The impact of neutrality, niche differentiation and species input on diversity and abundance distributions

Tommaso Zillio and Richard Condit
T. Zillio (tommaso.zillio@gmail.com) and R. Condit, Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, USA and National Center for Ecological Analysis and Synthesis, 735 State St. Santa Barbara, CA 93101, USA.

Hubbell (2001) showed that a species input process can maintain species diversity in the absence of any other diversifying forces, and he noted that species immigration as described in island biogeography (MacArthur and Wilson 1967) could be viewed as equivalent to the speciation process in an evolutionary sense. Both processes of species input can lead to high species diversity at equilibrium. The theory Hubbell refers to as ‘unified’ more often goes by ‘the neutral theory’, and its main contributions are quantitative predictions of diversity and species abundance derived from birth and death processes (Engen and Lande 1996). Since the dynamical introduction of species in the neutral model is not speciation in the usual (genetic) sense, we refer to it as species input, described by Chave and Leigh (2002) as: “a small but steady input of new species in the system, which represent immigration of novel types and speciation”.

Interest in the neutral model has been generated especially by its quantitative predictions of community structure. While there are patterns for which the neutral theory provides no explanation (Leigh et al. 2004, Tilman 2004, Dornelas et al. 2006), Hubbell (1997, 2001) and several other studies demonstrated how accurately it accounts for distributions in diverse communities. In particular the long tail of rare species so often observed (McGill 2003, Condit et al. 2005) is predicted quite closely by a neutral model. Analytical solutions for the equilibrium abundance distribution are now available for a neutral community (Volkov et al. 2003, Alonso and McKane 2004, McKane et al. 2004, Etienne and Olff 2004, Etienne 2005), or with symmetrical density-dependence (Volkov et al. 2005), and beta-diversity can be predicted under dispersal limitation (Chave and Leigh 2002, Zillio et al. 2005). However, the abundance distribution under dispersal limitation in a local community or with niche differences has not been derived.

We believe the dilemma – rejecting neutrality despite its accurate predictions – stems from confusion about the two key features of Hubbell’s model: neutrality and the speciation process. Here we examine which process in the model leads to the abundance distribution dominated by rare species. With simulations drawn from the neutral approach, we relax the assumption of identical species and consider 1) the rate...
of species input, 2) limited dispersal, and 3) niche
differences. We set aside the issue of diversity and
abundance at larger scales in order to focus on the
structure of local communities. Our goal is under-
standing which factors are important controls of
abundance and diversity and why the ‘neutral model’
can be successful.

In order to test how dispersal and niche partitioning
affect abundance distributions, we had to resort to
simulations, as did Loreau and Mouquet (1999),
Bastolla et al. (2001), Chave and Leigh (2002), Schwilk
and Ackerly (2005) and Gravel et al. (2006). Simula-
tion results are risky in that only a limited set of
parameters can be explored and noise can conceal
equilibria. Our simulations, though, include more
individuals than earlier studies, and we carefully
analyze multiple outputs for equilibrium behavior.
Dispersal, species input, and niche parameters can be
explored as widely as possible, but we cannot of course
consider all plausible combinations.

Methods

The model

We start with the classical voter model as developed by
geneticists (Kimura and Weiss 1964) describing DNA
mutation and independently rediscovered by matema-
ticians (Liggett 1985, 1999) describing how voting
preferences might spread; it has been studied subse-
tively in physics and chemistry (Frachebourg and
Krapivsky 1996). The original model describes an array
in d dimensions, but all ecological work has focused
on the plane, d = 2. Each site in the array is occupied
by one individual, and no empty sites are allowed,
producing constant density (the zero-sum game of
Hubbell 2001). Individuals are assigned numbers that
indicate the species (or vote). The simulation proceeds
by randomly choosing one individual and removing it.
The location is then immediately replaced by randomly
selecting another site in the array and assigning its
species to the vacated location. In the true voter model,
the parent is always one of the vacant site’s immediate
neighbors, which in ecological terms means poor
dispersal. (In ecology, removal—replacement is a
death—birth cycle, but the alternate versions describe
voter’s opinions or gene movement). In a finite system
and without the introduction of new species the only
stable equilibrium is mono-dominance, which is essen-
tially a statement of Gause’s (1934) principle. On the
other hand, in an infinite system with high dimension-
ality (d ≥3), stable coexistence is mathematically
possible even in the absence of speciation (Liggett
1985). Durrett and Levin (1996), Hubbell (1997,
2001), Houchmandzadeh and Vallade (2003), Volkov
et al. (2003, 2005) and Zillio et al. (2005) all explored
variants of the voter model in which a constant
speciation process was added. This allows species
diversity to be maintained indefinitely.

