Global Change Biology

Global Change Biology (2011) **17**, 3714–3723, doi: 10.1111/j.1365-2486.2011.02482.x

Nestedness patterns and the dual nature of community reassembly in California streams: a multivariate permutation-based approach

MARK NOVAK*, JONATHAN W. MOORE*† and ROBERT A. LEIDY‡

**Department of Ecology and Evolutionary Biology, Long Marine Laboratory, University of California, 100 Shaffer Rd., Santa Cruz, CA, 95060, USA, †Department of Biology, Simon Fraser, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada, ‡U.S. Environmental Protection Agency, 75 Hawthorne Street, San Francisco, CA, 94105, USA*

Abstract

Many factors contribute to the nonrandom processes of extinctions and invasions that are changing the structure of ecological communities worldwide. These factors include the attributes of the species, their abiotic environment, and the interactions and feedbacks between them. The relative importance of these factors has been difficult to quantify. We used nested subset theory and a novel permutation-based extension of gradient analysis to disentangle the direct and indirect pathways by which these factors affect the metacommunity structure of freshwater fishes inhabiting the streams tributary to the San Francisco Bay. Our analyses provide quantitative measures of how species and stream attributes may influence extinction vulnerability and invasion risk, highlight the need for considering the multiple interacting drivers of community change concurrently, and indicate that the ongoing disassembly and assembly of Bay Area freshwater fish communities are not fully symmetric processes. Fish communities are being taken apart and put back together in only partially analogous trajectories of extinction and invasion for which no single explanatory hypothesis is sufficient. Our study thereby contributes to the forecasting of continued community change and its effects on the functioning of freshwater ecosystems.

Keywords: body size, conservation, disturbance, extinction risk, freshwater stream fishes, invasion risk, metacommunity structure, multivariate gradient analysis, trait-based biogeography

Received 13 January 2011 and accepted 28 May 2011

Introduction

Natural and human-caused changes in the richness and composition of ecological communities are occurring by both species extinctions and species invasions, processes collectively referred to as community reassembly. Native community disassembly and nonnative community assembly are often nonrandom, but there remains great uncertainty as to the importance of their diverse potential drivers. Most hypotheses relate invasion success or extinction vulnerability either to the attributes of the species (e.g., the 'reckless invader' hypothesis) or their environment (e.g., the 'biotic acceptance' hypothesis), but few studies have considered these components of community change together (Catford *et al*., 2009). Furthermore, although many hypotheses consider the interactions between natives and nonnatives to be crucial determinants of community invasibility (e.g., the 'biotic resistance' hypothesis), few studies have investigated the potentially shared biotic and abiotic drivers of nonnative success and native vulnerability in unison (García-Berthou, 2007; Catford *et al*., 2009). Evidence suggests, however, that community reassembly is typically driven by the direct and indirect effects of multiple mechanisms acting simultaneously (e.g., Light & Marchetti, 2007).

Understanding the potentially complicated processes of community reassembly is a critical goal. Effective prioritization of conservation and restoration efforts requires disentangling the pathways by which altered environments and species assemblages interact to drive further community change (Olden *et al*., 2010). Furthermore, distinguishing the factors that influence extinction vulnerability and invasion success is a key for moving beyond random assembly experiments to illuminate the ecosystem consequences of ongoing biodiversity change (Gross *et al*., 2005; McIntyre *et al*., 2007; Zavaleta *et al*., 2009).

Case in point are the accelerating changes exhibited by freshwater fish communities (Dudgeon *et al*., 2006; Jelks *et al*., 2008) whose ecosystem-wide effects on primary productivity, food web structure, and nutrient dynamics are well known (Power, 1990; Schindler *et al*., 1997; Vanni, 2003). For example, previous analyses of local case histories in California suggest that nonnative establishment is

Correspondence: Mark Novak, tel. + 1 773 256 8645, fax + 1 831 459 3383, e-mail: mnovak1@ucsc.edu

driven primarily by the suitable matching of species physiological attributes and abiotic conditions, with success or failure being largely independent of the recipient community (Moyle & Light, 1996a,b). State-wide watershed-scale patterns of positively correlated native and nonnative richness have supported this inference (Marchetti *et al.*, 2004a,b). However, watershed-scale patterns also have suggested that the effects of altered abiotic conditions on native species are largely indirect, mediated less by environmental change *per se* than by the tendency of altered habitats to support nonnative species that consume or compete with natives (Light & Marchetti, 2007).

We employed nested subset theory to investigate the importance of the varied factors driving fish community structure in the streams tributary to the estuary of San Francisco Bay. This theory has long been recognized as a useful framework for identifying the mechanisms affecting changes in the composition of local communities (Patterson & Atmar, 1986) and has seen extensive application in the contexts of biogeography and conservation science (Fleishman *et al.*, 2007). We developed a novel permutation-based extension of nested subset gradient analysis that allowed us to disentangle the direct and indirect drivers of community composition by overcoming two limitations of previous analyses. Specifically, our method circumvents problems associated with rank order ties in explanatory variables and permits the explicit consideration of causal collinearity among putative drivers. It thereby enables the partitioning of the direct and indirect pathways by which both species and site attributes contribute to metacommunity structure, providing relative measures of their putative influence on the risks of extinction and invasion. Application of our approach to the fish communities of San Francisco Bay freshwater streams highlights the need for considering the multiple interacting drivers of community composition explicitly and concurrently and provides insight into the dual nature of community reassembly in this region.

Materials and methods

Nested subset theory

Nested metacommunity patterns occur when species present at species-poor sites are a subset of the species present at species-rich sites (site nestedness) or when the occurrences of species occupying few sites are a subset of the occurrences of species occupying many sites (species nestedness; Fig. S1). The presence of such patterns is interpreted as evidence for species- or site-specific variation in rates of colonization (invasion) and extinction (Lomolino, 1996; Taylor & Warren, 2001), with sites harboring fewer species, or species occupying fewer sites, inferred to experience higher rates of extinction or lower rates of colonization (Atmar & Patterson, 1993).

