Biodiversity on island chains: Neutral model simulations

Patrick B. Warren
Unilever R&D Port Sunlight, Bebington, Wirral CH63 3JW, United Kingdom

(Received 26 April 2010; revised manuscript received 24 September 2010; published 17 November 2010)

A neutral ecology model is simulated on an island chain, in which neighboring islands can exchange individuals but only the first island is able to receive immigrants from a metacommunity. It is found by several measures that α-diversity decreases along the chain. Subtle changes in taxon abundance distributions can be detected when islands in the chain are compared to diversity-matched single islands. The island chain is found to have unexpectedly rich dynamics. Significant β-diversity correlations are found between islands in the chain, which are absent between diversity-matched single islands. The results potentially apply to human microbial biodiversity and biogeography and suggest that measurements of interindividual and intraindividual β-diversity may give insights into microbial community assembly mechanisms.

DOI: 10.1103/PhysRevE.82.051922 PACS number(s): 87.23.−n, 87.10.Mn, 02.50.Ga

I. INTRODUCTION

It has recently been observed that human microbial biodiversity varies systematically between body sites [1,2], for example, phylogenetic diversity is higher for the palm of the hand and the sole of the foot than for the armpit and forehead (Fig. S14 in Ref. [2]). A high degree of interindividual variability is also observed to the point where the characterization of skin bacteria residues has even been proposed as a forensic tool [3]. The latter observation in particular supports the notion that *stochastic community assembly* may play a significant role in determining human microbial biodiversity and biogeography. Stochastic assembly is a key element of Hubbell's unified neutral model of biodiversity and biogeography [4], therefore the question arises as to whether neutral models can be applied to the human microbiome [5,6]. This is a hard problem and I do not claim to have solved it here, but it motivates the present study of community assembly mechanisms in the context of neutral theory.

The merits of neutral models have been debated extensively [7], and it is far from obvious that they should apply to microbial systems. However Hubbell's neutral model has recently been successfully applied to predict microbial diversity in tree holes [8,9]. Relevant to the present problem I also note that the dominant factor determining taxon abundancies has been argued to be the community assembly mechanism [10,11]. Other neutral model ideas such as the zero-sum constraint (single trophic level; community saturation) or speciation by point mutation seem to play a lesser role. If taxon abundancies are largely determined by community assembly therefore, a couple of limiting models (Fig. 1) present themselves to explain the observed variations in human microbial diversity. The first is a *variable immigration rate* model in which different body sites are envisaged as being microbial "islands" in contact with a microbial metacommunity but effectively isolated from each other. Here the variation in diversity corresponds to a variable immigration rate from the metacommunity. The second is an *island chain* model in which it is envisaged that migration can take place between islands but, *in extremis*, it is only the first island (e.g., the hand) that receives immigrants from the metacommunity. In this case one expects that diversity should decrease as one moves along the chain away from the island in contact with the metacommunity. This is confirmed in the present study.

Of course these models represent limiting cases and reality probably lies somewhere in between. A second question therefore is how one might distinguish between stochastic community assembly mechanisms. It is in this context that the present study may perhaps be most valuable. I find that both models lead to rather similar abundance distributions (α-diversity). This extends to both static and dynamic diversity characterizations. When this is conflated with other factors, such as deviations from the neutral model assumptions, it is probably going to be difficult to distinguish between assembly mechanisms on the basis of single-site measurements of microbial diversity. However I find that the diversity correlations between sites (β-diversity) are much more significant for the island chain model than for the variable immigration rate model. This raises the possibility that characterization of β-diversity might give insights into community assembly mechanisms.

Let me turn now to the details of the present study. The neutral model has been extensively analyzed for isolated islands in contact with a metacommunity [4,10,12–20]. Only for certain cases though has the neutral model been solved for multiple islands or "patches," which are able to exchange individuals [21–23]. In particular, the island chain problem has not been solved (i.e., where individuals can migrate between neighboring islands but immigration is restricted to

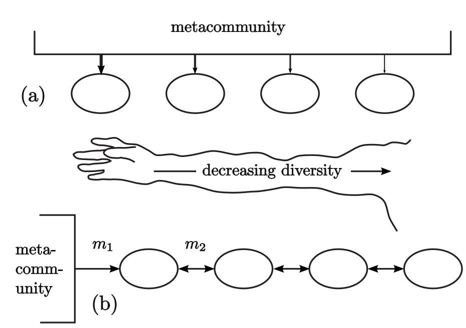

FIG. 1. Candidate explanations for a variation in human microbial diversity based on stochastic community assembly and the theory of island biogeography: (a) variable immigration rate model and (b) island chain model.

the first island in the chain). This problem can be described as an inhomogeneous linear "stepping stone" model. Homogeneous linear stepping stone models have been surveyed recently by Korolev *et al.* [24], and inhomogeneous models have been touched on by Powell and McKane [25]. The primary goal of the present study is to analyze this model, using the existing body of work on isolated islands as context. Although in principle one can approach this analytically, the experience of Vallade and Houchmandzadeh [18] for *two* islands suggests this may rapidly become effectively unmanageable. I therefore approach the problem by means of simulations.

II. DIVERSITY MEASURES

Let me begin with an overview of the mathematical characterization of diversity. The starting point is the partitioning of individuals in one or more populations into taxonomic groups (e.g., species). Suppose there are K taxa and N_i individuals in the ith taxon ($i=1,\ldots,K$), in a population of $J=\Sigma_{i=1}^{K}N_i$ individuals. The relative abundance of the ith taxon is defined to be $\omega_i=N_i/J$. The *taxon abundance distribution* is given by ϕ_k, the number of taxa containing k individuals. Given the set of N_i one can easily calculate ϕ_k, and one has $K=\Sigma_{k=1}^{\infty}\phi_k$ and $J=\Sigma_{k=1}^{\infty}k\phi_k$. Since no taxon can contain more individuals than there are in the community as a whole, $\phi_k=0$ for $k>J$. Similarly $\phi_J=1$ if and only if all the individuals belong to the same taxon (the "monodominated" state), otherwise $\phi_J=0$.