Our principal interest is in niche differences, or
species-specific habitat preference (Schwilc and Ackerly
2005). To consider habitat preferences, we assign a
habitat variable to every site, which we drew from a real
topographic map. Thus, each site in the array had an
elevation, although we could just as easily pretend the
habitat variable means soil moisture or phosphorus
concentration. The point of using a real topographic
map was to ensure that the spatial arrangement of
habitat follows at least one reality, and we know from
many studies on plant distributions that topography is
often relevant (Davies et al. 1998, Clark et al. 1999,

We implemented habitat preference by allowing
each species’ mortality to vary with the environmental
gradient. In one model, there are two discrete habitats,
while in the second, each species has a unique response
to the environment (“the continuous habitat model”).
In the discrete habitat model we divided the elevation
map in two parts, split on the median, with species
assigned a preference to one of the two zones. A single
parameter is required: the magnitude of the preference,
identical for all species, but with some species favor-
ing habitat 1 and others favoring habitat 2. In the
continuous habitat model the map is not partitioned,
but each species has a preferred elevation. The relative
mortality is modulated using a Gaussian habitat
preference centered on the species-specific preferred
elevation mean. The variance of the Gaussian is the
same for every species in the simulation. This contin-
uous model has thus two parameters which must
be fixed at the outset of a simulation: the variance of
the Gaussian habitat preference and the magnitude of
the preference, both identical for every species.

Diversity indicators

Hubbell (2001) demonstrated that the old diversity
parameter, Fisher’s z (Fisher et al. 1943), provides
simple predictions about diversity and abundances in
the neutral model with no dispersal limitation. Fisher
et al. defined z from:

\[ S = z \ln \left( 1 + \frac{N}{z} \right) \]  \hspace{1cm} (1)

and it turns out that \( z = (N-1)v/(1-v) \) (Etienne
2005), where \( N = \) community size and \( v = \) speciation
rate, which is conveniently approximated by \( z \sim Nv \)
when \( v < 1 \) and \( N \gg 1 \). The value given by Hubbell
(2001) is \( z = 2Nv \), with the factor 2 resulting if
multiple speciation events are allowed to occur within
one time-step (Etienne 2005). Under the neutral assumption, the number of singleton species (one individual) must also be $Nv$, because the extinction rate must equal the speciation rate at equilibrium. Moreover the same neutral community has a log-series species abundance distribution (Hubbell 2001, Pueyo 2006):

$$s(n) = \frac{x^n}{n}$$  

(2)

where $x < 1$ and $s(n), n = 1, \ldots, N$ is the number of species with $n$ individuals. The number of singletons is $s(1) = x$, but since in every meaningful case $x \approx 1$ with good precision, one expects that $s(1) \approx x$ at least when the dynamics is neutral. Since the log-series is a truncated power law with an exponent $\gamma$ equal to $-1$, it can be tested measuring the slope of rare species abundances on a log-log plot. This slope turns out to be a useful ecological indicator, telling if a community has more ($\gamma < -1$, curve is steeper) or less ($\gamma > -1$, curve is flatter) rare species than expected under neutral dynamics with species introduction. Since the log-series is expected to behave like a power law only for rare species, the slope $\gamma$ should be calculated only for abundances smaller than a given value; in our case, after having observed the results from the simulations, we chose to use only abundances in the range $1 \leq n \leq 100$. The calculation of $\gamma$ is performed starting from the abundance data $\{n_i\}, i = 1, \ldots, S$ by fitting a power law function $s(n) \sim n^\gamma$ by maximum likelihood, i.e. searching for the value of the exponent $\gamma$ that maximizes the likelihood:

$$\prod_{i=1}^{S} s(n_i) \theta(100 - n_i) = \prod_{i=1}^{S} k(\gamma) (n_i)^\gamma \theta(100 - n_i)$$  

(3)

where $n_i$ is the abundance of the $i$-th species, $\theta(x) = 1$ if $x \geq 0$, 0 otherwise, selecting thus only species with abundance less than 100 individuals, and $k(\gamma) = \sum_{j=1}^{100} j^\gamma$ is the normalization constant of the power law function up to abundance 100. This method does not require any binning or grouping of data and is guaranteed to yield the same results regardless of the particular way data are plotted.

