Biodiversity on island chains: neutral model simulations

Patrick B. Warren
Unilever R&D Port Sunlight, Bebington, Wirral, CH63 3JW, UK.
(Dated: April 6, 2010)

A neutral ecology model is simulated on an island chain, in which neighbouring islands can exchange individuals but only the first island is able to receive immigrants from a metacommunity. It is found by several measures that biodiversity decreases along the chain, being highest for the first island. Subtle changes in taxon abundance distributions can be detected when islands in the chain are compared to diversity-matched single islands. The results potentially apply to human microbial diversity, but highlight the difficulty of using static single-site taxon abundance distributions to discriminate between dispersal limitation mechanisms.

PACS numbers: 87.23.-n, 87.10.Mn, 02.50.Ga

It has recently been observed that human microbial diversity 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 inter-individual variability is also observed, to the point where the characterisation of residual skin bacteria has been proposed as a novel forensic tool [3]. The latter, in particular, supports the notion that stochastic dispersal limitation may play a significant role in determining microbial diversity. Stochastic dispersal limitation is a signature element of Hubbell’s unified neutral model of biodiversity and biogeography [4], and this motivates the question of whether neutral models can be applied to human microbial biodiversity and biogeography. This is a hard problem and I do not claim to have solved it here. Rather, the present study is restricted to exploring the role of dispersal mechanisms in the context of neutral theory, keeping in mind the possible application to the human microbiome.

The merits of neutral models have been debated extensively [5], and it is far from obvious that they should apply to human microbiota [6,7]. However Hubbell’s neutral model has recently been successfully applied to predict microbial diversity in tree holes [8,9]. In this context it is important to note that it has been argued dispersal limitation is the dominant factor determining taxon abundancies [10,11], with other neutral model ideas, such as the zero-sum constraint (single trophic level; community saturation) or speciation by point mutation, playing a lesser role.

If taxon abundancies are largely determined by stochastic dispersal limitation, then a couple of limiting hypotheses (Fig. 1) present themselves to explain the observed variations in human microbial diversity. The first is a variable-immigration-rate hypothesis in which different body sites are envisaged as being microbial ‘islands’ in contact with a microbial metacommunity but effectively isolated from each other. Here variation in diversity corresponds to a variable immigration rate from the metacommunity. The second hypothesis is an island-chain hypothesis 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, due to dispersal limitation. This is confirmed in the present study.

Of course these hypotheses represent limiting cases and, if dispersal limitation is relevant, reality probably lies somewhere in between. A second question therefore is whether one can use taxon abundance distributions to distinguish between dispersal mechanisms. Unfortunately, the present study finds that both hypotheses lead to rather similar abundance distributions. When this is conflated with other factors, such as deviations from neutral model dynamics [12], it is probably going to be difficult to distinguish between dispersal mechanisms on the basis of static single-site measurements of microbial diversity.

The neutral model has been extensively studied for isolated islands in contact with a metacommunity [4,10,13-15], but only for certain cases has it been solved for multiple islands, or ‘patches’, which are able to exchange...
individuals \[20, 21\]. In particular, the island chain problem has not been solved (i.e., where individuals can migrate between neighboring islands but immigration is restricted to the first island in the chain). The primary goal of the present study is to solve this problem. Although in principle one can approach this analytically, the experience of Vallade and Houchmandzadeh \[13\] for two islands suggests this will be effectively unmanageable. I therefore approach the problem by means of simulations.

