



Community differentiation on landscapes: drift, migration and speciation

Xin-Sheng Hu, Fangliang He and Stephen P. Hubbell

X.-S. Hu (*xin-sheng.hu@ualberta.ca*) and F. He, *Dept of Renewable Resources, 751 General Service Building, Univ. of Alberta, Edmonton, Alberta, T6G 2H1, Canada.* – S. P. Hubbell, *Dept of Ecology and Evolutionary Biology, 621 Charles E. Young Drive, Univ. of California, Los Angeles, CA 90095-1601, USA, and Smithsonian Tropical Research Inst., Unit 0948, APO AA 34002-0948, Panama.*

Theories of the differentiation of ecological communities on landscapes have typically not considered evolutionary dynamics. Here we analytically study the expected differentiation among local communities in a large metacommunity, undergoing speciation, ecological drift and intercommunity dispersal, in the context of neutral theory. We demonstrate that heterogeneity in species diversity and abundance arises among communities when local communities are small and intercommunity migration is infrequent. We propose a new measure to describe community differentiation, defined as the average correlation or the average probability (C_{st}) that two randomly sampled individuals of the same species within local communities are from the same ancestor. The effects of driving forces (migration, mutation, and ecological drift) are incorporated into the two-level hierarchical community structure in a finite island model of neutral communities. Community differentiation can increase the effective metacommunity size or the Hubbell's fundamental species diversity in the metacommunity by a factor $(1 - C_{st})^{-1}$. Significant community differentiation arises when $C_{st} \neq 0$. Intercommunity migration promotes species diversity in local communities but reduce species diversity in the metacommunity. In either the finite or infinite island case, one can estimate the number of intercommunity migrants by using multiple local community datasets when the speciation is negligible in the neutral local communities, or by using the metacommunity dataset when the speciation is included in the local neutral communities. These results highlight the significance of the evolutionary mechanisms in generating heterogeneous communities in the absence of complicated ecological processes on large landscapes.

How heterogeneity arises in the distribution of species and communities on landscapes is a fundamental ecological question for understanding the spatial structure of landscapes and biogeographic regions. Answers to this question bear on the century-old debate on community succession between the Clementian 'organismic' theory (convergence to a niche-stabilized equilibrium community) and the Gleasonian individualistic theory (divergence among non-equilibrium communities). Empirical studies have shown the presence of convergence (Sommer 1991, Wilson et al. 1995), divergence (McCune and Allen 1985, Fukami et al. 2005) or a mixture (Inouye and Tilman 1988). These observations have been variously interpreted in terms of ecological processes, such as natural disturbances, habitat variation, successional age, the source of propagules, and even the metrics used to measure the convergence or divergence of community structure (Lepš and Rejmánek 1991, Samuels and Drake 1997). Ecological processes are mostly invoked in attempts to interpret the origins of spatial landscape pattern (Tilman 1982, Turner 2005).

Evolutionary history and biogeographical processes can also play significant roles in producing community differentiation (Brown and Lomolino 2000). There is evidence of

this from the study of phylogenetic relationships among co-occurring species in local ecological communities, which are often more closely related than expected from random assemblages of species sampled from the metacommunity (Webb et al. 2002, Kembel and Hubbell 2006). Interest in the phylogenetic relationships of species in communities arose at the same time as the development of neutral theory in which speciation is the driving process in assembling metacommunity (Hubbell 2001). Studies in community phylogenetics point to the importance of understanding community differentiation in landscapes from an evolutionary perspective. However, the role of evolutionary forces in community differentiation is still largely unexplored since community ecology and evolutionary biology are typically studied on very different temporal and spatial scales (Johnson and Stinchcombe 2007). Combining an evolutionary perspective with consideration of ecological process and pattern is essential for studying the formation of heterogeneous landscapes.

Here we argue that evolutionary forces can have significant effects on community differentiation. We formulate a neutral theory of community differentiation (speciation, drift and migration); we present the theory

incorporating natural selection elsewhere. In our community differentiation theory, we make use of the concept of ‘effective community size’ (Orrock and Fletcher 2005, Hu et al. 2006). Lande et al. (2003) define the effective size of a neutral community as “the size of an ideal neutral community of constant size that would produce the same expected rate of loss of Simpson’s diversity” due to ecology drift. This definition is based on a comparison between the observed and expected losses of Simpson’s diversity under neutrality. Here an ideal neutral community can refer to a community in which individuals are randomly assembled and each individual has equal probability of birth and death. Instead of using the definition of Lande et al. (2003), we use an alternative concept ‘variance effective community size’, the size of an ideal neutral community that would produce the same increase rate of sampling variance per species per unit time due to ecological drift effects as the size of the real community under consideration. This definition is based on a comparison between the observed and expected increase rates in sampling variance of relative species abundance per unit time. This concept is logically analogous to the ‘variance effective population size’ in population genetics (Wright 1969), which we derive below.

We propose a novel measure to characterize community differentiation. This measure, denoted by C_{st} , analogous in concept to Cockerham and Weir’s (1987; called C&W hereafter) θ_2 , is defined as the average correlation of two randomly sampled individuals of the same species within local communities (Table 1). When identity-of-descent (IBD) measures are utilized for interpreting the correlation, C_{st} is the average probability of IBD for pairs of individuals

of the same species randomly sampled from the same local community, logically analogous to the interpretation of θ_2 in terms of IBD measures in population genetics (Cockerham and Weir 1987). $1-C_{st}$ is the probability of not IBD. Note that the notation C_{st} is used to distinguish this measure from Wright’s F_{st} that is used to describe population genetic differentiation (Wright 1969), where C stands for community, subscript s for subcommunity, and t for the total community, i.e. metacommunity. C_{st} is conceptually distinct from Hubbell’s F_2 that refers to the probability of randomly drawing two individuals of the same species (identity in state (IIS), not IBD!) within the metacommunity (Hubbell 2001, pp. 120–121). Community differentiation increases the probability that two individuals randomly sampled from the same community come from the same ancestors. Community structure can be analyzed through the analysis of variance (ANOVA) using community census data sets and individual-based indicator variables as described below. One can estimate C_{st} by the proportion of total species variation attributable to the variation among local communities (Weir 1996). A larger value of C_{st} indicates greater correlation among individuals within local communities and more differentiation among local communities. In the following, we begin by describing the concepts of two-level hierarchy of community structure. We then demonstrate the consequence of community differentiation from the changes in effective metacommunity size. We further derive C_{st} as a function of evolutionary forces (speciation, drift and migration) and investigate the effects of these forces on community differentiation represented by C_{st} . The analytical relationships between

Table 1. The mathematical symbols and their meanings in the proposed theory of community differentiation.