In standard neutral model dynamics, population sizes remain fixed while the number of taxa and the number of individuals per taxon fluctuates. I adopt the notation of Vallade and Houchmandzadeh [12,18] and write $\langle\cdots\rangle$ to indicate a quantity averaged over an ensemble of populations undergoing neutral model dynamics.

The information contained in $\langle\phi_k\rangle$ can also be represented by giving the ensemble-average probability $\langle p(\omega)\rangle$ that an individual belongs to a taxon of relative abundance ω [12,18,26]. For a community of a finite size, $\langle p(\omega)\rangle$ is a discrete array or "comb" of δ functions, even after ensemble-averaging, since ω can only take on discrete multiples of $1/J$. However as $J\to\infty$, $\langle p(\omega)\rangle$ becomes a continuous function. One can show that the continuum limit is $\langle p(\omega)\rangle=\lim_{J\to\infty}k\langle\phi_k\rangle$, where $k=\omega J$ [12].

The principal order parameter I shall use to measure α-diversity is the Simpson index [27], defined for a given set of taxon abundancies to be

$$D=1-\sum_{i=1}^{K}\omega_i^2. \tag{1}$$

It is related to the second moment of the taxon abundance distribution by

$$D=1-J^{-2}\sum_{k=1}^{\infty}k^2\phi_k. \tag{2}$$

From this it can easily be shown that in the continuum limit

$$\langle D\rangle=1-\int_0^1\omega\langle p(\omega)\rangle d\omega. \tag{3}$$

From the point of view of statistical physics the inclusion of "1−" in these definitions is "window dressing," designed to make D an increasing function of diversity. However we shall see below that there is also a technical advantage. The index satisfies $0\le\langle D\rangle\le1-1/K$. It takes the lower limit in the monodominated low diversity state, and the upper limit in the high diversity state where all individuals are uniformly distributed between taxa ($\omega_i=1/K$). There is a mild disadvantage in using $\langle D\rangle$ in that it loses sensitivity to the underlying abundance distribution as it approaches the limiting values.

In addition to the Simpson index, I shall occasionally use the ensemble-average number of taxa $\langle K\rangle=\Sigma_{k=1}^{J}\langle\phi_k\rangle$, and the ensemble-average monodominance probability $\langle\phi_J\rangle$ (as explained above ϕ_J is 1 or 0 according to whether or not all the individuals belong to the same taxon). As an order parameter, the Simpson index $\langle D\rangle$ has an important advantage over $\langle K\rangle$ and $\langle\phi_J\rangle$ in that it remains well defined in the continuum limit $J\to\infty$. Moreover there are some particularly simple theoretical expressions for $\langle D\rangle$ under neutral model dynamics.

The Simpson index generalizes naturally to spatial and temporal correlations. For the first of these, given two or more populations, I introduce the β-diversity or codominance index [22,23]

$$D^{(rs)}=1-\sum_{i=1}^{K}\omega_i^{(r)}\omega_i^{(s)}, \tag{4}$$

where $\omega_i^{(r)}$ is the relative abundance of the ith taxon in the rth population. If two populations have no species in common, then $D^{(rs)}=1$. The codominance index reduces to the Simpson index if $r=s$.

For time series data, the natural generalization of the Simpson index is the autocorrelation function [28]

$$D(t_0,t)=1-\sum_{i=1}^{K}\omega_i(t_0)\omega_i(t_0+t), \tag{5}$$

where $\omega_i(t)$ is the relative abundance at time t. In a steady-state situation the dependence on the initial time disappears and $D(t_0,t)=D(t)$; and $D(t)\to D$ as $t\to0$.

III. NEUTRAL THEORY

In this Sec. I describe neutral model dynamics in more detail. The overarching picture is that shown in Fig. 1. One considers one or more island populations, which are able to exchange individuals with each other and with a "metacommunity" which acts as a reservoir of biodiversity. The taxon distribution in the metacommunity slowly turns over due to speciation events.

In Hubbell's unified neutral model [4], the dynamics of the metacommunity and of the islands are closely related. In the metacommunity the dynamics is as follows. An individual is selected at random, and with probability $1-\nu$ is replaced with a copy of another individual drawn at random from the metacommunity or with probability ν is replaced by an individual belonging to a new taxon. Thus ν is a (point) speciation rate. For $\nu=0$ the metacommunity eventually falls into a monodominated state, in an ecological analog of the Matthew principle [29]. For $\nu>0$ the taxon abundance distribution is a balance between speciation and extinction.

An explicit expression for the taxon abundance distribution in a metacommunity of size J_M has been obtained by a number of workers [4,10,12,14,15,17]. Results are quoted as *metacommunity* (subscript "M") steady-state ensemble-averages:

$$\langle \phi_k \rangle_M = \frac{\theta\Gamma(J_M+1)\Gamma(J_M+\theta-k)}{k\Gamma(J_M+1-k)\Gamma(J_M+\theta)}, \tag{6}$$

where $\theta=(J_M-1)\nu/(1-\nu)$ is known as Hubbell's fundamental biodiversity parameter. In the usual limit $J_M \gg 1$ and $\nu \ll 1$ one has $\theta \approx J_M \nu$. The continuum limit ($J_M \to \infty$) of Eq. (6) is $\langle p(\omega)\rangle_M = \theta(1-\omega)^{\theta-1}$. Inserting Eq. (6) into Eq. (2) gives $\langle D\rangle_M = (1-J_M^{-1})\theta/(\theta+1)$. It follows that in the continuum limit

$$\langle D \rangle_M = \theta/(\theta+1). \tag{7}$$

This can also be obtained by inserting $\langle p(\omega)\rangle_M$ into Eq. (3). It was first noted by He and Hu by analogy to a similar problem in population genetics [30].

Neutral model dynamics on an island connected to the metacommunity is as follows. An individual is selected at random, and with probability $1-m$ is replaced with a copy of another individual drawn at random from the island, or with probability m is replaced by an individual drawn at random from the metacommunity. Thus m is the immigration rate (since the metacommunity is much larger than the island, emigration has essentially no effect and need not be considered). Similar to the metacommunity, the island community eventually falls into a monodominated state if $m=0$, whereas for $m>0$ a steady-state taxon abundance distribution arises as a balance between immigration and extinction.

