those features of natural communities that require explanation in terms of mechanisms acting differentially among species, and

---

[*] E-mail: gbell2@maclan.mcgill.ca.

Am. Nat. 2000. Vol. 155, pp. 606–617.
0003-0147/2000/15505-0002$03.00.

those that do not. Until this has been done, detailed discussions of the underlying causes of patterns of abundance and rarity may turn out to be pointless. A more complete account of neutral community models in relation to functional models will be published by Hubbell (in press).

In this article I shall describe some of the properties of a very simple neutral model representing a community of functionally identical species. The term "neutral" is used to mean that individuals possess the same properties, regardless of species membership; thus, individuals of different species are indistinguishable. This is invariably the sense in which the term is used in population genetics. It does not follow that species are equivalent, because abundant species have systematically different properties than rare species. The model is a Monte Carlo computer program that simulates the dynamics of communities drawn from a pool of N species. The community initially consists of i individuals of each species. The program then iterates the following four processes: immigration—a single individual is added to the community from each species in the pool with probability m; birth—each individual gives rise to a single offspring with probability b; death—each adult individual dies with probability d; density regulation—if the community exceeds its capacity of K individuals, excess individuals are removed at random, each individual having the same probability of being removed, until the community is reduced to exactly K individuals before the next cycle is begun. Note that it is individuals that are culled, and that all species are treated alike, so that each species is on average culled in proportion to its relative abundance alone. Species do not differ in their sensitivity to density, and there are no implicit or explicit interactions among species.

The diversity of the community that develops as a consequence of these operations can be characterized by the number of species S, by the form of the distribution of the number of individuals per species, s, and by the variance of species abundance q. Because q varies with the number of individuals in the community, it is most usefully expressed as the variance of species frequencies, $q\sim$. Many other parameters have been used by previous authors, and the Shannon-Weiner index, a measure derived from information theory, has been especially popular. I have not used it because it has few if any parallels in ecology, and its biological interpretation is obscure. The variance of abundance, on the other hand, is a familiar type of parameter with a direct connection to theory: in simple models, the variance is proportional to the rate of change in species composition within a group of ecologically equivalent species, just as the genetic variance is proportional to the rate of response to selection (Fisher 1930). It is closely related to Simpson's index. A useful brief review of diversity indexes (not including q) has been published by Smith and Wilson (1996).

The model can then be used to describe the relationship of diversity and abundance to the six parameters of the model: pool size N species, initial species populations i individuals, community capacity K individuals, immigration rate m per species per cycle, birth rate b, and death rate d per individual per cycle. It can then be determined whether the patterns exhibited by natural populations differ systematically from those generated by a finite stochastic birth-death-immigration process. For some purposes, it is necessary to study the properties of random samples from the community, in which case the number of individuals sampled from separately simulated communities n is a seventh parameter.

This neutral community model (NCM) differs from other kinds of neutral model in two respects. First, the probability of immigration is a constant that does not vary with the composition of the community; that is, the probability that an immigrant belonging to a given species will arrive does not depend on whether the species is already present in the community. This seems ecologically realistic. In other neutral models (e.g., Karlin and McGregor 1967 in population genetics; MacArthur and Wilson 1967 in community ecology) the corresponding parameter refers to the appearance of "new types," and it is therefore a variable whose value changes with diversity. Second, the manner of density regulation is different from other models, many of which are simple exponential processes with no regulation; it is most similar to the model developed by Ewens (1972).

Results

Frequency Distribution of Abundance

The diversity of an isolated community of fixed maximum size will continually decline as species are lost through stochastic extinction when rare. Eventually only a single species will remain, and in time this too will become extinct. The community is rescued by immigration. For low immigration rates of $m \ll 1/N$, abundance is approximately lognormally distributed with a mode in $s \gg 1$ (fig. 1*A*). It is skewed to the left by immigrants. These form a minor mode at $s = 1$, consisting predominantly of newly arrived individuals each constituting the sole representative of its species. Between these two modes lies a region where immigrants that are spreading and residents that are declining pile up to create an intermediate mode. This distribution appears to represent an equilibrium that is approached regardless of initial state i, although it may take hundreds of cycles to reach a quasi-stationary state.

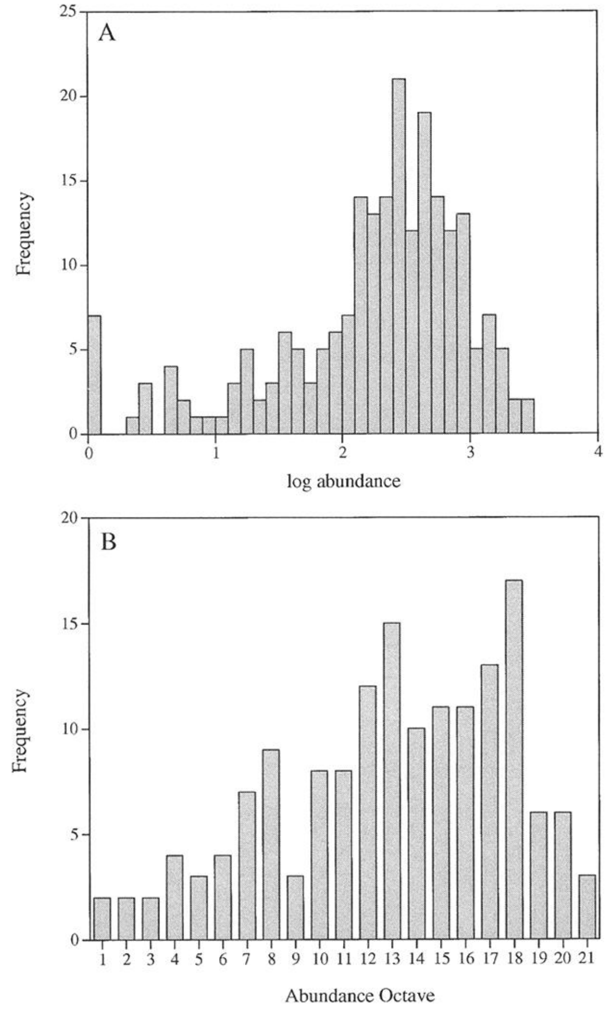

Figure 1: Frequency distribution of abundance. *A*, Neutral community model simulation after 2,000 generations with species pool $N = 1{,}000$; capacity $K = 100{,}000$; immigration $m = 0.0025$; birth $b = 0.5$; death $d = 0.5$; initial size $i = 100$. *B*, "Unveiled" distribution of abundance in the 146 species of bird breeding regularly on the British mainland. Occasional breeders are excluded, depressing the minor mode at the left. "Octaves" are powers of 2; their use is traditional. Redrawn from figure 4 of Nee et al. (1991).

