CrossMark

COMMUNITY ECOLOGY – ORIGINAL RESEARCH

Predicting occupancy and abundance by niche position, niche breadth and body size in stream organisms

Mariana P. Rocha[1,2,7] · **Luis M. Bini[3]** · **Tadeu Siqueira[4]** · **Jan Hjort[1]** ·
Mira Grönroos[5] · **Marja Lindholm[1]** · **Satu-Maaria Karjalainen[6]** · **Jani Heino[2]**

Abstract The regional occupancy and local abundance of species are thought to be strongly correlated to their body size, niche breadth and niche position. The strength of the relationships among these variables can also differ between different organismal groups. Here, we analyzed data on stream diatoms and insects from a high-latitude drainage basin to investigate these relationships. To generate measures of niche position and niche breadth for each species, we used sets of local environmental and catchment variables separately, applying the outlying mean index analysis. Beta regression and negative binomial generalized linear models were run to predict regional occupancy and mean local abundance, respectively. We found a positive occupancy–abundance relationship in both diatoms and insects, and that niche-based variables were the main predictors of variation in regional occupancy and local abundance. This finding was mainly due to local environmental niche position, whereas the effects of niche breadth on regional occupancy and local abundance were less important. We also found a relationship between body size and local abundance or regional occupancy of diatoms. Our results thus add to current macroecological research by emphasizing the strong importance of niche position rather than niche breadth and body size for regional occupancy and local abundance in rarely studied organisms (e.g., diatoms and insects) and ecosystems (i.e., wilderness streams).

Communicated by Bryan Brown.

✉ Mariana P. Rocha
mperezrocha@gmail.com

1 Geography Research Unit, University of Oulu, P.O. Box 3000, 90014 Oulu, Finland

2 Biodiversity, Natural Environment Centre, Finnish Environment Institute, P.O. Box 413, 90014 Oulu, Finland

3 Department of Ecology (ICB), Universidade Federal de Goiás (UFG), Goiânia, GO 74690-900, Brazil

4 Instituto de Biociências, Universidade Estadual Paulista (UNESP), Rio Claro, P.O. Box 199, São Paulo, Brazil

5 Department of Environmental Sciences, Section of Environmental Ecology, University of Helsinki, Niemenkatu 73, 15140 Lahti, Finland

6 Freshwater Center, Finnish Environment Institute, P.O. Box 413, 90014 Oulu, Finland

7 CAPES Foundation, Ministry of Education of Brazil, Brasília 70040-020, Brazil

Introduction

A positive relationship between regional occupancy (i.e., proportion of sites occupied) and local abundance (i.e., mean local abundance at occupied sites) of species is one of most extensively investigated patterns in macroecology (Gaston et al. 2000; Blackburn et al. 2006). This relationship has been reported for various groups of organisms (Gaston et al. 1998; Gaston and Blackburn 2000; Blackburn et al. 2006) and at different scales, ranging from broad biogeographic regions to small areas within drainage basins (Cowley et al. 2001; McGill and Collins 2003; Heino and Virtanen 2006; Heino and Gönroos 2014).

Another commonly observed macroecological pattern relates to the relationship between body size and range size (Gaston and Blackburn 2000). For example, some studies have demonstrated that algae (e.g., diatoms) are widely distributed (Soininen and Heino 2005), whereas insects are typically restricted to smaller regions (Hoffsten 2004). For macroorganisms (e.g., birds and mammals), large-bodied species are, however, typically more widely distributed than small-bodied species (Brown 1984; Gaston and Blackburn 2000). Thus, the strength of the relationship between body size and regional occupancy may vary depending on the organism group under study. Also, regional occupancy may or may not be related to body size at small spatial extents (Heino and Gönroos 2014). Rather, these relationships may be affected by dispersal modes (Passy 2012) or niche characteristics (Tales et al. 2004). For example, in passive dispersers (e.g., diatoms), large body size may limit dispersal and, therefore, their regional distribution (Passy 2007, 2012). On the other hand, in active dispersers (e.g., aquatic insects), large body size is often associated with a higher dispersal ability and colonization success (Hoffsten 2004; but see Saito et al. 2015). This latter pattern has been reported for insects, fishes, birds and mammals in general (Brown and Maurer 1987, 1989; Luiz et al. 2012).

Niche breadth and niche position are two other potential drivers of regional occupancy and local abundance patterns. The niche breadth hypothesis predicts that an increase in the variety of resources consumed and increased tolerance to environmental conditions would enhance population viability (Brown 1984). Thus, in a cross-species analysis, one would expect a positive relationship between niche breadth and both local abundance and regional occupancy. Similarly, the niche position hypothesis (Venier and Fahrig 1996) predicts that species consuming widespread resources and able to occur in average habitat conditions, with low values of niche position (as given by the method proposed by Doledéc et al. 2000), would be regionally widespread and locally abundant (Heino and Gönroos 2014; Tonkin et al. 2016). Thus, niche position may also account for a positive relationship between regional occupancy and local abundance of species (Tales et al. 2004; Heino 2005; Siqueira et al. 2009). Regional occupancy and local abundance of species are, therefore, expected to be simultaneously driven by their body size, niche breadth and niche position (Gaston et al. 2000; Tales et al. 2004; Heino and Gönroos 2014). Tests of the relationship between regional occupancy (or local abundance) and this set of potential predictors (i.e., body size, niche breadth and niche position) are, however, still scarce (Hurlbert and White 2007; Jenkins et al. 2007, Slatyer et al. 2013).

Strong support for the niche-based models in accounting for species regional occupancy or local abundance has been previously reported for aquatic organisms (Tales et al. 2004; Heino 2005; Siqueira et al. 2009; Heino and Gönroos 2014). However, few studies have compared niche breadth and niche position measures generated by different types of variables (i.e., local environmental versus catchment variables; but see Siqueira et al. 2009). The influence of catchment variables on abiotic features at local scales (e.g., water chemistry) and subsequently on aquatic organisms has not been studied thoroughly thus far (Soininen et al. 2015). One could expect a strong influence of catchment-scale variables on local environmental characteristics of streams (Hynes 1970; Corkum 1992). Thus, niche measures based on catchment variables would provide similar or even better explanations for the occupancy–abundance relationships than niche measures based on local environmental variables. This would be so because catchment-scale variables are more likely to integrate environmental changes in watersheds, which ultimately would affect local environmental factors and species communities over longer time scales, in contrast with noisier local environmental variables (Soininen et al. 2015). For example, Siqueira et al. (2009), in a study with tropical chironomids, found that niche position based on landscape variables better explained variation in regional occupancy and local abundance in comparison to those based on local environmental variables. In addition, analyzing niche variables based on different scales would be an interesting approach because broad-scale studies could rely on catchment-scale variables, which are easier to measure and less costly to obtain compared to local-scale variables (Siqueira et al. 2009; Soininen et al. 2015).