Mathematical symbols	Biological meanings
L	The number of local communities
m	The immigration rate per unit time for each local community
v	The speciation rate for each individual
S	The total number of species
J_L	Effective local community size
J_M	Effective metacommunity size
Y_{iuk}	An indicator variable, the kth individual of the uth species in the ith local community
Y_{uk}	An indicator variable, the kth individual of the uth species
p_{iu}	The relative abundance of the uth species in the ith local community
p_u	The relative abundance of the uth species in the metacommunity
n_i	The sample size of the ith local community
\bar{n}	Average sample size
C_{st}	The average correlation of two individuals of the same species randomly sampled within the same local communities
C_b	The average correlation of two individuals of the same species randomly sampled from different local communities
β	An estimable quantity ($= (C_{st} - C_b)/(1 - C_b)$) that can also be used to describe community differentiation
θ	Hubbell’s fundamental biodiversity number in each local community
Θ	Hubbell’s fundamental biodiversity number in the metacommunity
λ	The number of migrants to each local community
Λ_w	The sum of the expected probability that two different individuals within a local community are the same species
Λ_b	The sum of the expected probability that two different individuals, each from different local communities, are the same species.
Λ	The sum of the square of the relative abundance over all species in the metacommunity
D_M	Simpson diversity index for the metacommunity
$\sigma_{p_{iu}}^2$	The variance of relative abundances among local communities
$\sigma_{\delta p_{iu}}^2$	The average sampling variance for the uth species over all local communities
$V(\delta p_u)$	The sampling variance for the uth species due to ecological drift
$\bar{V}_{\delta p}$	The average sampling variance over all species
$\Phi(Q)$	The number of species with relative abundance Q in the metacommunity

C_{st} and Hubbell's fundamental biodiversity number are finally derived, showing how community differentiation shapes the species diversity in the metacommunity in the absence of complex ecological processes.

Two-level hierarchical community structure

Suppose that a metacommunity is subdivided into L (≥ 2) local communities, each with equal effective community size, J_L . The distribution of species relative abundances can be different among local communities owing to ecological drift effects. There are exchanges of migrants among local communities, with a constant migration rate m per unit time in each local community. Immigrants to each local community equally come from all other local communities. These assumptions are the same as the finite island model in population genetics (Takahata and Nei 1984).

Since the species names of all individuals are identified, the hierarchical structure considered here is individuals within local communities, and individuals between local communities (Fig. 1). Accordingly, two correlations can be defined: the average correlation between two individuals of the same species randomly sampled within the same local community (C_{st}), and the average correlation between two individuals of the same species sampled from different local communities (C_b). Note that C_{st} and C_b are analogous to C&W's θ_2 and θ_3 , respectively. No correlations within individuals are included, i.e. without the parameter similar to C&W's θ_1 , since individuals in the community are considered as the minimal operational unit (Hu et al. 2006) and genetic markers are not employed. In the absence of inter-community migration (local communities independently evolve), i.e. $m = 0$, then $C_b = 0$, but $C_{st} \neq 0$ due to the operation of ecological drift that generates correlation of pairs of individuals within communities.

Let indicator y_{iuk} ($i = 1, \dots, L; u = 1, \dots, S; k = 1, \dots, n_u$) be the k th individual of the u th species in the i th local community, where $y_{iuk} = 1$ if the k th individual is the u th species and $y_{iuk} = 0$ if the k th individual is not the u th species in the i th community. Let $\Lambda_w = \sum_{u=1}^S E(y_{iuk}y_{iuk'})$

($k \neq k'$), the sum of the expected probability that two different individuals within a local community are the same species, and $\Lambda_b = \sum_{u=1}^S E(y_{iuk}y_{i'uk'})$ ($i \neq i', k \neq k'$), the sum of the expected probability that two different individuals, each from different local communities, are the same species. Λ_w and Λ_b are analogous to C&W's Q_2 and Q_3 , respectively. According to the work of Weir (1996, pp. 180–181) and Cockerham (1969, 1973), $1 - \Lambda_w$ and $1 - \Lambda_b$ measure the sums of variances of between and within communities over all species, respectively. To avoid repeating expressions, estimation of $1 - \Lambda_w$ and $1 - \Lambda_b$ can be obtained using ANOVA (Weir 1996, pp. 170–176, pp. 180–182). Similar to the case in population genetics, one estimable quantity derived from the two-level hierarchical structure (C_{st} and C_b) is

$$\beta = \frac{\Lambda_w - \Lambda_b}{1 - \Lambda_b} = \frac{C_{st} - C_b}{1 - C_b} \quad (1)$$

In population genetics, β is analogous in concept to Wright's inbreeding coefficient (fixation index, F_{is}). The difference between β and F_{is} is that β is defined in terms of pairs of alleles within and between populations, while F_{is} is defined in terms of pairs of alleles within and between individuals of the same populations (Weir 1996, p. 181). Here, the fixation meaning of β is not stressed since individuals are considered as haploids in community ecology (Hu et al. 2006). When the number of local communities is large, or when the correlation between individuals between communities is small, β ($\approx C_{st}$) essentially describes the community differentiation (Cockerham and Weir 1987).