Exact results for the island taxon abundance distribution were obtained only recently [10,17] although partial results were obtained by previous authors [12–14]. The result is

$$\langle \phi_k \rangle = \binom{J}{k} \int_0^1 \frac{du}{u} \theta(1-u)^{\theta-1} \frac{(\mu u)_k [\mu(1-u)]_{J-k}}{(\mu)_J}, \tag{8}$$

where $(x)_n = \Gamma(x+n)/\Gamma(x)$ is the Pochhammer symbol, $\binom{J}{k} = \Gamma(J+1)/[\Gamma(k+1)\Gamma(J-k+1)]$ is the binomial coefficient, J is the island size, and $\mu = m(J-1)/(1-m)$ plays a role similar to θ for the metacommunity (note that in the ecology literature the parameter "μ" is often referred to as "I"). For $J \gg 1$ and $m \ll 1$ one has $\mu \approx Jm$. Note that J_M does not feature in this expression, in other words the island abundance distribution is insensitive to the metacommunity size. This point is discussed more thoroughly by Vallade and Houchmandzadeh [18]. The continuum limit of Eq. (8) is [10,12,17,31]

$$\langle p(\omega) \rangle = \mu\theta \int_0^1 du \binom{\mu}{\mu u} (1-\omega)^{\mu u-1} \omega^{\mu(1-u)} u^\theta. \tag{9}$$

Inserting Eq. (8) into Eq. (2) gives

$$\langle D \rangle = \frac{(1-J^{-1})\mu\theta}{(\mu+1)(\theta+1)}. \tag{10}$$

For $J \to \infty$ this reduces to a simple result previously reported by Etienne [16], which can also be found by inserting Eq. (9) into Eq. (3), namely,

$$\langle D \rangle = \frac{\mu\theta}{(\mu+1)(\theta+1)}. \tag{11}$$

We see the technical advantage of the "1−" in Eqs. (1)–(3) is that the diversity index factorizes into the product of the metacommunity diversity index $\langle D\rangle_M = \theta/(\theta+1)$ and a similar-looking island factor $\mu/(\mu+1)$.

Let me briefly mention that the ensemble-average number of taxa $\langle K\rangle = \sum_{k=1}^J \langle \phi_k\rangle$ can obviously be obtained from Eq. (8) but has to be evaluated numerically. For the monodominance probability, Eq. (8) simplifies to

$$\langle \phi_J \rangle = \int_0^1 \frac{du}{u} \theta(1-u)^{\theta-1} \frac{(\mu u)_J}{(\mu)_J}. \tag{12}$$

This also has to be evaluated numerically, although one can show that it depends strongly on all the relevant parameters and vanishes asymptotically for $J \to \infty$ at fixed μ as $\langle \phi_J\rangle \sim \theta\Gamma(\theta)(\mu \ln J)^{-\theta}$.

In the present work I assume that the (fast) island dynamics are decoupled from the (slow) metacommunity dynamics; in other words the metacommunity can be taken to have a "frozen" abundance distribution. This is because the metacommunity population turns over on a time scale of order $1/\nu$, whereas the island populations turn over on a time scale of order $1/m$, and typically one expects $\nu \ll m$ (i.e., the speciation rate is very much smaller than the immigration rate).

With this assumption, some results can be obtained for the generalized order parameters in Eqs. (4) and (5). First consider either two independent islands or one island at two well-separated points in time ($t \gg 1/m$). For a given metacommunity abundance distribution, one expects $\langle \omega_i^{(1)}\omega_i^{(2)}\rangle = \langle \omega_i^{(1)}\rangle\langle \omega_i^{(2)}\rangle$ where the ensemble averages are taken with respect to a frozen metacommunity abundance distribution. From the theory in, e.g., Ref. [14], it is known that the steady-state expectation value $\langle \omega_i\rangle$ is equal to the relative abundance of the ith taxon in the metacommunity. Hence $\langle \omega_i^{(1)}\rangle\langle \omega_i^{(2)}\rangle = [\omega_i^{(M)}]^2$. Averaging over metacommunity abundance distributions yields finally that $\langle D^{(rs)}\rangle = \langle D\rangle_M$ for $r \neq s$, *independent* of the island sizes or immigration rates; and $\langle D(t)\rangle \to \langle D\rangle_M$ as $t \to \infty$ (more precisely, for $1/m \ll t \ll 1/\nu$).

The first result implies that the codominance index for the variable immigration rate model in Fig. 1(a) is the same for all pairs of islands and is equal to the metacommunity Simpson index. This is the first hint that β-diversity might be a useful discriminant between community assembly mechanisms.

The second result, together with the additional insight that m is the sole relaxation time governing single island dynamics [14], means that a complete solution for the single-island autocorrelation function can be written down. It is

$$\langle D(t)\rangle = \langle D\rangle_M + (\langle D\rangle - \langle D\rangle_M)e^{-mt} = \frac{\theta}{\theta+1}\left(1 - \frac{e^{-mt}}{\mu+1}\right). \tag{13}$$

Note that this function contains enough information to be able to deduce in principle all of θ, μ, and m (and hence $J \approx \mu/m$), *without* having to make any measurements on the metacommunity.

Let me turn now to a brief discussion of the simulation methodology. For the simulations I generate a large number (10^4–10^5) of frozen steady-state metacommunity abundance distribution samples for given J_M and θ. I use these samples in the subsequent island and island chain simulations. I take $\theta = 10$ as a reference point motivated by Woodcock *et al.* [9] and $J_M = 10^5$ motivated by the requirement that $J_M \gg J \gg 1$ [18].

In reality, microbial population sizes are expected to be much bigger than the present values of J and J_M, hence the focus of attention is on quantities such as $\langle D\rangle$ and $\langle p(\omega)\rangle$ which are well defined in the continuum limit. Comparing Eq. (10) with Eq. (11) suggests that finite size corrections are of the order $1/J$. This is borne out by numerical investigation of $\langle p(\omega)\rangle$ and $k\langle \phi_k\rangle$. I find that the sampling errors are typically $\lesssim 1\%$ so it does not make much sense to increase J and J_M beyond the above values without additionally increasing the number of samples although clearly it does make sense to check for finite size effects from time to time.