The typical goal of nestedness analysis is to not only establish whether a nested metacommunity pattern exists, but also to infer its potential drivers. This is done using gradient analysis, where the sites or species of the metacommunity matrix are ordered by each of many one-dimensional variables (or some multivariate ordination thereof) hypothesized to affect colonization or extinction rates (Ulrich *et al.*, 2009). The site or species trait that maximizes nestedness is considered the dominant driver of extinction or colonization (Ulrich *et al.*, 2009). For some metrics of nestedness [e.g., nestedness metric based on paired overlap and decreasing fill (NODF), Almeida-Neto *et al.*, 2008; see next], this procedure is equivalent to determining the site (or species) attribute exhibiting the strongest correlation to the species richness of sites (or the site occupancy of species). For other metrics, more indirect correlative inferences are required (e.g., matrix temperature; see Almeida-Neto *et al.*, 2007). Of course, other causes such as sampling intensity and cross-site differences in habitat quality also can contribute to nestedness (Cam *et al.*, 2000; Ulrich *et al.*, 2009) and may be particularly probable when sites cannot be treated as independent replicates of one another or when species occurrences are phylogenetically influenced. Useful methods for dealing with such issues have recently been proposed (e.g., Leibold *et al.*, 2010; Peres-Neto & Legendre, 2010), although not in the context of nested subset theory.

A significant limitation of the univariate approach to gradient analysis is an assumption that putative explanatory traits are themselves either uncorrelated or causally independent. Collinearity among predictors may be alleviated in part by calculating partial or semipartial correlations between predictor and response variables. However, predictors still are assumed to be causally independent. In many circumstances, such an assumption is unwarranted. For example, the oft-observed relationship between vulnerability and body size may be offset to an unknown degree by the counteracting relationship between body size and fecundity that may decrease a species' vulnerability (Reynolds *et al.*, 2005). Disentangling such correlations is facilitated by the explicit consideration of hypothesized causal relationships among all predictor and response variables.

A second limitation is introduced by the occurrence of ties in the rank ordering of an attribute. When such ties are present, as is commonly the case for species attributes such as trophic level or habitat affinity, the inferred explanatory power of the attribute may be considerably altered by the ordering of secondary attributes (Appendix S1). The effects of these secondary attributes remain hidden by univariate methods, potentially leading to incorrect inferences regarding the importance of putative reassembly drivers.

The method described next circumvents these limitations and enabled us to disentangle the manner by which a suite of species- and site attributes can explain the structure of a threatened freshwater fish community.

California fish communities

The region surrounding San Francisco Bay is a juxtaposition of degradation and diversity, exhibiting a rich fauna of fishes and high human population densities. The composition of stream fish assemblages was characterized at 275 sites within 23 watersheds tributary to the estuary from 1993 to 2004 (94% prior to 1998; Appendix S2). Sites were stratified to maximize the diversity of representative habitat types (i.e., riffle, run, pool) in different geomorphic settings (high-elevation, high-gradient, bedrock to low-elevation, and low-gradient unconsolidated substrate). Fishes were sampled primarily by single-pass electro-fishing, although in deeper (>1 m) or shallower (<5 cm) habitats, respectively, seines or dip nets and visual surveys also were employed. A minimum section of 30 m stream length was sampled at all sites. This typically included at least two pool-riffle-run units. Sections of >30 m length were surveyed when few or no fishes were collected within the first 30 m. Habitats within a reach were sampled with equal effort, although those immediately adjacent to stream banks often received more intensive sampling as they typically provided the most heterogeneous habitat (see Leidy, 2007 for further details).

Site attributes

Information on a total of 18 site-specific variables was obtained for each site. These included measures of physical, biotic, and water quality conditions (Leidy, 2007). We focused on 10 of these variables (Table 1, Fig. 1), discarding others due to their high correlation with included variables (e.g., water vs. air temperature), or because they were not measured at all sites, or because no specific hypothesis regarding their effect on species occupancy was proposed. Each site also was subjectively rated on the extent to which human activities had visibly altered the form and physical structure of the stream channel, water quality, and the riparian habitat. More specifically, this rating included assessments of the intactness of the riparian habitat, levels of siltation, turbidity and apparent pollution, changes in substrate, and the degree of stream channelization (see Leidy & Fiedler, 1985 for details).

Species attributes

We assembled information on 10 different species attributes hypothesized to affect extinction vulnerability and invasion success from the literature and public databases (Moyle, 2002; http://www.fishbase.org). These included attributes indicative of their ecology, reproductive potential, and physiological tolerance, and whether or not a species was native or nonnative to the streams of the estuary (Table 1, Fig. 1).

Nested subset analysis

To estimate metacommunity nestedness, we used the metric based on paired overlap and decreasing fill, NODF (Almeida-Neto *et al.*, 2008; Oksanen *et al.*, 2010). This metric is calculated from the percentage of species occurring in sites having lower richness that overlap with the species occurring at sites having higher richness and the percentage of sites occupied by species occupying few sites that overlap with the sites occupied by species occupying many sites. Thus, unlike other metrics, nestedness may not only be calculated for the whole incidence matrix (NODF_M), but for species (NODF_R) and sites (NODF_S) individually as well. The NODF metric is also less sensitive to matrix size and shape, and less prone to Type I error, than other commonly used metrics (Almeida-Neto *et al.*, 2008). Sites or species with equal species richness or occupancy do not contribute to the nestedness measured by NODF (Almeida-Neto *et al.*, 2008). Based on the NODF metric, nestedness is maximized when the columns and rows of the incidence matrix are ordered by descending marginal totals (i.e., decreasing richness and occupancy).

We also estimated the degree to which species exhibited checkerboard patterns in their co-occurrence using the *C*-score metric (Stone & Roberts, 1990; Almeida-Neto *et al.*, 2007; Oksanen *et al.*, 2010) because some observed levels of nestedness were less than expected by chance (see Results). Checkerboard patterns occur when two or more species occupy mutually exclusive sites and are considered indicative of negative interactions (e.g., competition or predation) when sites are equally suitable for all species. The *C*-score is the average number of checkerboards exhibited across all species pairs and is invariant with respect to matrix sorting.

We used the null model algorithm based on constrained fixed marginal totals implemented in Almeida-Neto & Ulrich (2011) with a sequential swap algorithm to determine the probability that nestedness and checkerboard estimates could be obtained by chance (Gotelli, 2000; Ulrich *et al.*, 2009). Simulations have suggested that this is the best-performing, although most conservative, of the available algorithms (Ulrich & Gotelli, 2007; Ulrich *et al.*, 2009). We used 100 000 swaps to minimize Type I and Type II error rates (Fayle & Manica, 2010) and estimated probabilities using 10 000 simulated matrices (Almeida-Neto & Ulrich, 2011).

