



META-ANALYSIS

Ranked species occupancy curves reveal common patterns among diverse metacommunities

David G. Jenkins*

Department of Biology, University of Central Florida, Orlando, FL 32816-2369, USA

ABSTRACT

Aim Community ecologists often compare assemblages. Alternatively, one may compare species distributions among assemblages for macroecological comparisons of species niche traits and dispersal abilities, which are consistent with meta-community theory and a regional community concept. The aim of this meta-analysis is to use regressions of ranked species occupancy curves (RSOCs) among diverse metacommunities and to consider the common patterns observed.

Location Diverse data sets from four continents are analysed.

Methods Six regression models were translated from traditional occupancy frequency distributions (OFDs) and are distributed among four equation families. Each regression model was fitted to each of 24 data sets and compared using the Akaike information criterion. The analysed data sets encompass a wide range of spatial scales (5 cm–50 km grain, 2–7000 km extent), study scales (11–590 species, 6–5114 sites) and taxa. Observed RSOC regressions were tested for the differences in scale and taxa.

Results Three RSOC models within two equation families (exponential and sigmoidal) are required to describe the very different data sets. This result is generally consistent with OFD research, but unlike OFD-based expectations the simple RSOC patterns are not related to spatial scale or other factors. Species occupancy in diverse metacommunities is efficiently summarized with RSOCs, and multi-model inference reliably distinguishes among alternative RSOCs.

Main conclusions RSOCs are simple to generate and analyse and clearly identified surprisingly similar patterns among very different metacommunities. Species-specific hypotheses (e.g. niche-based factors and dispersal abilities) that depend on spatial scale may not translate to diverse metacommunities that sample regional communities. A novel set of three metacommunity succession and disturbance hypotheses potentially explain RSOC patterns and should be tested in subsequent research. RSOCs are an operational approach to the regional community concept and should be useful in macroecology and metacommunity ecology.

Keywords

Distribution, disturbance, macroecology, metacommunity, occupancy, presence, regional community, succession.

*Correspondence: David G. Jenkins, Department of Biology, University of Central Florida, Orlando, FL 32816-2369, USA. E-mail: dgjenkin@mail.ucf.edu

INTRODUCTION

Local community concepts have long dominated community ecology (Lawton, 1999; Ricklefs, 2008). Consequently, many ecologists have analysed within and among assemblages using familiar statistics (e.g. relative abundance, similarity indices, alpha and beta diversity; Magurran, 2004; Dornelas *et al.*, 2009). Analyses within or among assemblages in a species-by-sites

matrix (where rows = species and columns = sites) are called Q-mode analyses (Gotelli & Graves, 1996) and correspond to a figurative, vertical perspective of community ecology in which species 'pile up' in a site (Ricklefs, 2008). An unfortunate side-effect of a vertical, assemblage-centric focus is that the influence of regional processes (i.e. processes occurring at scales beyond the sites) on local assemblages has remained less clear, though community ecology is increasing its spatial and temporal scales

(e.g. Ricklefs & Schluter, 1993; Holyoak *et al.*, 2005; Ricklefs, 2008). As community ecology expands its spatial and temporal scales to approach those of biogeography, additional approaches are needed.

Ricklefs (2008) proposed a regional community concept that emphasizes populations distributed across ecological and geographical gradients. One approach to the regional community concept is to use an R-mode analysis, which compares species (rows in a species-by-sites matrix; Gotelli & Graves, 1996) distributed among localities and is orthogonal to a Q-mode analysis. In addition to the regional community concept, the R-mode perspective is consistent with studies of comparative ecology (e.g. Westoby *et al.*, 1996), macroecology (Brown, 1995), adaptive evolution (e.g. Martins, 2000) and niche (Chase & Leibold, 2003).

Species distribution data are recorded as presence and absence among multiple sites in either R- or Q-mode analyses. To consider both presence and absence is to analyse incidence (Rita & Ranta, 1993). Analyses that include absence require that absence is reliably demonstrated, which may be problematic for some purposes (Brotons *et al.*, 2004; Elith *et al.*, 2006; Lobo *et al.*, 2010). To focus on presence alone is to analyse occupancy or occurrence, analogous to ecological analyses of relative abundance (e.g. Dornelas *et al.*, 2009) in that absent species have no abundance. The term 'occupancy' is often applied to local- or regional-scale analyses of species distributions (Gotelli & Simberloff, 1987; Tokeshi, 1992; Collins & Glenn, 1997; McGeoch & Gaston, 2002) whereas 'occurrence' is often applied to range-scale analyses (Scott *et al.*, 2002). I use the term occupancy here.

Just as Q-mode analyses may be hierarchical (e.g. the Jaccard index analyses detail while species richness summarizes), R-mode analyses may also be viewed as detailed or summary. The details of R-mode pattern can be analysed by species co-occurrence (Gotelli & McCabe, 2002) and nestedness (Rodríguez-Gironés & Santamaria, 2006). To date, occupancy frequency distributions (hereafter OFDs) most closely approach a summary R-mode analysis. Multiple OFD shapes are possible (eleven by Tokeshi, 1992; four by Collins & Glenn, 1997; eight by McGeoch & Gaston, 2002) but have problems. As in all frequency distributions, data are sorted into categorical bins, and the choice of bin widths may alter interpretations (Gray *et al.*, 2006). An OFD is not information rich (*sensu* Tuft, 2001); if published in the absence of a species-by-sites matrix an OFD does not enable subsequent analyses of inter-specific differences, detailed change through time (e.g. invasive or endangered species, response to disturbance), meta-analyses (e.g. this work) or phylogenetic community analyses that require species identity (e.g. Webb *et al.*, 2002). Most importantly, OFDs can be difficult to analyse (McGeoch & Gaston, 2002). A method that yields results that are more readily analysed and more informative will help to fully analyse species occupancy data for regional community structure, spatial or temporal change, or the potential effects of phylogeny.

Ranked species occupancy curves

Here I apply a multi-model inference approach to regressions of empirical ranked species occupancy curves (RSOCs) as a new, efficient summary R-mode analysis that relates directly to OFDs (Fig. 1). The ranking of occupancy itself is not new; Willis (1922) ranked species by area, and ranked occupancy is a component of some modern diversity analyses (e.g. Gardezi & Gonzalez, 2008; Heino, 2008; Muneerpeerakul *et al.*, 2008). However, ranked occupancy has not been quantitatively examined for general patterns among diverse empirical systems. Moreover, ranked occupancy has not been related to OFDs and their associated hypotheses, let alone diverse other concepts.