Estimation of C_{st} can be obtained using the method similar to the one for estimating F_{st} with haploid genetic markers (Weir 1996, pp. 166–176). In the case of fixed number of local communities (L), C_{st} for the u th species can be calculated by $C_{st} = \sum_{i=1}^L n_i (p_{iu} - p_u)^2 / (\bar{n}(L-1)p_u(1-p_u))$ where n_i is the sample size of the i th local community, p_{iu} is the relative abundance of the u th species in the i th local community sample, $p_u = \sum_{i=1}^L n_i p_{iu} / \sum_{i=1}^L n_i$, and $\bar{n} = \sum_{i=1}^L n_i / L$ (Weir 1996, pp. 166–167). Multiple species estimate can be calculated by $C_{st} = \sum_{u=1}^S \sum_{i=1}^L n_i (p_{iu} - p_u)^2 / \sum_{u=1}^S (\bar{n}(L-1)p_u(1-p_u))$. In the case of random model of local communities, C_{st} can be estimated using ANOVA by introducing an indicator variable to mark each individual in the community (Weir 1996, pp. 169–176). These steps are not repeated further.

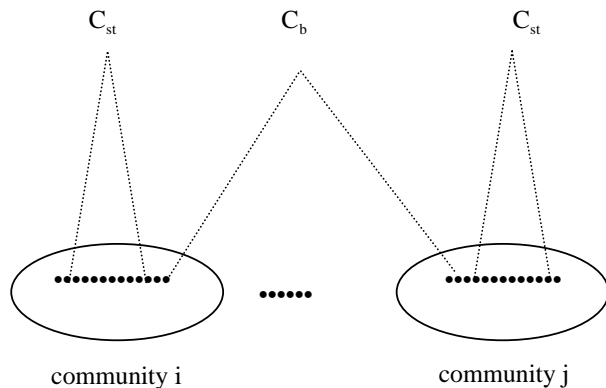


Figure 1. Two-level hierarchical community structure: correlation between pairs of randomly sampled individuals of the same species within the same community (C_{st}) and among different communities (C_b). A large community differentiation can increase the correlation within a local community, C_{st} , but reduce the correlation among communities, C_b .

Effective metacommunity size

The following derivation is based on the average increase rate of the sampling variance of relative species abundance per unit time, and this approach is analogous to the one used by Wright (1943) in deriving his formula for sampling variance of gene frequency in a subdivided population. Let p_u be the relative abundance of the u th species in the metacommunity ($\sum_{u=1}^S p_u = 1$). Let y_{uk} ($u = 1, \dots, S; k = 1, \dots, n_u$) be the indicator variable for the k th individual in the u th species in the metacommunity. The indicator variable is $y_{uk} = 1$ when the k th individual is the u th species, and $y_{uk} = 0$ otherwise. The expectation of the product for two individuals of the same species randomly

sampled from the metacommunity, i.e. the expectation of conspecific pair, $E(y_{uk}y_{uk'})$, can be written as

$$E(y_{uk}y_{uk'}) = p_u^2(1 - C_{st}) + p_u C_{st} = p_u^2 + p_u(1 - p_u)C_{st}, \quad (2)$$

similar to C&W's expression but different in biological meaning. The above equation clearly indicates the biological meaning of C_{st} , referring to the average correlation per species or the average probability that two individuals randomly sampled from the metacommunity come from common ancestors. This analysis is distinct from the community phylogenetic analysis where the common ancestors among species and branch lengths are the focus (Hubbell 2001, Webb et al. 2002, Kembel and Hubbell 2006). It is necessary to emphasize that the lengths of generation time vary with species and C_{st} may differ from one species to another. We only consider the average correlation coefficient or the average IBD probability of pairwise conspecific individuals although species-based C_{st} 's can be separately estimated.

Let $\sigma_{p_{iu}}^2$ be the variance of p_{iu} ($i = 1, \dots, L$; $u = 1, 2, \dots, S$) among local communities ($= \sum_{i=1}^L (p_{iu} - p_u)^2 / L$). In statistical language, the component of variance between communities (community differentiation) can be described by the covariance of individuals within a community (Cockerham 1973, Weir 1996, p. 47). Thus, we obtained:

$$\sigma_{p_{iu}}^2 = E(y_{iuk}y_{iuk'}) - E(y_{iuk})E(y_{iuk'}) = p_u(1 - p_u)C_{st} \quad (3)$$

The above equation also gives the definition of C_{st} from a different aspect, $C_{st} = \sigma_{p_{iu}}^2 / p_u(1 - p_u)$, similar in concept to the definition of F_{st} in population genetics (Wright 1969).

Let J_M be the variance effective size of the neutral metacommunity. Ecological drift causes the relative species frequencies to fluctuate per unit time. The sampling variance for the u th species due to ecological drift is:

$$V(\delta p_u) = \frac{p_u(1 - p_u)}{J_M} \quad (4)$$

Let $\Lambda = \sum_{u=1}^S p_u^2$ and $D_M = 1 - \Lambda$, the Simpson diversity index in the metacommunity. The average over all species for Eq. 2 can be readily derived as:

$$\overline{E(y_{uk}y_{uk'})} = \frac{1}{S} \sum_{i=1}^S E(y_{uk}y_{uk'}) = \frac{\Lambda + (1 - \Lambda)C_{st}}{S} \quad (5)$$

Thus, the average of sampling variance over all species becomes:

$$\overline{V_{\delta p}} = \frac{1}{S} \sum_{u=1}^S V(\delta p_u) = \frac{1 - \Lambda}{SJ_M} \quad (6)$$

We now consider the effects of community differentiation on the variance effective metacommunity size. In each local community, the sampling variance for the relative abundance of the u th species in the i th community (ecological drift) is $p_{iu}(1 - p_{iu})/J_L$, derived from the binomial distribution. The average sampling variance for the u th species over all local communities, denoted by $\sigma_{\delta p_{iu}}^2$, is:

$$\sigma_{\delta p_{iu}}^2 = \frac{1}{LJ_L} \sum_{i=1}^L p_{iu}(1 - p_{iu}) \quad (7)$$