Let me start by summarising the mathematical characterisation of taxon abundance distributions. Suppose there are \( K \) taxa and \( N_i \) individuals in the \( i \)-th taxon \((i = 1 \ldots K)\), in a population of \( J = \sum_{i=1}^{K} N_i \) individuals. The relative abundance of the \( i \)-th taxon is defined to be \( \omega_i = N_i/J \). The taxon abundance distribution is characterised by \( \phi_k \), the number of taxa containing \( k \) individuals. Formally \( \phi_k = \sum_{i=1}^{K} \delta_{k, N_i} \) where \( \delta_{n,m} \) takes the value unity if \( n = m \) and is zero otherwise. Given the set of \( N_i \) one can easily calculate \( \phi_k \). One has \( K = \sum_{k=1}^{\infty} \phi_k \) and \( J = \sum_{k=1}^{\infty} k \phi_k \). 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 (saturated) and are specified at the outset, whilst the number of taxa and the number of individuals per taxon fluctuates. I adopt the notation of Vallade and Houchmandzadeh \[13, 18\] and write \((\cdot \cdot \cdot )\) to indicate a quantity averaged over an ensemble of populations undergoing neutral model dynamics. The information in \( \langle \phi_k \rangle \) is conveniently represented by giving the ensemble-average probability \( p(\omega) \) that an individual belongs to a taxon of relative abundance \( \omega \). For a community of a finite size, \( p(\omega) \) is a discrete array or 'comb' of \( \delta \)-functions, even after ensemble-averaging, since \( \omega \) can only take on discrete multiples of \( 1/J \). However as \( J \to \infty \), \( p(\omega) \) becomes a continuous function. One can show that the continuum limit is \( p(\omega) = \lim_{J \to \infty} k \langle \phi_k \rangle \) where \( k = \omega J \). \[13\]

I shall additionally use several ensemble-average measures of diversity. The principal one of these is the Simpson diversity index \[23\], defined for a given set of taxon abundances to be \( D = 1 - \sum_{k=1}^{K} \omega_i^2 \). It is related to the second moment of the taxon abundance distribution by \( D = 1 - J^{-2} \sum_{k=1}^{\infty} k^2 \phi_k \). From this it can easily be shown that in the continuum limit

\[
\langle D \rangle = 1 - \int_{0}^{1} \omega p(\omega) d\omega. \tag{1}
\]

The second diversity measure is the ensemble-average number of taxa \( \langle K \rangle = \sum_{k=1}^{\infty} \phi_k \). The third is the ensemble-average monodominance probability \( \langle \phi_J \rangle \) —as explained above \( \phi_J = 1 \) 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 some advantages over \( \langle K \rangle \) and \( \langle \phi_J \rangle \); it remains well defined in the continuum limit \( J \to \infty \), there are some particularly simple theoretical expressions for \( \langle D \rangle \) under neutral model dynamics, and in particular there is a prediction (confirmed by simulation) that \( \langle D \rangle \) factorises into a product of the metacommunity diversity index, and an island factor. Also the Simpson diversity index generalises naturally to a measure of \( \beta \)-diversity \[20\], and to time-series data \[24\]. The index satisfies \( 0 \leq \langle D \rangle \leq 1 - 1/K \). There is a mild disadvantage in that \( \langle D \rangle \) loses sensitivity to the underlying abundance distribution at the limiting values.

Let me next summarise neutral model dynamics. In the metacommunity it 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 \( \nu \) is replaced by an individual belonging to a new taxon. Thus \( \nu \) is the speciation rate. For \( \nu = 0 \) the metacommunity eventually falls into a monodominated state, in an ecological analog of the Matthew principle \[22\]. 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 \[4, 10, 13, 17\]. Results are quoted as metacommunity (subscript 'M') steady-state ensemble-averages:

\[
\langle \phi_k \rangle_M = \frac{\theta \Gamma(J_M + 1)}{k \Gamma(J_M + 1 - k)} \frac{\Gamma(J_M + \theta - k)}{\Gamma(J_M + \theta)} \tag{2}
\]

where \( \theta = (J_M - 1)\nu/(1 - \nu) \). One has \( \theta \approx J_M \nu \) for \( J_M \gg 1 \) and \( \nu \ll 1 \). It can be shown that \( \langle J \rangle_M = \sum_{k=1}^{J_M} k \langle \phi_k \rangle_M = J_M / \theta \) (an identity), and \( \langle K \rangle_M = \sum_{k=1}^{J_M} \langle \phi_k \rangle_M = J_M / \theta - 1 + k \). The continuum limit of Eq. \(2\) can be obtained using Stirling’s approximation. One finds \( p(\omega) = \theta (1 - \omega)^{\theta - 1} \). It follows that the metacommunity diversity order parameter in the continuum
The results were obtained by previous authors [13–15]. The metacommunity can be taken to have a static abundance distribution (upper curves; flat means the island abundance distribution turns over on a time scale of order 1/nu). This is because the metacommunity dynamics are decoupled (discussed in more detail below). Therefore I generate a large number (10^3–10^5) of metacommunity abundance distribution samples for given J_M and θ, equilibrating for 10^3/m approximately 100 replacement steps between samples to ensure statistical independence. I use these samples in subsequent island and island chain simulations. As a reference point, I shall use θ = 10, motivated by Woodcock et al. [4], and J_M = 10^5, motivated not so much by time scale considerations (see later) but by the requirement that J_M ≫ J > 1 [18]. Except where otherwise stated, averages are over 10^3 samples.