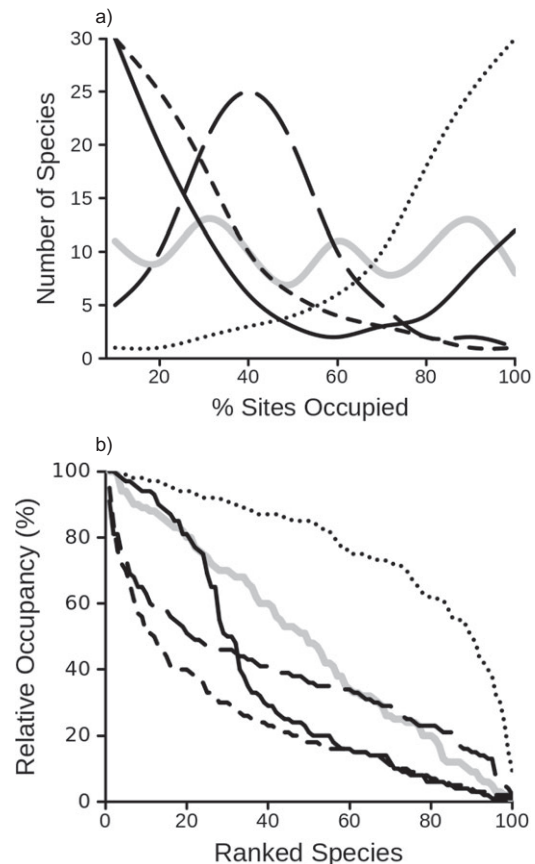


Figure 1 Occupancy frequency distributions (OFDs) translate to ranked species occupancy curves (RSOCs) (also see Table 1). (a) OFDs are histograms (shown here as lines for consistency with part b below), where species are binned into categories of the percentage of sites occupied. OFDs depicted here mimic those of Collins & Glenn (1997) plus a random OFD (McGeoch & Gaston, 2002), in which species were distributed among bins (mean = 10, SD = 2) to obtain multiple modes (see Table 1). The solid black line = bimodal-core dominant; long-dashed line = unimodal-central mode; short-dashed line = unimodal-satellite dominant; dotted line = unimodal-core dominant; grey line = random (based on the terminology of McGeoch & Gaston, 2002). (b) RSOCs are based on the same data as the OFDs in (a) but rank the species rather than assign them to histogram bins. Line patterns in (b) match those in (a).

Table 1 Ranked species occupancy curves (RSOCs), hierarchically listed by model family and equation. Each potential RSOC equation corresponds to a previously described occupancy frequency distribution (OFD), listed here by three major papers on OFDs.

RSOC model family	RSOC equation*	McGeoch & Gaston (2002)	Tokeshi (1992)§	Collins & Glenn (1997)
A. Exponential	Concave	Unimodal-satellite mode	Strongly unimodal (L) Unimodal (L) Weakly unimodal (L)	Regional distance scaling or Poor organismal dispersal scaling
	Convex	Unimodal-core mode	Strongly modal (R) Unimodal (R) Weakly unimodal (R)	Strong organismal dispersal scaling
B. Normal	Lognormal†	Unimodal-central mode		Moderate organismal dispersal scaling
C. Sigmoidal	Symmetric	Bimodal-symmetrical	Strongly bimodal Bimodal	
	Asymmetric (with long tail) Asymmetric (with long shoulder)	Bimodal-satellite mode dominant Bimodal-core mode dominant	Weakly bimodal (L) Weakly bimodal (R)	Local distance scaling
D. Linear	Linear; small, random residuals	Uniform	Uniform/other	
	Linear; large, random residuals	Random‡		

*Equations used here are defined below, where O_i is the occupancy of species i , R_i is the ranked order of species i , and y_0 , a , b and c are coefficients:

Exponential concave: $O_i = y_0 + a \exp(-bR_i)$ where initial parameters were $y_0 = 0.01$, $a = 1.0$, $b = 0.01$.

Exponential convex: $O_i = O_1 + y_0 [1 - a \exp(bR_i)]$ where initial parameters were $y_0 = 0.01$, $a = 1.0$, $b = 0.01$.

Lognormal function (Loehle & Hansen, 2005): $\ln(O_i) = aR_i + b$.

Sigmoidal symmetric (logistic function; Tjørve, 2003): $O_i = a/(1 + \exp(-bR_i + c))$, where initial parameters were $a = 1.0$, $b = -0.1$, $c = -1.0$.

Sigmoidal asymmetric (cumulative Weibull; Tjørve, 2003): $O_i = a[1 - \exp(-bR_i^c)]$ where initial parameters were $a = 1.0$, $b = -1.0$, $c = -1.0$. Tails and shoulders were evaluated graphically.

Linear: $O_i = aR_i + b$.

†Lognormal matched graphical depictions of unimodal-central (McGeoch & Gaston, 2002) and moderate organismal dispersal scaling (Collins & Glenn, 1997).

‡McGeoch & Gaston (2002) described a frequency distribution with variable and multiple modes as random (i.e. no unimodal or bimodal pattern), which should not be confused with a random process (e.g. dispersal) that may lead to a normal distribution around a central value and which is described by the lognormal equation.

§R, regional; L, local.

RSOCs translate directly from previously described OFDs (Fig. 1, Table 1), in much the same way that ranked species-abundance distributions are related to frequency distributions of abundance (e.g. MacArthur, 1957; Hubbell, 2001; McGill *et al.*, 2007). For example, an asymmetric bimodal OFD (solid line in Fig. 1a) translates to an asymmetric sigmoidal RSOC (solid line in Fig. 1b) because the large OFD mode of low-occupancy (i.e. satellite) species comprises species in the long RSOC tail, and the weaker OFD mode of high-occupancy (core) species comprises species in the upper RSOC shoulder. Likewise, a random OFD pattern (i.e. no clear modality; grey line in Fig. 1a) translates to a linear RSOC (grey line in Fig. 1b).

Each RSOC can be fitted to one of four standard model families (exponential, normal, sigmoidal, linear) that correspond to described OFDs (Table 1, Fig. 1). Finally, specific regression models within or among model families can be compared using multi-model inference (Anderson *et al.*, 2000). I emphasize comparisons of model families here, though sigmoidal models are analysed as either symmetric or asymmetric to match OFDs (Table 1, Fig. 1).

Multi-model inference of RSOC regressions permits definitive tests of alternative OFD-based hypotheses. For example,

OFDs at scales less than those of species geographic ranges are expected to be right-skewed (i.e. many species are infrequently observed), whereas bimodal OFDs are expected at the scale of species geographic ranges (McGeoch & Gaston, 2002; Hui & McGeoch, 2007a; Heino, 2008). I tested for the artefactual effects of spatial scale (spatial grain and extent; Wiens, 1989) and study scale (number of species, number of sites; Jenkins, 2006) on RSOCs. In addition to artefactual scale effects (Wilson, 2008), species-specific biological mechanisms (e.g. metapopulation processes, niche-related factors, dispersal ability) have also been invoked to explain OFDs (e.g. Hanski, 1982; Brown, 1984; Nee *et al.*, 1991; Tokeshi, 1992; Collins & Glenn, 1997; van Rensburg *et al.*, 2000; McGeoch & Gaston, 2002; Hui & McGeoch, 2007a,b). However, to invoke a biological mechanism for an OFD or RSOC is to assume that a species-specific effect (e.g. rapid dispersal) is consistent across all species in the metacommunity (i.e. multiple communities linked by dispersal; Leibold *et al.*, 2004). I treat a metacommunity here as a sample of a regional community, and I consider the assumption that species of a metacommunity are similar for a trait (e.g. dispersal) to be at odds with an analysis that compares diverse species with diverse dispersal potentials and life histories.

Table 2 Analysed data sets, sorted in increasing order of number of species. Spatial grain is the size of the sample (e.g. plot width) and spatial extent is the greatest distance between sample sites (Wiens, 1989). Grain and extent were estimated from information in cited papers or personal contact with study authors.