The sampling variance for the mean relative abundance of the u th species over all local communities, i.e. the sampling variance for $p_u = \sum_{i=1}^L p_{iu} / L$, can be derived as:

$$V(\delta p_u) = \frac{1}{L} \sigma_{\delta p_{iu}}^2 = \frac{1}{L^2 J_L} \sum_{i=1}^L p_{iu}(1 - p_{iu}) \quad (8)$$

Expanding Eq. 8 yields the following equation:

$$V(\delta p_u) = \frac{1}{LJ_L} (p_u(1 - p_u) - \sigma_{p_{iu}}^2) \quad (9)$$

Application of Eq. 3 to 9 reads:

$$V(\delta p_u) = \frac{p_u(1 - p_u)(1 - C_{st})}{LJ_L} \quad (10)$$

The average of sampling variance over all species becomes:

$$\overline{V_{\delta p}} = \frac{1}{S} \sum_{u=1}^S V(\delta p_u) = \frac{(1 - \Lambda)(1 - C_{st})}{SJ_L} \quad (11)$$

From Eq. 6 and 11, the variance effective size of the metacommunity is given by:

$$J_M = \frac{LJ_L}{1 - C_{st}} \quad (12)$$

The above expression indicates that the effective metacommunity size is increased by a factor $(1 - C_{st})^{-1}$. Figure 2 shows that the effective metacommunity size increases as the local communities become more differentiated, and it doubles when $C_{st} = 0.5$, i.e. $\frac{J_M - LJ_L}{LJ_L} = \frac{C_{st}}{1 - C_{st}} = 1$.

From Eq. 12, the effective metacommunity size can change considerably ($J_M \gg LJ_L$) when $C_{st} \gg 0.5$. The ecological relevance to this result is that the average sampling variance per species over all communities reduces when community differentiation becomes large. This can occur in the case of the same demography in each local community and no variation in effective among local communities.

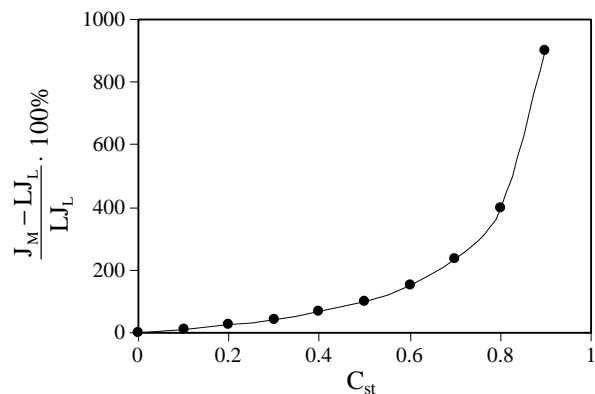


Figure 2. Community differentiation (C_{st}) increases the effective metacommunity size (J_M). The results indicate that the increment percentage in J_M , $(J_M - LJ_L) / LJ_L \times 100\%$, can be substantial when local communities approach fixation ($C_{st} \rightarrow 1$).

Linking community structure to evolutionary forces

In this section, we proceed to incorporate the effects of driving forces (ecological drift, speciation, and intercommunity migration) into the two-level hierarchy community structure parameters (C_{st} , C_b and β) under the neutrality assumption in the finite island model. Let v be the speciation rate for each individual. Suppose that exchanges of migrants among communities occur, with m proportions of J_L individuals replaced by immigrants for each local community. Note that m is the per capita immigration rate. Immigrants to each local community equally come from all other $L-1$ communities, the same as the island model mentioned in the preceding section. Let $a = (1 - m)^2 +$

$\frac{m^2}{L-1}$, $b = \frac{1-a}{L-1}$, $\rho = (1-v)^2$, $\gamma = 1 - \frac{1}{J_L}$, $\alpha = \frac{L}{L-1}$, $d = (1 - m\alpha)^2$, and $W = 1 - \rho + \rho b - \rho\alpha\gamma + \rho^2\gamma d$. Based on C&W's work, the recurrent equations for the changes in C_{st} and C_b can be immediately obtained (results not repeated here; Cockerham and Weir 1987, p. 8513). A balance can be reached between the effects of ecological drift that increase community differentiation and the effects of intercommunity migration and speciation that reduce community differentiation. The steady state expressions are:

$$C_{st} = \frac{\rho(a - \rho d)}{J_L W} \quad (13a)$$

$$C_b = \frac{\rho b}{J_L W} \quad (13b)$$

$$\beta = \frac{\rho d}{J_L(1 - \rho d) + \rho d} \quad (13c)$$

Figure 3a shows that an increase in migration reduces the correlation of individuals within communities (C_{st}) but increases the correlation of individuals among communities (C_b). When the number of local communities increases, β approaches C_{st} (Fig. 3b).

In the case of small speciation (v) and migration rate (m), Eq. 13c can be approximated as:

$$\beta \approx \frac{1}{1 + \theta + 2\lambda} \quad (14)$$

where $\theta = 2J_L v$ (Hubbell's fundamental biodiversity number in the local communities) and $\lambda = J_L m$, the number of migrants (Hu et al. 2007). 2λ is equivalent to μ in Vallade and Houchmandzadeh (2003), I in Etienne and Alonso's (2005) and γ in Volkov et al. (2003). However, it is difficult to derive a simple approximation for C_{st} . Equation 14 is analogous to C&W's β (Cockerham and Weir 1987, p. 8513) and β essentially describes community differentiation.