Effect of Increased Immigration

An interesting special case is provided by $b = d$, where each species barely replaces itself through reproduction. The effect of varying the absolute values of b and d while retaining their equality is self-evident. If $b = d = 0$ or $b = d = 1$, then the model is no longer stochastic, since birth and death are impossible in the one case and inevitable in the latter. Species number is therefore maximal and variance minimal at these points; S falls and $q{\sim}$ increases as b and d increase from 0 or decrease from 1, S becoming minimal and $q{\sim}$ maximal at $b = d = 0.5$. The effect of raising the birth rate above the death rate is generally similar to reducing the immigration rate: it creates a situation in which a greater proportion of the population are surviving residents, or the offspring of residents, rather than incomers.

As immigration rates increase from very low values, the total number of species S increases. This is caused primarily through an increase in the number of rare species in the minor and intermediate modes, so the distribution becomes increasingly skewed to the left, and for high rates of immigration comes to resemble a log-series curve (fig. 2). It can readily be demonstrated that this is the consequence of the balance between long-resident lineages and the lineages descending from recent immigrants. The long-resident lineages retain a skewed lognormal distribution whose mode is not very sensitive to the migration rate. This distribution becomes more symmetrical at high migration rates, but at the same time, individuals descending from recent immigrants become much more abundant at high migration rates. As the immigration rate approaches the birth rate, very rare species that have immigrated re-

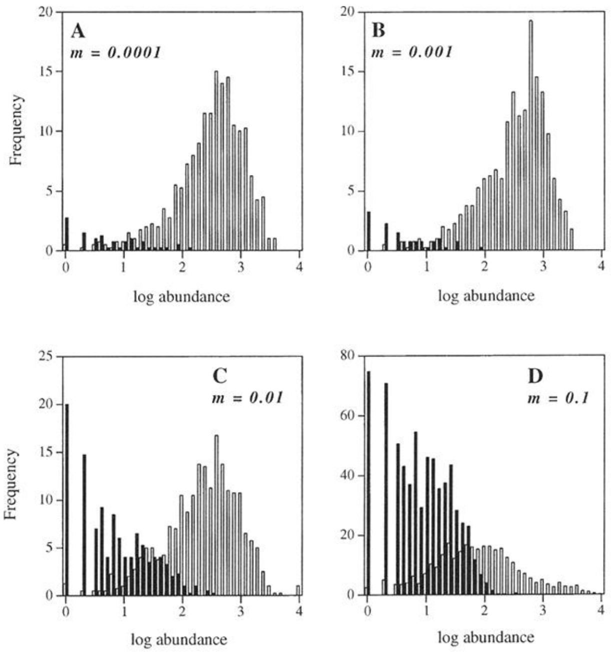

Figure 2: The effect of migration on the distribution of abundance. Lineages appearing before cycle 2,000 were classed as "long-resident" and those descending from individuals immigrating between cycles 2,000 and 2,100 as "immigrant." Open bars represent long-resident lineages; solid bars, immigrant lineages. Each diagram shows the mean of four replicate simulations. In all cases, $N = 1{,}000$, $K = 100{,}000$, $b = d = 0.5$, $i = 100$. *A*, Migration $m = 0.0001$. *B*, Migration $m = 0.001$. *C*, Migration $m = 0.01$. *D*, Migration $m = 0.1$.

cently and are represented by only one or two individuals come to dominate the population.

Diversity in Relation to Immigration and Community Size

These changes in the elevation and shape of the distribution drive changes in diversity parameters. For a given pool N, species number S increases monotonically with immigration m and community capacity K (fig. 3*A*). The behavior of the variance of abundance q is more complex (fig. 3*B*). Because abundance tends to a lognormal distribution, the appropriate parameter is the variance of log s, which I shall designate $q_\sim$. This is equivalent to the variance of frequency. If K is small, then $q_\sim$ is minimal when m is very low or very high. For low m, a single species is present most of the time, between rare events of successful immigration. For high m, the community tends to consist of K species each with a single representative. In either case, the variance of abundance among species actually present is very small. At intermediate rates of immigration abundance varies among species, so that $q_\sim$ at first increases and then decreases with m. A similar pattern can be traced when K is large. At low migration rates the community is dominated by the major mode. As migration increases the minor mode waxes, the distribution becomes more skewed and the variance increases because the major mode remains in nearly the same position while the number of rare species is rising. If m increases too far, however, the major mode moves to the left, and the community shrinks to a nearly J-shaped distribution on the left, causing the variance to fall. For any given K, therefore, $q_\sim$ is maximal at intermediate m. Moreover, a similar line of reasoning shows how $q_\sim$ is maximal for intermediate K at any given value of m. These interactions between m and K give rise to the complex topography shown in figure 3*B*.

In most real cases, m and K will be unknown, whereas the relationship between S and $q_\sim$ can be investigated from survey data. Taking sections through figure 3, it can be seen that the neutral model does not, in general, predict any consistent correlation between S and $q_\sim$. For fixed K, S and $q_\sim$ will vary with m (fig. 4*A*). When K is not much greater than N, S and $q_\sim$ will appear to be negatively correlated. This is because S increases and $q_\sim$ decreases sharply with increased immigration, provided that m is sufficiently high, whereas if $m < 1/N$, then neither S nor $q_\sim$ respond much to variation in m. For greater values of K, however, both S and $q_\sim$ rise as immigration increases at low values of m; thus, the graph of $q_\sim$ on S has a pronounced hump. Alternatively, m may be fixed and K allowed to vary (fig. 4*B*). In this case, S and $q_\sim$ are positively correlated if $m > 1/N$; at low immigration rates, $q_\sim$ decreases at large K while S continues to increase, pro-