No.	Study system	Spatial grain (km)	Spatial extent (km)	No. of species	No. of sites	Data source
1	Bumblebees in montane meadows	0.15	50	11	12	Durrer & Schmid-Hempel (1995)*
2	Fishes in depressional wetlands	0.1	30	12	6	Snodgrass <i>et al.</i> (2000)
3	Monogenean parasites of <i>Loboe coubie</i>	0.0002	1500	13	35	Guégan & Hugueny (1994)
4	Small mammals in forest fragments	0.015	16	15	35	Nupp & Swihart (2000)
5	Fishes in beaver-influenced streams	0.1	12	16	23	Schlosser & Kallemeyn (2000)
6	Helminth parasites of <i>Sorus araneus</i>	0.0001	2	17	114	Haukialmi & Henttonen (1993)*
7	Cladocera in new ponds	0.000055	200	20	25	Louette & De Meester (2005)
8	Plants in urban fragments	0.05	27	22	423	Bastin & Thomas (1999)
9	Amphibians in depressional wetlands	0.1	30	25	22	Snodgrass <i>et al.</i> (2000)
10	Bumblebees in northern Spain	0.1	220	28	27	Obeso (1992)*†
11	Endemic plants in serpentine landscape	0.015	45	29	24	Harrison (1999)
12	Alien plants in serpentine landscape	0.015	45	33	24	Harrison (1999)
13	Helminth parasites of <i>Hydromys chrysogaster</i>	0.0004	1550	45	34	Smales & Cribb (1997)*‡
14	Herbaceous plants in secondary forest	5.6	27.5	49	10	Vellend (2004)
15	Ants in coastal southern California	0.02	43	49	40	Suarez <i>et al.</i> (1998)
16	Cumulative zooplankton in new ponds	0.01	0.112	61	12	Jenkins & Buikema (1998)
17	Macroinvertebrates in alpine ponds	0.01	1.6	62	25	Oertli <i>et al.</i> (2008)
18	Plants in US remnant prairies	0.0005	2700	64	63	Diamond & Smeins (1988)
19	Herbaceous plants in primary forests	5.6	27.5	70	17	Vellend (2004)
20	Pteridophytes in Amazonian rainforest	0.1	600	84	28	Tuomisto & Poulsen (1996)
21	Trichoptera in Danish streams	0.05	435	109	157	Wiberg-Larsen <i>et al.</i> (2000)
22	Mammals in Europe	50	4500	119	2183	Heikinheimo <i>et al.</i> (2007)
23	Plants in Konza Prairie	0.0035	5	165	1040	Hartnett PVC02 (2007)§
24	Breeding birds in USA and Canada	39.4	7141	590	5114	Breeding Bird Survey (2000)¶

*Also discussed in McGeoch & Gaston (2002). No. 1 was identified as a bimodal, core-dominated OFD, no. 6 as a bimodal OFD, no. 10 as either unimodal or lognormal and no. 13 as unimodal. As in Table 1 (above), these should correspond to: no. 1 sigmoidal, no. 6 sigmoidal, no. 10 either exponential or lognormal, no. 13 exponential, respectively.

†The most inclusive 1988 data were analysed here.

‡Table 1 data were analysed here.

§Konza Prairie LTER Data Set PVC02, David C. Hartnett, principal investigator. Data analysed were for transects A–D in 2007, all watersheds, all soil types, all treatments. Data obtained 23 July 2008 from <http://www.konza.ksu.edu/datasets/knzdsdetail.aspx?currMenu=0&datasetcode=PVC02>

¶The North American Breeding Bird Survey's C50y1 to C50y10 data sets for the year 2000 were analysed. Data represent all of Canada and the USA except Hawaii. Data obtained 2 May 2008 from <ftp://ftpext.usgs.gov/pub/er/md/laurel/BBS/DataFiles/>

METHODS

I collected 24 data sets from the peer-reviewed scientific literature and publicly available web sites (North American Breeding Bird Survey and Konza Prairie, KS, USA) to represent a variety of taxa, spatial scales and study scales (Table 2). Three data sets were split and then analysed to account for inherent differences in data. The data of Snodgrass *et al.* (2000) consisted of markedly different taxa (fish and amphibians) with different distributions in the study system (Table 2). Likewise, endemic plants were analysed separately from alien plants in the serpentine soils data set of Harrison (1999), and data for primary forests were analysed separately from data for secondary (i.e. former agriculture) forests in the data set of Vellend (2004). Of the 25 OFD studies reviewed by McGeoch & Gaston (2002), only four provided complete occupancy data in a form that could be extracted per species and thus enable comparison with McGeoch &

Gaston's review (Table 2) – a problem that confirmed the need for a more information-rich (Tufte, 2001) approach.

Each data set (species listed in rows and sites listed in columns) was handled as described below. Steps 1–4 can be followed for any given data set, while step 5 compares data sets and evaluates the effects of other variables.

1. Count the number of study sites in which each species was observed (occupancy). Calculate relative occupancy per species as (occupancy/total number of sites in a data set). This step ensures equitable comparisons among data sets; absolute occupancy may be suitable for analyses within any one data set. All analyses described below are based on relative occupancy.

2. Sort species by relative occupancy values, in decreasing order. The species observed in the most sites is ranked first and the species observed in fewest sites is ranked last. The plot of species relative occupancy as a function of rank is a RSOC.

3. Compute nonlinear and linear regressions of RSOCs (see Table 1 for equations). Note that the primary goal for regressions here was to determine which model family (i.e. exponential, sigmoidal, normal or linear) fits best. For example, a concave exponential equation may be expressed in different forms that vary slightly in fit, but this finer point is secondary to first discovering that a concave exponential is better than another family of models. Nonlinear regressions used the Levenberg–Marquardt algorithm (< 999 iterations) and both linear and nonlinear regressions were computed by ordinary least squares (OLS) in *SPSS* v.16. Because OLS regressions assume normality, homogeneity of variance and independent error terms, residuals were also evaluated graphically (Quinn & Keough, 2002).

4. Compute the Akaike information criterion (AIC) scores and weights (w_i ; Anderson *et al.*, 2000) for each regression equation on each data set. The regression equation with the greatest w_i represents the model that best retained information, adjusted for the number of parameters in the equation (range = 2–4 for equations in Table 1).

5. Evaluate curve families and models for differences in spatial scale (i.e. grain, extent; Wiens, 1989), study scale (number of species, number of sites; Jenkins, 2006), taxonomic groups and dispersal modes. Grain was estimated for each metacommunity as the linear distance of a sample unit (e.g. transect length, plot diameter), and extent was estimated as the greatest distance among sample sites, based on information in the analysed papers or other sources. Curve families and models were compared for differences in $\log_{10}(\text{grain})$, $\log_{10}(\text{extent})$, $\log_{10}(\text{number of species})$ and $\log_{10}(\text{number of sites})$ by analysis of variance (ANOVA); log-transformed data met parametric assumptions. Two of the 24 data sets were not clearly identified as fitting one model, so these two data sets were omitted from ANOVAs for clarity. The distributions of curve families and models among coarse taxonomic groups (plants, invertebrates, vertebrates) and dispersal modes (active or passive) were tested by χ^2 analyses. Statistical tests were computed with *SPSS* v.18.