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 for the steady-state properties (equilibrating for 10^3/m replacement steps between samples). For example Fig. 2 compares theory and simulation results for the Simpson diversity index.

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 inter-island migration rate m_2. This is illustrated in Fig. 1(b). 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 neighbouring 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 $2m_2$. If the chosen individual lies on the terminal island, it is replaced with a copy of another individual on the island with probability $m_2$, or with a migrant from the neighbouring island with probability $m_2$. Migrants are copies of individuals chosen at random on neighbouring islands.

Fig. 3 shows that the approach to steady state of a

FIG. 4: Aspects of diversity on island chains of length $n$: (a) diversity index $\langle D \rangle$, (b) number of taxa $\langle K \rangle$, and (c) monodominance probability $\langle \phi \rangle$. The dashed lines in (b) and (c) are the $\langle D \rangle$-matched single island results for $n = 9$ (see text).

metacommunity with probability $m_1$, or a migrant from the neighbouring 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 $2m_2$. If the chosen individual lies on the terminal island, it is replaced with a copy of another individual on the island with probability $m_2$, or with a migrant from the neighbouring island with probability $m_2$. Migrants are copies of individuals chosen at random on neighbouring islands.

Fig. 5 shows that the approach to steady state of a

FIG. 5: Diversity index $\langle D \rangle$, normalised by the theoretical metacommunity diversity index $\langle D \rangle_M = \theta/(\theta + 1)$, along a chain of $n = 9$ islands: (a) varying inter-island migration rate $m_2$ only, (b) varying migration and immigration rates together at fixed island size, (c) varying migration and immigration rates inversely with island size, (d) varying metacommunity diversity parameter.
suggests is that there is a spectrum of relaxation modes, which are excited differently according to the initialisation protocol, and which are subsequently mixed up by the non-linear dynamics. A more detailed exploration of this is left for future work.

Representative steady-state results for the island chain simulations are shown in Figs. (a) and (b). The first conclusion (Fig. (a)) is that diversity decreases, by whatever measure, as one moves away from the island in contact with the metacommunity. Fig. (b) shows how island diversity varies with immigration and migration rates $m_1$ and $m_2$, island size $J$, and the value of $\theta$. Increasing the inter-island migration rate $m_2$ (Fig. (b) (a)) has the effect of increasing the diversity along the island chain apart from the first island. Additionally increasing the metacommunity immigration rate $m_1$ (Fig. (b) (b)) leads to increased diversity along the whole chain. Fig. (c) supports the notion that the island diversity is governed by the combinations $J m_1$ and $J m_2$ rather than the individual values of $J$, $m_1$ and $m_2$, in close analogy to the theory for the single island. Similarly Fig. (d) strongly suggests that the Simpson diversity index continues to be factorisable into the metacommunity diversity index multiplied by a contribution from the structure of the island chain, again in close analogy to the single island result.

I next compare islands in the chain to ‘$(D)$-matched’ single islands. Here $(D)$-matching means a value for $\mu$ is inferred from Eq. (7) (i.e. $\mu = \langle D \rangle/(\langle D \rangle - \langle D \rangle_M)$ where $\langle D \rangle_M = \theta/(\theta + 1)$), and used to calculate values of $\langle K \rangle = \sum_{k=1}^{J} \langle \phi_k \rangle$ and $\langle \phi_j \rangle$ from Eqs. (4) and (5). I assume the island size $J$ is fixed. The procedure amounts to matching the first and second moments of $\langle \phi_k \rangle$. The dashed lines in Fig. (b) and (c) show systematically that the ensemble-average number of taxa is reduced and the monodominance probability is increased, comparing an island in the island chain with its $(D)$-matched single island counterpart. Thus there is a tendency towards fewer, larger taxa, when islands in a chain are compared to $(D)$-matched single islands.