In this study, we expected that the relationships between body size and regional occupancy would differ between stream diatoms and insects (Fig. 1a). On the one hand, the regional occupancy of insects should increase with increasing body size because large insects can disperse actively and attain larger distribution than small species at the scale of a drainage basin (Hoffsten 2004; Heino 2013). On the other hand, the regional occupancy of diatoms should decrease with increasing body size because passive dispersal rates of larger cells are lower than those of smaller cells (Snoeijs et al. 2002; Passy 2007). In general, small-bodied species are locally more abundant than large species (Hutchinson and MacArthur 1959; Brown and Maurer 1989). Thus, for both insects and diatoms, we predicted a negative relationship between local abundance and body size. Also, the degree of regional occupancy and local abundance should correlate negatively with niche position (considering that the higher the measure proposed by (Doledéc et al. 2000), the higher the specificity of the niche) and positively with niche breadth for both organismal groups (Fig. 1b). Based on previous findings from streams (Heino 2005; Tales et al. 2004; Tonkin et al. 2016), we expected that niche position, in comparison with niche breadth, would be a better predictor of regional occupancy and local abundance, yet different findings have
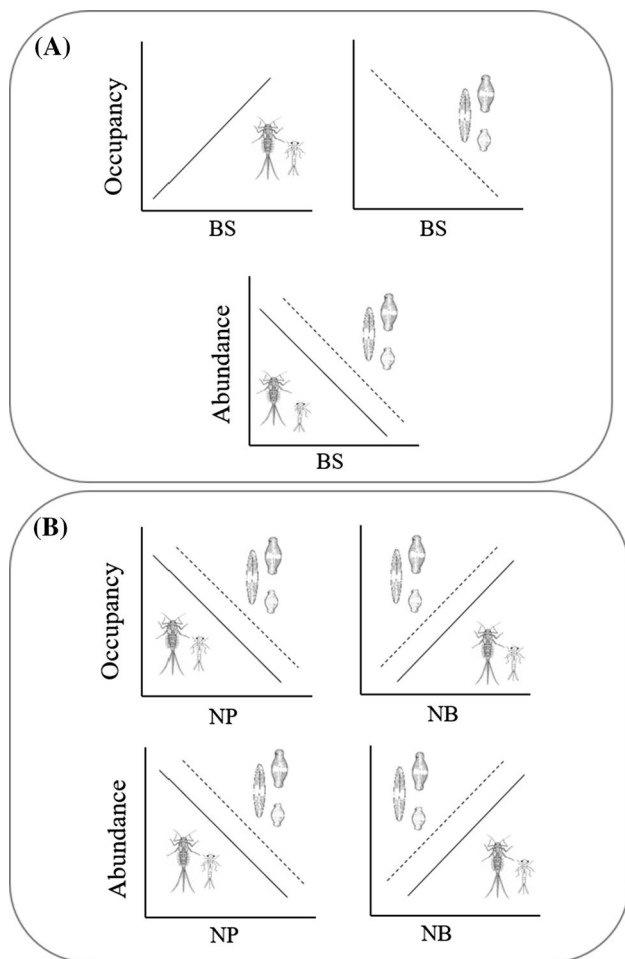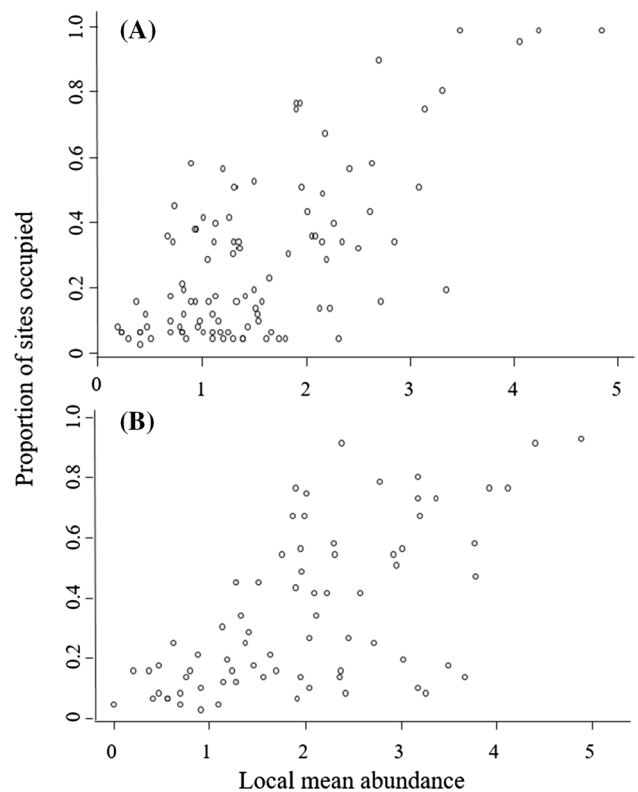
Fig. 1 Expected relationships between niche position (NP), niche breadth (NB) and body size (BS) in predicting occupancy and abundance (**a** and **b**) of stream diatoms and insects. The scenario B is predicted to be the same for both taxonomic groups. Continuous line = insects; dashed line = diatoms

Fig. 2 Relationships between regional occupancy and mean local abundance at occupied sites. **a** = Diatoms; model statistics: pseudo-R^2 = 0.498, $P < 0.001$. **b** = Insects; model statistics: pseudo-R^2 = 0.429, $P < 0.001$

also been obtained (Siqueira et al. 2009; Slatyer et al. 2013). When exploring how local environmental or catchment niches were linked with variation in regional occupancy and local abundance, we assumed that catchment niches would outperform local environmental niches when accounting for regional occupancy and local abundance (Siqueira et al. 2009).

Materials and methods

Study area

We used diatom and insect data collected in 54 wilderness streams in the Tenojoki drainage basin (centered on 70°N, 27°E; total basin area: 16,386 km^2) in Finland (Fig. S1). The surveys were conducted in 2012. Stream waters in this

drainage basin are highly oligotrophic, and pH ranges from 6.3 to 6.7 (Heino 2013; Schmera et al. 2013). The study area is very sparsely human-populated, and human activities (e.g., forestry and agriculture) are uncommon. Consequently, streams in our study area are pristine or near-pristine.

Insect and diatom data

Insects

Each of the 54 stream sites surveyed in this study was sampled using a 3-min kick-net (net mesh size: 0.3 mm) sampling effort. The sample for each site consisted of six 30-s subsamples divided in reaches with differences in depth, current velocity, particle size and moss cover found within ca. 50 m^2. The six subsamples were pooled into a composite sample representing a single sample for each stream and immediately preserved in 70% alcohol in the field. The use of pooled samples is recommendable to cover, as well as possible, the environmental variation (Johnson et al. 2004; Sandin and Johnson 2004) and record most of the species present in a site (Mykrä et al. 2006) (Fig. 2).

Insects were identified to the species level, but early larval stages were identified to genus level because some individuals did not show sufficient morphological characteristics to allow identification to species level. In total, we obtained 74 taxa that occurred at more than two sites. For simplicity, we hereafter call these taxa as species because out of the 74 taxa only 13 were identified to genus level.

Based on maximum larval length, the following ranks were used to represent insect body size (BS): (1) 0–0.25 cm; (2) 0.25–0.50 cm; (3) 0.5–1 cm; (4) 1–2 cm and (5) 2–4 cm (see Supporting Information Fig. S5 c and d for insect size class distributions). This information was provided by personal communication with S. Doledéc (Université Lyon, France), J. Ilmonen (Metsähallitus, Natural Heritage Services, Finland) and L. Paasivirta (Salo, Finland), along with data from Tachet et al. (2010).

Diatoms

The same 54 stream sites surveyed for insects were also surveyed for diatoms. Five to 10 stones (ranging in size from 10 to 30 cm) were haphazardly selected at each site. After, algae were scraped off from the 10 stones to have a total area of 250 cm^2 in each site. In the laboratory, the samples were treated using acid combustion (HNO_3: H_2SO_4, 2:1) to oxidize frustules (SFS-EN 14407 2005). Thereafter, diatom slides were prepared with Naphrax mounting medium. Subsamples of about 500 valves per site were identified and counted using differential interference microscopy with 1000 × magnification. Such a counting effort is typical in studies of stream diatom communities (e.g., Kelly et al., 1998). We used the following diatom cell size ranks based on Rimet and Bouchez (2012): (1) 0–99 μm^3; (2) 100–299 μm^3, (3) 300–599 μm^3; (4) 600–1499 μm^3; (5) ≥ 1500 μm^3 (see Supporting Information Fig. S5 a and b for diatom size class distributions). Most diatoms were identified to species level (ca. 98%) and few to genus level (ca. 2%), totalizing 101 taxa that occurred at more than two sites. Again, for simplicity, we will call all taxa as species.