RESULTS

Though the 24 data sets were quite diverse (Table 2), only exponential and sigmoidal RSOC models were observed (Table 3). Linear and lognormal models never represented data more effectively than exponential or sigmoidal models. Twenty-two of the 24 analysed data sets were clearly classified as either exponential (10 RSOCs) or sigmoidal (12 RSOCs); 2 of the 24 data sets were identified as being either an exponential or a sigmoidal RSOC (Table 3). Thus, about half of the data sets (sigmoidal RSOCs) were consistent with a bimodal OFD largely composed of core and satellite species, but about half of the data sets were consistent with a unimodal-satellite OFD, largely composed of satellite species. Two of the four data sets also analysed by McGeoch & Gaston (2002) did not correspond to their designated corresponding OFD shape (Table 3).

Within the exponential family of models, only concave RSOCs were observed; all 10 exponential data sets had a satellite mode. Sigmoidal models were divided into asymmetric (eight data sets) and symmetric distributions (six data sets; Table 3). Symmetry was strictly defined because it is required for the logistic equation, whereas asymmetry was typically apparent for data that were fitted best by the cumulative Weibull model because those data included numerous satellite species (which made a long tail in the distribution).

RSOCs were not significantly different for spatial scale and did not clearly differ for study scale. Data sets with exponential and sigmoidal RSOCs were not significantly different for spatial grain ($F_{1,20} = 0.03$; $P = 0.85$) or spatial extent ($F_{1,20} = 0.78$; $P = 0.39$). Likewise, data sets in the three observed models were not significantly different for spatial grain ($F_{2,19} = 0.37$; $P = 0.70$) or spatial extent ($F_{2,19} = 0.91$; $P = 0.42$). Data sets with exponential RSOCs had marginally significantly more species ($F_{1,20} = 3.9$; $P = 0.06$) than data sets with sigmoidal RSOCs, but this effect was due to the species-rich Breeding Bird Survey data set (when this one data set was removed from analysis the significant difference was removed). In addition, the three models were not significantly different for the number of species ($F_{2,19} = 2.0$; $P = 0.16$). The number of study sites was not significantly different between model families ($F_{1,20} = 0.39$; $P = 0.53$) or between the three models ($F_{2,19} = 0.29$; $P = 0.74$).

RSOCs were also not sorted by taxon or dispersal mode. Plants, invertebrates and vertebrates were randomly distributed among model families ($P = 0.82$, χ^2) and models ($P = 0.94$, χ^2), as were active and passive dispersers (model families $P = 0.90$; equations $P = 0.30$, both χ^2).

Two data sets were split as an example of comparisons that are possible with RSOCs. Endemic and alien species of serpentine soils (Harrison, 1999) had different RSOC curves (Table 3, Fig. 2). Endemic species were best described by an exponential curve while alien species were best described by an asymmetric sigmoidal curve with a long tail. Endemic and alien species were comparable in species richness (29 and 33, respectively) but alien species were more likely to be limited in distribution than endemic species in this system and at the time samples were collected.

Primary and secondary forest herbaceous plants (Vellend, 2004) also differed. Primary forest vegetation was best described by an exponential curve while secondary forest vegetation was best described by a symmetrical sigmoidal curve (Table 3, Fig. 3). Herbaceous vegetation in secondary forests had fewer total species and the RSOC declined more rapidly than in primary forest because multiple species were more narrowly distributed. In addition, many species changed position in their occupancy rank and relative occupancy values (Fig. 3). Mean change in relative occupancy was -0.14 (i.e. on average, forest herbs occupied 14% fewer sites in secondary forests than in primary forests), and 23 species may have been extirpated in secondary forests as a result of land-use practices (Fig. 3).

Table 3 Results summary for species occupancy distribution analyses. Numbered data sets are those listed in Table 2. All possible models (see Table 1) were analysed for each data set, but only models identified as retaining the most information by Akaike information criterion (AIC) are listed here. R^2 is the coefficient of determination from nonlinear regression models, AICc is the corrected AIC coefficient, and w_i is the Akaike weight of that model (Anderson *et al.*, 2000). Values of w_i indicate the probability that model i is the best model among those tested.

Model	No.	Study system	R^2	AICc	w_i
Exponential concave RSOCs (exponential concave function)					
	3	Monogenean parasites of <i>Lobos coubie</i> *	0.908	-59.9	0.373
	7	Cladocera in new ponds	0.991	-137.2	0.935
	9	Amphibians in depressional wetlands	0.976	-166.5	0.904
	10	Bumblebees in northern Spain	0.983	-196.7	0.885
	11	Endemic plants in serpentine landscape	0.965	-172.6	0.705
	13	Helminth parasites of <i>Hydromys chrysogaster</i> †	0.964	-392.5	0.492
	17	Macroinvertebrates in alpine ponds	0.975	-431.6	1.000
	19	Pteridophytes in Amazonian rainforest	0.990	-689.2	0.994
	20	Herbaceous plants in primary forests	0.988	-478.7	1.000
	21	Trichoptera in Danish streams	0.979	-805.3	0.999
	23	Plants in Konza Prairie	0.977	-1292.2	1.000
	24	Breeding birds in USA and Canada	0.988	-5508.3	1.000
Asymmetric sigmoidal RSOCs (cumulative Weibull function)					
	1	Bumblebees in montane meadows	0.970	-50.6	0.954
	2	Fishes in depressional wetlands	0.976	-64.5	0.964
	4	Small mammals in forest fragments	0.985	-84.4	0.990
	8	Plants in urban fragments	0.949	-135.1	0.921
	12	Alien plants in serpentine landscape	0.965	-204.9	0.969
	13	Helminth parasites of <i>Hydromys chrysogaster</i> †	0.959	-392.5	0.492
	15	Ants in coastal southern California	0.974	-323.5	1.000
	18	Plants in US remnant prairies	0.973	-436.5	1.000
Symmetric sigmoidal RSOCs (logistic function)					
	3	Monogenean parasites of <i>Lobos coubie</i> *	0.907	-59.8	0.347
	5	Fishes in beaver-influenced streams	0.934	-72.0	0.859
	6	Helminth parasites of <i>Sorus araneus</i>	0.980	-108.7	0.895
	14	Herbaceous plants in secondary forests	0.989	-328.3	0.992
	16	Cumulative zooplankton in new ponds	0.990	-401.6	1.000
	22	Mammals in Europe	0.99	-925.3	1.000

*This data set is listed in both exponential concave and symmetric sigmoidal based on its AIC w_i values.

†This data set is listed in both exponential concave and asymmetric sigmoidal based on its AIC w_i values.

Table 4 Hypothesized ecological mechanisms for occupancy frequency distributions (OFDs), which correspond to ranked species occupancy curves (RSOCs). Species-specific hypotheses have already been published, and some have been the subject of much research. Metacommunity hypotheses are novel to this work.

Occupancy pattern	Species-specific mechanisms	Metacommunity mechanisms
Unimodal satellite-dominant OFD and exponential RSOC	Niche-based local factors (Brown, 1984) Transient species (Nee <i>et al.</i> , 1991) Dispersal scaling (Collins & Glenn, 1997)	Early succession (recruitment limitation) and/or strong disturbance Late succession (strong biotic dominance) and/or low disturbance
Bimodal OFD and sigmoidal RSOC	Metapopulation processes (core/satellite, rescue effects; Hanski, 1982) Generalists and specialists (Brown, 1984)	Mid-succession (approaching quorum) and/or intermediate disturbance

DISCUSSION

Only two general shapes of RSOCs described 24 very different metacommunities. However, the consistent patterns were not

related to spatial scale as expected from hypotheses based on species-specific processes and that derive from OFDs. Processes operating at metacommunity scales (Table 4) may be required to explain the common patterns observed among diverse meta-

communities; a set of three hypotheses is proposed below. More practically, analysing RSOs by regression and multi-model inference is simple, adds to the existing method for analysing OFDs (Tokeshi, 1992) and readily permits a more information-rich (Tufté, 2001) outcome (e.g. Fig. 3).