Equation 13 describes the structure with finite number of local communities ($L \geq 2$). In the extreme case where the number of local communities is infinite ($L \rightarrow \infty$), analogous in concept to the Wright's infinite island model in population genetics (Wright 1969), we obtain $b = 0$, $W = (1 - \rho)(1 - \rho\gamma a)$, and $C_b = 0$ from Eq. 13b. Note that d in Eq. 13 can also be expressed as $d = a - b$. Under this condition, Eq. 13a is simplified as $C_{st} = \rho\gamma / (J_L(1 - \rho\gamma a)) \approx 1/(1 + 2\lambda + \theta)$. Similarly, Eq. 13c can be approximated

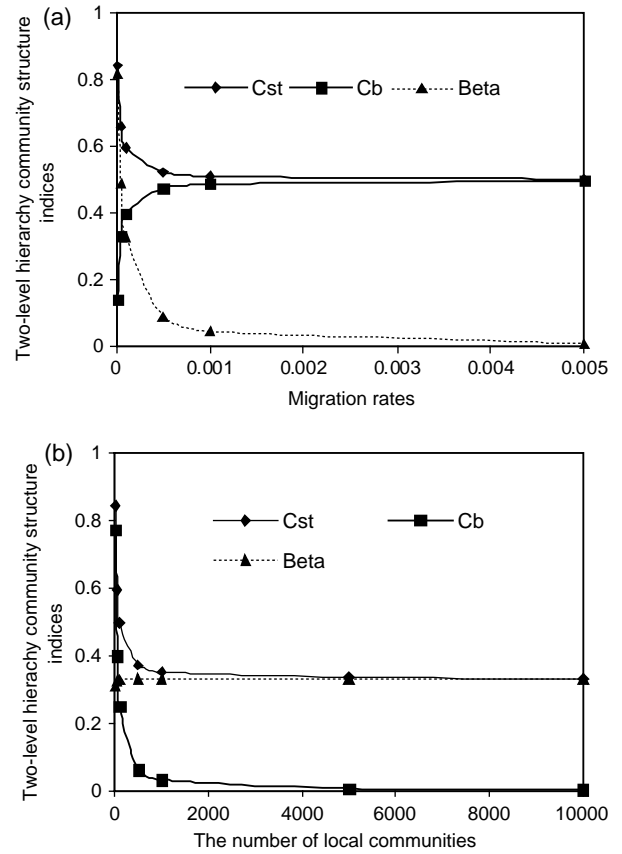


Figure 3. Changes in three parameters describing two-level hierarchical community structure: (a) effects of migration rates, (b) effects of the number of local communities. Results are calculated according to Eq. 13. The parameters used in (a) are the number of local communities $L = 50$, speciation rate $v = 10^{-6}$, and the effective local community size $J_L = 10^4$. The parameters used in (b) are the migration rate $m = 10^{-4}$ and the same speciation rate and the effective local community size as in (a).

by $\beta \approx 1/(1 + 2\lambda + \theta)$. Thus, in the case of infinite number of local communities, we obtain:

$$C_{st} = \beta = \frac{1}{1 + 2\lambda + \theta} \quad (15)$$

Equation 15 reduces to $1/(1 + 2\lambda)$ when the effects of speciation are very small on the specie abundance distribution in local communities. In the case of $\theta = 0$, $C_{st} = 1/(1 + 2\lambda)$ is the average correlation or the average IBD probability of pairwise conspecific individuals that is entirely determined by the same immigrants. Exchanges of migrants (λ) reduce community differentiation (C_{st}). Equation 15 indicates that speciation (θ) has the same function as migration (λ) in affecting community differentiation, implying that the community with a larger species richness owing to speciation tends to have a smaller extent of community differentiation.

Intercommunity migration

Let $\Theta = 2J_{Mv}$ be Hubbell's fundamental biodiversity number for the metacommunity. From Eq. 12, the following relationship between Θ and θ can be derived as

$$\Theta = \frac{2LJ_{LV}}{1 - C_{st}} = \frac{L\theta}{1 - C_{st}} \quad (16)$$

indicating that the fundamental biodiversity number for the metacommunity is increased by a factor $(1 - C_{st})^{-1}$. Figure 4 shows that a smaller number of intercommunity exchanges facilitate the increase of the fundamental biodiversity number in the metacommunity.

Let $\Phi(Q)$ be the number of species with the relative abundance Q in the metacommunity so that $\Phi(Q)dQ$ represents the expected number of species whose relative abundances fall within the range of $(Q, Q + dQ)$. Under the hypothesis of the infinite species model for the neutral metacommunity (Hubbell 2001), $\Phi(Q)$ has the following expression (Kimura and Crow 1964):

$$\Phi(Q) = \frac{\Theta(1 - Q)^{\Theta-1}}{Q} \quad (17)$$

The relationship between the Simpson diversity index ($D_M = 1 - \int_0^1 Q^2 \Phi(Q) dp$ in the continuous form) and Θ in the metacommunity is given by $D_M = \Theta/(1 + \Theta)$ (Etienne 2005, He and Hu 2005). Θ can be simply estimated by $\hat{\Theta} = \hat{D}_M/(1 - \hat{D}_M)$ and its variance $V(\hat{\Theta}) = 2\hat{\Theta}(1 + \hat{\Theta})^2/((2 + \hat{\Theta})(3 + \hat{\Theta}))$. Note that this method requires the strict neutral assumption, which otherwise produces biased estimates (Kimura 1983, He and Hu 2005). Equation 16 indicates that Θ in Eq. 17 is amplified ($> 2LJ_{LV}$) when the metacommunity is subdivided into L local communities that are linked through intercommunity migration. Such amplification has not been highlighted in previous studies (Etienne 2005, 2007, He 2005) and deserves attentions in our empirical analyses. Substitution of Eq. 16 into D_M yields the expression for estimating θ , i.e.:

$$\theta = \frac{D_M(1 - C_{st})}{L(1 - D_M)} = \frac{\Theta(1 - C_{st})}{L} \quad (18)$$