Simulation of neutral model dynamics in the forward time direction is very straightforward, however a neat trick to obtain steady-state samples is contained in the "coalescence" algorithms which propagate backward in time. These algorithms are described by Mckane *et al.* [14], Etienne [16], and Rosindell *et al.* [32]. For the island chain simulations a hybrid of the Etienne and Rosindell *et al.* algorithms was developed. Typically, I use a coalescence algorithm to generate steady-state samples and forward-time simulations for dynamics such as the autocorrelation functions discussed below. I undertook a number of single island simulations to build confidence in the simulation and analysis methodologies. I find excellent agreement between these simulations and the theoretical predictions.

IV. ISLAND CHAINS

I now turn to the key question in the present study, namely, the behavior of the island chain model in Fig. 1(b). The island chain simulations are performed similarly to the single island simulations. I introduce an immigration rate m_1 for the first island, and an interisland migration rate m_2. Specifically, the dynamics are as follows. An individual is selected at random. If the chosen individual lies on the first island, it is replaced with a copy of another individual on the island with probability $1-m_1-m_2$, an immigrant from the metacommunity with probability m_1, or a migrant from the neighboring island with probability m_2. If the chosen individual lies on an island interior to the chain, it is replaced with a copy of another individual on the island with probability $1-2m_2$ or with a migrant from one of the neighboring islands (selected at random) with probability $2m_2$. If the chosen individual lies on the last island, it is replaced with a copy of another individual on the island with probability $1-m_2$ or with a migrant from neighboring island with probability m_2. Migrants are copies of individuals chosen at random on neighboring islands. The immigration and migration probabilities are constrained by $0 < m_1 + m_2 < 1$ and $0 < 2m_2 < 1$.

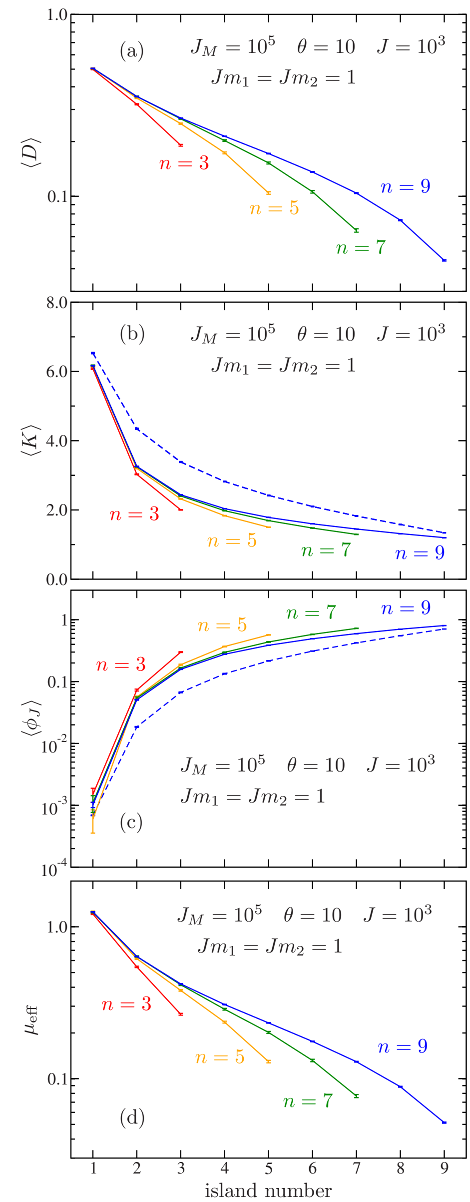

FIG. 2. (Color online) Aspects of diversity on island chains of length n: (a) diversity index $\langle D\rangle$, (b) number of taxa $\langle K\rangle$, (c) monodominance probability $\langle \phi_J\rangle$, and (d) $\mu_{\rm eff}$ for equivalent single islands calculated from Eq. (14). The dashed lines in (b) and (c) are the $\langle D\rangle$-matched single island results for $n=9$.

A. Steady-state α-diversity

Representative steady-state results are shown in Figs. 2–5, for island chains of varying lengths n. The first conclusion (Fig. 2) is that diversity decreases, by whatever measure, as one moves away from the island in contact with the meta-

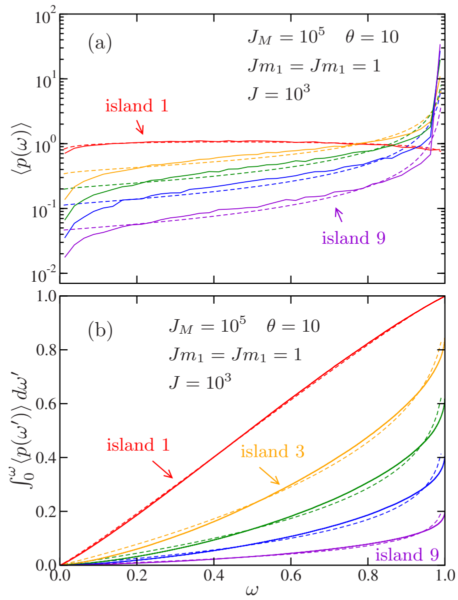

FIG. 3. (Color online) Abundance distributions for the odd-numbered islands in an island chain of length $n=9$: (a) the probability $\langle p(\omega) \rangle$ that a randomly selected individual belongs to a taxon of relative abundance ω and (b) the cumulative distribution function of the same. The results (solid lines) are compared to the theoretical expectations for $\langle D \rangle$-matched single islands calculated from Eq. (9) (dashed lines).

community. Hence the island chain model is a viable candidate to explain the diversity variations noted for the human microbiome. The second conclusion is that away from the last island (i.e., away from the island farthest from the metacommunity) there is a degree of convergence of the behavior.