The RSO regression approach used here accurately described 22 of 24 (92%) occupancy data sets, and the modest ambiguity observed for the remaining two data sets greatly sim-

plified OFD-based possibilities. The results obtained here are consistent with the observation that two OFD shapes are most common (McGeoch & Gaston, 2002). Beyond that simplicity, two sigmoidal RSO forms were detected, which helps delineate differences among data sets.

Spatial scale is expected to be important for OFD shape (McGeoch & Gaston, 2002; Hui & McGeoch, 2007a). Translated to RSOs, larger spatial scales should obtain exponential RSOs and smaller spatial scales should obtain sigmoidal RSOs. The data sets analysed here varied widely in spatial scale and study scale (Table 2) but exponential and sigmoidal RSOs were not significantly different for grain or extent (Wiens, 1989). Spatial scale depends in part on organismal dispersal scales (Collins & Glenn, 1997), which must differ widely in a diverse regional community (*sensu* Ricklefs, 2008) and are often not well quantified. Thus, heterogeneity in regional community dispersal scales may nullify scale effects expected by OFD theory and based on species-specific mechanisms. In addition, investigators already scale their studies to their study system; the spatial grain of migratory bird occupancy data necessarily exceeds that of data for intestinal parasites. Regardless, the fact that spatial scale and study scale did not differ among RSOs indicates that RSO analyses apply broadly to diverse study systems.

Three RSO models in two model families were observed, but linear, convex and lognormal models were *not* observed here. A linear model is consistent with uniform or random OFDs. Given the diversity of data sets analysed here, linear RSOs may be rarely observed, and then may be transient, such as early in the colonization history of similar habitats. A convex model is consistent with a metacommunity that approaches a single assem-

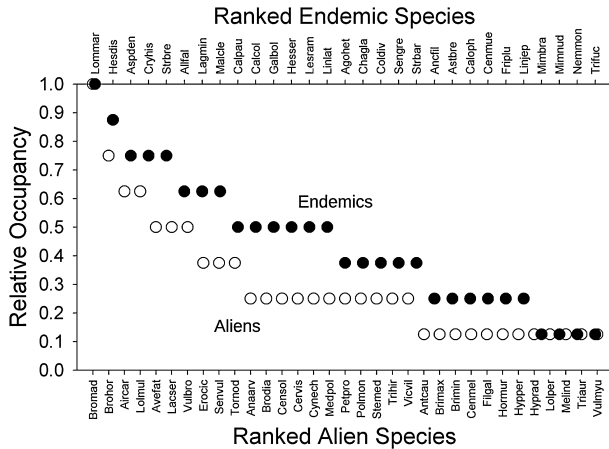


Figure 2 Ranked species occupancy curves (RSOCs) for data of Harrison (1999), where solid circles represent endemic plants and open circles alien plants. Endemic plants were best described by an exponential RSO, while alien plants were best described by an asymmetric sigmoidal RSO (also see Table 3). Species codes are the first three letters of the genus plus first three letters of the species name (see Tables 1 & 2 in Harrison, 1999, for full names).

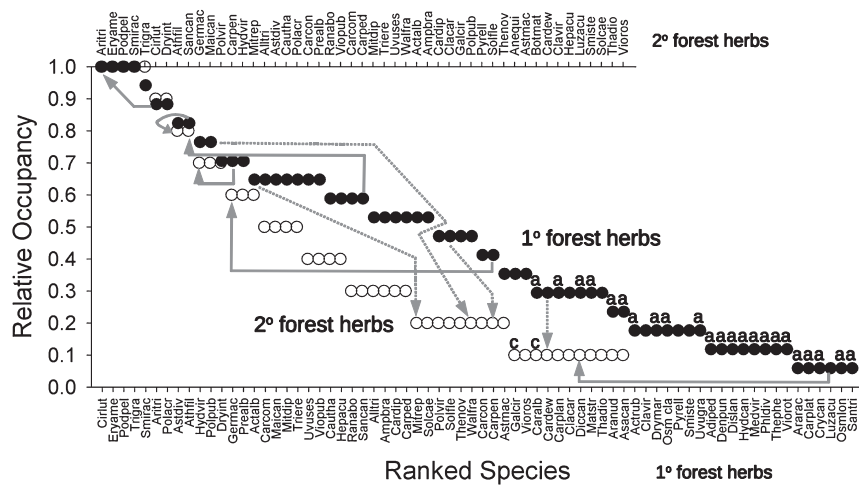


Figure 3 Ranked species occupancy curves (RSOCs) for data of Vellend (2004); solid circles represent primary forest herbaceous vegetation and were best described by an exponential RSO, open circles are secondary forest herbaceous vegetation and were best described by a symmetric sigmoidal RSO (also see Table 3). Solid grey arrows track selected species that increased in rank and/or relative occupancy between primary and secondary forests, whereas dotted grey arrows track selected species that decreased in rank and/or relative occupancy. Mean change in relative occupancy was -0.14 between primary and secondary forests. Species present in primary forest but absent in secondary forests are indicated with a letter a, and two species that apparently colonized secondary forests are indicated with a letter c. Species codes are first three letters of the genus plus the first three letters of the species name (see Vellend, 2004, Appendix C for full names).

blage dominated by multiple strong dispersers (Collins & Glenn, 1997; Kneitel & Miller, 2003). The absence of convex models among diverse data sets analysed here suggests that dispersal limitation and/or species sorting processes (Leibold *et al.*, 2004) are prevalent in multiple metacommunities. Convex RSOCs may also occur if distributions are actually sigmoidal RSOCs with truncated tails: that this was not the case indicates that the analysed data sets were sufficiently sampled to detect rare species in RSOC tails. Conversely, a convex RSOC observed in the future would indicate that the sampling regime be evaluated for its potential inability to detect rare species. Finally, the absence of lognormal distributions here indicates that no metacommunity was composed of species roughly similar in occupancy. Thus RSOCs would appear to reflect diverse historical and deterministic processes (Ricklefs & Schluter, 1993), operating at multiple spatial and temporal scales on multiple species. This outcome does not seem consistent with neutral theory (Hubbell, 2001), though neutral theory does not explicitly address occupancy patterns.

Two data sets (Harrison, 1999, and Vellend, 2004) were dissected for details. The difference between endemic and alien plant RSOCs in serpentine soils (Harrison, 1999) is consistent with an ongoing alien invasion, especially because endemics did not resist invasion (Harrison *et al.*, 2006). In the Vellend (2004) data set, logging and agriculture apparently extirpated multiple populations of herbaceous forest species and fragmented the remaining populations in subsequent secondary forests. This result suggests long-lasting effects on seed banks and/or growth conditions in secondary forests for herbaceous plants.