Algorithm

The algorithm starts on a grid initialized with a single species occupying all sites, but the number of species grows because of species input. At each time step, a random grid location is chosen, and the species identity at the site is targeted for replacement. Whether a death event is carried out depends stochastically on habitat preference.

First, consider replacement.

1. With probability $v$ (the rate of introduction), the site is assigned a species identity not currently present in the plot.
2. With probability $1-v$, the site is assigned the identity of a neighbor chosen at random from all locations closer than $d_l$ grid points away. In other words, the potential parents are in the square centered on the vacant site with side $2 \times d_l + 1$.

The dispersal kernel is thus a step function, with the probability of drawing a parent from inside the square equal for all parents, and the probability of drawing from outside the square always zero. The square kernel is vastly easier and faster to simulate, and closely approximates the more realistic circular kernel (Zillio et al. 2005). We see a posteriori that the scale of the kernel has a negligible impact on our results, so the details of the shape of the kernel are not important to our conclusions.

Now consider habitat preference. After targeting a grid location for death, the species at the site and its habitat preference are considered before proceeding. With a survival probability $p_s$, the individual can be spared, with $p_s$ depending on the habitat preference. If the individual survives, another site is chosen at random from the entire grid, and again tested for survival. A time-step is only counted when a death takes place.

In the model of discrete habitat, $p_s$ take only two values. If a species is on its preferred half of the terrain, $p_s = h$; otherwise, $p_s = 0$. That is, if an individual is targeted for a death event outside its preferred terrain its death is certain, but on its preferred terrain it has a chance at survival. So $h$ is a rescue probability: a small value means a slight advantage on one habitat, whereas a value close to one means species seldom die on their favored terrain. Again, all species have the same $h$, but some species prefer low elevation, other species high.

In the continuous habitat case, the survival probability $p_s$ is calculated as a Gaussian. Let $P$ be the mean preference of the species targeted for death, $\sigma^2$ the variance, and $E$ the elevation at the chosen point. Then:

$$p_s = h \exp \left( -\frac{(P - E)^2}{2\sigma^2} \right)$$  

(4)

The strength of the preference is $h$, as in the discrete model. Here, a species achieves its maximum rescue effect only at elevation $P$, the center of its Gaussian. $\sigma$ is the breadth of the tolerance, expressed in meters,
indicating how far from the center a species gains any benefit. Small values of $\sigma$ mean species could be rescued only over a narrow range of elevations, but a large enough $\sigma$ removes any habitat impact at all. This mimics the standard Gaussian response to environmental gradients (Whittaker 1960, Gauch 1982, Faith et al. 1987).

At the outset of any one simulation, $h$ and $\sigma$ are fixed. In the discrete model, every species is assigned its preferred half at random. In the continuous model, every species is assigned $P$ at random, drawn on a uniform distribution between $\min(El) - \sigma$ and $\max(El) + \sigma$, where $\min(El)$ and $\max(El)$ are the minimum and maximum elevation of the map used, respectively. We do not maintain a pool of outside species as a metacommunity. The introduction process represents either completely novel species, or simply re-introduction of a species that has gone extinct.

With no habitat preference (i.e. $h = 0$) our model is neutral; if the dispersal $d_l$ is set to a value greater than the community extent, space is eliminated and we recover Hubbell’s original drift model.

**Simulations**

All simulations were run on a square 1000 $\times$ 1000 grid, so $10^6$ individuals. As typical for these models, it was wrapped at the edges, meaning that a dispersal event off the right edge lands on the left edge. The elevation of the grid was taken from a 2.2 $\times$ 2.2 km section of the Barro Colorado Island, Panama, digital terrain map (R. Stallard, unpubl.). Locations on the grid are thus spaced by 2.2 m, which corresponds to the average distance between trees $\geq$ 2.6 cm dbh in the Barro Colorado forest (data from http://ctfs.si.edu/datasets). The elevation of this section ranges from 27 m to 170 m ASL, with a median of 96.4 m (Fig. 1).

Given $N$ individual in the community, a generation is an update of $N$ timesteps (so that, on average, each individual is updated once). All simulations were run for 10000 generations, aimed at attaining equilibrium. Subsequently, at every 100th generation, we computed the number of species with $n = 1, \ldots, N$ individuals, $s(n)$, and the total number of species $S$. All analyses were based on the mean of 100 such configurations taken from the final 10000 generations. We used the slope of a graphed of $\log(s(n))$ vs $\log(n)$ to characterize the form of the abundance distribution, as described in Methods. Fisher’s $\alpha$ was calculated by solving Eq. 1.