Permutation gradient analysis

Our method for inferring the effects of species and site attributes on the nestedness of the metacommunity is founded on a permutation approach (Appendix S3). Let $\boldsymbol{R}$ denote the matrix of all species attributes and $\boldsymbol{C}$ the matrix of all site attributes. The rows of $\boldsymbol{R}$ are the species and those of $\boldsymbol{C}$ the sites, with their columns containing the respective attributes. The rows (r, species) and columns (c, sites) of the metacommunity incidence matrix are then randomly shuffled and the rows of $\boldsymbol{R}$ and $\boldsymbol{C}$ ordered correspondingly. This is repeated a large number of times (10 000 permutations). After each shuffle, we estimated NODF_M, NODF_R, NODF_C of the incidence matrix, as well as the rank order correlation of each attribute (as ordered in $\boldsymbol{R}$ or $\boldsymbol{C}$) with a number sequence of length equal to the number of species (r for the $\boldsymbol{R}$ attributes) or sites (c for the $\boldsymbol{C}$ attributes) in descending order. Rows and columns may be shuffled simultaneously because shuffling columns (or rows) has no effect on species (or site) nestedness. With a sufficient number of permutations the resulting distributions of attribute-specific rank order correlation coefficients becomes normally distributed regardless of the data-types or distributions of the original attribute values by virtue of the central limit theorem.

The relative contributions (β) of species and site nestedness to the nestedness of the overall metacommunity may be esti-

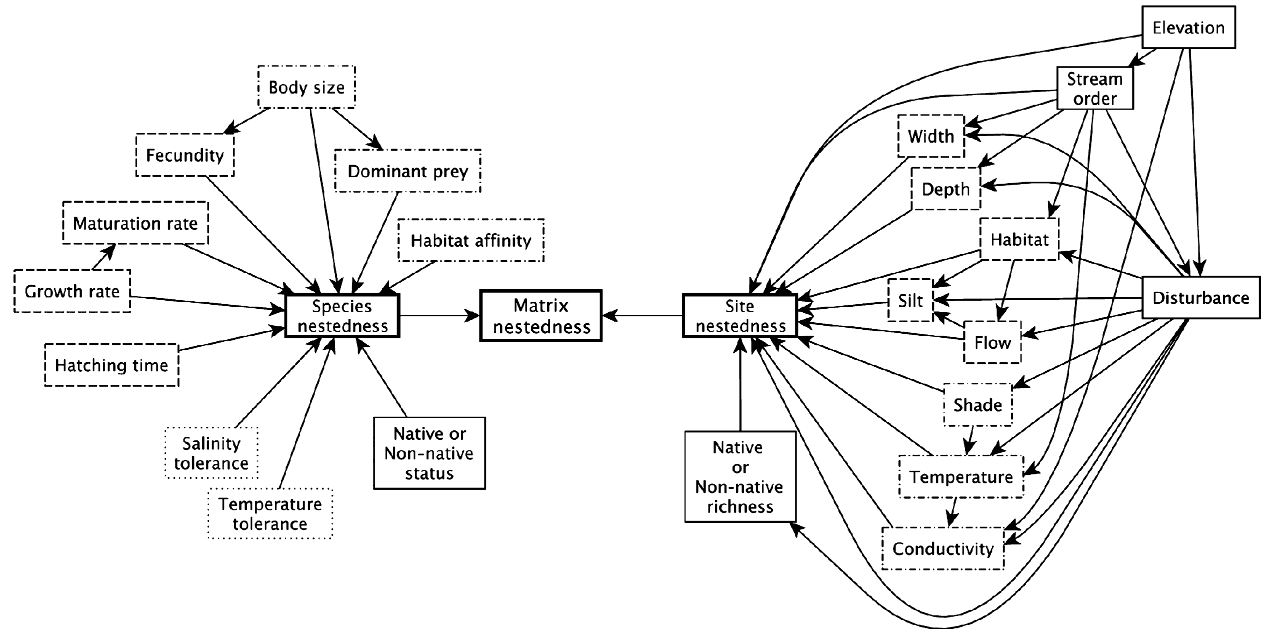

Fig. 1 The causal pathways, as specified using structural equations, by which species and site attributes were hypothesized to directly and indirectly affect the nested metacommunity structure of fishes in streams tributary to the San Francisco Bay.

mated from the resulting permutations by multiple regression ($\mathrm{NODF_M} \sim 0 + \beta_R \cdot \mathrm{NODF_R} + \beta_C \cdot \mathrm{N\ ODF_C}$). This is equivalent to

$$\beta_i = \frac{i(i-1)}{c(c-1)+r(r-1)},$$

where r is the number of species, c is the number of sites, and i is either r or c for estimating β_R or β_C, respectively.

Under an assumption of independence among attributes the effect of each species- (or site-) attribute on species (or site) nestedness may be estimated from the permutations by standard multiple regression (e.g., $\mathrm{NODF_R} \sim \beta_0 + \Sigma\beta_i\ x_i$). This assumption was not defensible for our dataset as many attributes were expected to have causal relationships (e.g., body size, diet, and fecundity). We therefore used structural equation models (path analysis) to explicitly account for these dependencies and thereby tease apart from the total effect the direct and indirect effects that each attribute had on nestedness (Fig. 1; Rosseel, 2010). Applied to the permutations, the resulting unstandardized direct effect coefficients (β_D) represent the maximum change in species (site) nestedness expected when the species (sites) of the incidence matrix are ordered by a focal attribute if there were no ties and the effects of all other attributes are removed. More specifically, they represent the change in nestedness expected when an attribute is reordered from being randomly ordered ($\rho = 0$) to fully ordered, with all other attributes held constant. Total effect coefficient (β_T) represents the change in nestedness expected when all attributes, excluding those downstream of the focal attribute's causal pathway, are held constant. For instance, positive coefficients reflect increases in nestedness when an attribute is ordered in descending order.