Local stream environmental and catchment variables

For the local stream environmental variables, we measured current velocity (m s^{-1}) and depth (cm) at 30 random spots in a riffle. Also, we measured the mean width (m) of the sampling sites based on five cross-channel measurements. Moss cover (%) and streambed particle size distribution (i.e., % of particle size classes) were visually estimated at 10 haphazardly selected plots of 1 m^2 in each riffle site. For streambed particle size, we applied a modified Wentworth's scale (Wentworth 1992) as follows: sand (0.25–2 mm), gravel (2–19 mm), pebble (16–64 mm), cobble (64–256 mm) and boulder (256–1024 mm). The mean values for each site

were used in subsequent statistical analysis. Shading (%) by riparian vegetation was visually estimated at each sampling site by a person standing in the center of the stream channel. Additionally, we measured pH and conductivity (μS/cm^{-1}) in the field. Also, water samples were analyzed for total nitrogen (μg/L^{-1}), color (Pt–Co mg/L), iron (μg/L^{-1}) and manganese (μg/L^{-1}) in the laboratory, following the Finnish national standards (National Board of Waters and the Environment 1981).

A total of seven catchment-scale variables for the entire catchment area above each sampling site was computed using ArcGIS 10.1 software (ESRI, Redlands, USA). The drainage area (km^2), the distance to the nearest upstream lake (km), and proportions of mire and lakes (%) were determined using data sets obtained from the National Land Survey of Finland (2010a, b). In a situation where there was no lake upstream of a site, a value representing two-times the longest measured distance to an upstream lake was utilized to make the variable usable in statistical analyses. Mean of tasseled cap greenness, a measure of vegetation abundance (Crist and Cicone 1984), was computed from a Landsat 7 ETM+ satellite image (see Hjort and Luoto 2006). Mean of slope angle (°) was calculated from a 25 m-resolution digital elevation model (DEMs; National Land Survey of Finland 2000c). Topography-derived moisture conditions of the catchments were calculated using the DEMs and topographic wetness index (TWI; Beven and Kirkby 1979). For descriptive statistics of environmental and catchment variables see Table S2 and S3, respectively.

Data analysis: niche measurements and models

We used the outlying mean index (OMI) analysis (Doledéc et al. 2000) to obtain the niche position (NP) and niche breadth (NB) for each species. We ran an OMI analysis separately for each set of environmental variables (i.e., local and catchment-scale variables) to obtain our "local environmental niches" and "catchment niches". Thus, we used the acronyms NP-env and NB-env for the niche characteristics derived from local environmental variables. Similarly, we used the acronyms NP-cat and NB-cat for the niche characteristics derived from catchment-scale variables. We calculated niches measures based on the full set of environmental and catchment variables (see Supplementary Information Table S2 and S3, respectively). OMI is an eigenanalysis-based method that measures NP by calculating the distance from the mean environmental conditions (centroid) used by a species and the mean environmental conditions (niche hyperspace) of the entire surveyed area. Species with high values of NP tend to occur is less common habitats, and those with low values tend to occur in habitats with high availability in the area under study. OMI also provides a measure of niche breath (or tolerance). Species with high

values of NB occur across a broad range of environmental or catchment conditions, whereas species with low values of NB occur in a limited range of conditions. We used the R package "*ade4*" to run the OMI analyses (Chessel et al. 2012; R Development Core Team 2013).

Our first response variable was the proportion of sites occupied. For each group of organisms, we used two beta regression models to analyse the relationship between this response variable and the predictor variables. The first model, which we called "local environmental" model, included NP-env, NB-env, and body size as explanatory variables. The second model, which we called "catchment model", included NP-cat, NB-cat, and body size as explanatory variables. We compared the performance of these models (i.e., "local environmental" and "catchment" models) using the Akaike information criterion (Burnham and Anderson 2002). We used simple bivariate scatter plots (plotting the response against each of the explanatory variables) as a criterion to judge the importance of variables within models. We chose this modeling approach (i.e., beta regression) because the proportion of occupied sites assumes values in the interval [0, 1]. In these cases, the use of beta regression is advisable because the alternative approach, which consists of logit transforming the data and using ordinary least square regression, has drawbacks. In addition to heteroscedasticity, for example, proportions and rates are typically beta distributed. Thus, assuming a normal distribution for estimation and hypothesis testing purposes, even after logit transformation, may be inaccurate, especially when sample size is small (Ferrari and Cribari-Neto, 2004).

Our second response variable, mean local abundance at the occupied sites (after rounding the mean values to integers), was also regressed against NP-env, NB-env and body size ("local environmental" model) and against NP-cat, NB-cat and body size ("catchment" model). As before, we compared the performance of these models using an information-theoretic approach (Burnham and Anderson 2002). Due to data overdispersion, we used negative binomial generalized linear models (Zeileis et al. 2008) to test the relationships between mean local abundance and the explanatory variables.

For all models described above, diatom and insect body sizes (BS) were used as ranked variables. Also, to improve model fit, NP and NB were log-transformed prior to analyses. In short, the four models analyzed, for diatoms and insects, were: (a) Occupancy ~ NP-env + NB-env + BS; (b) Occupancy ~ NP-cat + NB-cat + BS; (c) Mean local abundance ~ NP-env + NB-env + BS; (d) Mean local abundance ~ NP-cat + NB-cat + BS. We decided to run these pairs of competing models to be able to formally compare the predictive power of the two types of explanatory variables (i.e., environmental vs catchment). Also, by doing so, we avoided collinearity problems caused by the high

relationship between NP-env and NP-cat ($r = 0.79$ and $r = 0.81$, for insect and diatom datasets, respectively).

To support our models and avoid biased results, we performed a re-sampling analysis following a previous study (Siqueira et al. 2009). This is because the OMI analysis uses information on abundance and distribution, potentially leading to circularity when NB and NP are used as predictors of abundance and distribution. We thus randomly divided the whole dataset in two subsets; one for estimating regional occupancy, and the other for niche measurements (NP and NB). Thus, with each of those acquired subsets, we ran occupancy models as described above. The random split and models were repeated 1000 times for both insect and diatom data subsets. We reported both overall pseudo-R^2 values and P values associated with the explanatory variables that resulted from the randomizations. All the analysis was performed in the R environment (R Development Core Team 2013). However, due to estimation problems with the negative binomial generalized linear models when using the resampling approach (the algorithm did not converge in many cases), this procedure was adopted for the beta regression models only.

In general, species cannot be considered as independent data points as they are hierarchically structured in a phylogeny (Felsenstein 1985). Thus, depending on the level of phylogenetic signal, cross-species statistical analyses may exhibit high rates of type I error. No correction for phylogeny was used for the above-mentioned analyses, as phylogenies are not currently available for all the species included in this study. However, using taxonomic orders as a proxy for phylogeny in a graphical inspection, we did not find any indication that taxonomy could account for significant variation in those species datasets (see Supporting Information Fig. S2). Additionally, previous studies have found that the occupancy–abundance relationship remains largely unchanged had a phylogeny correction been conducted or not (Blackburn et al. 1997; Cowley et al. 2001; Holt and Gaston 2003; Tales et al. 2004).