A more detailed examination of the abundance distributions shows that there is a subtle and non-trivial redistribution of the taxon abundances. When compared to the $(D)$-matched single islands, Fig. (b) shows that $p(\omega)$ 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. (b). This corresponds to the increased monodominance probability. At first sight this is at odds with with the redistribution towards mid-range abundancies, nevertheless it is a real effect and indeed is the reason why monodominance was separately studied.

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 \(\langle D\rangle\)). The loss of the high abundance taxa and the increased monodominance probability are more mysterious and I do not at present have a clear mechanistic explanation. Possibly what is happening for islands in a chain, compared to \(\langle D\rangle\)-matched single islands, is that the monodominated state \((\omega = 1)\) has become ‘stickier’ in dynamical terms, without actually becoming an adsorbing state. In the monodominated state there is of course only one taxon, with \(\omega = 1\), and this may come at the expense of the high abundance taxa with \(0.8 \lesssim \omega < 1\).

For \(\langle D\rangle \lesssim 0.5\) 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 \(p(\omega) = \theta(1-\omega)^{\theta-1}\) with \(\theta = \langle D\rangle/(1-\langle D\rangle)\) is a very good approximation to \(p(\omega)\) from Eq. [3]. 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 \(J_M\).

Despite these subtleties, it is clear from Fig. 6 that the taxon abundance distributions on an island chain are quite well approximated by \(\langle D\rangle\)-matched single islands. This is the origin of the claim in the introduction that it is probably going to be difficult to distinguish between dispersal mechanisms on the basis of static single-site measurements of the taxon abundances. To resolve this question, or indeed to distinguish between dispersal-limitation and niche-adaptation [7], probably requires more detailed examination of the \(\beta\)-diversity [20, 21, 28], and dynamics [13, 14, 24]. In this context it may be useful to explore spatial correlations [21] and temporal correlations [24], which are natural generalisations of the Simpson index \(D\).

I thank Mike Cates, Chris Quince, Bill Sloan, and Dave Taylor, for helpful comments.

[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. (USA) (2010), doi:10.1073/pnas.100162107.
[4] S. P. Hubbell, The unified neutral theory of biodiversity and biogeography (Princeton University press, Princeton, NJ, 2001).
[5] G. Bell, Science 293, 2413 (2001); B. J. McGill, Nature 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 441, E1 (2006); E. G. Leigh, Jr., J. Evol. Biol. 20, 2075 (2007).
[6] B. Foxman, et al., Interdisc. Perspect. Infect. Diseases (2008), doi:10.1155/2008/613079.
[7] For human microbiota it seems quite likely that the most dominant taxon are niche-adapted, as for example Staphylococcus epidermidis on skin. However there is a large tail of rarer taxa for which dispersal limitation may be relevant.
[8] T. Bell, et al., Science 308, 1884 (2005); see also Science 309, Letters (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] G. Bianconi, L. Ferretti, and S. Franz, Europhys. Lett. 87, 28001 (2009).
[13] M. Vallade and B. Houchmandzadeh, Phys. Rev. E 68, 061902 (2003).
[14] I. Volkov, et al., Nature 424, 1035 (2003).
[15] A. J. McKane, D. Alonso, and R. V. Solé, Theor. Pop. Biol. 65, 67 (2004).
[16] F. He, Functional Ecol. 19, 187 (2005).
[17] R. S. Etienne and D. Alonso, J. Stat. Phys. 128, 485 (2006).
[18] M. Vallade and B. Houchmandzadeh, Phys. Rev. E 74, 051914 (2006).
[19] P. Babak, Phys. Rev. E 74, 021902 (2006).
[20] J. Chave and E. G. Leigh, Jr., Theor. Pop. Biol. 62, 153 (2002).
[21] E. Condit, et al., Science 295, 666 (2002).
[22] The probability distribution function \(p(\omega)\) has also been advocated for the non-parametric Kolmogorov-Smirnov test to compare taxon abundance distributions; see J. C. Tipper, Paleobiology 5, 423 (1979).
[23] E. H. Simpson, Nature 163, 688 (1949).
[24] S. Azaele, et al., Nature 444, 926 (2006).
[25] 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.”.
[26] F. He and X.-S. Hu, Ecol. Lett. 8, 386 (2005).
[27] D. Alonso and A. J. McKane, Ecol. Lett. 7, 901 (2004).
[28] M. Dornelas, S. R. Connolly, and T. P. Hughes, Nature 440, 80 (2006).