Native vs. nonnative gradient analysis

We repeated all analyses for native and nonnative species separately. To facilitate the comparison of checkerboard patterns exhibited by native and nonnative species, we standardized their *C*-scores by the number of sites each species-pair occupied,

$$\text{std. } C\text{-score} = \frac{\sum_{i,j}((c_i - c_{ij})(c_j - c_{ij}))/(c_i + c_j - c_{ij})}{r(r-1)/2},$$

where c_i and c_j are the number of species i and j's occurrences and c_{ij} is their number of co-occurrences. For the gradient analysis, we removed native status from the list of potential species attributes, and added the richness of natives (nonnatives) to the site attributes in the analysis of nonnative (native) species. The addition of the other group's richness as a putative explanatory variable allowed us to quantify the support for the biotic resistance and environmental acceptance hypotheses (Catford *et al.*, 2009).

Results

A total of 33 species were identified at 256 sites. Eight surveyed sites at which no fishes were documented, repeat surveys of the same site, those with missing attributes, and those performed in saline sloughs were removed prior to analysis. An unidentifiable sunfish individual also was removed prior to analysis. The resulting incidence matrix was composed of 15 native species observed at 253 sites and 18 nonnative species observed at 71 sites (Appendix S2).

Species of the metacommunity exhibited a stronger degree of nestedness than expected by chance (Table 2; $NODF_R = 26.4$, $P < 0.001$) indicating that infrequently observed species were present at a nonrandom subset of the sites occupied by species observed at many sites. In contrast, sites were less nested than expected by chance ($NODF_C = 42.4$, $P = 0.03$). As a result of the large ratio of surveyed sites to the total number of observed species, site nestedness dominated the signal of the whole incidence matrix ($\beta_R = 0.016$, $\beta_C = 0.984$) which therefore also exhibited less nestedness than expected by chance ($NODF_M = 42.1$, $P = 0.04$). Both species and sites exhibited significantly more checkerboarding than expected by chance (Table 2).

Our hypothesized causal model of species attributes explained 46.1% of the permutation-derived variation in species nestedness, whereas the model of site attributes explained 31.4% of the variation in site nestedness (Fig. 2). The attribute with the largest total effect on metacommunity nestedness was the distinction between native and nonnative species ($\beta_T = 9.8$); that is, ordering the rows of the metacommunity with native species above nonnative species increased nestedness by 9.8 units. Ordering of species by their temperature tolerance from highest to lowest ($\beta_T = 7.9$) and ordering of sites by their stream order from highest to lowest ($\beta_T = 6.7$) had the next largest total effect. As a result of indirect effects, the total effects of stream order and of body size ($\beta_T = -4.2$) were more than twice as large as their direct effects ($\beta_D = 2.8$ and -1.8, respectively). The latter occurred because body size exhibited positive relationships with a species' fecundity ($\beta_D = 0.59$) and dominant

Table 1 Site- and species-specific attributes included in the analysis of fish metacommunities in streams tributary to the San Francisco Bay

Attribute	Description
Site (column) attributes	
Native/ nonnative richness	Continuous; number of observed species
Elevation	Continuous; from digitized USGS 7.5′ scale topographic maps (m)
Stream order	Ordinal
Disturbance	Ordinal; integrated visual rating of direct anthropogenic impact (0. Pristine–5. Channelized); see Leidy & Fiedler (1985)
Stream width	Continuous; wetted channel (m)*
Stream depth	Continuous; mean water depth (cm)†
Dominant habitat	Nominal; 1. pool/2. pool-riffle/3. riffle*†‡
Flow rate	Continuous; discharge (cfs)
Fine sediment	Continuous; % substrate silt/clay/mud according to Wentworth particle-size scale*†‡
Riparian shade	Continuous; % wetted channel covered by a vertical projection of the riparian vegetation
Water temperature	Continuous (°C)
Conductivity	Continuous (μmho)
Species (row) attributes	
Native status	Binary; native or nonnative to San Francisco Bay tributaries§
Body size	Continuous; maximum body length (cm)§
Dominant prey	Ordinal; 1. vertebrates/2. vertebrates–invertebrates/3. invertebrates/4. invertebrates–algae/5. algae–detritus§
Habitat affinity	Nominal; primary water-column position: 1. benthic/2. benthic–suspension/3. suspension§
Fecundity	Continuous; maximum eggs/individual§
Growth rate	Continuous; maximum growth in first year (mm)§
Hatching time	Continuous; minimum time to hatch (days)§
Maturation rate	Continuous; mean age at first reproduction (years)§
Temperature tolerance	Continuous; maximum habitat temperature (°C)§
Salinity tolerance	Continuous; maximum habitat salinity (ppt)§

*Mean of 3–5 transects placed perpendicular to stream flow.

†Mean of 9–15 point estimates taken equidistantly along replicate transects.

‡Estimated visually within the 1 m^2 quadrat surrounding each sampling point.

§From Moyle (2002) and http://www.FishBase.org.

Table 2 Descriptive and null-model-derived statistics of the nested and checkerboard patterns exhibited by freshwater fishes in streams tributary to the San Francisco Bay

	Fill	Contribution*	NODF	*P*-value	*Z*-score†	*C*-score‡	*P*-value	*Z*-score†	Standardized *C*-score§
All species									
Matrix	0.099	–	42.14	0.04	−1.80	130 536¶	–	–	–
Sites	–	0.984	42.40	0.03	−1.86	247.23	<0.001	13.70	37.70
Species	–	0.016	26.44	<0.001	3.51	4.00	<0.001	14.07	4.58×10^{-2}
Natives only									
Matrix	0.186	–	44.84	<0.001	−3.61	73 324¶	–	–	–
Sites	–	0.997	44.87	<0.001	−3.64	698.32	<0.001	8.08	139.50
Species	–	0.003	35.47	0.16	1.00	2.30	<0.001	7.57	1.72×10^{-2}
Nonnatives only									
Matrix	0.103	–	25.48	0.03	1.47	4481¶	–	–	–
Sites	–	0.942	25.52	0.10	1.05	29.29	0.19	0.90	7.57
Species	–	0.058	24.74	0.05	1.64	1.80	0.18	0.89	9.63×10^{-2}

*The proportion of whole matrix nestedness ($NODF_M$) explained by nestedness among sites ($NODF_C$) and species ($NODF_R$).

†Z-score = $(x - \mu_{simul.})/\sigma_{simul.}$, where x is the observed NODF or C-score value, and $\mu_{simul.}$ and $\sigma_{simul.}$ are the mean and standard deviation, respectively, of 10 000 simulated matrices. Positive Z-scores indicate higher x than expected by chance, whereas negative Z-scores indicate lower x than expected by chance.