That only exponential and sigmoidal (asymmetric or symmetric) RSOCs were observed among diverse data sets begs explanation. Spatial scale has been hypothesized to drive OFD patterns (van Rensburg *et al.*, 2000; McGeoch & Gaston, 2002; Hui & McGeoch, 2007a; Heino, 2008) but was not a significant factor in RSOC analyses here. All ecological hypotheses on OFDs to date have focused on species-specific mechanisms (e.g. niche-based factors and dispersal abilities). However, an OFD or RSOC is an aggregate, emerging property of multiple species-specific mechanisms. Alternatively, hypotheses based on processes operating at the regional scales analysed by OFDs and RSOCs may be required to explain metacommunity patterns. Below I briefly outline a set of three metacommunity-scale hypotheses that may explain common patterns among diverse metacommunities. To be clear, observations reported above are not also used to 'test' the subsequent hypotheses in order to avoid post-hoc retrofitting and circular reasoning. Instead, hypotheses will best be tested with additional analyses of systems with known successional and disturbance regimes. Finally, the hypotheses proposed here represent alternative, conditional models that can serve as the basis for highly rigorous inference in subsequent tests (McGill *et al.*, 2006).

Metacommunity hypotheses for RSOC patterns

Three hypotheses represent major stages in a successional sequence or disturbance gradient (Fig. 4) in which RSOCs shift

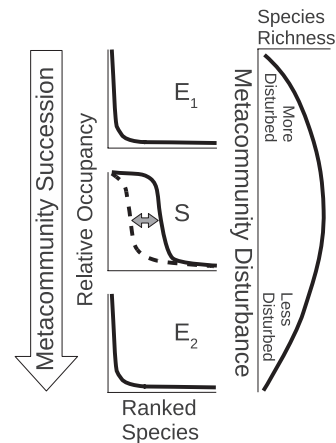


Figure 4 Hypotheses to explain exponential and sigmoidal RSOCs among diverse metacommunities. In metacommunity succession and disturbance regimes, the exponential RSOC (E_1) shifts to a sigmoidal RSOC (S) and then back to a second exponential RSOC (E_2). Within the sigmoidal RSOC lies a transition from an asymmetric sigmoidal RSOC (dashed line) toward the symmetric RSOC (solid line, if attained), and then back again to an asymmetric sigmoidal RSOC. See text for further explanation.

from exponential (E_1) to sigmoidal (S) and back to exponential (E_2). Hereafter the entire series is referred to as ESE, which itself is an overall hypothesis. The ESE hypotheses build on rich histories of research on succession and disturbance (e.g. Pickett & White, 1986; Glenn-Lewin *et al.*, 1992; Mackey & Currie, 2001; Svensson *et al.*, 2009). Succession and disturbance are interdependent in time, but succession is a temporal process, while persistent disturbance gradients can also be observed in space. Thus, the ESE series may be considered a successional sequence or a spatial disturbance gradient for a metacommunity. Disturbance in the ESE gradient is related to the intermediate disturbance hypothesis (Connell, 1978) which has been much-discussed but partially supported (Mackey & Currie, 2001), in part because 'disturbance' can have various meanings (Svensson *et al.*, 2009). Disturbance is defined here as an exogenous change in conditions which alters the occupancy of at least some species in a metacommunity, where spatial extent is standardized to the spatial extent of the metacommunity and so frequency (disturbance events per unit time) and severity (Σ change in occupancy) of the disturbance are important. The hypotheses below should apply most easily to synchronous succession and consistent disturbance across a metacommunity – heterogeneous succession and disturbance should promote the chance of E_1 and S RSOCs and reduce the chance of E_2 RSOCs.

An RSOC pattern *alone* cannot diagnose the mix of regional and local processes in a metacommunity because: (1) pattern does not prove process (Cale *et al.*, 1989) and (2) the ESE transition involves different mechanisms to obtain an exponential RSOC before (E_1) and after (E_2) a sigmoidal (S) RSOC in metacommunity succession. Teasing apart the processes requires detailed analyses of information beyond the summary RSOC

pattern. For example, a species in the RSOC tail may be dispersal limited (E_1) or it may be a fugitive species (Horn & MacArthur, 1972) with distributions driven by competitive interactions (E_2).

Hypothesis E1: recruitment limitation early in succession or strong/frequent disturbance causes an exponential RSOC

A metacommunity may be recruitment limited if spatial isolation of habitats exceeds organismal dispersal distances (Collins & Glenn, 1997) or if the metacommunity is early in succession. Recruitment limitation may also occur with large-scale disturbance, including the cumulative effects of multiple smaller-scale events (e.g. human land use; McKinney, 2002). Severe and/or frequent disturbance that is roughly consistent across the metacommunity should extirpate some species from some habitats and inhibit their recolonization of other habitats. Other species may resist or be resilient to the frequent and/or intense disturbance across multiple habitats. The result is an exponential RSOC (Fig. 4), where resistant/resilient species are most prevalent. If all species potentially access all habitats of the metacommunity, then any one assemblage represents a random sample of the disturbance regime and the metacommunity species pool. In that case, relative occupancy of a species represents the probability that the species will be sampled in any one assemblage and a local assemblage 'sample' from a metacommunity's exponential concave RSOC will obtain relatively low species richness because most species are regionally rare.

Hypothesis S: a mixture of local niche limitation and regional recruitment limitation at intermediate successional time and/or disturbance causes a sigmoidal RSOC

In this case, recruitment rates are sufficient through time and space to permit more species to colonize more habitats. This hypothesis assumes an intermediate successional interval and thus intermediate disturbance frequency, and assumes that metacommunity processes (e.g. species sorting; Leibold *et al.*, 2004) have become more important. The metacommunity thus shifts from regional recruitment limitation (E_1) through a mixture of local niche-based processes and neutral processes that may typify some metacommunities (Gravel *et al.*, 2006) toward local niche-based processes as primary drivers of occupancy patterns (i.e. a quorum effect; Jenkins & Buikema, 1998; Jenkins, 2006). An asymmetric (long tail) sigmoidal RSOC is hypothesized here as transitional between exponential and symmetric sigmoidal RSOCs (Fig. 4) and can range from nearly exponential to nearly symmetric to represent a broad band of transitional RSOC shapes. A symmetric sigmoidal RSOC (Fig. 4) represents an upper limit to an occupancy distribution for a metacommunity, given that an exponential convex RSOC would indicate a single assemblage rather than a metacommunity (Kneitel & Miller, 2003). An asymmetric sigmoidal RSOC may be obtained if keystone predation is consistently mitigated to a narrow sphere of influence (Brose *et al.* 2005) in trophic networks across the metacommunity. Metacommunities with predator-mediated competitive coexistence or strong habitat

heterogeneity should obtain a symmetric sigmoidal RSOC, given that these factors enhance the probability that more species can occupy more metacommunity habitats. A local sample from a sigmoidal RSOC will obtain greater species richness than from an exponential RSOC because more species are more regionally common to be 'sampled' by the local habitat.

Hypothesis E2: strong niche-based interactions late in succession and/or with weak/infrequent disturbance cause an exponential RSOC

Hypothesis E_2 relies on niche-based metacommunity theory, especially species-sorting processes such as competition and predation (Chase & Leibold, 2003; Leibold *et al.*, 2004) and should apply to metacommunities with little or no trophic cascade driven by *specialist* predators; otherwise predator-mediated coexistence among competitors throughout the metacommunity should contribute to a sigmoidal RSOC. Also, hypothesis E_2 is expected with strong, wide-ranging *generalist* predators and prey species that either resist (high occupancy) or are vulnerable (low occupancy) to predation. As for hypothesis E_1 , a local sample from an E_2 RSOC again obtains low species richness, but importantly, accumulated knowledge of species listed in the RSOC can diagnose E_1 versus E_2 . For example, slow-dispersing, competitive dominant species that are ranked highly in an exponential RSOC would be consistent with hypothesis E_2 but not hypothesis E_1 . If habitats are generally similar throughout the metacommunity, then more highly ranked competitors (Keddy & Shipley, 1989) should occupy more sites. Otherwise, habitat heterogeneity should disrupt a correlation between competitive rank and occupancy as well as the transition from a symmetric sigmoidal RSOC to asymmetric sigmoidal and then exponential RSOCs (Fig. 4).