We had three sets of parameters to test: the rate of species input ($v$), the dispersal distance ($d_l$), and the habitat preference ($h$ for the magnitude and $\sigma$ for the breadth).

---

1. In the Barro Colorado 50-ha plot in Panama, new species have appeared at a rate approximately $10^{-4}$ per recruit. Thus we used $v = 10^{-3}$, $10^{-4}$ and $10^{-5}$ to explore variation.
2. Muller-Landau (2001) estimated seed dispersal kernels from seed collection sites within the Barro Colorado plot, and reported a mean dispersal distance of 39 m, equivalent to 18 steps in our model’s 2.2 m grid. We thus explored \( d_l = 5, 15 \) and 45.

3. We have no basis for how tree mortality varies with topography at Barro Colorado, indeed, direct measures of habitat preference are scarce (Gaston 2003). Thus we considered \( h = 0, 0.01, 0.1, 0.5 \), covering no habitat preference to a very strong habitat preference, and \( \sigma = 2, 10, 50 \), spanning a very narrow tolerance to very broad.

We did not run every combination of all parameters. In particular, we discovered that the dispersal parameter had little impact on any result, and so in the continuous habitat model, we set \( d_l = 30 \).

An additional series of simulations were used for examining convergence to the equilibria. We used a continuous habitat with \( d_l = 30; \sigma = 2; v = 10^{-2}, 10^{-3}, 10^{-4} \) and \( 10^{-5} \); \( h = 0, 0.1, \text{ and } 0.5 \). These simulations were performed for 20,000 generations, and every 10 generations from the start we registered the total number of species \( S \), the number of species with one individual \( s(1) \), and the abundance of the initial species, \( s(I) \) (Fig. 2).

A last series of simulations was used to explore the behavior of the system when \( n = 0 \). These simulations started from an equilibrium configuration obtained with \( v = 10^{-4} \) and were followed for 20,000 generations after removing species input.

**Results**

A 10-fold increase in the rate of species introduction \( v \) caused a 7-to-10-fold increase in the number of species \( S \) present after 10,000 generations (Table 1, 2), echoing conclusions in Chave et al. (2002). Variation in dispersal had no impact on \( S \). Surprisingly, the stronger the habitat preference, the fewer the species, while the breadth \( \sigma \) of the response had a negligible impact.

Dispersal had no impact on the abundance distribution, since curves from different dispersal distances but the same input rate and habitat parameters were superimposed (Fig. 3). In contrast, a 10-fold change in the input parameter moved the curves by close to 10-fold, but the slopes were unchanged (Table 1, 2). Niche differences had only a slight impact on the abundance distribution.

The absolute number of singletons was close to the neutral prediction of \( s(1) = Nv \) when habitat preference was absent or weak. In most simulations this meant 10–20% of the species were singletons, with a moderately higher fraction when the species input rate was higher (Table 1, 2). A near constant proportion of singletons follows because \( S \) increased nearly linearly with \( v \) (Chave et al. 2002).

With \( v = 10^{-3} \), communities reached equilibria in all features we measured within 20,000 generations. The number of singleton species saturated most rapidly, the total number of species more slowly, while the abundance of the initial species was the slowest to attain equilibrium. At \( v = 10^{-4} \) or \( 10^{-5} \), the number of singletons was at equilibrium by 20,000 generations (not shown), but the initial species \( I \) was still abundant. If its abundance \( s(I) \) was plotted versus the number of generations \( t \) multiplied by the immigration rate \( v \), the curves of the various simulations collapsed (Fig. 2), showing that \( 1/v \) is interpretable as a time scale of the system. When habitat preference was added, all quantities displayed shorter equilibrium times (not shown).

In the absence of species introduction, diversity rapidly decayed (Fig. 4). With no niche differences...
Table 1. Simulations with two discrete habitats. v: species input rate. \( d_l \): maximum dispersal allowed. h: habitat preference strength. S: average number of species. s(1): average number of species with only one individual. \( a \): Fisher's parameter of diversity. \( g \): slope of the first two orders of magnitude in abundance of the species abundance curve on a log–log plot.