I next use Eq. (11) to define an effective or $\langle D \rangle$-matched single island parameter

$$\mu_{\text{eff}} \equiv \langle D \rangle / (\langle D \rangle_M - \langle D \rangle), \tag{14}$$

where $\langle D \rangle_M = \theta/(\theta+1)$. This is shown in Fig. 2(d). This quantity can be used to calculate $\langle D \rangle$-matched single island values of $\langle K \rangle = \sum_{k=1}^{J} \langle \phi_k \rangle$ and $\langle \phi_J \rangle$ from Eqs. (8) and (12), respectively, assuming the effective island size is $J_{\text{eff}}=J$. For $n=9$ these are shown as the dashed lines in Figs. 2(b) and

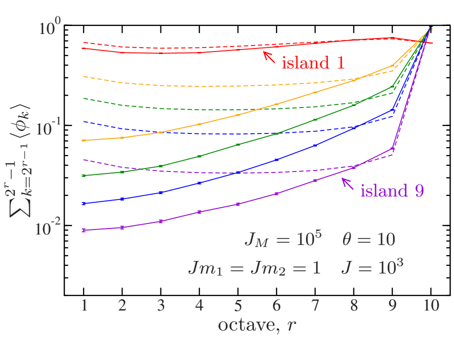

FIG. 4. (Color online) Preston plots of $\langle \phi_k \rangle$, binned by octave, for odd-numbered islands in an island chain of length $n=9$. The results (solid lines) are compared to the theoretical expectations for $\langle D \rangle$-matched single islands calculated from Eq. (8) (dashed lines).

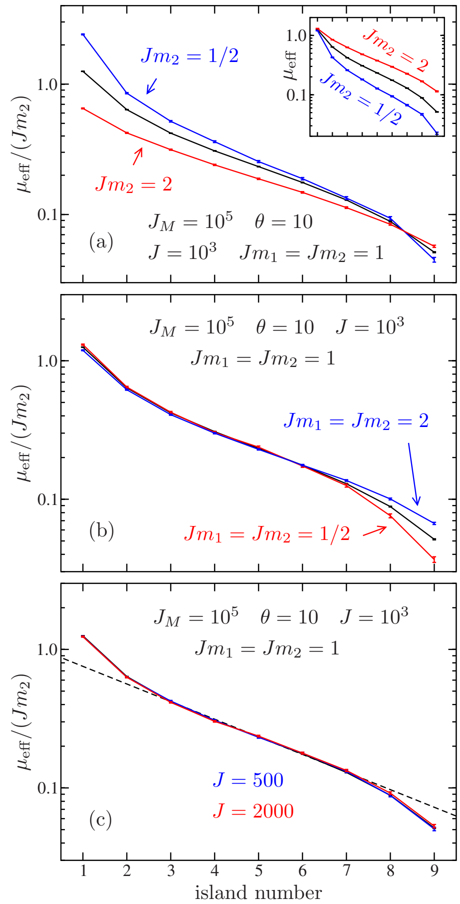

FIG. 5. (Color online) Diversity variation along a chain of $n=9$ islands as measured by μ_{eff} normalized by Jm_2: (a) varying m_2 only at fixed J (the inset shows the same data without the normalization), (b) varying $m_1=m_2$ at fixed J, and (c) varying $m_1=m_2=1/J$. The dashed line in (c) is $\mu_{\text{eff}}=e^{-0.3n}$.

2(c). It is clear that, systematically, $\langle K \rangle$ is reduced and $\langle \phi_J \rangle$ is increased when an island in the island chain is compared with its $\langle D \rangle$-matched single island counterpart. Thus on the island chain there is a tendency toward *fewer* and *larger* taxa.

I will show below that the effective immigration rate, in so far as it can be defined, is not equal to μ_{eff}/J. In the light of this it is perhaps questionable to rely on $J_{\text{eff}}=J$. I therefore computed $\langle p(\omega) \rangle$ for island chains of length $n=9$ (this requires typically 10^5 samples) and compared to $\langle D \rangle$-matched single islands calculated using Eq. (9). This does not require J_{eff}. The comparison is shown in Fig. 3. It is clear that there is a subtle and nontrivial redistribution of the taxon abundancies: $\langle p(\omega) \rangle$ is reduced for $\omega \lesssim 0.2$ and $\omega \gtrsim 0.8$, but increased for $0.2 \lesssim \omega \lesssim 0.8$. This means that the number of taxa with intermediate abundancies is increased at the expense of the very rare taxa and the high abundance taxa. But, in addition, the cumulative distribution function jumps up at $\omega=1$, as shown clearly in Fig. 3(b). This corresponds to the increased monodominance probability. At first sight this is at odds with the redistribution toward midrange abundancies, nevertheless it is a real effect and indeed is the reason why monodomination was separately studied. For completeness and to make contact with more conventional ecological studies, Preston plots corresponding to this data are shown in Fig. 4. The reduction in the number of rare taxa is very clear in this representation but the subtleties for the more common taxa are lost by the logarithmic binning procedure.

The loss of the very rare taxa can perhaps be attributed to the filtering properties of the island chain. These taxa are already rare in the metacommunity and it could simply be that a representative from a rare taxon is less likely to arrive via migration along an island chain than via direct immigration from the metacommunity (at matched $\langle D\rangle$). Note that Eq. (3) implies the mean value $\int_0^1 d\omega\omega\langle p(\omega)\rangle$ is fixed under $\langle D\rangle$ matching. This means that the redistribution of abundancies cannot be asymmetric and explains why there has to be a concomitant reduction in the high abundance taxa. The increased monodominance probability is more mysterious and I do not at present have a clear mechanistic explanation. Possibly what is happening is that the monodominated state ($\omega=1$) has become "stickier" in dynamical terms, without actually becoming an adsorbing state. This is partially confirmed by the β-diversity analysis below.