‡Site C-scores calculated on the transposed incidence matrix.

§C-score standardized by sites occupied (see main text).

¶Checkerboard units; the total number of 2-by-2 checkerboards occurring in the matrix (Gotelli, 2000).

prey ($\beta_D = 0.02$), which also had negative effects on species nestedness ($\beta_D = -4.01$ and -2.09, respectively).

The degree of species nestedness observed among native species had a 16% probability of occurring by chance alone (Table 2; $NODF_R = 35.5$, $P = 0.16$). Native-only sites, however, were significantly less nested than expected by chance alone ($NODF_C = 44.9$, $P < 0.001$). Again the large ratio of occupied sites to the total number of observed native species caused site nestedness to dominate the signal of the whole incidence matrix ($\beta_R = 0.003$, $\beta_C = 0.997$). Thus, the matrix was also less nested than expected by chance ($NODF_M = 44.8$, $P < 0.001$). Both native species and the sites they occupied exhibited significantly more checkerboarding than expected by chance (Table 2). In contrast, nonnative species exhibited significantly stronger matrix-wide nestedness than expected by chance (Table 2; $NODF_M = 25.5$, $P = 0.03$) because both species ($NODF_R = 24.7$, $P < 0.05$) and their occupied sites ($NODF_C = 25.5$, $P = 0.10$) exhibited nestedness ($\beta_R = 0.058$, $\beta_C = 0.942$). Neither the nonnative species nor their sites exhibited significant levels of checkerboarding (Table 2). Nevertheless, when standardized by the number of occupied sites, nonnative species exhibited considerably more checkerboarding among themselves than they did with native species and did native species among themselves (Table 2).

Our hypothesized causal model of species attributes explained 61.4% and 20.6% of the variation in species nestedness for native and nonnatives, respectively, whereas our model of site attributes explained 32.7% and 40.1% of the variation in site nestedness for natives and nonnatives, respectively (Fig. 3). Species attributes had considerably larger effects on native species than on nonnatives, with the ordering of native species by their temperature tolerance ($\beta_T = 19.1$), fecundity ($\beta_T = -15.3$), body size ($\beta_T = -9.36$), and time to hatch ($\beta_T = 7.6$) having the largest total effects. The total effect of body size on native species nestedness was opposite in sign to its direct effect ($\beta_D = 2.51$) due to its indirect effects via dominant prey and habitat affinity. The ordering of sites by their stream order ($\beta_T = 7.8$), conductivity ($\beta_T = -4.9$), and disturbance ($\beta_T = -4.8$) had the largest total effects on native nestedness. The ordering of sites by conductivity ($\beta_T = 6.2$) and disturbance ($\beta_T = 4.1$) also had large total effects on nonnative nestedness, but their effect was positive rather than negative. Stream depth had a large and positive total effect on both natives ($\beta_T = 4.2$) and nonnatives ($\beta_T = 4.8$). Richness of natives and nonnatives had only weak effects on the other's nestedness ($\beta_T = 0.2$ and 1.7, respectively). While few site attributes had their direct effects on nestedness reversed via indirect effects, some attributes (e.g., stream order and elevation) had total effects that were almost twice as large, or twice as small, as their direct effects.

Our results were robust to seasonal variation in community structure and potential biases associated with <100% detection probabilities (Appendix S4). The quantitative importance of attributes as inferred by our

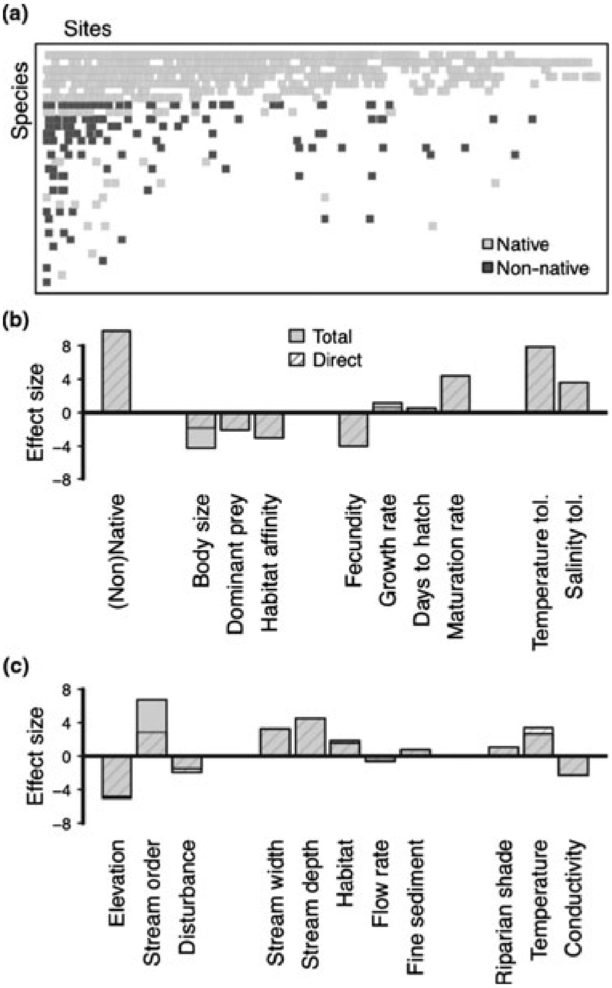

Fig. 2 (a) The incidence matrix of stream fish metacommunity, ordered by occupancy and richness, highlighting the importance of distinguishing between native and nonnative species. Below, total (β_T) and direct (β_D) effects of (b) species- and (c) site attributes on the species- and site nestedness of the metacommunity as revealed by the explicit consideration of their causal correlations using permutation-based gradient analysis. See Table 1 for variable explanations. Effect sizes indicate the change in nestedness expected when rows or columns are ordered by a focal variable in descending order, accounting for the ordering of secondary variables.

approach corresponded little with those of standard univariate analyses (Appendix S5)

Discussion

Attributes of extinction and invasion risk

Our application of the permutation-based gradient analysis to the fish communities of streams tributary to the San Francisco Bay underscores how an integrated view of community structure, and by inference of species extinctions (community disassembly) and invasions (community assembly), is necessary. Of all the hypotheses reviewed by Catford *et al.* (2009) and Olden *et al.*(2010), no single hypothesis was consistent with all the important drivers of extinction or invasion risk inferred by our analysis (Fig. 3).