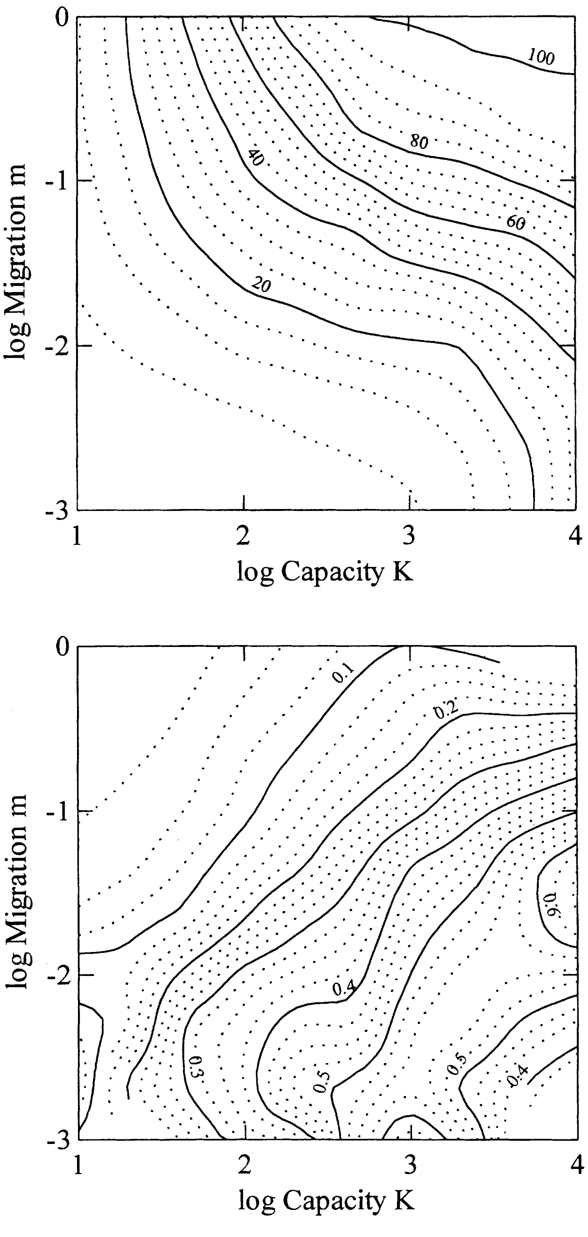

Figure 3: The effect of migration and community size on diversity. Simulations were performed at 10 nearly equal log intervals from m = 0.001 to 0.9 and from K = 10 to 10,000. The population parameters were N = 100, $b = d = 0.5$, $i = K/\mathrm{N}$. Each of the 100 values in each graph is the mean of a varying number of replicates; more replicates were used for parameter combinations (involving low K or low m), yielding more variable results. A total of 368 simulations was performed. These means were then used to draw a contour plot using SYSTAT 7.0.1. *A*, Species number S. *B*, Variance of log abundance $q_\sim$.

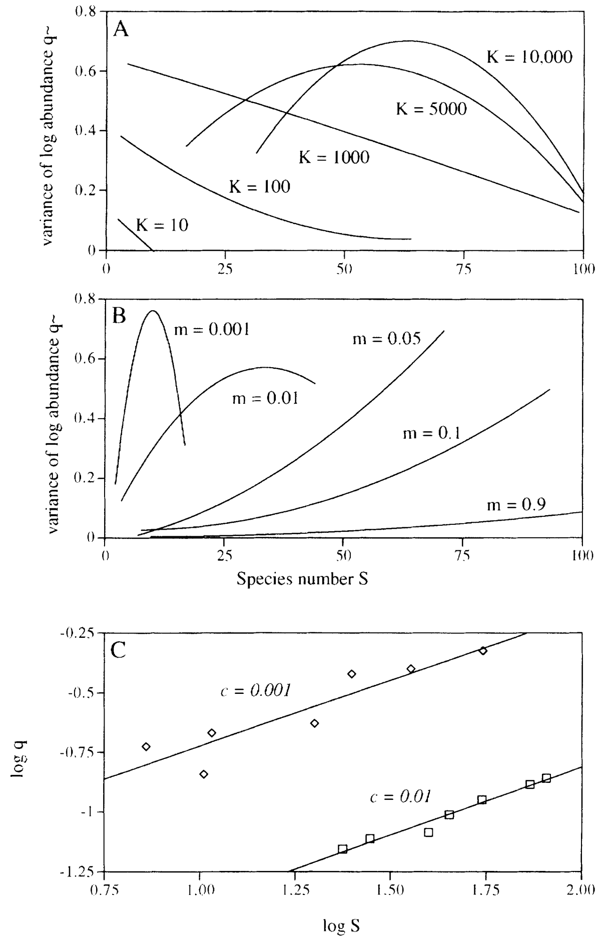

Figure 4: Relationship between species number S and variance of log abundance $q\sim$. Curves are second-order polynomials fitting the values derived for figure 3. *A*, K fixed. S and $q\sim$ vary with m, lower values of m being to the left. *B*, m fixed. S and $q\sim$ vary with K, lower values of K being to the left. *C*, K and m constrained by $m = c\sqrt{M}$. Linear regression models are $c = 0.001 : \log q\sim = 0.555 \log S - 1.278 (r^2 = 0.84)$; $c = 0.01 : \log q\sim = 0.571 \log S - 1.953 (r^2 = 0.97)$.

ducing a hump-shaped curve. Despite these complexities, however, two general rules will hold provided that migration rates are not very low: S and $q\sim$ will tend to be negatively correlated among communities of similar capacity whose immigration rates differ; whereas S and $q\sim$ will tend to be positively correlated among communities with similar rates of immigration but differing capacities.

Two special cases that give rise to contrasting patterns can be recognized.

The first is mK = constant, and especially $mK = 1$. The community then receives on average a single immigrant in each cycle. This is an interesting case because in the neutral theory of population genetics the number of alleles maintained in a population is governed by a quantity similar to mK. Species number S at first increases with K, but when m approaches $1/N$ the supply of immigrants becomes insufficient to prevent the domination of the community by a few species, and the pattern reverses. Thus, the relationship between S and K is hump shaped, with a peak in $K < N$. The variance of abundance $q\sim$, however, increases monotonically with K, and I speculate that $mK = 1$ is the only locus for which this is true. Essentially, the same results were obtained for $b > d$ as for $b = d$.

The second case is of more ecological interest. As the spatial dimension l of a community increases, the number of immigrants is expected to increase with perimeter, and the capacity with area; thus, m should increase directly with l and K with l^2, so that we can define $m = c\sqrt{K}$, where c is a scaling constant. In this case, S increases monotonically with K, whereas the graph of $q\sim$ on K is a hump-shaped curve. However, $q\sim$ declines only when the community is nearly saturated, with S approaching N. There is therefore a positive relationship between S and $q\sim$ for most reasonable communities that do not include most of the species in the available pool. This yields a linear regression of log $q\sim$ on log S, whose slope is independent of c (fig. 4*C*).