<table>
<thead>
<tr>
<th>( d_l )</th>
<th>h</th>
<th>( v = 10^{-3} )</th>
<th>( v = 10^{-4} )</th>
<th>( v = 10^{-5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S )</td>
<td>( s(1) )</td>
<td>( a )</td>
<td>( g )</td>
<td>( S )</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>7118</td>
<td>1001 (14.0%)</td>
<td>1036</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>6998</td>
<td>996 (14.2%)</td>
<td>1015</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>6952</td>
<td>1003 (14.4%)</td>
<td>1007</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>6387</td>
<td>995 (15.6%)</td>
<td>912</td>
</tr>
<tr>
<td>15</td>
<td>0.01</td>
<td>6391</td>
<td>996 (15.6%)</td>
<td>913</td>
</tr>
<tr>
<td>45</td>
<td>0.01</td>
<td>6767</td>
<td>998 (14.8%)</td>
<td>976</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>5188</td>
<td>951 (18.3%)</td>
<td>716</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>5215</td>
<td>952 (18.3%)</td>
<td>721</td>
</tr>
<tr>
<td>45</td>
<td>0.1</td>
<td>5546</td>
<td>972 (17.5%)</td>
<td>774</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>4341</td>
<td>757 (17.4%)</td>
<td>583</td>
</tr>
<tr>
<td>15</td>
<td>0.5</td>
<td>4314</td>
<td>772 (17.9%)</td>
<td>579</td>
</tr>
<tr>
<td>45</td>
<td>0.5</td>
<td>4499</td>
<td>799 (17.8%)</td>
<td>607</td>
</tr>
</tbody>
</table>

Discussion

We have shown that there are circumstances where niche differences among species have little impact on the dominant control. In these circumstances, neutral predictions on diversity, Fisher's \( \alpha \) and the entire community structure, while an input of new species is the dominant control. In these circumstances, neutral diversity can be seen as analogous to the theory of ideal gases. Gas molecules of different 'species' (oxygen vs nitrogen) can be seen as analogous to the theory of ideal gases.
nitrogen etc.) are very different in many ways, nevertheless, properties of an ideal gas can be predicted by random movements of particles whose detailed chemical differences are ignored. In both cases the theory is obviously “wrong”, ignoring important details of the system, but is able to yield several important results that would have been impossible to derive otherwise. As we do not expect the theory of ideal gases to agree with empirical results in all respects (it does not predict the liquid-gas phase transition, for instance), likewise we should not expect the neutral theory to predict everything. The theories’ values lie in their simplicity, in their ability to provide a manageable way to calculate some properties of the system, and as a starting points to develop more refined theories (quasi-ideal gases or quasi-neutral models).

Ricklefs (2005) wondered if neutral theories, with their long extinction times, can arrive at equilibria in realistic times. In this work we have started our simulations with a single species occupying the entire region, which is in this sense the worst case scenario (Cox 1989), i.e. the situation with the longest expected

<table>
<thead>
<tr>
<th>σ</th>
<th>h</th>
<th>n = 10⁻³</th>
<th>S</th>
<th>s(1)</th>
<th>α</th>
<th>γ</th>
<th>n = 10⁻⁴</th>
<th>S</th>
<th>s(1)</th>
<th>α</th>
<th>γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.01</td>
<td>6893</td>
<td>1001 (14.5%)</td>
<td>997</td>
<td>−1.01</td>
<td>848</td>
<td>101 (11.9%)</td>
<td>91</td>
<td>−1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.03</td>
<td>6720</td>
<td>1004 (14.9%)</td>
<td>968</td>
<td>−1.02</td>
<td>798</td>
<td>99 (12.4%)</td>
<td>85</td>
<td>−1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>5609</td>
<td>993 (17.7%)</td>
<td>784</td>
<td>−1.08</td>
<td>618</td>
<td>97 (15.7%)</td>
<td>64</td>
<td>−1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>6643</td>
<td>1002 (15.1%)</td>
<td>955</td>
<td>−1.02</td>
<td>786</td>
<td>101 (12.8%)</td>
<td>84</td>
<td>−1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.03</td>
<td>5848</td>
<td>999 (17.1%)</td>
<td>823</td>
<td>−1.07</td>
<td>649</td>
<td>99 (15.2%)</td>
<td>68</td>
<td>−1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>4327</td>
<td>962 (22.2%)</td>
<td>581</td>
<td>−1.24</td>
<td>484</td>
<td>95 (19.6%)</td>
<td>49</td>
<td>−1.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.01</td>
<td>6219</td>
<td>997 (16.0%)</td>
<td>884</td>
<td>−1.04</td>
<td>698</td>
<td>101 (14.5%)</td>
<td>73</td>
<td>−1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.03</td>
<td>5465</td>
<td>993 (18.2%)</td>
<td>761</td>
<td>−1.10</td>
<td>604</td>
<td>99 (16.4%)</td>
<td>62</td>
<td>−1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>4379</td>
<td>963 (22.0%)</td>
<td>589</td>
<td>−1.23</td>
<td>494</td>
<td>95 (19.2%)</td>
<td>50</td>
<td>−1.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Simulations with a continuous habitat. σ: species input rate. h: habitat preference width. h: habitat preference strength. S: average number of species. s(1): average number of species with only one individual. α: Fisher’s parameter of diversity. γ: slope of the first two orders of magnitude in abundance of the species abundance curve on a log–log plot.