Returning now to the $\langle D\rangle$-matched single islands, Fig. 5 shows how μ_{eff} varies with immigration and migration rates m_1 and m_2 and the island size J. Figure 5(a) shows that increasing the interisland migration rate m_2 has the effect of increasing the diversity along the island chain apart from the first island (inset). The effect of additionally increasing the metacommunity immigration rate (so that $m_1=m_2=m$) is shown in Fig. 5(b). This leads to increased diversity along the whole chain and, at least away from the last couple of islands, one has quite accurately $\mu_{\text{eff}}\propto Jm$. Figure 5(c) shows the effect of varying $m_1=m_2=m$ and J simultaneously, keeping Jm fixed. The highly accurate data collapse strongly supports the notion that the island diversity is governed by the combinations Jm_1 and Jm_2 rather than the individual values of J, m_1, and m_2. This is in close analogy to the theory for the single island and confirms that finite-size corrections are negligible. I also investigated varying θ and found a similarly highly accurate data collapse to that shown in Fig. 5(c) so that it appears the effect of the Hubbell parameter is removed when $\langle D\rangle$ is converted to μ_{eff}. Finally, I found essentially no influence of J_M on the results within the measured accuracy.

The semilogarithmic plots in Fig. 5 show that μ_{eff} for the intermediate islands is quite accurately represented by a geometric progression. To summarize, at least for $m_1=m_2=m$ and intermediate islands, all the data can be quite accurately represented by

$$\langle D\rangle=\frac{\theta\mu_{\text{eff}}}{(\theta+1)(\mu_{\text{eff}}+1)},\quad \mu_{\text{eff}}\approx Jm\times A^r, \tag{15}$$

where $A=0.75\pm0.01$ and $r=1,\ldots,n$ labels the islands.

For $\langle D\rangle\lesssim0.5$ the abundance distribution for a $\langle D\rangle$-matched metacommunity is almost exactly the same as that for a $\langle D\rangle$-matched single island. By this I mean that $\langle p(\omega)\rangle_M=\theta(1-\omega)^{\theta-1}$ with $\theta=\langle D\rangle/(1-\langle D\rangle)$ is a very good approximation to $\langle p(\omega)\rangle$ from Eq. (9). However a complete comparison with an equivalent metacommunity is frustrated by the residual dependence of $\langle K\rangle_M$ and $\langle\phi_J\rangle_M$ on the metacommunity size J_M.

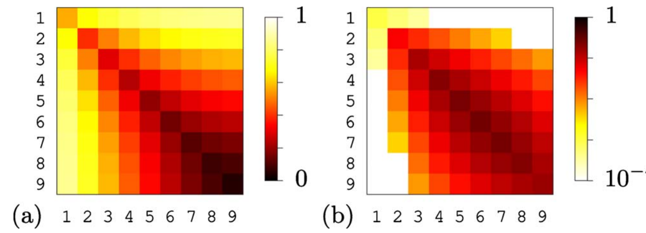

FIG. 6. (Color online) Heat maps of two measures of diversity correlations between pairs of islands in a chain of length $n=9$: (a) the normalized codominance index $\langle D^{(rs)}\rangle/\langle D\rangle_M$ and (b) the comonodominance correlation function defined in Eq. (16) (note the logarithmic color scale). Other parameters are as for Fig. 3. In (b) a cell is colored white if there are 10 or fewer instances that are comonodominated out of 10^5 samples.

B. Steady-state β-diversity

Figure 6(a) shows a heat map of the codominance index described in Eq. (4) for pairs of islands in an island chain of length $n=9$ for the parameter set used for Fig. 3. Note that the diagonal values in such a representation reduce to the Simpson indices of individual islands. A corresponding heat map for the analogous set of individual islands would therefore have the same colors along the diagonal as a consequence of the $\langle D\rangle$ matching, but the off-diagonal cells would be completely white. This is because, as already described, the codominance index between pairs of islands which are isolated from each other is $\langle D\rangle_M$. (I have checked this in simulations.)

Figure 6(b) shows a heat map for a comonodominance correlation function defined by

$$\langle\phi_J^{(r)}\phi_J^{(s)}\rangle-\langle\phi_J^{(r)}\rangle\langle\phi_J^{(s)}\rangle. \tag{16}$$

In the absence of correlations the off-diagonal components of the heat map would be completely white (again I have checked this in simulations). The diagonal elements reduce to $\langle\phi_J\rangle(1-\langle\phi_J\rangle)$ (note that this is nonmonotonic in $\langle\phi_J\rangle$). The comonodominance correlations lend support to the notion mentioned above that monodominance is dynamically "stickier" for island chains: if an island in a chain is in the monodominated state, the neighboring islands are more likely to be in a monodominated state than would be expected from pure chance.

C. Dynamics

Finally, I embarked on a preliminary study of the dynamics of the island chain, again using the parameter set for Fig. 3, although in this case I also examined the effect of varying the chain length n. The dynamics is unexpectedly feature rich. My starting point is to determine the autocorrelation functions defined in Eq. (5) for the individual islands in the chain. In reporting the results, the time t is defined to be the number of replacement steps divided by the total island chain population size nJ. This is appropriate to the interpretation of neutral model dynamics as a continuous-time Markov process.

First I observe that at long times one always has $\langle D(t)\rangle\to\langle D\rangle_M$ (while there is a proof of this for single islands, I do

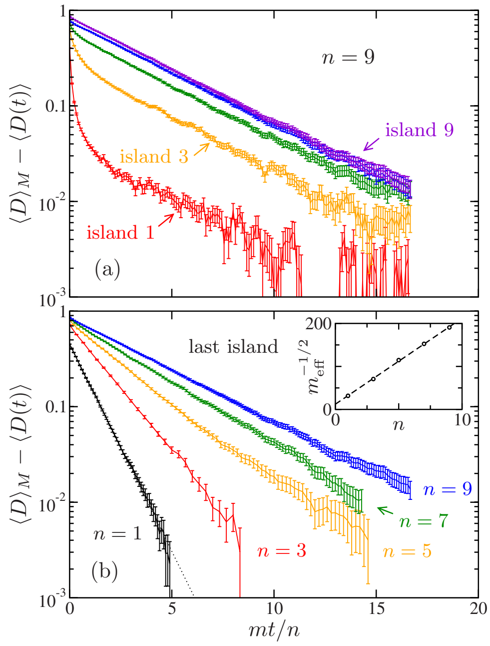

FIG. 7. (Color online) Autocorrelation functions on island chains: (a) for the odd-numbered islands in a chain of length $n=9$ and (b) for the *last* island in chains of varying length n. A factor $1/n$ is included in the time scale for convenience. The dotted line in (b) is Eq. (13) for a single island. The inset in (b) shows $m_{\rm eff}^{-1/2}$ as a function of n, where $m_{\rm eff}$ is determined by fitting the data in the main plot to an exponential decay; see also Table I. Other parameters are as for Fig. 3.