Results

Diatoms

The OMI analysis showed that the most important local environmental variables related to the distribution of diatoms across the studied area were moss and shading (OMI axis 1), pebble, conductivity, and boulder (OMI axis 2) (Supporting information, Fig. s3a). Of the catchment variables, mean slope and lake percentage (OMI axis 1), greenness and mire (OMI axis 2) were the most important variables (Supporting information, Fig. s3b).

We found a positive relationship between regional occupancy and local mean abundance of diatoms (pseudo $R^2 = 0.498$) (Fig. 2a). The regional occupancy of diatoms was significantly and negatively related with niche position and body size. We also found a positive relationship between this response variable and niche breadth. These patterns of the regression coefficients were found independently of the variables used to define niche characteristics (i.e., local environmental or catchment variables; with pseudo $R^2 = 0.74$ and pseudo $R^2 = 0.57$, respectively). Judging by the patterns of relationships depicted in the scatterplots, niche position, as compared to niche breadth (or body size; Fig. S5), was the main predictor variable (compare Fig. 3a with Fig. 3b and Fig. 3c with Fig. 3d), and the model based on local environmental niches was clearly superior to the one based on catchment niches (delta AIC = 67.25; Table 1).

For diatom mean local abundance, we also found that, by comparing the niche metrics, niche position was the main explanatory variable (Fig. 4), and that the model based on local environmental niches was superior to the one based on catchment niches (delta AIC = 8.98). NB-env was not a significant predictor of diatom mean local abundance (Table 1).

Insects

The OMI analysis showed that the most important local environmental variables for insect distributions were shading, moss and stream width (OMI axis 1), and pebble, cobble and boulder (OMI axis 2). The most important catchment characteristics affecting insect distribution were mean slope and lake percentage (OMI axis 1), and catchment area and distance to an upstream lake (OMI axis 2) (see Supporting information, Fig. S4A, B).

We found a positive relationship between regional occupancy and local abundance (pseudo $R^2 = 0.429$) (Fig. 2b). Most of the variation in regional occupancy among the 74 insect species was explained by our beta regression models (local environmental niche model: pseudo $R^2 = 0.73$; catchment niche model: pseudo $R^2 = 0.60$). Insect occupancy was significantly and positively related with niche breadth and negatively related with niche position. Niche position was the main predictor of insect occupancy (Fig. 5), and the model based on catchment variables clearly had less support than the model based on local environmental (delta AIC = 26.35; Table 1).
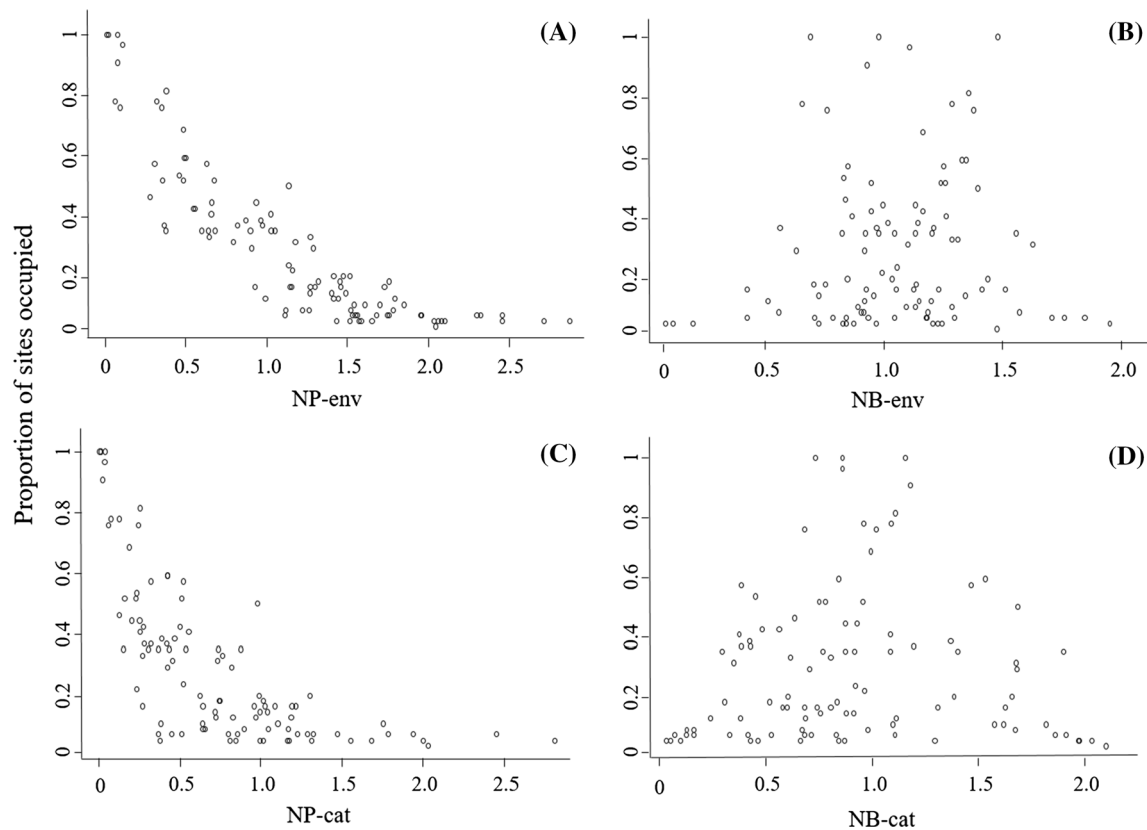
Fig. 3 Relationships between regional occupancy (proportion of sites occupied) and niche metrics of diatoms. Shown are the results for niche metrics based on local environmental variables (**a**, **b**) and catchment variables (**c**, **d**). (Legend: *y*-axis for all panels represents 'Proportion of sites occupied')

Table 1 Modelling results accounting for variation in the occupancy and abundance of diatoms and insects. Generalized liner models (assuming a beta distribution) were used to model occupancy and GLMs (assuming a negative binomial distribution) were used to model local mean abundance

| Group	Response	Model	Explanatories	Coefficient	SE	Z	P	AIC	Pseudo R^2	Exp	Obs
Diatoms	Occupancy	Local	(Intercept)	1.04	0.28	3.76	0.000	− 185.92	0.74		
			NP-env	− 2.13	0.12	− 18.47	0.000			−	Yes
			NB-env	0.58	0.20	2.87	0.004			+	Yes
			Body size	− 0.11	0.05	− 2.22	0.026			−	Yes
		Catchment	(Intercept)	0.49	0.24	2.05	0.040	− 118.67	0.57		
			NP-cat	− 2.24	0.18	− 12.42	0.000			−	Yes
			NB-cat	0.85	0.17	4.95	0.000			+	Yes
			Body size	− 0.22	0.07	− 3.41	0.001			−	Yes
Diatoms	Abundance	Local	(Intercept)	3.30	0.36	9.17	0.000	597.98	0.43		
			NP-env	− 0.83	0.14	− 6.14	0.000			−	Yes
			NB-env	0.04	0.26	0.17	0.867			+	Ns
			Body size	− 0.19	0.07	− 2.65	0.008			−	Yes
		Catchment	(Intercept)	2.83	0.27	10.68	0.000	606.96	0.37		
			NP-cat	− 0.99	0.18	− 5.43	0.000			−	Yes
			NB-cat	0.55	0.19	2.92	0.004			+	Yes
			Body size	− 0.26	0.07	− 3.50	0.000			−	Yes
Insects	Occupancy	Local	(Intercept)	− 0.02	0.43	− 0.04	0.967	− 104.97	0.73		
			NP-env	− 1.71	0.14	− 12.65	0.000			−	Yes
			NB-env	0.77	0.25	3.05	0.002			+	Yes
			Body size	0.13	0.08	1.63	0.103			+	Ns
		Catchment	(Intercept)	− 0.40	0.34	− 1.17	0.242	− 78.62	0.60		
			NP-cat	− 2.24	0.21	− 10.82	0.000			−	Yes
			NB-cat	1.24	0.25	4.90	0.000			+	Yes
			Body size	0.08	0.09	0.92	0.360			+	Ns
Insects	Abundance	Local	(Intercept)	2.94	0.65	4.51	0.000	534.60	0.31		
			NP-env	− 0.82	0.18	− 4.59	0.000			−	Yes
			NB-env	0.92	0.37	2.53	0.011			+	Yes
			Body size	− 0.15	0.12	− 1.26	0.200			−	Ns
		Catchment	(Intercept)	3.36	0.48	7.06	0.000	547.77	0.18		
			NP-cat	− 0.88	0.25	− 3.45	0.001			−	Yes
			NB-cat	0.42	0.34	1.22	0.222			+	Ns
			Body size	− 0.16	0.13	− 1.27	0.203			−	Ns