Summary

Regressions of RSOCs improve on OFDs by retaining more detail while providing more definitive analyses of occupancy patterns. Thus, RSOCs should be useful in ecology and biogeography, including palaeontological and microbiological studies that often rely on occupancy data (e.g. Noguez *et al.*, 2005; Foote *et al.*, 2007). Outcomes of RSOC regressions reveal surprising commonality among very different metacommunities, but OFD-based expectations that spatial scale causes such commonality were not supported. Species-specific hypotheses already proposed to explain OFD patterns may not apply to metacommunities composed of diverse species; instead, metacommunity-scale hypotheses may be needed to explain commonality among metacommunities in occupancy distributions. Three hypotheses based on metacommunity succession and disturbance are proposed. Going forward, tests of these hypotheses will require metacommunities with documented successional and disturbance regimes, which should be readily available.

Regressions of RSOCs should be useful for multiple approaches in ecology and biogeography. For example, RSOC regressions provide a summary analysis to accompany detailed

R-mode analyses of occupancy patterns based on co-occurrence (Gotelli & McCabe, 2002) and nestedness (Rodríguez-Gironés & Santamaría, 2006). RSOC regressions are also consistent with macroecology (Brown 1995), metacommunity ecology (Holyoak *et al.*, 2005) and a regional community concept (Ricklefs, 2008). Ricklefs (2008) discussed a regional community concept in which a list of species occupy an average length v of an ecological gradient V . The probability v/V that a species occupies a site is the same as relative occupancy analysed here.

ACKNOWLEDGEMENTS

This work was improved by comments from Kim Medley, Pedro Quintana-Ascencio, Steve DeClerck, Joost Vanoverbeke, Brian McGill, David Currie and anonymous referees.

REFERENCES

- Anderson, D.R., Burnham, K.P. & Thompson, W.L. (2000) Null hypothesis testing: problems, prevalence, and an alternative. *The Journal of Wildlife Management*, **64**, 912–923.
- Bastin, L. & Thomas, C.D. (1999) The distribution of plant species in urban vegetation fragments. *Landscape Ecology*, **14**, 493–507.
- Brose, U., Berlow, E.L. & Martinez, N.D. (2005) Scaling up keystone effects from sample to complex ecological networks. *Ecology Letters*, **8**, 1317–1325.
- Brotans, L., Thuiller, W., Araújo, M.B. & Hirzel, A.H. (2004) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.
- Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124**, 255–279.
- Brown, J.H. (1995) *Macroecology*. University of Chicago Press, Chicago.
- Cale, W.G., Henebry, G.M. & Yeakley, J.A. (1989) Inferring process from pattern in natural communities. *BioScience*, **39**, 600–605.
- Chase, J.M. & Leibold, M.A. (2003) *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press, Chicago.
- Collins, S.L. & Glenn, S.M. (1997) Effects of organismal and distance scaling on analysis of species distribution and abundance. *Ecological Applications*, **7**, 543–551.
- Connell, J. (1978) Diversity in tropical rain forests and coral reefs. *Science*, **199**, 1302–1310.
- Diamond, D. & Smeins, F. (1988) Gradient analysis of remnant true and upper coastal prairie grasslands of North America. *Canadian Journal of Botany*, **66**, 2152–2161.
- Dornelas, M., Moonen, A.C., Magurran, A.E. & Barberi, P. (2009) Species abundance distributions reveal environmental heterogeneity in modified landscapes. *Journal of Applied Ecology*, **46**, 666–672.
- Durrer, S. & Schmid-Hempel, P. (1995) Parasites and the regional distribution of bumblebee species. *Ecography*, **18**, 114–122.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Foote, M., Crampton, J.S., Beu, A.G., Marshall, B.A., Cooper, R.A., Maxwell, P.A. & Matcham, I. (2007) Rise and fall of species occupancy in Cenozoic fossil mollusks. *Science*, **318**, 1131–1134.
- Gardezi, T. & Gonzalez, A. (2008) Scale dependence of species–energy relationships: evidence from fishes in thousands of lakes. *The American Naturalist*, **171**, 800–815.
- Glenn-Lewin, D.C., Veblen, T.T. & Peet, R.K. (eds) (1992) *Plant succession. Theory and prediction*. Chapman & Hall, London.
- Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology*. Smithsonian Institution Press, Washington, DC.
- Gotelli, N.J. & McCabe, D.J. (2002) Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology*, **83**, 2091–2096.
- Gotelli, N.J. & Simberloff, D. (1987) The distribution and abundance of tallgrass prairie plants: a test of the core-satellite hypothesis. *The American Naturalist*, **130**, 18–35.
- Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters*, **9**, 399–409.
- Gray, J.S., Bjørgesæter, A. & Ugland, K.I. (2006) On plotting species abundance distributions. *Journal of Animal Ecology*, **75**, 752–756.
- Guégan, J.F. & Hugué, B. (1994) A nested parasite species subset pattern in tropical fish: host as major determinant of parasite infracommunity structure. *Oecologia*, **100**, 184–189.
- Hanski, I. (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos*, **38**, 210–221.
- Harrison, S. (1999) Local and regional diversity in a patchy landscape: native, alien, and endemic herbs on serpentine. *Ecology*, **80**, 70–80.
- Harrison, S., Grace, J.B., Davies, K.E., Safford, H.D. & Viers, J.H. (2006) Invasion in a diversity hotspot: exotic cover and native richness in the Californian serpentine flora. *Ecology*, **87**, 695–703.
- Haukiskalmi, V. & Henttonen, H. (1993) Coexistence in helminths of the bank vole *Clethrionomys glareolus*. II. Intestinal distribution and interspecific interactions. *Journal of Animal Ecology*, **62**, 230–238.
- Heikinheimo, H., Fortelius, M., Eronen, J. & Mannila, H. (2007) Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *Journal of Biogeography*, **34**, 1053–1064.
- Heino, J. (2008) Temporally stable abundance–occupancy relationships and occupancy frequency patterns in stream insects. *Oecologia*, **157**, 337–347.
- Holyoak, M., Leibold, M.A. & Holt, R.D. (eds) (2005) *Metacommunities: spatial dynamics and ecological communities*. University of Chicago Press, Chicago.
- Horn, H.S. & MacArthur, R.H. (1972) Competition among fugitive species in a harlequin environment. *Ecology*, **53**, 749–752.

- Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Princeton University Press, Princeton, NJ.
- Hui, C. & McGeoch, M.A. (2007a) Modeling species distributions by breaking the assumption of self-similarity. *Oikos*, **116**, 2097–2107.
- Hui, C. & McGeoch, M.A. (2007b) A self-similarity model for the occupancy frequency distribution. *Theoretical Population Biology*, **71**, 61–70.
- Jenkins, D.G. (2006) In search of quorum effects in metacommunity structure: species co-occurrence analyses. *Ecology*, **87**, 1523–1531.
- Jenkins, D.G. & Buikema, A.L., Jr (1998) Do similar communities develop in similar sites?: a test with zooplankton structure and function. *Ecological Monographs*, **68**, 421–443.
- Keddy, P.A. & Shipley, B. (1989) Competitive hierarchies in herbaceous plant communities. *Oikos*, **54**, 234–241.
- Kneitel, J. & Miller, T. (2003) Dispersal rates affect species composition in metacommunities of *Sarracenia purpurea* inquilines. *The American Naturalist*, **162**, 165–171.
- Lawton, J.H. (1999) Are there general laws in ecology? *Oikos*, **84**, 177–192.
- Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.
- Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.
- Loehle, C. & Hansen, A. (2005) Community structure and scaling relations for the avifauna of the US Pacific and inland northwest. *Ecological Complexity*, **2**, 59–70.
- Louette, G. & De Meester, L. (2005) High dispersal capacity of cladoceran zooplankton in newly founded communities. *Ecology*, **86**, 353–359.
- MacArthur, R.H. (1957) On the relative abundance of bird species. *Proceedings of the National Academy of Sciences USA*, **43**, 293–295.
- McGeoch, M.A. & Gaston, K.J. (2002) Occupancy frequency distributions: patterns, artefacts and mechanisms. *Biological Reviews*, **77**, 311–331.
- McGill, B.J., Maurer, B.A. & Weiser, M.D. (2006) Empirical evaluation of neutral theory. *Ecology*, **87**, 1411–1423.
- McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I. & White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, **10**, 995–1015.
- Mackey, R.L. & Currie, D.J. (2001) The diversity–disturbance relationship: is it generally strong and peaked? *Ecology*, **82**, 3479–3492.
- McKinney, M.L. (2002) Urbanization, biodiversity, and conservation. *BioScience*, **52**, 883–890.
- Magurran, A.E. (2004) *Measuring biological diversity*. Blackwell, Oxford.
- Martins, E.P. (2000) Adaptation and the comparative method. *Trends in Ecology and Evolution*, **15**, 296–299.
- Muneepeerakul, R., Bertuzzo, E., Lynch, H.J., Fagan, W.F., Rinaldo, A. & Rodriguez-Iturbe, I. (2008) Neutral meta-community models predict fish diversity patterns in the Mississippi–Missouri basin. *Nature*, **453**, 220–222.
- Nee, S., Gregory, R.D. & May, R.M. (1991) Core and satellite species: theory and artefacts. *Oikos*, **62**, 83–87.
- Noguez, A.M., Arita, H., Escalante, A.E., Forney, L., García-Oliva, F. & Souza, V. (2005) Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecology and Biogeography*, **14**, 241–248.
- Nupp, T.E. & Swihart, R.K. (2000) Landscape-level correlates of small-mammal assemblages in forest fragments of farmland. *Journal of Mammalogy*, **81**, 512–526.
- Obeso, J.R. (1992) Geographic distribution and community structure of bumblebees in the northern Iberian peninsula. *Oecologia*, **89**, 244–252.
- Oertli, B., Indermuehle, N., Angélibert, S., Hinden, H. & Stoll, A. (2008) Macroinvertebrate assemblages in 25 high alpine ponds of the Swiss National Park (Cirque of Macun) and relation to environmental variables. *Hydrobiologia*, **597**, 29–41.
- Pickett, S.T. & White, P.S. (1986) *The ecology of natural disturbance and patch dynamics*. Academic Press, San Diego, CA.
- Quinn, G.P. & Keough, M.J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge.
- van Rensburg, B.J., McGeoch, M.J., Matthews, W., Chown, S.L. & van Jaarsveld, A.S. (2000) Testing generalities in the shape of patch occupancy frequency distributions. *Ecology*, **81**, 3163–3177.
- Ricklefs, R.E. (2008) Disintegration of the ecological community. *The American Naturalist*, **172**, 741–750.
- Ricklefs, R.E. & Schluter, D. (eds) (1993) *Species diversity in ecological communities: historical and geographical perspectives*. University of Chicago Press, Chicago.
- Rita, H. & Ranta, E. (1993) On analysing species incidence. *Annales Zoologici Fennici*, **30**, 173–176.
- Rodriguez-Gironés, A.M. & Santamaria, L. (2006) A new algorithm to calculate the nestedness temperature of presence–absence matrices. *Journal of Biogeography*, **33**, 924–935.
- Schlosser, I.J. & Kallemeyn, L.W. (2000) Spatial variation in fish assemblages across a beaver-influenced successional landscape. *Ecology*, **81**, 1371–1382.
- Scott, J.M., Heglund, P.J. & Morrison, M.L. (2002) *Predicting species occurrences*. Island Press, Washington, DC.
- Smales, L.R. & Cribb, T.H. (1997) Helminth parasite communities of the water-rat, *Hydromys chrysogaster*, from Queensland. *Wildlife Research*, **24**, 445–457.
- Snodgrass, J., Bryan, A., Jr & Burger, J. (2000) Development of expectations of larval amphibian assemblage structure in southeastern depression wetlands. *Ecological Applications*, **10**, 1219–1229.

- Suarez, A.V., Bolger, D.T. & Case, T.J. (1998) Effects of fragmentation and invasion on native ant communities in coastal southern California. *Ecology*, **79**, 2041–2056.
- Svensson, J.R., Lindgarth, M. & Pavia, H. (2009) Equal rates of disturbance cause different patterns of diversity. *Ecology*, **90**, 496–505.
- Tjørve, E. (2003) Shapes and functions of species–area curves: a review of possible models. *Journal of Biogeography*, **30**, 827–835.
- Tokeshi, M. (1992) Dynamics of distribution in animal communities: theory and analysis. *Researches on Population Ecology*, **34**, 249–273.
- Tufte, E.R. (2001) *The visual display of quantitative information*, 2nd edn. Graphics Press, Cheshire, CT.
- Tuomisto, H. & Poulsen, A.D. (1996) Influence of edaphic specialization on pteridophyte distribution in Neotropical rain forests. *Journal of Biogeography*, **23**, 283–293.
- Vellend, M. (2004) Parallel effects of land-use history on species diversity and genetic diversity of forest herbs. *Ecology*, **85**, 3043–3055.
- Webb, C.O., Ackerly, D.D., McPeck, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.
- Westoby, M., Leishman, M., Lord, J., Poorter, H. & Schoen, D.J. (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **351**, 1309–1318.
- Wiberg-Larsen, P., Brodersen, K., Birkholm, S. & Grøn & Skriver, J. (2000) Species richness and assemblage structure of Trichoptera in Danish streams. *Freshwater Biology*, **43**, 633–647.
- Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, **3**, 385–397.
- Willis, J.C. (1922) *Age and area: a study in geographical distributions and origin of species*. Cambridge University Press, Cambridge.
- Wilson, P.D. (2008) The pervasive influence of sampling and methodological artefacts on a macroecological pattern: the abundance–occupancy relationship. *Global Ecology and Biogeography*, **17**, 457–464.

BIOSKETCH

Dave Jenkins is interested in too many topics for his own good, but one theme is the analysis of pattern to indicate the relative roles of regional and local processes in ecology.

Editor: Brian McGill