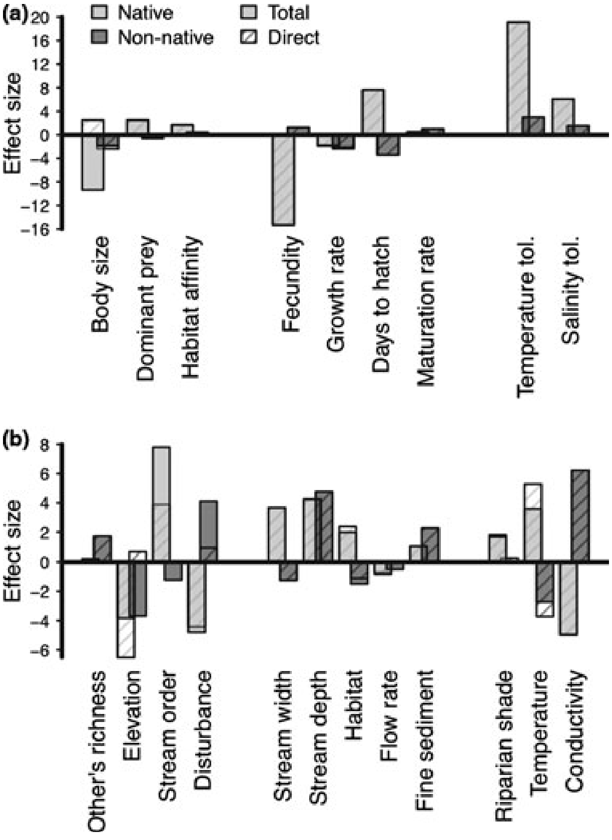

Fig. 3 The total and direct effects of (a) species- and (b) site attributes on the nested community structure of native and nonnative fishes as revealed by the explicit consideration of their causal correlations using the permutation-based gradient analysis.

For example, supporting the ‘biotic acceptance’ hypothesis that posits that the same environmental factors that reduce native vulnerability promote nonnative invasion, elevation and stream depth both had similarly strong associations with natives and nonnatives. This is to say that lower elevation and deeper stream reaches corresponded to both decreased native vulnerability and increased nonnative invasion risk, given our hypothesized causal model. However, while high order streams were also associated with reduced native vulnerability through both direct and equally large indirect effects, they were no more at risk of being invaded than were low order streams. Furthermore, while disturbance, conductivity, and temperature all had associations of similar magnitude for both natives and nonnatives, the direction of their effects on natives and nonnatives nestedness were of opposite sign – more disturbed reaches and those with higher conductivity were associated with increased native vulnerability and increased nonnative invasion risk, and those with higher temperatures were associated with decreased native vulnerability and decreased nonnative invasion risk.

Patterns of site nestedness suggest no support for the 'biotic resistance' hypothesis that posits that communities with many native species show reduced risk of nonnative invasion. Indeed the opposite was true, with sites having higher native richness showing increased nonnative nestedness than sites with lower native richness, consistent with other studies demonstrating positive spatial correlations in the richness of native and nonnative fishes (Marchetti *et al.*, 2004a). The observation that standardized checkerboard scores were lower between natives and nonnatives than within nonnatives themselves (Table 2, Fig. S3) similarly suggests a relatively low influence of negative species interactions (Baltz & Moyle, 1993; Moyle & Light, 1996b). Among the species' attributes contributing to metacommunity patterns, fecundity and temperature tolerance exhibited the strongest associations with natives (Fig. 3a). Higher temperature tolerance thus corresponded to decreased native vulnerability, consistent with the effect of stream temperature on their site nestedness. Temperature tolerance was not an important attribute for nonnative species.

The weak effect of fecundity on nonnative species suggests little support for the 'propagule pressure' hypothesis (Catford *et al.*, 2009). In fact, despite fecundity's large inferred effect on natives, this hypothesis' underlying principle – which posits that a high supply or frequency of offspring increases invasion success – also was not supported for native species. Higher fecundity was associated with an increase in the inferred vulnerability of natives, presumably due to an unconsidered correlated trait. The inferred strength of fecundity's negative effect on vulnerability affected a net reversal of body size's direct effect to decrease vulnerability. The strong positive association of body size with fecundity meant that its indirect effect on nestedness was negative and stronger than its positive direct effect. Only because of this indirect effect was large body size inferred to exhibit a positive association with native vulnerability, as is typically observed (Reynolds *et al.*, 2005).

More generally, given our assumed causal model of between-attribute relationships (Fig. 1), the power of species attributes to explain the metacommunity structure of native and nonnative species differed remarkably. Nonnative metacommunity structure was weak (Table 2) and far less well explained by their species attributes than was the structure of natives (Fig. 3a). This observation is consistent with the idiosyncratic nature of characteristics associated with successfully invading nonnative fishes in California (Marchetti *et al.*, 2004a,c) and around the globe (Blanchet *et al.*, 2009), and suggests that nonnative life-histories may play less of a role in their success than do other factors influencing the choices of humans in their selective introduction of fishes (see also Dill & Cordone, 1997; Moyle, 2002). The high standardized checkerboard scores observed among nonnatives (Table 2, Fig. S3) may similarly have more to do with the consequences of these human-made choices than with the competitive interactions between species. Understanding the control and establishment of nonnative fishes will thus entail integrating sociology with the study of fish biology.

Implication for nestedness analysis

Our analyses highlight the necessity of distinguishing the species patterns of site occurrence from the site patterns of species richness in the study of metacommunity structure. The shape of the metacommunity (the number of species observed vs. the number of sites surveyed) exhibited strong control over the inferred contribution of species and sites to the nestedness of the entire metacommunity. For example, the strong nested pattern observed among species of the entire fish metacommunity would have been masked by the lower than expected degree of nestedness observed among sites due to the high ratio of sites to species in this dataset (Table 2). It is therefore unlikely that a gradient analysis utilizing a metric of matrix-wide metacommunity nestedness such as matrix temperature (Atmar & Patterson, 1993) would have discerned the importance of distinguishing native and nonnative species in explaining metacommunity structure (Fig. 2a and b). The importance of disturbance, conductivity, and days-to-hatch would thus have remained obscured by the counteracting sign of their effects on native and nonnative species (Figs 2 and 3). We suggest that future analyses of metacommunity structure continue to use measures of nestedness, such as the NODF metric, which can distinguish between the row- and column-nestedness of a metacommunity.

