Author(s): Vilmi, Annika; Tolonen, Kimmo; Karjalainen, Satu Maaria; Heino, Jani

Title: Niche position drives interspecific variation in occupancy and abundance in a highly-connected lake system

Year: 2019

Version: Accepted version (Final draft)

Please cite the original version:

Vilmi, A., Tolonen, K., Karjalainen, S. M., & Heino, J. (2019). Niche position drives interspecific variation in occupancy and abundance in a highly-connected lake system. Ecological Indicators, 99, 159-166. https://doi.org/10.1016/j.ecolind.2018.12.029

Accepted to Ecological Indicators

Niche position drives interspecific variation in occupancy and abundance in a highly-connected lake system

Annika Vilmi*[,1,2], Kimmo T. Tolonen[3], Satu Maaria Karjalainen[2], Jani Heino[4]

[1]State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, 18 North Linshan Road, Qixia, Nanjing 210046, China
[2]Freshwater Centre, Finnish Environment Institute
[3]Department of Biological and Environmental Science, University of Jyväskylä, Finland
[4]Biodiversity Centre, Finnish Environment Institute

*Corresponding author, e-mail: annika.vilmi@gmail.com, phone: +86 18351006316

Author Contributions: AV and JH designed the study. AV and KT performed the species identifications. AV and JH analyzed the data. AV, KT, SK and JH wrote the manuscript.

Abstract

We examined how niche position, niche breadth, biological traits and taxonomic relatedness affect interspecific variation in occupancy and abundance of two commonly-used indicator groups, i.e. diatoms and macroinvertebrates. We studied 327 diatom and 117 macroinvertebrate species that occupied the littoral zones of a large (305 km^2) highly-connected freshwater system. We collated information on the biological traits and taxonomic relatedness of each species. Using principal coordinates analysis, we formed biological trait and taxonomic vectors describing distances between species and used the resulting vectors as predictor variables. As environmental data, we had site-specific physico-chemical variables, which were used in outlying mean index analyses to determine the niche position and niche breadth of a species. We used linear models to study if and how these two niche parameters and biological traits as well as taxonomic relatedness affected occupancy and abundance. We observed positive occupancy-abundance relationships for both diatoms and macroinvertebrates. We further found that, for both groups, occupancy was better explained by the predictor variables compared with abundance. We also observed that niche parameters, especially niche position, were the main determinants of variation in occupancy and abundance for both diatoms and macroinvertebrates. Local abundances of diatom and macroinvertebrate species were also, to a small degree, affected by biological traits or taxonomic relatedness. We further saw that the relationship between niche position and occupancy was negative, indicating that the more marginal the niche position, the rarer a species is. Our findings provide support for the use of diatoms and macroinvertebrates as ecological indicators as their occupancies and abundances were affected by niche parameters, which is not necessarily always clear in challenging study systems with high connectivity (i.e. high movement of material and species) among sites. These findings also suggest that indices using information on species' occupancy, abundance and niche requirements are useful in environmental assessment.

Keywords

Diatoms, macroinvertebrates, niche parameters, biological traits, taxonomic relationships

1. Introduction

Species occupancy, abundance and their relationships are amongst the most widely-studied topics in macroecology and biogeography (Brown 1984; Gaston et al. 2000), and researchers have described the occupancy-abundance patterns in a variety of ecosystems across the globe (Venier and Fahrig 1996; Soininen and Heino 2005; Foggo et al. 2007; Roney et al. 2015; Tonkin et al. 2016). The message has been relatively consistent: occupancy and abundance tend to be positively and often strongly associated with each other. Species with low local abundances tend to show limited distributions, whereas locally abundant species usually are widespread in a region (Gaston and Blackburn 2000; Gaston et al. 2000). These ideas are important for meaningful biodiversity conservation (Gaston et al. 2000), especially in an era when human impacts are greater than before (Lewis and Maslin 2015; Waters et al. 2016). As freshwaters are one of the most threatened ecosystems in the world (Heino et al. 2009; Vörösmarty et al. 2010; Vilmi et al. 2017), they also deserve special attention in the context of the interspecific occupancy-abundance relationship (Gaston et al. 2000; Heino and Tolonen 2018).

An increasing number of attempts have been made to detect various factors that account for variation in species occupancy and abundance. In recent years, the niche breadth (Brown 1984) and niche position (Hanski et al. 1993; Venier and Fahrig 1996) hypotheses have been employed to investigate how species' niche characteristics (i.e. their relationships to the environment) account for their regional occupancy and local abundance (Tales et al. 2004; Faulks et al. 2015; Tonkin et al. 2016). Ultimately, species optima are located at different parts of a continuum of environmental conditions (Fig. 1). Some species have the same niche positions, while niche positions of other species are located at different parts of the continuum. Some niches may be positioned at the very end of the environmental range, while other niches may be positioned in average environmental conditions. Although species may have similar niche positions, their niche breadths can strongly differ from each other at the same time. Species with large niche breadths can tolerate a wide variety of environmental conditions, while species with small niche breadths are very specialized to certain environmental conditions (Fig. 1, Brown 1984; Venier and Fahrig 1996; Heino and de Mendoza 2016).