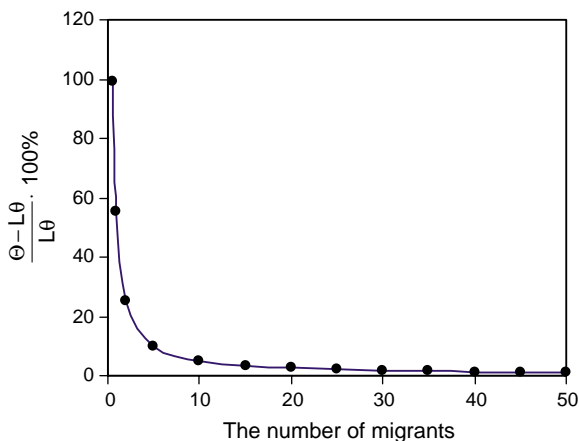


Figure 4. Effects of intercommunity migration on Hubbell's fundamental diversity number for the metacommunity (Θ). The increment percentage, $((\Theta - L\theta)/L\theta) \times 100\%$, gradually decreases when the number of intercommunity exchanges (λ) increases. The results indicate that local community differentiation increases Hubbell's fundamental diversity number for the metacommunity.

In the case of finite number of local communities, the expression for the number of migrants per unit time is derived from Eq. 14 and 18:

$$\lambda = \frac{1}{2\alpha} \left(\frac{1}{\beta} - \frac{D_M(1 - C_{st})}{L(1 - D_M)} - 1 \right) \quad (19)$$

Note that Eq. 19 is the result for the case of finite number of local communities under the condition of small speciation and migration rate. The above expression includes the effect of speciation on the estimation of λ in the local communities, but this estimation in reality requires a dataset for the entire metacommunity (Hu et al. 2007).

In the case of infinite number of local communities ($\beta = C_{st}$ when L approaches infinite), Eq. 19 becomes:

$$\lambda = \frac{1}{2} \left(\frac{1}{C_{st}} - \theta - 1 \right) \quad (20)$$

The above equation can also be directly obtained from Eq. 15 for the case of infinite number of local communities. Furthermore, when speciation is negligible in the local communities, Eq. 20 simplifies to the approximation $\lambda \approx (C_{st}^{-1} - 1)/2$, in which the metacommunity dataset is no longer required for estimating λ .

Discussion

Different from the traditional ecological explanations for landscape heterogeneity (Turner 2005), we have analytically shown that heterogeneity among local communities within a metacommunity landscape can arise solely from drift, migration and speciation. Ecological processes such as disturbances and habitat heterogeneity may not be necessary for community differentiation, although they could accelerate differentiation. The differentiation of communities is particularly evident when local communities are small and exchange of migrants among communities is infrequent. The present study provides a novel evolutionary perspective for understanding the formation of landscape heterogeneity (Johnson and Stinchcombe 2007).

The proposed community differentiation measure (C_{st}) has a clear biological meaning and reflects the average correlation of two randomly sampled individuals of the same species from the same local communities. The explicit relationship between C_{st} (or β) and speciation, drift and dispersal under the neutral hypothesis makes this measure a very useful tool for describing community structure. The idea linking β -diversity in community ecology and other diversity indices to evolutionary forces has recently been explored in the context of neutral theory in macroecology (Economo and Timothy 2008). These studies made use of the probability that two individuals sampled randomly at separation distance x would be of the same species given a certain dispersal rate, denoted as $F(x)$, and so dealt with populations having continuous distributions in space (Malécot 1969, Nagylaki 1974, Hubbell 2001, Chave and Leigh 2002, Condit et al. 2002). This type of β -diversity index is distinct from C_{st} in two aspects. One is that C_{st} is used for discrete local communities, while $F(x)$ is used for the continuously distributed community. The second distinction is that the correlation of two randomly sampled

individuals refers to the probability of IBD in C_{st} but the probability of IIS in $F(x)$. Two conspecific individuals that are IIS may not be IBD, but two conspecific individuals that are IBD must be IIS.

In comparison with other β -diversity studies (MacArthur 1965, Levins 1968, Lewontin 1972, Pielou 1975, Routledge 1977, Koleff et al. 2003, Lande et al. 2003, Magurran 2004), one common feature in analysis is the linear partition of species diversity into additive components (within and between communities). The advantage with C_{st} measure is that the single parameter can be used for comparing community structures in different studies, which is difficult using the statistical tests in previous community structure analyses (Lande et al. 2003).

C_{st} is related to Simpson's indices within and between communities, the components of the expected mean squares in ANOVA (Weir 1996, pp. 166–176, pp. 180–182). This ANOVA can be easily extended to multiple-level hierarchical community analysis. Correspondingly, more parameters need to be defined in terms of the average correlation of two randomly sampled individuals at different hierarchical levels, rather than the single hierarchical statistical test (Lande et al. 2003), as implied from a recent work by Hardy and Senterre (2007) from the perspective of partitioning phylogenetic structure. Community differentiation at different hierarchy levels can be simultaneously assessed in terms of the average correlation at the corresponding levels.

Several practical applications are suggested by the present theoretical results in the context of neutral theory. The first is the use of C_{st} (or β) to estimate the number of migrants among communities. This is similar to the case of using F_{st} to estimate the number of migrants between populations in population genetics even if restrictive assumptions are required (Whitlock and McCauley 1999). Although alternative statistic methods are need for estimating the number of inter-community migrants, use of C_{st} ($\lambda \approx (C_{st}^{-1} - 1)/2$ for the case of infinite number of local communities, or Eq. 19 for the case of finite number of local communities) could be potentially useful, and a related exploration is recently proposed (Munoz et al. 2008). Note that the inter-community migrants are the two directional exchanges of individuals, different from the case studied by Hu et al. (2007) who explored theory for unidirectional migrants (from metacommunity to a local community). The critical assumption for the use of the present method is neutrality, particularly that per capita rates of dispersal are unaffected by species differences.