Species Number in Samples

If a random sample of n individuals is taken from each simulated community, the number of species recorded S increases with the number of individuals sampled n. Because S cannot exceed the pool N, sampling from the basically lognormal distribution arising from the neutral model generates a sigmoid relationship between S and n, as noted by Preston (1948). This curve represents a species-area graph, if it is allowed that n will generally increase proportionately with area. A graph of log S on log n for the replacement model ($b = d$) is therefore convex upward, regardless of m and K (fig. 5). (Sampling from the log-series distributions generated by Caswell's models would yield a semilogarithmic relationship and would therefore also lead to a graph of log S on log n that is convex upward [Watterson 1974].) This indicates that the distribution of abundance in the neutral model is not necessarily "canonical," in the sense that the highest abundance class should contain more individuals than any other class, that is, that the product of the number of species and the number of individuals per species should peak at the extreme right-hand end of the distribution, for the category with the most individuals per species. Instead, the "individuals curve" (distribution of this product) generally peaks at intermediate values of abundance. A nearly canonical lognormal is generated by the neutral model if $m < 1/S$ and $b > d$ (fig. 5*C*). A few species then increase rapidly to high abundance without being challenged very

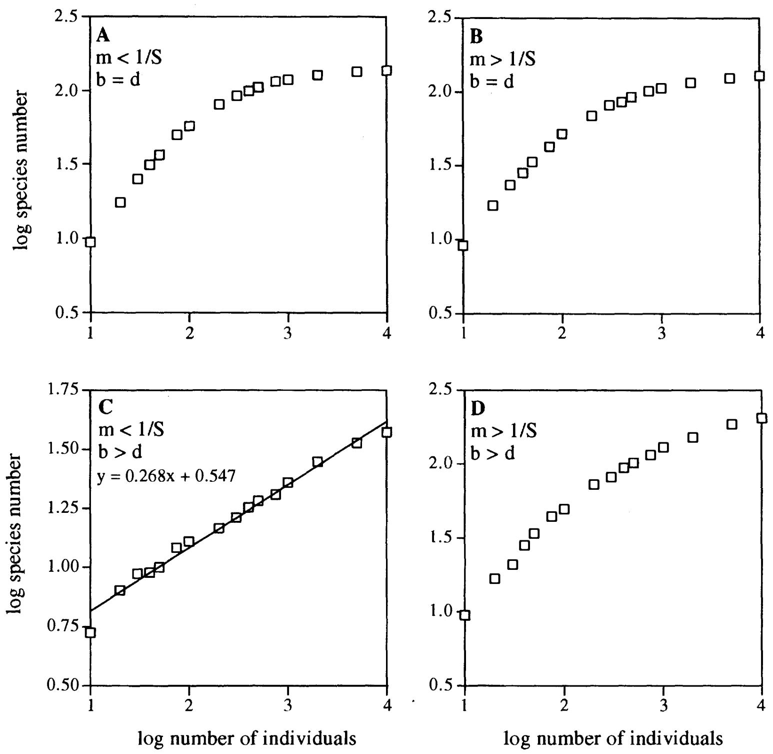

Figure 5: Species diversity and sample size. Species pool $N = 100$, community size $K = 100{,}000$. *A*, $m = 0.001$; $b = d = 0.5$. *B*, $m = 0.1$; $b = d = 0.5$. *C*, $m = 0.001$; $b = 0.5$; $d = 0.25$. *D*, $m = 0.1$; $b = 0.5$; $d = 0.25$.

often by the spread of immigrant lineages, so that the major mode moves to the right. This generates a linear or nearly linear regression of log S on log n. In the examples that I have computed, this regression has a slope of about +0.25, as expected for the canonical lognormal (Preston 1962).

Increase of Diversity through Time

In communities that are initially empty ($i = 0$) both S and $q_\sim$ increase continuously through time, eventually approaching equilibrium values. The length of time that passes before a quasi-stationary state is attained depends on rates of immigration and growth, but if $m \geq 1/N$ and $b \approx d$, then it is usually of the order of 200–500 cycles and seems never to exceed 2,000 cycles. Samples taken while the community is still changing yield a positive regression of $q_\sim$ on S, from low values early in succession to high values in saturated communities (fig. 6). This is attributable in large part to the increase in the number of individuals through time.

Species Turnover through Time

The species number characteristic of a certain combination of parameter values does not represent a community of fixed composition, since species are indistinguishable. Community composition must therefore change through time, while species number remains more or less the same. Successive communities are not independent, because each is constituted largely of the survivors and offspring of its predecessor. Thus, the composition of a community will be highly correlated with that of its immediate successor, whereas this correlation will fall through time, because of extinction and immigration, approaching 0 for its successors in the distant future. Figure 7 shows how the correlation of composition declines in a sigmoid fashion through log time. The rate of decline depends on K and m, being least when K is large and m is small. This is because the history of a community is less readily effaced by demographic stochasticity in large communities where immigrants are rare. The correlation may, indeed, remain high for long periods of time. The longevity of an isolated logistic population founded by i individuals is roughly $10i$

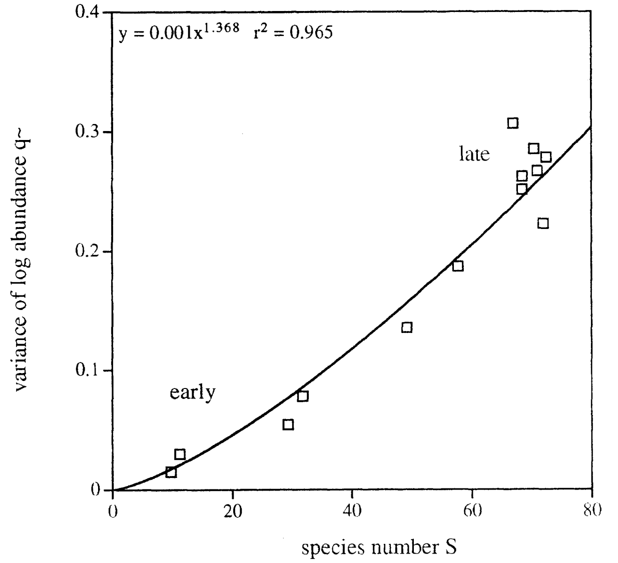

Figure 6: Increase of diversity during succession. Population parameters $N = 100$, $K = 1{,}000$, $b = d = 0.5$, $m = 0.1$. Plotted points are means of four independent replicates at time intervals between 1 and 10,000 generations, with $i = 0$.

generations (Bell 1988), so that at low migration rates one might expect the correlation of composition in the NCM to begin to decay appreciably at about 10 (K/N) cycles, which seems to provide a reasonable rule of thumb. For example, communities with 10,000 individuals distributed among 100 species, and which receive one immigrant or fewer per generation, will generally retain nearly the same composition for 100 generations or more. They would be perceived to be extremely stable on the timescale of almost all ecological studies.