not know of a more general proof). Figure 7 shows $\langle D\rangle_M - \langle D(t)\rangle = \Sigma_{i=1}^{K}\langle\Delta\omega_i(0)\Delta\omega_i(t)\rangle$, plotted semilogarithmically as a function of time. Generally, I find the behavior shown in Fig. 7(a) is repeated for all chain lengths, namely: the first island shows quite a strong nonexponential behavior; the last island is quite accurately monoexponential from early times; the intermediate islands show intermediate behavior; and all the islands ultimately decay with a common "terminal" relaxation time which I shall define to be $1/m_{\rm eff}$. This relaxation time can most easily be found by fitting an exponential to the autocorrelation function for the last island. This is shown in Fig. 7(b) and the results are collected in Table I.

Island hopping can be viewed as a diffusion-like process and one might expect $m_{\rm eff}\sim n^{-2}$. In fact, as is common for diffusion problems, the inset to Fig. 7(b) shows that $m_{\rm eff}$ can be accurately represented by an *offset* diffusion law, $m_{\rm eff} = A(n+n_0)^{-2}$ where $A=(2.5\pm 0.1)\times 10^{-3}$ and $n_0=0.6\pm 0.1$.

Table I extends the $\langle D\rangle$-matching procedure to obtain an estimate of the effective island size $J_{\rm eff}$ for the last island in chains of varying length n. This is done by first calculating $\mu_{\rm eff}$ from the Simpson index $\langle D\rangle$, and then using $m_{\rm eff}$ determined from the decay of the autocorrelation function to compute $J_{\rm eff}=\mu_{\rm eff}/m_{\rm eff}$. One sees that the effective island size is a weakly increasing function of n (in particular it is clearly not proportional to the total population of the islands nJ). Hence most of the reduction in the diversity can be attributed to a reduced effective immigration rate.

Figure 8 shows the behavior of the first island in chains of variable length in some more detail. Observe that the autocorrelation functions initially follow a universal decay curve, only to fall off this as they encounter the n-dependent terminal relaxation time $1/m_{\rm eff}$. For $t\gtrsim 1$ the universal behavior appears to follow a power law with an exponent -0.7 ± 0.1 [note that this cannot be continued to $t=0$ since $\langle D\rangle_M - \langle D(t)\rangle$ must remain bounded].

TABLE I. Properties of the last island in an island chain of length n: the Simpson index $\langle D\rangle$, the effective single island $\mu_{\rm eff}$ calculated from Eq. (14), the relaxation rate $m_{\rm eff}$ determined from Fig. 7(b), and the effective island size $J_{\rm eff}=\mu_{\rm eff}/m_{\rm eff}$. The first row are the exact results for a single island ($n=1$). Other parameters are as for Fig. 3. A figure in brackets is an estimate of the error in the final digit.

n	$\langle D\rangle$	$\mu_{\rm eff}$	$m_{\rm eff}\times 10^3$	$J_{\rm eff}\times 10^{-3}$
1	$\equiv 5/11$	$\equiv 1$	$\equiv 1$	$\equiv 1$
3	0.191(2)	0.265(4)	0.200(5)	1.33(7)
5	0.104(2)	0.130(2)	0.076(2)	1.7(1)
7	0.065(2)	0.077(2)	0.043(1)	1.8(1)
9	0.0445(5)	0.0515(5)	0.0275(5)	1.90(5)

I also examined briefly the approach to the steady state for islands in an island chain. This is illustrated in Fig. 9. I find the time scale is more or less set by the terminal relaxation time $1/m_{\rm eff}$ although one can see that aspects of the behavior are distinctly nonexponential, for example, the early-time recovery from the monodominated state for $n=9$. I expect that this corresponds to a spectrum of relaxation modes, which are excited differently according to the initialization protocol and which are subsequently mixed up by the nonlinear dynamics. A more detailed exploration of this is left for future work.

According to the offset diffusion law noted above, for $n\gtrsim 4.4$ the assumed time scale separation between $1/m_{\rm eff}$ and the notional metacommunity relaxation time $1/\nu\approx J_M/\theta\approx 10^4$ no longer holds. Nevertheless the assumption of a frozen metacommunity abundance distribution is still valid since in reality the actual speciation rate is much smaller than the notional speciation rate.

V. DISCUSSION

From the point of view of statistical physics, the present study leaves open a large number of questions. For example,

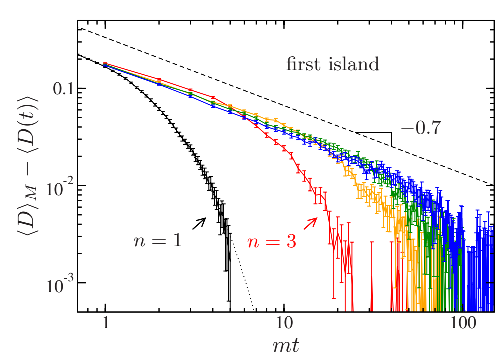

FIG. 8. (Color online) Autocorrelation functions for the first island in island chains of varying length n. The dotted line is Eq. (13) for a single island. Other parameters are as for Fig. 3.