Previous research has provided strong evidence that the two niche parameters, niche position and niche breadth, strongly affect occupancy and abundance of different types of organisms (Passy 2012; Heino and Tolonen 2018). For instance, when studying stream macroinvertebrates in a tropical region, Tonkin et al. (2016) found that the two niche parameters explained well variation in occupancy, but it was not correlated with mean local abundance. A study on fish species across boreal lakes suggested that intraspecific niche variation and a positive abundance-occupancy relationship are connected to each other (Faulks et al. 2015). Tales et al. (2004) found support for the niche position hypothesis as the largest contributor to regional occupancy and local abundance in temperate river fishes. Recent studies on subarctic stream diatoms and insects (Rocha et al. 2018) as well as boreal lake macroinvertebrates (Heino and Tolonen 2018) have also shown that especially niche position has strong effects on regional occupancy and local abundance.

Species characteristics other than niche parameters may also affect occupancy, abundance and their relationships (Tales et al. 2004; Heino and de Mendoza 2016; Heino and Tolonen 2018; Rocha et al. 2018). Body size measures and resource use features are typical examples of species characteristics, i.e. biological traits (e.g. Passy 2012). The importance of biological traits in affecting occupancy, abundance and their relationships has been reported, for instance, for riverine fishes (Tales et al. 2004), for aquatic macroinvertebrates across lentic waterbodies (Verberk et al. 2010) and for diatoms across streams (Rocha et al. 2018). In addition to biological traits of species, also taxonomic relatedness as a proxy for evolutionary aspects has been considered in the same context. Heino and Tolonen (2018) did not find strong influences of taxonomic relatedness nor traits similarity on occupancy and abundance of macroinvertebrates across a set of lakes, while the effects of niche parameters, particularly that of niche position, were clear.

Previous research centering on the topic of occupancy and abundance and the factors determining them have covered different types of freshwater ecosystems, ranging from across-streams (Rocha et al. 2018) to across-lakes (Heino and Tolonen 2018) and to across-waterbodies (Verberk et al. 2010) studies. However, there is a knowledge gap for information from systems with high connectivity, such as large lake systems. We previously showed that diatom and macroinvertebrate communities exhibited pure spatial patterns (e.g. dispersal-related processes) and that local environmental conditions only comparatively little affected community structures in such a highly-connected system (Vilmi et al. 2016; Tolonen et al. 2017). These findings were not in line with main assumptions of current bioassessment methods that assume a high importance of local environment on species compositions (Heino 2013). The high connectivity of a system is thus a characteristic which can enhance the effects of dispersal and other spatial processes (Foggo et al. 2007). Thus, in that sense, it is worth studying how interspecific occupancy and abundance, the building blocks of ecological indicators, are formed in these sorts of open, large freshwater systems.

Here, we investigated if and how niche parameters, biological traits and taxonomic relatedness affect occupancy and abundance of freshwater diatoms and macroinvertebrates in a large lake system which contains no apparent barriers for dispersal. For each species, we calculated niche position and niche breadth (based on an extensive dataset of local environmental variables) and determined biological traits and taxonomic relatedness. Using species as data points, we asked the following questions: (1) What are the relationships between occupancy and abundance of freshwater diatom and macroinvertebrate species within a large lake system? (2) Which factors (niche parameters, traits, and taxonomy) best predict occupancy, abundance and their relationship of diatom and macroinvertebrate species in such a study system? (3) Are the findings similar for the two distinct groups of organisms? (4) Are the findings similar to studies examining patterns across waterbodies?

2. Material and Methods

2.1. Field sampling and laboratory analyses

We used data on diatom and macroinvertebrate taxa to explore our research questions. The biological data were collected in early autumn 2013 from a large (305 km^2) lake system of Lake Kitkajärvi. The originally oligotrophic lake system is located in north-eastern Finland. Due to land use and inflow of purified municipal waste water, some parts of the lake system have shown signs of eutrophication (e.g. Vilmi et al. 2015). We collected the diatom and macroinvertebrate samples from 81 similar, stony littoral sites around the lake system (see map of study area in Fig. A.1). In the laboratory, for diatoms, we identified approximately 500 valves from each site, and for macroinvertebrates, we identified all individuals that were captured in a site-specific kick-net sampling. We identified the diatoms and macroinvertebrates to the lowest taxonomic level possible, which was in most cases species level, although some valves or individuals were assigned to genus level. Thus, from now on, we refer to the studied taxa here as 'species'.

We gathered a broad set of site-specific local environmental variables. In the field, we visually assessed the particle sizes of the benthic substratum and measured the slope of bottom. We also collected water samples, which were analyzed in the laboratory. We used fetch, calculated with the Wind Fetch Model (Rohweder et al. 2008), as a proxy for wave disturbance at each site. Further details on the sampling and laboratory methods are thoroughly presented in our earlier publications on the effects of local environmental variables on diatom and macroinvertebrate community structures (Vilmi et al. 2016; Tolonen et al. 2017).