The second application is in estimating species diversity in the metacommunity. Our results indicate that the effective size of a metacommunity is greater by a factor of $1/(1-C_{st})$ than what is expected if dispersal rate is unlimited. When the intercommunity migration rate becomes more limited, the effective metacommunity size increases, as does the fundamental metacommunity diversity (Θ). This prediction is consistent with Hubbell's (2001) findings about the effect of intercommunity migration on the species diversity in the metacommunity. Hubbell (2001, chapter 7) indicates that the intercommunity migration plays different roles. A higher rate of dispersal facilitates the accumulation of species diversity in the local communities but reduces species diversity in the

metacommunity; a lower rate of dispersal does not effectively counteract the effects of ecological drift, but facilitates the species diversity in the metacommunity (Fig. 4). This can be inferred from Eq. 16, showing that the species diversity in the metacommunity increases as the migration rate reduces. The increase in the effective metacommunity size is because in the island model, each local community has the same demography and no variation exists among communities. The average sampling variance per species over all communities becomes smaller and the effective metacommunity size increases.

Note that local communities are all the same sizes in our island model. This is not the case in the real-world, nor are they equally connected by migration. A more complicated model that relaxes these conditions under neutrality is needed. However, the general expectation is that the qualitative relationships among speciation, migration, and ecological drift should remain the same as those explored here under the assumption of neutral community.

If we relax the neutral assumption and include the force of natural selection in formulating $C_{st}(\neq 0)$, then different outcomes are possible. When natural selection is very divergent among local communities, such as heterogeneous habitats due to the effects of environmental factors or due to varying interspecific interactions (Tilman 1982), the variation in the species richness and abundance distribution due to demographic divergence can be increased among communities. As a consequence, the sampling variance for the metacommunity decreases, compared to the results under neutrality (Eq. 4). In population genetics, it can be shown that the effective metapopulation size decreases in the presence of variation in fitness among populations (Whitlock and Barton 1997, Orrick and Fletcher 2005). This can also be implied from the relationship between effective metapopulation and local population sizes (Whitlock and Barton 1997): $N_{e(meta)} = nN_{e(local)}((1 + V)(1 - F_{st}) + 2N_{e(local)}F_{st}V)^{-1}$ (V is the variance in fitness among local populations, $N_{e(meta)}$ is the effective whole population size, and $N_{e(local)}$ is the effective subpopulation size). Under this case, the effective metacommunity size can be smaller than the sum of local effective community size, i.e. $J_m < LJ_l$, contrast to the neutral case ($J_m > LJ_l$; Eq. 12). This means, seemingly paradoxically, that the fundamental biodiversity number for the metacommunity can sometimes be less than the sum of all local biodiversity numbers, i.e. $\Theta < L\theta$. This situation can occur when there are differences in demography due to natural selection among communities.

On the other hand, when natural selection is convergent among all local communities, e.g. directional selection, variation due to demography divergence approximates to zero. Under this situation, community differentiation is mainly generated by ecological drift, and the same consequence as the solely neutral process can be predicted, i.e. $J_m > LJ_l$, as implied from Eq. 12. This result similar to this in population genetics is $N_{e(meta)} = N_{e(local)}/(1-F_{st})$ in the absence of divergence in natural selection among local populations ($V = 0$) (Wright 1943). Thus, a theoretical comparison between Θ and $L\theta$ provides an assessment whether divergent natural selection is involved in shaping community structure or not (neutral hypothesis: $\Theta \geq L\theta$; niche hypothesis: $\Theta < L\theta$; Hutchinson 1957). An important

theoretical challenge is to explore a method to estimate θ from the local communities, for which ecologists are much more likely to have data. One extreme case is that the metacommunity is not subdivided ($C_{st}=0$), and the relationship $\Theta=L\theta$ holds in the context of neutral community ecology.

Using C_{st} or β to describe community differentiation does not require neutral assumptions as long as all communities are at the same trophic level. We expect that a community scale-dependent property for the change in C_{st} or β may exist due to the effects of ecological drift. Small local communities result in a large degree of community differentiation. We also expect that empirically there will be a negative relationship between community differentiation and the distance separating communities due to dispersal limitation, as is implied by Eq. 14 for the case of finite number of local communities or Eq. 15 for the case of infinite number of local communities. Subsequent work should incorporate the heterogeneity in migration between long and short distance dispersal, with a larger probability for dispersal from nearby than far away communities. Such a negative relationship is also explicitly demonstrated in previous β -diversity studies in tropical forest with the continuous spatial model (Chave and Leigh 2002, Condit et al. 2002). The statistical test for such a negative relationship can inform us of the presence or absence of an isolation-by-distance effect. In addition, testing the effect of environmental factors (e.g. climate) on C_{st} or β can inform us about which factors do or do not contribute to community differentiation. This test does not require the neutral assumptions to be met, but such test is only a necessary step for demonstrating the absence of selection given the difficulty of examining all environmental factors.

Acknowledgements – We appreciate T. Wiegand and L. Svensson for useful comments and suggestions that improved the earlier version of this article. The work was supported by the Natural Sciences and Engineering Research Council of Canada (FH) and National Science Foundation (SPH).