Only successive pairs of AIC values are comparable

Exp expected relationship [negative (−) or positive (+)], *obs* observed as expected? [yes or non-significant (ns)]

The pseudo coefficients of determination of the models predicting insect mean local abundance were substantially lower than those predicting occupancy (local environmental niche model: pseudo $R^2 = 0.31$; catchment niche model: pseudo $R^2 = 0.18$). Considering the local environmental model, we found a negative relationship between niche position and insect abundance. This response variable was also positively correlated with niche breadth. However, only niche position was significantly and negatively correlated with insect mean local abundance (Fig. 6) when the catchment variables were used to estimate niche characteristics (delta AIC = 13.17; Table 1).

The resampling approach

The results of the resampling approach applied to both diatom and insect occupancy were similar to those obtained with the full datasets. First, the regression coefficients associated with niche position were always significant and, second, those coefficients associated with niche breadth and body size, for both local and catchment models, were so only in a few cases (Fig. S6 and S7; see Fig. S8 for the pseudo coefficients of determination). Thus, our models were robust to the combinations of sampling sites used in the analyses and, at the same time, they show that our models (using the
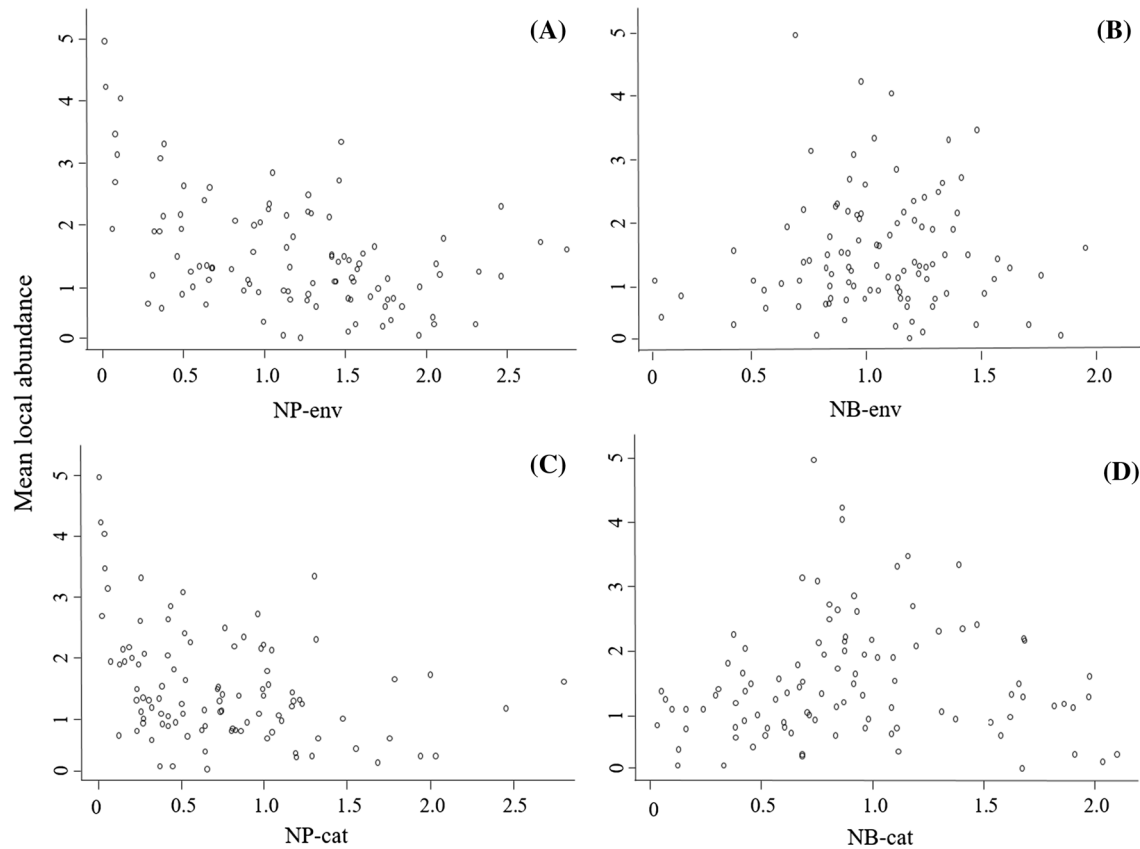
Fig. 4 Relationships between mean local abundance and niche metrics of diatoms. Shown are the results for niche metrics based on local environmental variables (**a**, **b**) and catchment variables (**c**, **d**). (Legend: y-axis for all panels represent 'Mean local abundance')

entire dataset) were not biased by the use of abundance and distribution in the OMI analysis. Given that the resampling procedure applied to regional occupancy model generated results that agreed with those obtained with the entire dataset and that the relationship between regional occupancy and local abundance (for both diatoms and insects) is strong, we assume that the local abundance models are not biased either.

Discussion

We found that both diatoms and insects of high-latitude streams exhibited positive occupancy–abundance relationships (i.e., locally abundant species tended to be regionally widespread). Also, we found that niche position was a consistent predictor of regional occupancy and local mean abundance of the organisms studied, whereas body size was a significant predictor of only diatom regional occupancy and local abundance. The recurrence of niche position as a significant predictor is an interesting result, as it shows that the degree of species marginality is important for their regional occupancy and local mean abundance.

We expected that body size would explain a significant amount of variance in regional occupancy and local abundance of stream organisms given that body size is a proxy for different attributes, including resource utilization (White et al. 2007; Borregaard and Rahbek 2010; Passy 2012) and dispersal ability (Hoffsten 2004; Passy 2007; Heino 2013). Our findings pointed out that body size was, as predicted, a negative correlate of diatom regional occupancy and local abundance, consistent with what is expected for small organisms (Passy 2012). However, contrary to our expectation and previous research (Hoffsten 2004), body size was not a significant correlate of regional occupancy and mean local abundance of insects. The relative range of diatom body size was larger than that for insect body size and, thus, it may help to explain why the relationships between our response variables and body size were significant for diatoms only. Finally, it may be that niche characteristics are simply more important than body size in determining the regional occupancy and local abundance of insect species in high-latitude streams.