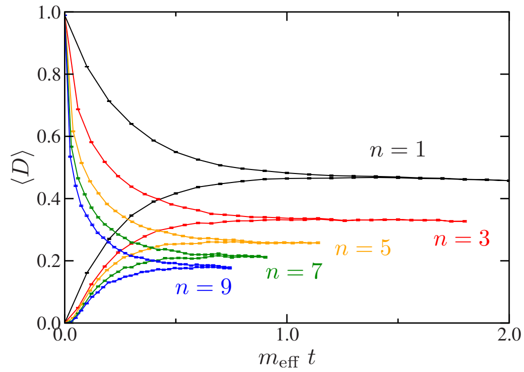

FIG. 9. (Color online) Approach to steady state for the central island of a chain of length n islands starting from either a completely monodominated state (lower curves) or a flat abundance distribution (upper curves; flat means the N_i are equalized subject to $\Sigma_{i=1}^{K} N_i = J$). The $n=1$ case is for a single island. Time is scaled by the terminal relaxation time $1/m_{\rm eff}$ (see Table I). Other parameters are as for Fig. 3.

the rich dynamics (e.g., power law behavior) begs for deeper theoretical understanding. There is clearly plenty of scope for further work, but here let me address the question which motivated the present study, namely, the possibility of discriminating between stochastic community assembly mechanisms such as those illustrated in Fig. 1. First, it is clear that as far as the steady-state abundance distributions on individual islands are concerned, an island chain can be quite well represented by a collection of $\langle D \rangle$-matched single islands (Fig. 3). Hence it appears that it will be difficult to distinguish between assembly mechanisms on the basis of static measurements of α-diversity.

On the other hand, as Figs. 7–9 show, the dynamics is very rich. There is a nontrivial spectrum of relaxation modes. Sufficiently refined measurements ought to be able to distinguish between this, and the dynamics of collections of isolated islands. But note that broadly speaking the overall time scale seems to be set by an effective immigration rate (Table I). Hence it seems difficult to use time-dependent α-diversity measurements to distinguish between assembly mechanisms.

The most promising avenue seems to lie in the β-diversity. My simulations indicate that significant correlations emerge in the diversity of systems of islands which are able to exchange individuals, compared to sets of single islands for which there are essentially no correlations. Moreover as Fig. 6(a) shows, islands with closer ties have stronger correlations. In terms of measurement protocols, assessment of β-diversity seems to be almost as straightforward as assessment of α-diversity (and almost certainly much easier than assessing the time dependence). Thus, for example, one might use the interindividual β-diversity to determine the properties of the metacommunity, and the intraindividual β-diversity to discriminate between assembly mechanisms [33].

ACKNOWLEDGMENTS

I thank Mike Cates, Chris Quince, Alan McKane, Bill Sloan, and Dave Taylor for helpful discussions.

[1] E. A. Grice *et al.*, Science **324**, 1190 (2009).
[2] E. K. Costello *et al.*, Science **326**, 1694 (2009).
[3] N. Fierer *et al.*, Proc. Natl. Acad. Sci. U.S.A. **107** 6477 (2010).
[4] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, Princeton, NJ, 2001).
[5] B. Foxman, *et al.*, Interdisc. Perspect. Infect. Diseases 2008, 613979.
[6] For human microbiota it seems quite likely that the most dominant taxa are niche adapted (e.g., *Staphylococcus epidermidis* on skin). However there is a large tail of rarer taxa for which stochastic assembly may be relevant. There is also the question of microhabitat variations.
[7] G. Bell, Science **293**, 2413 (2001); B. J. McGill, Nature (London) **422**, 881 (2003); J. Chave, Ecol. Lett. **7**, 241 (2004); D. Alonso, R. S. Etienne, and A. J. McKane, Trends Ecol. Evol. **21**, 451 (2006); J. Chave, D. Alonso, and R. S. Etienne, Nature (London) **441**, E1 (2006); E. G. Leigh, Jr., J. Evol. Biol. **20**, 2075 (2007).
[8] T. Bell, *et al.*, Science **308**, 1884 (2005); see also **309**, 1997 (2005).
[9] S. Woodcock *et al.*, FEMS Microbiol. Ecol. **62**, 171 (2007).
[10] R. S. Etienne and D. Alonso, Ecol. Lett. **8**, 1147 (2005).
[11] R. S. Etienne, D. Alonso, and A. J. McKane, J. Theor. Biol. **248**, 522 (2007).
[12] M. Vallade and B. Houchmandzadeh, Phys. Rev. E **68**, 061902 (2003).
[13] I. Volkov *et al.*, Nature (London) **424**, 1035 (2003).
[14] A. J. Mckane, D. Alonso, and R. V. Solé, Theor Popul. Biol. **65**, 67 (2004).
[15] F. He, Funct. Ecol. **19**, 187 (2005).
[16] R. S. Etienne, Ecol. Lett. **8**, 253 (2005).
[17] R. S. Etienne and D. Alonso, J. Stat. Phys. **128**, 485 (2007).
[18] M. Vallade and B. Houchmandzadeh, Phys. Rev. E **74**, 051914 (2006).
[19] P. Babak, Phys. Rev. E **74**, 021902 (2006).
[20] R. S. Etienne, Ecol. Lett. **10**, 608 (2007).
[21] E. Condit *et al.*, Science **295**, 666 (2002).
[22] J. Chave and E. G. Leigh, Jr., Theor Popul. Biol. **62**, 153 (2002).
[23] E. P. Economo and T. H. Keitt, Ecol. Lett. **11**, 52 (2008).
[24] K. S. Korolev *et al.*, Rev. Mod. Phys. **82**, 1691 (2010).
[25] C. R. Powell and A. J. McKane, Ecol. Complexity **6**, 316 (2009).
[26] J. C. Tipper, Paleobiology **5**, 423 (1979).
[27] E. H. Simpson, Nature (London) **163**, 688 (1949).
[28] S. Azaele *et al.*, Nature (London) **444**, 926 (2006).
[29] Matthew 25:29 (King James Version): “For unto every one that hath shall be given, and he shall have *abundance*: but from

him that hath not shall be taken away even that which he hath.” (my italics).

[30] F. He and X.-S. Hu, Ecol. Lett. **8**, 386 (2005).

[31] D. Alonso and A. J. McKane, Ecol. Lett. **7**, 901 (2004).

[32] J. Rosindell, Y. Wong, and R. S. Etienne, Ecol. Inform. **3**, 259 (2008).

[33] Note that, in principle, the interindividual β-diversity yields θ; and, given θ, the α-diversity yields μ_{eff}.