Fig. 3. Species–abundance distributions. The x axis is the abundance n of a species, and the y axis is the average number of species with a given abundance. (a) Discrete habitats with h = 0. (b) Discrete habitats with h = 0.5. (c) Continuous habitat preference, with ν = 10⁻³.
extinction time. Our results show that many ecological indicators (total number of species, number of singletons, shape of the species abundance) attain equilibrium in a short time compared to the average extinction time. If we assume that our model describes trees, and that one generation (the average life span of an individual) is approximately 100 years, then 10,000 generations become 1 million years, a reasonable time in the history of tropical forests. Unlike extinction time, the time to reach equilibrium diversity and abundance does not depend on the size of the community.

We are not commenting on what maintains diversity at wider scales, arguing only that it may be irrelevant at the local scale. Species could be maintained by strong stabilizing forces at wider scales (Chesson 2000), or perhaps by a low rate of true speciation (Hubbell 2001). Other evidence is needed to test these possibilities. We only suggest that stabilizing forces are irrelevant at a small scale and that local diversity is controlled by regional diversity (Ricklefs 1987).

The idea that species input might affect community structure has been raised (among others) by Schmidt and Wilson (1985) and Pulliam (1998). Hanski and Gyllenberg (1993) and Loreau and Mouquet (1999) examined the abundance distribution in models with species input, but both omitted spatial explicitness at the individual level in the local community; both did demonstrate, though, a long tail of rare species. In Tilman’s (2004) model of stochastic niches, species input is considered, but Tilman dismissed its relevance. His model, though, is strongly biased in favor of abundant species already occupying their niches; we suggest that this is why it does not detect a major role of species input on diversity or abundance. The relative importance of the local and regional processes has been discussed by Zobel (1997) and Ricklefs (1987), and, more recently, by Leibold et al. (2004) and in the book of Holyoa et al. (2005).

Another explanation for rarity is the theory of “tourist species”: (Magurnan and Henderson 2003, Ulrich and Ollik 2004, Ulrich and Zalewski 2006) that consider the excess of rare species in communities relative to a log-normal to be due to species not typical of the community and with short persistence times. There certainly is a correlation between persistence times and abundances: species have short persistence times due to the stochastic dynamics. At every given moment a rare species will have on average a shorter extinction time than a common species, making it appear as a tourist species. But in our models and any model with stochastic births and deaths coupled with species input, there is no distinction between tourist and resident species.

There has been a tendency to confound neutrality with the species input process, but it appears to be simply a historical accident that the two ideas are linked. Speciation obviously has nothing to do with neutrality: a non-neutral model can include species input, while a neutral model need not. The latter is generally considered uninteresting, but we suggest that the former is. Tropical rainforests, marine abyssal communities, and coral reefs (Karlson et al. 2004, Rex et al. 2005) appear to be dominated by input from a highly diverse metacommunity. On the other hand, low-diversity communities of vertebrates are most likely not.

We conclude that precise predictions for the abundance distribution and species richness are possible in communities dominated by species input. The neutral model as quantified by Hubbell (2001) and Volkov et al. (2003) works because it accurately treats the dominant force.

Acknowledgements – The National Center for Ecological Analysis and Synthesis provided us space and computer facilities for the analysis, and sabbatical support to RC. The Center for Tropical Forest Science supported TZ to carry out this work. The Smithsonian Tropical Research Institute supports the Barro Colorado Island research site, and we thank Robert Stallard for a copy of his digital terrain map of the island.

References
Gause, G. 1934. The struggle for existence. – William & Wilkins.