Niche position and breadth have been regarded as important factors underlying variation in regional occupancy and local abundance among species (Tales et al. 2004; Heino
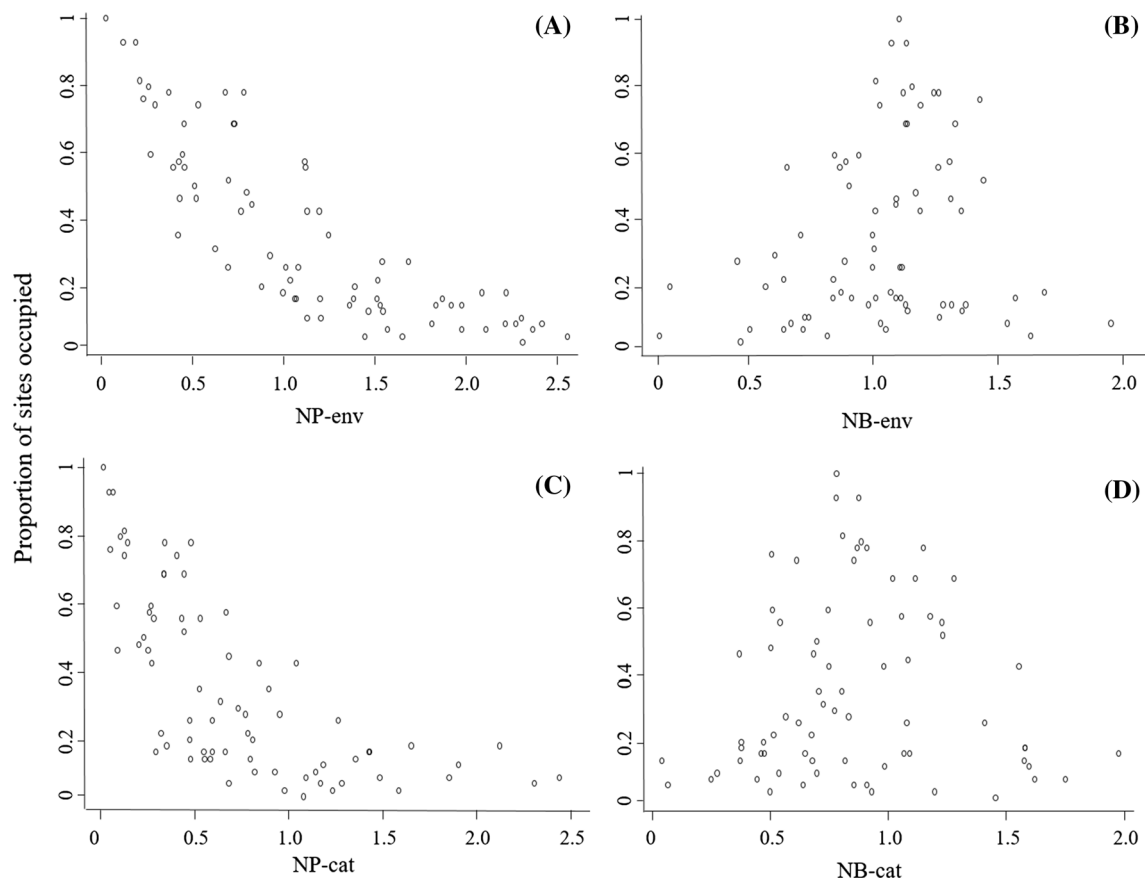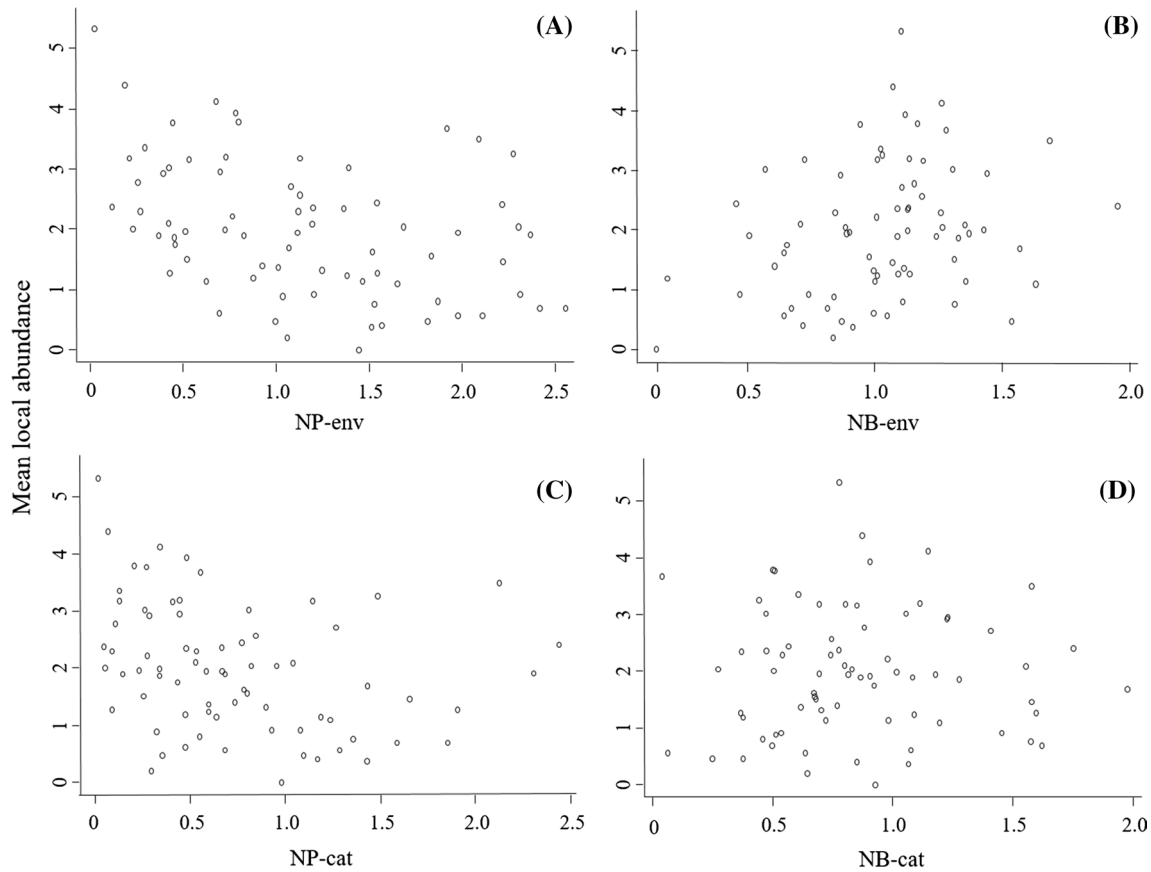
Fig. 5 Relationships between regional occupancy (proportion of sites occupied) and niche metrics of insects. Shown are the results for niche metrics based on local environmental variables (**a**, **b**) and catchment variables (**c**, **d**). (Legend: *y*-axis for all panels represents 'Proportion of sites occupied')

2005; Siqueira et al. 2009; Heino and Gönroos 2014; Tonkin et al. 2016). Our models revealed that niche position, especially when based on local environmental variables, was the most important factor affecting the regional occupancy and local abundance of stream organisms. Thus, species with marginal niche positions (i.e., with the highest values) tended to have lower abundances and to be less widespread that those with non-marginal niche positions (i.e., with the lowest values for this variable). Thus, independently of the group of organisms, our results provided consistent support for the habitat availability hypothesis. This finding agrees with recent studies conducted in stream environments (Tales et al. 2004; Heino and Gönroos 2014; Tonkin et al. 2016). Our results, indicating the importance of niche position, together with those obtained by Slatyer et al. (2013), who showed a positive relationship between niche breadth and geographical range size, clearly indicates the need of a meta-analysis to test the relative importance of these hypotheses in a comparative context.