Discussion

Distribution of Abundance

The properties of the neutral community model that I have described resemble those of natural communities in many important respects. In particular, the distribution of abundance among species is always approximately lognormal, skewed to the left by immigration. This is strikingly similar to many natural communities (see Gaston 1994). For example, Nee et al. (1991) present the distribution of abundance for the native breeding birds of Great Britain as a rare example of a community in which all species, even the rarest, have been adequately censused. It is a left-skewed lognormal resembling a neutral model with low immigration (fig. 1*B*), perhaps because species reported as breeding infrequently were omitted.

The neutral models described previously by Caswell (1976) do not lead to lognormal distributions of abundance. They are versions of differential-equation models developed previously to analyze neutral theory in population genetics, and resemble the model that I have used except that the constant rate of immigration refers to new species only. Thus, the probability that a new species, not already present, will arrive in a given interval of time is independent of the number of species already present. This is equivalent to assuming that the pool of available variants is effectively infinite, which may be a more appropriate simplification for gene mutations than for species. Different versions of this model all lead to log-series distributions of abundance. With very high rates of immigration the right-hand end of the distribution may resemble a

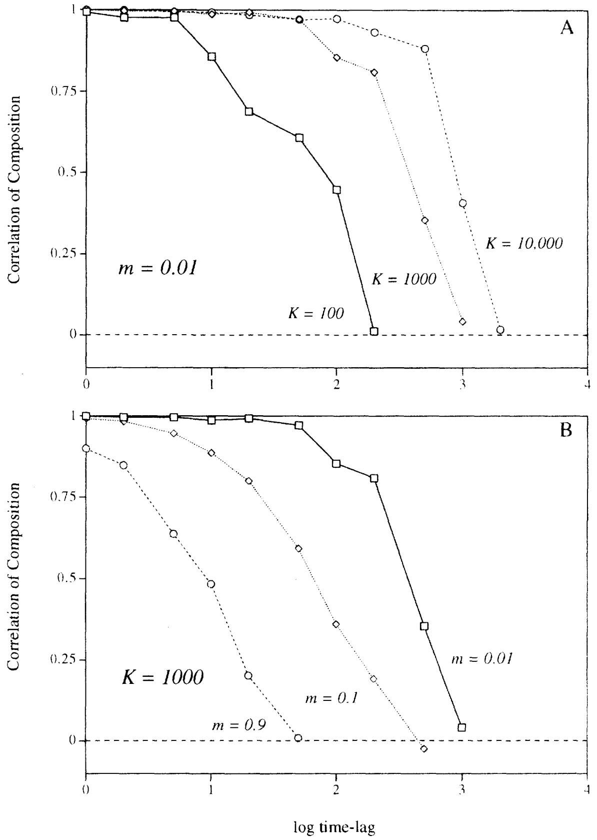

Figure 7: Decay of correlation of community composition through time. Plotted points are means of two to six independent replicates, each run out to 2,000 generations and then run for a number of subsequent generations, representing the time lag. Community composition (numbers of individuals of each species, including species absences) was recorded after 2,000 generations and then again after the lag, and the Pearson correlation coefficient of the two communities was calculated. Standard parameter values were $N = 100$, $b = d = 0.5$, $i = K/N$. *A*, Immigration rate m fixed, community size K varies. *B*, Community size K fixed, immigration rate m varies.

lognormal, but the distribution as a whole does not (see fig. 2 in Caswell 1976). The neutral model described by Hubbell (1995), however, generates S-shaped dominance-diversity curves indicative of lognormal distributions skewed to the left by immigration. The immigration model is not specified precisely, but seems to involve drawing a fixed number of individuals at random from a pool of fixed composition. When the species in this pool have equal frequencies, this is nearly equivalent to the model I have used, in which each species has a fixed and equal probability of receiving a single migrant in each cycle.

Neutral models may therefore generate either log-series or lognormal distributions, depending on the pattern of immigration. In either case, demographic stochasticity tends to create a lognormal distribution among common species. This is skewed to the left by immigration, but with a constant immigration rate the underlying lognormal is clearly apparent, unless the immigration rate is very high. When every migrant represents a new species the immigration rate is effectively very high, and the skew becomes so extreme that the distribution approaches the log series, in which the most frequent abundance category comprises species represented by a single individual. This is likely to be an extreme case, and in more plausible circumstances neutral models will tend to generate the moderately skewed lognormal distributions characteristic of natural communities. Immigration rates are likely to be high, relative to birth rates, in small areas or among motile organisms; thus, a community of plants censused in a region much larger than average seed dispersal distances is expected to resemble the upper two panels of figure 2, showing a clear lognormal pattern, whereas moths censused at a light trap will tend to resemble the lower two panels of figure 2, showing an approach to a log-series distributions.

Neutral Models and Neutral Communities

Whether or not one can infer that natural communities are structured by little more than scale effects and stochastic variation in vital rates is less straightforward (Harvey et al. 1983). Certainly, any such inference can be entertained only if the term "community" refers to a set of species that are ecologically or functionally equivalent in some sense, and "assemblage" or some other term might be preferable. The term "structure" is perhaps equally questionable, given that the only elements of structure are the number of species and the distribution of abundance: the neutral model does not apply to organized communities containing coupled species such as predator and prey, parasite and host, or mutualists. Even the notion of ecological equivalence is rather vague; I shall take it to refer to a set of species for each member of which no interaction with another member is positive. If community structure is determined to some extent by competition, then at least one interaction for each member is negative; the neutral model is the limiting case in which all interactions are negative and equal. This definition excludes sets such as "British birds" (because some members prey on others) and "herons" (because species may live on different continents and could never interact). Needless to say, this criterion is difficult to apply precisely in practice, especially because the greater the degree of equivalence, the fewer species there will be (K. J. Gaston, unpublished). It also ignores the important point that simple counting may be inadequate to express the diversity of communities whose members vary in size. For sympatric species of similar habits and size, then, it is possible to argue that because lognormal distributions emerge as a general consequence of a simple neutral model there is no need to seek any more specialized or onerous hypothesis to explain the general form of the distribution of abundance. The main reason for rejecting this argument, I believe, is that the same species tend to be abundant and the same species rare, at different sites and at different times, and it follows that there is some systematic source of variation among species that must be responsible for the overall distribution of abundance.