References

- Brown, J. H. and Lomolino, M. V. 2000. Concluding remarks: historical perspective and the future of island biogeography theory. – *Global Ecol. Biogeogr.* 9: 87–92.
- Chave, J. and Leigh, E. G. 2002. A spatially explicit neutral model of β -diversity in tropical forests. – *Theor. Popul. Biol.* 62: 153–168.
- Cockerham, C. C. 1969. Variance of gene frequencies. – *Evolution* 23: 72–84.
- Cockerham, C. C. 1973. Analysis of gene frequencies. – *Genetics* 74: 679–700.
- Cockerham, C. C. and Weir, B. S. 1987. Correlation, descent measures: drift with migration and mutation. – *Proc. Natl Acad. Sci. USA* 84: 8512–8514.
- Condit, R. et al. 2002. Beta-diversity in tropical forest trees. – *Science* 295: 666–669.
- Economou, E. P. and Timothy, H. K. 2008. Species diversity in neutral metacommunities: a network approach. – *Ecol. Lett.* 11: 52–62.
- Etienne, R. S. 2005. A new sampling formula for neutral biodiversity. – *Ecol. Lett.* 8: 253–260.
- Etienne, R. S. 2007. A neutral sampling formula for multiple samples and an exact test of neutrality. – *Ecol. Lett.* 10: 608–618.
- Etienne, R. S. and Alonso, D. 2005. A dispersal-limited sampling theory for species and alleles. – *Ecol. Lett.* 8: 1147–1156.
- Fukami, T. et al. 2005. Species divergence and trait convergence in experimental plant community assembly. – *Ecol. Lett.* 8: 1283–1290.
- Hardy, O. J. and Senterre, B. 2007. Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. – *J. Ecol.* 95: 493–506.
- He, F. 2005. Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. – *Funct. Ecol.* 19: 187–193.
- He, F. and Hu, X.S. 2005. Hubbell's fundamental biodiversity parameter and the Simpson diversity index. – *Ecol. Lett.* 5: 386–390.
- Hu, X. S. et al. 2006. Neutral theory in macroecology and population genetics. – *Oikos* 113: 548–556.
- Hu, X. S. et al. 2007. Species diversity in local neutral communities. – *Am. Nat.* 170: 844–853.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. – Princeton Univ. Press.
- Hutchinson, G. E. 1957. Concluding remarks. – *Cold Spring Harbor Symp. Quant. Biol.* 22: 415–427.
- Inouye, R. S. and Tilman, D. 1988. Convergence and divergence of old-field plant communities along experimental nitrogen gradients. – *Ecology* 69: 995–1004.
- Johnson, M. T. J. and Stinchcombe, J. R. 2007. An emerging synthesis between community ecology and evolutionary biology. – *Trends Ecol. Evol.* 22: 250–257.
- Kembel, S. and Hubbell, S. P. 2006. The phylogenetic structure of a neotropical forest tree community. – *Ecology* 87: S86–S99.
- Kimura, M. 1983. The neutral theory of molecular evolution. – Cambridge Univ. Press.
- Kimura, M. and Crow, J. F. 1964. The number of alleles that can be maintained in a finite population. – *Genetics* 49: 725–738.
- Koleff, P. et al. 2003. Measuring beta diversity for presence-absence data. – *J. Anim. Ecol.* 72: 367–38
- Lande, R. et al. 2003. Stochastic population dynamics in ecology and conservation. – Oxford Univ. Press.
- Lepš, J. and Rejmánek, M. 1991. Convergence or divergence: what should we expect from vegetation succession? – *Oikos* 62: 261–264.
- Levins, R. 1968. Evolution in changing environment. – Princeton Univ. Press.
- Lewontin, R. C. 1972. The apportionment of human diversity. – *Evol. Biol.* 6: 381–398.
- MacArthur, R.H. 1965. Patterns of species diversity. – *Biol. Rev.* 40: 510–533.
- Magurran, A. E. 2004. Measuring biological diversity. – Blackwell.
- Malécot, G. 1969. The mathematics of heredity. Revised, edited and translated by D. M. Yermanos. – W. H. Freeman.
- McCune, B. and Allen, T. F. H. 1985. Will similar forests develop on similar sites? – *Can. J. Bot.* 63: 367–376.
- Munoz, F. et al. 2008. Beta diversity in spatially implicit neutral models: a new way to assess species migration. – *Am. Nat.* 172: 116–127.
- Nagyaki, T. 1974. The decay of genetic variability in geographically structured populations. – *Proc. Natl Acad. Sci. USA* 71: 2932–2936.
- Orrock, J. L. and Fletcher, R. J. 2005. Changes in community size affect the outcome of competition. – *Am. Nat.* 166: 107–111.
- Pielou, E. C. 1975. Ecological diversity. – Wiley.
- Routledge, R. D. 1977. On Whittaker's components of diversity. – *Ecology* 58: 1120–1127.
- Samuels, C. L. and Drake, J. A. 1997. Divergent perspectives on community convergence. – *Trends Ecol. Evol.* 12: 427–432.

- Sommer, U. 1991. Convergence succession of phytoplankton in microcosms with different inoculum species composition. – *Oecologia* 87: 171–179.
- Takahata, N. and Nei, M. 1984. F_{st} and G_{st} statistics in the finite island model. – *Genetics* 107: 501–504.
- Tilman, D. 1982. Resource competition and community structure. Monogr. Popul. Biol. – Princeton Univ. Press.
- Turner, M. G. 2005. Landscape ecology: what is the state of the science? – *Annu. Rev. Ecol. Evol. Syst.* 36: 319–344.
- Vallade, M. and Houchmandzadeh, B. 2003. Analytical solution of a neutral model of biodiversity. – *Phys. Rev. E.* 68: 061902–161905.
- Volkov, I. et al. 2003. Neutral theory and relative species abundance in ecology. – *Nature* 424: 1035–1037.
- Webb, C. O. et al. 2002. Phylogenies and community ecology. – *Annu. Rev. Ecol. Syst.* 33: 475–505.
- Weir, B. S. 1996. Genetic data analysis II. – Sinauer.
- Whitlock, M. and Barton, N. H. 1997. The effective population size with migration and extinction. – *Genetics* 146: 427–441.
- Whitlock, M. C. and McCauley, D. E. 1999. Indirect measures of gene flow and migration: $F_{st} \neq 1/(4Nm + 1)$. – *Heredity* 82: 117–125.
- Wilson, J. B. et al. 1995. An assembly rule in the ground and herbaceous strata of a New Zealand rain forest. – *Funct. Ecol.* 9: 61–64.
- Wright, S. 1943. Isolation by distance. – *Genetics* 28: 114–138.
- Wright, S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. – Univ. of Chicago Press.