A previous study on tropical streams showed that niche breadth was the main predictor of regional occupancy and/ or local abundance (Siqueira et al. 2009). The difference with our result (i.e., niche position more important than niche breadth) may be related to the features of high-latitude streams, which are characterized by low temperatures, short growing seasons and harsh winter conditions (Wrona et al. 2013; Tolonen et al. 2017). These factors may impose that species niche position along harsh environmental gradients is important in determining species regional occupancy and local abundance. The lower level of support for the niche breadth hypothesis has also been reported in previous studies (Gregory and Gaston 2000; Cowley et al. 2001; Tales et al. 2004), which might in part be related to difficulties in generating adequate niche breadth measurements for species (Gaston 1994). These difficulties may be related to missing important variables describing niche breadth and the multidimensional nature of the niche (Hutchinson 1957). However, we measured the most influential local environmental and catchment variables affecting the distributions of species in northern streams (Heino and Soininen 2006; Heino and de Mendoza 2016), suggesting that our niche breadth measures should not be worse than the niche position measures.

Fig. 6 Relationships between mean local abundance and niche metrics of insects. Shown are the results for niche metrics based on local environmental variables (**a**, **b**) and catchment variables (**c**, **d**). (Legend: y-axis for all panels represents 'Mean local abundance')

We found that models based on local environmental niches performed better than those based on catchment niches in predicting local abundance and regional occupancy of both diatoms and insects (see AICs in Table 1). The relationships between regional occupancy, local abundance and niche characteristics thus seem to be affected by the types of variables used to generate niche metrics. We speculate that local environmental niches, as compared to catchment scale niches, are better predictors of species regional occupancy and local abundance, as it is at the local scale where species are ultimately selected to occur in certain sites. However, environmental variation at larger scales, such as that at the catchment scale, may also affect regional occupancy and local abundance (Poff 1997; Vinson and Hawkins 1998). For example, Siqueira et al. (2009) found that niche indices defined by landscape variables were important to explain insect regional occupancy and local abundance variations across tropical streams. Because the environmental templates (e.g., temperature and seasonality) differ between tropical and high-latitude regions, we might also expect that different drivers are important for the occupancy and abundance of species. Thus, before generalizing from these

two case studies (Siqueira et al. 2009 and the present study), we believe that the relative importance of local environmental and catchment niches in predicting the distributions of aquatic organisms (local abundance and regional occupancy) should be evaluated in further studies.

Conclusions

Our results showed a strong positive occupancy–abundance relationship in both stream diatoms and insects. We also showed that niche-based metrics are important predictors of regional occupancy and local abundance of diatoms and insects across high-latitude streams. A consistent finding was that mainly the niche position hypothesis (Hanski et al. 1993) was supported in our study as compared to the weaker support for the niche breadth hypothesis (Brown 1984). These findings did not change substantially when the species niches were characterized by local ("local environmental niches") instead of catchment variables ("catchment niches"). However, models based on local environmental variables were superior to those based on catchment

variables. Body size was related to mean local abundance and regional occupancy of diatoms, but it was not important for insects. Our results thus add to current macroecological research by showing the strong importance of niche position for occupancy and abundance in this context of rarely studied organisms (e.g., diatoms and insects) and ecosystems (i.e., high-latitude wilderness streams). Also, they suggest avenues for further research concerning, at least, two themes: first, the relative role of niche position and niche breadth in a meta-analytic context (akin to Slatyer et al. 2013) and, second, the relative predictive power of niche metrics derived from landscape and local scale variables.

References

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrolog Sci Bull 24:43–69

Blackburn TM, Gaston KJ, Quinn RM, Arnold H, Gregory RD (1997) Of mice and wrens: the relation between abundance and geographic range size in British mammals and birds. Philos Trans R Soc Ser B 352:419–427

Blackburn TM, Cassey P, Gaston KJ (2006) Variations on a theme: sources of heterogeneity in the form of the interspecific relationship between abundance and distribution. J Anim Ecol 75:1426–1439

Borregaard MK, Rahbek C (2010) Causality of the relationship between geographic distribution and species abundance. Q Rev Biol 85:3–25

Brown JH (1984) On the relationship between abundance and distribution of species. Am Nat 124:255–279

Brown JH, Maurer BA (1987) Evolution of species assemblages: effects of energetic constraints and species dynamics on the diversification of the North American avifauna. Am Nat 130:1–17

Brown JH, Maurer BA (1989) Macroecology: the division of food and space among species on continents. Science 243:1145–1150

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Colorado State University, Fort Collins

Chessel D, Dufour AB and Dray S (2012) ADE-4: analysis of ecological data. exploratory and euclidean methods in environmental sciences. In: http://CRAN.R-project.org/package=ADE4. Accessed 22 Sep 2016

Corkum LD (1992) Spatial distribution patterns of macroinvertebrates along rivers within and among biomes. Hydrobiologia 239:101–114

Cowley MJR, Thomas CD, Wilson RJ, Léon-Cortés JL, Gutiérrez D, Bulman CR (2001) Density-distribution relationships in British butterflies: II. An assessment of mechanisms. J Anim Ecol 70:426–441

Crist EP, Cicone RC (1984) A physically-based transformation of thematic mapper data—the TM tasseled cap. IEEE Trans. Geiosci Remote Sens 3:256–263

Dolédec S, Chessel D, Gimaret-Carpentier C (2000) Niche separation in community analysis: a new method. Ecology 81:2914–2927

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Ferrari SLP, Cribari-Neto F (2004) Beta Regression for modelling rates and proportions. J. Appl. Stat. 31(7):799–815

Gaston KJ (1994) Rarity. Chapman & Hall, London

Gaston KJ, Blackburn TM (2000) Pattern and process in macroecology. Blackwell, Oxford

Gaston KJ, Blackburn TM, Gregory RD, Greenwood JJD (1998) The anatomy of the interspecific abundance-range size relationship for the British avifauna: I. Spatial patterns Ecol Lett 1:38–46

Gaston KJ, Blackburn TM, Greenwood JJD, Gregory RD, Quinn RM, Lawton JH (2000) Abundance-occupancy relationships. J Appl Ecol 37:39–59

Gregory RD, Gaston KJ (2000) Explanations for commonness and rarity in British breeding birds: separating resource use and resource availability. Oikos 88:515–526

Hanski I, Kouki K, Halkka A (1993) Three explanations of the positive relationship between distribution and abundance of species. In: Ricklefs RE, Schluter D (eds) Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, pp 108–116

Heino J (2005) Positive relationship between regional distribution and local abundance in stream insects: a consequence of niche breadth or niche position? Ecography 28:345–354

Heino J (2013) Environmental heterogeneity, dispersal mode and co-occurrence in stream macroinvertebrates. Ecol Evol 3:344–355

Heino J, de Mendoza G (2016) Predictability of stream insect distributions is dependent on niche position, but not on biological traits or taxonomic relatedness of species. Ecography 39:001–011

Heino J, Gönroos M (2014) Untangling the relationships among regional occupancy, species traits, and niche characteristics in stream invertebrates. Ecol Evol 4:1931–1942

Heino J, Soininen J (2006) Regional occupancy in unicellular eukaryotes: a reflection of niche breadth, habitat availability or size-related dispersal capacity? Freshw Biol 5:672–685