Assumptions and limitations

Our analysis did not account explicitly for potential autocorrelation in species and site attributes associated with their shared evolutionary histories or lack of spatial independence. It is possible, for example, that phylogenetically controlled traits not included in our analysis are driving the counter-intuitive effect of fecundity on native nestedness, although visual inspection of the data suggests that this is not the case. Furthermore, our use of the NODF metric affected a discounting of site pairs with matching species sets

(Almeida-Neto *et al.*, 2008). This served to reduce the influence of first-order spatial autocorrelation potentially introduced by the nonrandom sampling of the region's environments. We therefore consider these potential sources of biased inference to be weak.

A further nontrivial limitation of our permutation-based approach to gradient analysis also remains. The interpretation of permutation-based attribute effects is less intuitive than the interpretation of rank order correlation coefficients between traits and the incidence matrix's marginal totals (richness and occupancy). Both types of approaches rely on an assumption that species and site attributes exhibit monotonic relationships to nestedness in the case of the permutation-based approach, or to richness and occupancy in the case of simple correlative approach, although for either approach these relationships may be nonlinear. The counteracting benefits of the permutation-based approach, in addition to those described above, are that (i) traits are related to nestedness structure directly rather than via richness, occupancy, or the species order of the maximally packed matrix that are themselves only correlates of nestedness, (ii) it requires no *post hoc* adjustment to control for Type I errors introduced by multiple comparisons of the same data, and that (iii) putative explanatory traits need not be continuous variables and may exhibit a wide variety of nonnormal distributions. Correlations between nestedness and the ordering of each trait derived from the permutation-based approach are normally distributed regardless of the trait's individual distribution and thereby satisfy the assumptions of subsequent parametric analysis (e.g., structural equation modeling) without additional transformation or the use of generalized linear models.

Ultimately, of course, the utility of nestedness theory rests on statistical correlations. Many underlying assumptions are often not met (Atmar & Patterson, 1993). Indeed, the most critical assumption unaddressed by our study is that spatial patterns of species occurrences do in fact reflect the intrinsically temporal processes of community reassembly. Analyses relying on snapshot surveys accumulated across sites may fail to reflect the true temporal dynamics of species or the traits to which they are compared (Donlan *et al.*, 2005). As a consequence, the true underlying drivers and consequences of community change can remain obscured (White & Kerr, 2006). The robustness of the relationships between nestedness patterns and temporal rates of extinction and invasion has itself only recently seen some support in fish and plant communities (Taylor & Warren, 2001; Elmendorf & Harrison, 2009); additional temporal tests of nested subset theory are needed (Donlan *et al.*, 2005). Procedures to incorporate temporal perspectives in the analysis of biogeography and macroecology are increasingly available, but their application remains largely rare due to limited temporal data (Kerr *et al.*, 2007; Fisher *et al.*, 2010). Thus, as in most nonexperimental studies, our inferred effects of species and site attributes on extinction and invasion risk should be cautiously treated as hypotheses with which to refine further investigations (Fleishman *et al.*, 2007).

Conclusions

In their synthesis of biogeography's relevance to the conservation of freshwater fishes, Olden *et al.* (2010) advised for the trait-based quantification of extinction and invasion risk and a recognition of the interactive effects of multiple stressors in freshwater ecosystems. Methods that advance our ability to address these needs are equally necessary in other ecosystems and taxonomic groups (Catford *et al.*, 2009; Sutherland *et al.*, 2009). Our permutation-based analyses allow such an integration of the multiple potential drivers of community structure and, by inference, community reassembly. By allowing the explicit consideration of causal collinearity among putative drivers, our method enables the partitioning of the direct and indirect pathways by which both species traits and environmental variables contribute to metacommunity structure. It also provides standardized inferences regarding each attribute's influence on this structure. Such measures are needed to progress beyond the forecasting of best- and worst-case scenarios of biodiversity change and how they will affect the functioning of ecosystems (Zavaleta *et al.*, 2009). Our analysis of San Francisco Bay freshwater stream fishes suggests that their communities are being taken apart and put back together with both shared and nonshared rules of disassembly and assembly. Future work is needed to determine how strongly the attributes inferred to confer native vulnerability and nonnative invasion risk correlate with the attributes that contribute to the varied ecosystem processes that fishes affect (Gross *et al.*, 2005).

Acknowledgements

We thank P. Guimarães Jr. and J. Byrnes for their insights, R. Thill and R. Terry for assistance in compiling the databases, P. L. Fiedler and three anonymous reviewers for comments on the manuscript, and the numerous individuals who assisted with the field surveys. MN acknowledges the support of the USGS Western Ecological Research Center and the freshwater ecology research group at UCSC.

References

Almeida-Neto M, Ulrich W (2011) A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software*, **26**, 173–178.

Almeida-Neto M, Guimarães PR Jr, Lewinsohn TM (2007) On nestedness analyses: rethinking matrix temperature and anti-nestedness. *Oikos*, **116**, 716–722.

Almeida-Neto M, Guimarães P, Guimarães PR Jr, Loyola RD, Ulrich W (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239.

Atmar W, Patterson B (1993) The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, **96**, 373–382.

Baltz DM, Moyle PB (1993) Invasion resistance to introduced species by a native assemblage of California stream fishes. *Ecological Applications*, **3**, 246–255.

Blanchet S, Leprieur F, Beauchard O, Staes J, Oberdorff T, Brosse S (2009) Broad-scale determinants of non-native fish species richness are context-dependent. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2385–2394.

Cam E, Nichols JD, Hines JE, Sauer JR (2000) Inferences about nested subsets structure when not all species are detected. *Oikos*, **91**, 428–434.

Catford JA, Jansson R, Nilsson C (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions*, **15**, 22–40.

Dill WA, Cordone AJ (1997) History and status of introduced fishes in California, 1871-1996: conclusions. *Fisheries*, **22**, 15–18.

Donlan CJ, Knowlton J, Doak DF, Biavaschi N (2005) Nested communities, invasive species and holocene extinctions: evaluating the power of a potential conservation tool. *Oecologia*, **145**, 475–485.