Although this argument is very persuasive, it is nevertheless not entirely out of the question that a neutral model could explain both the general form of the distribution and the specific identity of rare and common species. Suppose that the entire community of individuals belonging to some set of species is described by the NCM and therefore exhibits a lognormal distribution of abundance, in which some species will be very abundant and others rare. The community at any particular site will also develop a lognormal structure. It is impossible to predict which species will become common and which rare under the conditions used in the simulations described here, where each species has the same rate of immigration; but when a relatively small local community is embedded within the much larger global community, then the species that are globally the most common will have the greater rates of immigration and will therefore tend to be locally the most common also. This spatial covariance would in turn generate a positive range-abundance relationship, in agreement with most reports (Gaston 1996).

I think that most ecologists would be uncomfortable with this interpretation, because it would contradict the widely held view, so familiar that it is rarely enunciated, that the readily observed differences among species must have functional consequences that will cause some to become more abundant than others. For many types of character, however, this point of view has a serious weakness. Suppose that some prominent morphological character, of the sort that is often used to distinguish one species from another, has the effect of increasing the abundance of species that possess it. Species with dissected leaves, say, tend to increase in abundance whereas those with entire leaves decrease. The consequence will be a steady increase in the proportion of species with dissected leaves in the community, until those with entire leaves have been eliminated. At this point, the variation among species is quasi-neutral and the NCM will apply. For this reason, there is little prospect of finding consistent correlations between abundance and characters of this sort.

It is more likely that abundance will be related to the degree of specialization, so that species that are able to grow successfully in a wider range of sites will be more abundant than more narrowly specialized types (Brown 1984; but see Gaston et al. 1997 for a critique). Given local density regulation, this will generate frequency-dependent selection so that all species, whatever their abundance, have equal fitness at equilibrium. The NCM does not apply because any departure from equilibrium causes the reappearance of variation in fitness. This approach is capable of explaining the distribution of abundance through the underlying distribution of the conditions of growth to which species are differentially adapted. The models of MacArthur (1957) and Sugihara (1980), for example, are attempts to interpret abundance in this way through creating general theories of environmental structure. The weaknesses of the approach are that the simple physical analogies that the models appeal to have no known relationship to the structure of real environments and that there is no general account of the distribution of "conditions of growth."

The final possibility is perhaps the most appealing. There may indeed be systematic differences among species in their inherent tendency to become abundant or rare. This tendency has an unknown distribution but is rather weak. The history of any particular site is dominated by the accidents of arrival, establishment and proliferation, and the distribution of abundance therefore tends toward the lognormal generated by demographic stochasticity.

Setting these complications aside, we can at least conclude that the lognormal form of most abundance distributions can no longer be represented as evidence for one or another theory of community structure, because it is generated by a simple neutral model. The form of this distribution can also be predicted to some extent. For instance, variance $q\sim$ will increase as the distribution becomes more left skewed at sites where immigration is greater relative to birth. A good example is provided by the difference between moth communities on small rocky skerries and larger, more heavily vegetated islands in the Baltic (Nieminen and Hanski 1998). The converse is also true: where a distribution differs markedly from the lognormal it cannot be explained in this way and therefore requires a functional explanation. For example, communities that are dominated by two or three nearly equally abundant species, and that lack species of moderate abundance, are inconsistent with the NCM. Hanski (1982) has described cases in which the number of sites at which a species occurs is distributed bimodally among species, and this may represent a situation of this sort.

Community Structure in Space and Time

There is an important distinction between dynamic null hypotheses and statistical null hypotheses. Dynamic null hypotheses, such as the NCM, assert that the input to the system varies stochastically, whereas statistical null hypotheses assert that the output varies stochastically. Consider, for example, the distribution of species number S over a range of sites surveyed through a number of years. The simplest statistical H_0 would be that each species is distributed at random among sites and years, given its overall abundance. The community would then be unstructured both in space and in time, with no tendency for neighboring sites or consecutive years to resemble one another. This lack of structure, however, does not by any means follow from a dynamic H_0. The species that are actually present at a given place and time are determined, not by chance alone but also by history. The accidents of immigration and colonization, followed by local growth, persistence, and dispersal, will ensure that neutral models create correlations among sites (Hubbell 1995) and among years (fig. 7). These correlations can be strong, extensive and persistent, especially when the effects of history are preserved from the ravages of chance by large community size or low immigration rate. Where such correlations are found to exist, the overthrow of the statistical H_0 does not in the least imply that the dynamic H_0 is inadequate. Indeed, it is unlikely that the dynamic H_0 can ever be convincingly refuted merely by an inspection of the patterns of abundance and diversity. It can be refuted only by demonstrating that these patterns correspond to spatiotemporal variation in other features of the landscape, such as productivity, and this requires information in addition to the biological survey itself.

Measures of Diversity

It is generally recognized that there are two aspects of diversity. One is the number of distinct entities, the number of alleles in population genetics, or the number of species S in community ecology. The second is a measure of the disparity in abundance of these entities. The genetic variance is used for this purpose in population genetics, because of its theoretical link to the rate of response to selection and because it can be analyzed (partitioned into component variances) by standard procedures. The same reasons justify the use of the variance of log abundance $q\sim$ in community ecology, in place of the rich variety of other measures suggested piecemeal by previous authors (see Smith and Wilson 1996). These two quantities are different aspects of diversity; they are not different components of diversity, because there is no quantity of interest that can be partitioned into a range and a variance. Attempts to conflate them in any single "index of diversity" are therefore difficult to justify. Rather, the relationship between the two remains to be studied.

Observed patterns in species number and the variance of abundance can be compared with those generated by the NCM. The most general of these is that the two main measures of diversity, species number S and the variance of log abundance $q\sim$, will tend to be positively correlated when they are estimated at sites that vary in size or age. This appears to be the case when completely different communities are compared (Sugihara 1980). I have not found any compilation of data for similar communities at different sites, but negative correlations have been reported for sedges by G. Bell, M. Lechowicz, and M. Waterway (unpublished manuscript) and for moths by Cook and Graham (1996).

Regulation of Diversity

The variation of species number S with sample size n and community size K constitute two versions of the species-area relationship. Increasing sample size is akin to sampling successively larger areas of a mainland region, whereas varying community size is more nearly comparable with sampling islands of different size. In most cases, log S increases linearly with log area, so that the species-area relationship can be described by a power law; in other cases, however, S increases directly with log area (see Connor and McCoy 1979).