Heino J, Virtanen R (2006) Relationships between distribution and abundance vary with spatial scale and ecological group in stream bryophytes. Freshw Biol 51:1879–1889

Hjort J, Luoto M (2006) Modelling patterned ground distribution in Finnish Lapland: an integration of topographical, ground and remote sensing information. Geogr Ann 88:19–29

Hoffsten PO (2004) Site-occupancy in relation to flight morphology in caddisflies. Freshw Biol 49:810–817

Holt AR, Gaston KJ (2003) Interspecific abundance–occupancy relationships of British mammals and birds: is it possible to explain the residual variation? Glob Ecol Biogeogr 12:37–46

Hurlbert AH, White EP (2007) Ecological correlates of geographical range occupancy in North American birds. Glob Ecol Biogeogr 16:764–773

Hutchinson GE (1957) Concluding remarks. Cold Spring Harb Symp Quant Biol 22:145–159

Hutchinson GE, MacArthur RH (1959) A theoretical ecological model of size distributions among species of animals. Am Nat 93:117–125

Hynes HBN (1970) The Ecology of Running Waters. University of Toronto Press, Toronto

Jenkins DG, Brescacin CR, Duxbury CV, Elliott JA, Evans JA, Grablow KR, Hillegass M, Lyon BN, Metzger GA, Olandese ML, Pepe D, Silvers GA, Suresch HN, Thompson TN, Trexler CM, Williams GE, Williams NC, Williams SE (2007) Does size matter for dispersal distance? Glob Ecol Biogeogr 16:415–425

Johnson RK, Goedkoop W, Sandin L (2004) Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. Freshw Biol 49:1179–1194

Kelly MG, Cazaubon A, Coring E, Dell'Uomo A, Ector L, Goldsmith B, Guasch H, Hürlimann J, Jarlman A, Kawecka B, Kwandrans J, Laugaste R, Lindstrom EA, Leitao M, Marvan P, Padisak J, Pipp E, Prygiel J, Rott E, Sabater S, Van Dam H, Vizinet J (1998) Recommendations for the routine sampling of diatoms for water quality assessments in Europe. J Appl Phycol 10:215–224

Luiz OJ, Madin JS, Robertson DR, Rocha LA, Wirtz P, Floeter SR (2012) Ecological traits influencing range expansion across large oceanic dispersal barriers: insights from tropical Atlantic reef fishes. Proc Roy Soc B 279:1033–1040

McGill B, Collins C (2003) A unified theory for macroecology based on spatial patterns of abundance. Evol Ecol Res 5:469–492

Mykrä H, Ruokonen T, Muotka T (2006) The effect of sample duration on the efficiency of kick-sampling in two streams with contrasting substratum heterogeneity. Verh Internat Verein Limnol 29:1351–1355

National Board of Waters and the Environment (1981) Vesihallinnonanalyysimenetelmät. Publ Nat Board Wat Finland Report 213:1–136

National Land Survey of Finland (2000c) Digital elevation model. PaITuli spatial data service. In: https://research.csc.fi/paituli. Accessed 15 July 2016

National Land Survey of Finland (2010a) Basic map 1: 20 000. PaITuli spatial data service. In: https://research.csc.fi/paituli. Accessed 15 July 2016

National Land Survey of Finland (2010b) Topographic database 1: 10 000. PaITuli spatial data service. In: https://research.csc.fi/paituli. Accessed 15 July 2016

Passy SI (2007) Differential cell size optimization strategies produce distinct diatom richness–body size relationships in stream benthos and plankton. J Ecol 95:745–754

Passy SI (2012) A hierarchical theory of macroecology. Ecol Lett 15:923–934

Poff NL (1997) Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. J North Am Benthol Soc 16:391–409

R Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. Available at: http://www.R-project.org/. Accessed 10 June 2016

Rimet F, Bouchez A (2012) Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowl Manag Aquat Ecosyst 406:1–14

Saito VS, Soininen J, Fonseca Gessner AA, Siqueira T (2015) Dispersal traits drive the phylogenetic distance decay of similarity in neotropical stream metacommunities. J Biogeogr 42:2101–2111

Sandin L, Johnson RK (2004) Local, landscape and regional factors structuring benthic macroinvertebrate assemblages in Swedish streams. Landscape Ecol 19:501–514

Schmera D, Eros T, Heino J (2013) Habitat filtering determines spatial variation of macroinvertebrate community traits in northern headwater streams. Comm Ecol 14:77–88

SFS-EN 14407 (2005) Water quality. Guidance standard for the identification, enumeration and interpretation of benthic diatom samples from running waters. Suomen standardisoimisliitto SFS ry, Helsinki

Siqueira T, Bini LM, Cianciaruso MV, Roque FO, Trivinho-Strixino S (2009) The role of niche measures in explaining the abundance-distribution relationship in tropical lotic chironomids. Hydrobiologia 636:163–172

Slatyer RA, Hirst M, Sexton JS (2013) Niche breadth predicts geographical range size: a general ecological pattern. Ecol Lett 16:1104–1114

Snoeijs P, Busse S, Potapova M (2002) The importance of diatom size in community analysis. J Phycol 38:265–272

Soininen J, Heino J (2005) Relationships between local population persistence, local abundance and regional occupancy of species: patterns in diatoms of boreal streams. J Biogeogr 32:1971–1978

Soininen J, Bartels P, Heino J, Luoto M, Hillebrand H (2015) Toward more integrated ecosystem research in aquatic and terrestrial environments. Bioscience 216:1–9

Tachet H, Richoux P, Bourneau M, Usseglio-Polatera P (2010) Freshwater Invertebrates, Systematics, Biology, Ecology. CNRS Editions, Paris

Tales E, Keith P, Oberdorff T (2004) Density-range size relationship in French riverine fishes. Oecologia 138:360–370

Tolonen KE, Leinonen K, Marttila H, Erkinaro J, Heino J (2017) Environmental predictability of taxonomic and functional community composition in high-latitude streams. Freshw Biol 62:1–16

Tonkin JD, Arimoro FO, Haase P (2016) Exploring stream communities in a tropical biodiversity hotspot: biodiversity, regional occupancy, niche characteristics and environmental correlates. Biodivers Conserv 25:975–993

Venier LA, Fahrig L (1996) Habitat availability causes the species distribution-abundance relationship. Oikos 76:564–570

Vinson MR, Hawkins CP (1998) Biodiversity of stream insects: variation at local, basin, and regional spatial scales. Annu Rev Entomol 43:271–293

Wentworth CK (1992) A scale of grade and class terms for clastic sediments. J Geology 30:377–392

White EP, Ernest SKM, Kerkhoff AJ, Enquist BJ (2007) Relationships between body size and abundance in ecology. TREE 22:323–330

Wrona FJ, Reist JD, Amundsen Arne, Chambers PA, Christoffersen K, Culp JM, di Cenzo PD, Forsström L, Hammar J, Heino J, Heikkinen RK, Kahilainen KK, Lesack L, Lehtonen H, Lento J, Luoto M, Marsh P, Marcogliese DJ, Moquin PA, Mustonen T, Prowse TD, Power M, Rautio M, Swanson H, Thompson M, Toivonen H, Vasiliev V, Virkkala R, Zavalko S (2013) Freshwater ecosystems. In: Meltofte H (ed) Arctic Biodiversity Assessment. Status and Trends in Arctic Biodiversity. Conservation of Arctic Flora and Fauna, Arctic Council, Iceland, pp 335–377

Zeileis A, Kleiber C, Jackman S (2008) Regression models for count data in R. J Stat Software 27:1–25