Dudgeon D, Arthington AH, Gessner MO *et al.* (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, **81**, 163–182.

Elmendorf SC, Harrison SP (2009) Temporal variability and nestedness in California grassland species composition. *Ecology*, **90**, 1492–1497.

Fayle TM, Manica A (2010) Reducing over-reporting of deterministic co-occurrence patterns in biotic communities. *Ecological Modelling*, **221**, 2237–2242.

Fisher JAD, Frank KT, Leggett WC (2010) Dynamic macroecology on ecological timescales. *Global Ecology and Biogeography*, **19**, 1–15.

Fleishman E, Donnelly R, Fay JP, Reeves R (2007) Applications of nestedness analyses to biodiversity conservation in developing landscapes. *Landscape and Urban Planning*, **81**, 271–281.

García-Berthou E (2007) The characteristics of invasive fishes: what has been learned so far? *Journal of Fish Biology*, **71**, 33–55.

Gotelli NJ (2000) Null model analysis of species co-occurence patterns. *Ecology*, **81**, 2606–2621.

Gross K, Cardinale BJ, Bascompte J (2005) The functional consequences of random vs. ordered species extinctions. *Ecology Letters*, **8**, 409–418.

Jelks HL, Walsh SJ, Burkhead NM *et al.* (2008) Conservation status of imperiled North American freshwater and diadromous fishes. *Fisheries*, **33**, 372–407.

Kerr JT, Kharouba HM, Currie DJ (2007) The macroecological contribution to global change solutions. *Science*, **316**, 1581–1584.

Leibold MA, Economo EP, Peres-Neto P (2010) Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. *Ecology Letters*, **13**, 1290–1299.

Leidy RA (2007) *Ecology, Assemblage Structure, Distribution, and Status of Fishes in Streams Tributary to the San Francisco Estuary, California*. San Francisco Estuary Institute, U.S. Environmental Protection Agency 530, 198.

Leidy RA, Fiedler PL (1985) Human disturbance and patterns of fish species diversity in the San Francisco Bay drainage, California. *Biological Conservation*, **33**, 247–267.

Light T, Marchetti MP (2007) Distinguishing between invasions and habitat changes as drivers of diversity loss among California's freshwater fishes. *Conservation Biology*, **21**, 434–446.

Lomolino MV (1996) Investigating causality of nestedness of insular communities: selective immigrations or extinctions? *Journal of Biogeography*, **23**, 699–703.

Marchetti MP, Light T, Moyle PB, Viers JH (2004a) Fish invasions in California watersheds: testing hypotheses using landscape patterns. *Ecological Applications*, **14**, 1507–1525.

Marchetti MP, Moyle PB, Levine R (2004b) Alien fishes in California watersheds: characteristics of successful and failed invaders. *Ecological Applications*, **14**, 587–596.

Marchetti MP, Moyle PB, Levine R (2004c) Invasive species profiling? Exploring the characteristics of non-native fishes across invasion stages in California. *Freshwater Biology*, **49**, 646–661.

McIntyre PB, Jones LE, Flecker AS, Vanni MJ (2007) Fish extinctions alter nutrient recycling in tropical freshwaters. *PNAS*, **104**, 4461–4466.

Moyle PB (2002) *Inland Fishes of California*. University of California Press, Berkeley, CA.

Moyle PB, Light T (1996a) Biological invasions of fresh water: empirical rules and assembly theory. *Biological Conservation*, **78**, 149–161.

Moyle PB, Light T (1996b) Fish invasions in California: do abiotic factors determine success? *Ecology*, **77**, 1666–1670.

Oksanen J, Blanchet FG, Kindt R *et al.* (2010) *vegan: Community Ecology Package v1.17-4*. R Foundation for Statistical Computing, Vienna.

Olden JD, Kennard MJ, Leprieur F, Tedesco PA, Winemiller KO, García-Berthou E (2010) Conservation biogeography of freshwater fishes: recent progress and future challenges. *Diversity and Distributions*, **16**, 496–513.

Patterson BD, Atmar W (1986) Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biological Journal of the Linnean Society*, **28**, 65–82.

Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecology and Biogeography*, **19**, 174–184.

Power ME (1990) Effects of fish in river food webs. *Science*, **250**, 811–814.

Reynolds JD, Dulvy NK, Goodwin NB, Hutchings JA (2005) Biology of extinction risk in marine fishes. *Proceedings: Biological Sciences*, **272**, 2337–2344.

Rosseel Y (2010) *lavaan: Latent Variable Analysis, v.0.3-1*. R Foundation for Statistical Computing, Vienna.

Schindler DE, Carpenter SR, Cole JJ, Kitchell JF, Pace ML (1997) Influence of food web structure on carbon exchange between lakes and the atmosphere. *Science*, **277**, 1302–1306.

Stone L, Roberts A (1990) The checkerboard score and species distributions. *Oecologia*, **85**, 74–79.

Sutherland WJ, Adams WM, Aronson RB *et al.* (2009) One hundred questions of importance to the conservation of global biological diversity. *Conservation Biology*, **23**, 557–567.

Taylor CM, Warren ML (2001) Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. *Ecology*, **82**, 2320–2330.

Ulrich W, Gotelli NJ (2007) Null model analysis of species nestedness patterns. *Ecology*, **88**, 1824–1831.

Ulrich W, Almeida-Neto M, Gotelli NJ (2009) A consumer's guide to nestedness analysis. *Oikos*, **118**, 3–17.

Vanni MJ (2003) Nutrient cycling by animals in freshwater ecosystems. *Annual Review of Ecology and Systematics*, **33**, 341–370.

White P, Kerr JT (2006) Contrasting spatial and temporal global change impacts on butterfly species richness during the 20th century. *Ecography*, **29**, 908–918.

Zavaleta ES, Pasaria J, Moore J, Hernandez D, Suttle KB, Wilmers C (2009) Ecosystem responses to community disassembly. *Annals of the New York Academy of Sciences*, **1162**, 311–333.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Effects of ties on the inference of attribute importance.
Appendix S2. Site locations and species-specific site-occupancies.
Appendix S3. R-code and simple example of the permutation-based gradient analysis.
Appendix S4. Robustness of results to false absences and the structure of the summer-only metacommunity.
Appendix S5. Contrast to univariate correlation and partial-correlation approaches.