The log-series distributions generated by the models studied by Caswell (1976) give rise to semilog relationships of the form $S = a \log n + c$, where a is a migration rate. Lognormal distributions yield a sigmoid curve, which is roughly linear in the central part of the plot (Preston 1948). In either case, the graph of log S on log n is convex upward. The NCM can give rise to linear log-log graphs, however, when birth rate is high and immigration rate low. In these circumstances a canonical lognormal is generated, and a power law with an exponent of about 0.27 follows, as expected (Preston 1962). Leitner and Rosenzweig (1997) have shown that this exponent should be about 0.77 for nested samples of a region within which species distributions have explicit spatial locations; this spatial structure is not present in the NCM.

Caswell's models predict a similar semilog relationship between species number S and community size K, so that the log-log graph will be convex upward. He suggests that a linear log-log plot might be the consequence of an increase of immigration rate with species number. Lognormal distributions of abundance generally give rise to linear log-log graphs. Under the NCM, the graph of log S on log K is convex upward for $m > 1/s$, and concave upward for $m < 1/s$; for $m = 1/s$ the graph is linear with a slope of about 0.3 (see fig. 3).

Neutral models are thus capable of generating species-area relationships that resemble those found in nature. In particular, the characteristic linearity of log-log plots emerges from the NCM with plausible combinations of parameter values. Moreover, the NCM can generate power laws with exponents close to those characteristic of natural systems. Indeed, the dependence of the shape of the relationship on the magnitude of immigration rate relative to birth rate can be given a simple and natural interpretation. In small areas the perimeter is large relative to the interior, and communities will contain a large proportion of recent immigrants; in these circumstances the log-log graph will be convex upward. In large areas the perimeter is much smaller relative to the interior, and most individuals will be natives, born within the area; the log-log graph will then tend to be linear. The neutral model thus gives rise to a biphasic species-area relationship that resembles that of natural regions within a single biogeographic province (Rosenzweig 1995).

More generally, an important lesson of the neutralist-selectionist controversy in population genetics was that prolonged contemplation of frequency distributions is unlikely to provide decisive evidence about ecological or evolutionary mechanisms. It will be rare indeed that a particular pattern of abundance or diversity cannot be explained both by a neutral model, given the appropriate combination of parameters, and equally by some functional hypothesis. Sugihara's sequential-breakage model, for example, generates canonical lognormal abundance distributions through a procedure that is taken to represent an ecological mechanism of habitat partitioning (Sugihara 1980), and the self-similarity of species distributions may by itself lead to skewed lognormal distributions (Harte et al. 1999). One of the most important roles of neutral models in community ecology is merely cautionary. It might be that lognormal distributions of abundance indicate competition, that species-area relationships reveal the degree of community saturation, or that the spatial or temporal structure of diversity reflects patterns of heterogeneity and disturbance. These possibilities should be evaluated, however, in relation to the simpler explanations offered by chance and history.

Acknowledgments

This work was supported by research grants from the Natural Science and Engineering Research Council of Canada and Fonds pour les Chercheurs et à l'Aide de la Recherche of Québec. I am grateful to K. Gaston and an anonymous reviewer for useful criticisms of the original manuscript. The final version of the manuscript benefited from discussions with S. Hubbell, who has independently reached very similar conclusions.

Literature Cited

Bell, G. 1988. Sex and death in protozoa. Cambridge University Press, Cambridge.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 122: 295–299.

Caswell, H. 1976. Community structure: a neutral model analysis. Ecological Monographs 46:327–354.

Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. American Naturalist 113:791–833.

Cook, L. M., and C. S. Graham. 1996. Evenness and species number in some moth populations. Biological Journal of the Linnean Society 58:75–84.

Ewens, W. J. 1972. The sampling theory of selectively neutral alleles. Theoretical Population Biology 3:87–112.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon, Oxford.

Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12:42–58.

Gaston, K. J. 1994. Rarity. Chapman & Hall, London.

———. 1996. The multiple forms of the interspecific abundance-distribution relationship. Oikos 76:211–220.

Gaston, K. J., T. M. Blackburn, and J. H. Lawton 1997. Interspecific abundance-range size relationships: an appraisal of mechanisms. Journal of Animal Ecology 66: 579–601.

Hanski, I. 1982. Dynamics of regional distribution: the core and satellite species hypothesis. Oikos 38:210–221.

Harte, J., A. Kinzig, and J. Green. 1999. Self-similarity in the distribution and abundance of species. Science (Washington, D.C.) 284:334–336.

Harvey, P. H., R. K. Colwell, J. W. Silvertown, and R. M. May. 1983. Null models in ecology. Annual Review of Ecology and Systematics 14:189–211.

Hubbell, S. P. 1979. Tree dispersion, abundance and diversity in a tropical dry forest. Science (Washington, D.C.) 203:1299–1309.

———. 1995. Towards a theory of biodiversity and biogeography on continuous landscapes. Pages 171–199 *in* G. R. Carmichael, G. E. Folk, and J. L. Schnoor, eds. Preparing for global change: a midwestern perspective. SPB Academic, Amsterdam.

———. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. Coral Reefs 16(suppl.): S9–S21.

———. 1998. The maintenance of diversity in a neotropical tree community: conceptual issues, current evidence and challenges ahead. Pages 17–44 *in* F. Dallmeier and J. A. Comisky, eds. Biodiversity research, monitoring and modeling. Man in the Biosphere. UNESCO and Parthenon, Paris.

———. In press. A unified theory of biodiversity and biogeography. Princeton University Press, Princeton, N.J.

Hubbell, S. P., and R. B. Foster. 1986. Biology, chance and history, and the structure of tropical tree communities. Pages 314–329 *in* J. M. Diamond and T. J. Case, eds. Community ecology. Harper & Row, New York.

Hughes, R. G. 1986. Theories and models of species abundance. American Naturalist 128:879–899.

Karlin, S., and J. McGregor. 1967. The number of mutant forms maintained in a population. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, pp. 415–438.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

Leitner, W. A., and M. L. Rosenzweig. 1997. Nested species-area curves and stochastic sampling: a new theory. Oikos 79:503–512.

MacArthur, R. H. 1957. On the relative abundance of bird species. Proceedings of the National Academy of Sciences of the USA 43:293–295.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, N.J.

May, R. M. 1975. Patterns of species abundance and diversity. Pages 81–120 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Harvard University Press, Cambridge, Mass.

Nee, S., P. H. Harvey, and R. May. 1991. Lifting the veil on abundance patterns. Proceedings of the Royal Society of London B, Biological Sciences 243:161–163.

Nieminen, M., and I. Hanski. 1998. Metapopulations of moths on islands: a test of two contrasting models. Journal of Animal Ecology 67:149–160.

Preston, F. W. 1948. On the commonness and rarity of species. Ecology 29:254–283.

———. 1962. The canonical distribution of commonness and rarity. Ecology 43:185–215.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.

Smith, B., and J. B. Wilson. 1996. A consumer's guide to evenness indices. Oikos 76:70–82.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. American Naturalist 133:458–463.

Tokeshi, M. 1993. Species abundance patterns and community structure. Advances in Ecological Research 24: 111–186.

Watterson, G. A. 1974. The sampling theory of selectively neutral alleles. Advances in Applied Probability 6: 463–488.

Associate Editor: Dolph Schluter

LINKED CITATIONS

- Page 1 of 3 -

You have printed the following article:

The Distribution of Abundance in Neutral Communities
Graham Bell
The American Naturalist, Vol. 155, No. 5. (May, 2000), pp. 606-617.
Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28200005%29155%3A5%3C606%3ATDOAIN%3E2.0.CO%3B2-1

Literature Cited

Community Structure: A Neutral Model Analysis
Hal Caswell
Ecological Monographs, Vol. 46, No. 3. (Summer, 1976), pp. 327-354.
Stable URL:
http://links.jstor.org/sici?sici=0012-9615%28197622%2946%3A3%3C327%3ACSANMA%3E2.0.CO%3B2-Q

The Statistics and Biology of the Species-Area Relationship
Edward F. Connor; Earl D. McCoy
The American Naturalist, Vol. 113, No. 6. (Jun., 1979), pp. 791-833.
Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28197906%29113%3A6%3C791%3ATSABOT%3E2.0.CO%3B2-7

The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population
R. A. Fisher; A. Steven Corbet; C. B. Williams
The Journal of Animal Ecology, Vol. 12, No. 1. (May, 1943), pp. 42-58.
Stable URL:
http://links.jstor.org/sici?sici=0021-8790%28194305%2912%3A1%3C42%3ATRBTNO%3E2.0.CO%3B2-U

The Multiple Forms of the Interspecific Abundance-Distribution Relationship
Kevin J. Gaston
Oikos, Vol. 76, No. 2. (Jun., 1996), pp. 211-220.
Stable URL:
http://links.jstor.org/sici?sici=0030-1299%28199606%2976%3A2%3C211%3ATMFOTI%3E2.0.CO%3B2-J

LINKED CITATIONS

- Page 2 of 3 -

Interspecific Abundance-Range Size Relationships: An Appraisal of Mechanisms
Kevin J. Gaston; Tim M. Blackburn; John H. Lawton
The Journal of Animal Ecology, Vol. 66, No. 4. (Jul., 1997), pp. 579-601.
Stable URL:
http://links.jstor.org/sici?sici=0021-8790%28199707%2966%3A4%3C579%3AIASRAA%3E2.0.CO%3B2-Y

Dynamics of Regional Distribution: The Core and Satellite Species Hypothesis
Ilkka Hanski
Oikos, Vol. 38, No. 2. (Mar., 1982), pp. 210-221.
Stable URL:
http://links.jstor.org/sici?sici=0030-1299%28198203%2938%3A2%3C210%3ADORDTC%3E2.0.CO%3B2-R

Null Models in Ecology
Paul H. Harvey; Robert K. Colwell; Jonathan W. Silvertown; Robert M. May
Annual Review of Ecology and Systematics, Vol. 14. (1983), pp. 189-211.
Stable URL:
http://links.jstor.org/sici?sici=0066-4162%281983%2914%3C189%3ANMIE%3E2.0.CO%3B2-U

Theories and Models of Species Abundance
R. G. Hughes
The American Naturalist, Vol. 128, No. 6. (Dec., 1986), pp. 879-899.
Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28198612%29128%3A6%3C879%3ATAMOSA%3E2.0.CO%3B2-G

Nested Species-Area Curves and Stochastic Sampling: A New Theory
Wade A. Leitner; Michael L. Rosenzweig
Oikos, Vol. 79, No. 3. (Sep., 1997), pp. 503-512.
Stable URL:
http://links.jstor.org/sici?sici=0030-1299%28199709%2979%3A3%3C503%3ANSCASS%3E2.0.CO%3B2-5

On the Relative Abundance of Bird Species
Robert H. MacArthur
Proceedings of the National Academy of Sciences of the United States of America, Vol. 43, No. 3. (Mar. 15, 1957), pp. 293-295.
Stable URL:
http://links.jstor.org/sici?sici=0027-8424%2819570315%2943%3A3%3C293%3AOTRAOB%3E2.0.CO%3B2-L

LINKED CITATIONS

- Page 3 of 3 -

Metapopulations of Moths on Islands: A Test of Two Contrasting Models
Marko Nieminen; Ilkka Hanski
The Journal of Animal Ecology, Vol. 67, No. 1. (Jan., 1998), pp. 149-160.
Stable URL:
http://links.jstor.org/sici?sici=0021-8790%28199801%2967%3A1%3C149%3AMOMOIA%3E2.0.CO%3B2-P

The Commonness, And Rarity, of Species
F. W. Preston
Ecology, Vol. 29, No. 3. (Jul., 1948), pp. 254-283.
Stable URL:
http://links.jstor.org/sici?sici=0012-9658%28194807%2929%3A3%3C254%3ATCAROS%3E2.0.CO%3B2-U

The Canonical Distribution of Commonness and Rarity: Part I
F. W. Preston
Ecology, Vol. 43, No. 2. (Apr., 1962), pp. 185-215.
Stable URL:
http://links.jstor.org/sici?sici=0012-9658%28196204%2943%3A2%3C185%3ATCDOCA%3E2.0.CO%3B2-8

A Consumer's Guide to Evenness Indices
Benjamin Smith; J. Bastow Wilson
Oikos, Vol. 76, No. 1. (May, 1996), pp. 70-82.
Stable URL:
http://links.jstor.org/sici?sici=0030-1299%28199605%2976%3A1%3C70%3AACGTEI%3E2.0.CO%3B2-T

The Sampling Theory of Selectively Neutral Alleles
G. A. Watterson
Advances in Applied Probability, Vol. 6, No. 3. (Sep., 1974), pp. 463-488.
Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28197409%296%3A3%3C463%3ATSTOSN%3E2.0.CO%3B2-Q