In this study, we used the following variables as local environmental variables: electrical conductivity, saturation of oxygen, suspended solids, slope, fetch, mean particle size, and particle size diversity, as well as concentrations of aluminium, boron, manganese, NH_4-N, NO_2+NO_3-N, oxygen, PO_4-P, silicon, sodium, soluble total nitrogen, soluble total phosphorus, and zinc. These variables were not highly correlated with each other and showed considerable among-site variation within the data.

2.2. Niche position and niche breadth

We first determined niche position (OMI values) and niche breadth (Tolerance values; Tales et al. 2004) of each species using the outlying mean index (OMI) analysis (Dolédec et al. 2000). The analysis basically measures the marginality of species habitat distributions by the distances between mean environmental conditions used by a species and average environmental conditions that are available among the study sites. The ecological interpretation of the OMI and Tolerance values, which the analysis produces, is as follows: species with high OMI values have marginal niches and species with low OMI values have non-marginal niches. Species with high Tolerance values occur in a variety of environmental conditions thus having large niche breadths. Species with low Tolerance values are present only in certain environmental conditions and they have smaller niches.

For the OMI analysis, a site-by-species abundance data matrix is needed, as well as an environmental variables data matrix. We made logarithmic transformations for the environmental variables and standardized them. Before proceeding to the OMI analyses, we excluded species that were present only at one site. After doing so, we had a set of 327 diatom and 117 macroinvertebrate species to investigate. We log-transformed (log x + 1) the species abundance matrices. We used the R package ade4 (Dray et al. 2018) for conducting the OMI analyses.

2.3. Biological traits

We collated information on biological traits for all 327 diatom and 117 macroinvertebrate species. For diatoms, we followed Rimet and Bouchez's (2012) work and gathered information on sizes, ecological guilds and colonial formation of species. The sizes are reported as biovolume classes: 0-99 μm^3 (class S1), 100-299 μm^3 (S2), 300-599 μm^3 (S3), 600-1499 μm^3 (S4), and > 1500 μm^3 (S5). Diatom species were assigned to four guilds: low profile, high profile, motile and planktonic guilds. The different ecological guilds differ from each other by, for instance, resource use and motility (Passy 2007). As a third biological trait, we distinguished colonial and single cell species. This aspect is important in terms of, for example, resource use (e.g. light or nutrients) or potential grazing pressure. Although the biological trait information was mainly collected from Rimet and Bouchez (2012), not all of our species were in their list. We hence used OMNIDIA software (Lecointe et al. 1993) and its databases to find out the missing information. In some cases, we made trait assignments based on characteristics of similar species belonging to the same genus.

For macroinvertebrates, we also considered a measure of size as a biological trait. Here, the size is actually the dry mass of each species, and it was divided to five classes: 0-0.99 mg (class DM1), 1-3.49 mg (DM2), 3.5-9.99 mg (DM3), 10-34.99 mg (DM4), and > 35 mg (DM5) to facilitate comparisons with the diatom data. Further, we assigned each species to functional feeding groups (FFG), which were shredders, scrapers, predators, piercers, collector-gatherers and filterers. The third biological trait for macroinvertebrates was their substrate-association type. In our data, there were crawlers, swimmers, burrowers, sessiles and semisessiles. The biological trait classifications used here were earlier given in a previous study by Tolonen et al. (2017) using the same macroinvertebrate data. The three trait groups we used are directly related to species' vulnerability to predation and resource acquisition and habitat use (Merritt and Cummins 1996; Tolonen et al. 2003; Schmera et al. 2015).

We calculated trait distances between species using function gowdis in the R package FD (Laliberté et al. 2014). Using the resulting trait distances, we performed principal coordinates analysis (PCoA) with function pco from the R package labdsv (Roberts 2016) to form trait vectors. Using trait vectors was appropriate, because species are not a manifestation of only one trait, but, instead, a summary of several traits (e.g. Verberk et al. 2013).

2.4. Phylogenetic relatedness

Without true phylogenies at hand, we used taxonomic information as proxies for phylogenetic information. For each of the diatom and macroinvertebrate species, we determined multiple taxonomic levels above species. In the diatom dataset, we had genus, family, order, subphylum and phylum levels. These were collected from Rimet and Bouchez (2012), but updated when necessary, using new literature (e.g. Lange-Bertalot et al. 2017) and AlgaeBase (www.algaebase.org). For macroinvertebrates, we had genus, family, suborder, order, class, phylum, superphylum and kingdom levels. This information was collected from Fauna Europaea (www.fauna-eu.org).

To account for taxonomic relatedness, we calculated taxonomic distances between species using the function taxa2dist from the R package vegan (Oksanen et al. 2018). We then formed taxonomic vectors using principal coordinates analysis (PCoA) using the function pco from the R package labdsv (Roberts 2016).

2.5. Statistical methods

First, we explored the relationship between occupancy and abundance of diatom and macroinvertebrate species using linear regression models. In practice, we used logit-transformed proportion of sites occupied as 'occupancy'. As 'abundance', we used log-transformed mean abundance at occupied sites. We performed the linear models and drew the plots with the R packages stats (R Core Team 2018) and ggplot2 (Wickham 2016).

We then used linear regression models to study the explanatory power of the two niche parameters (i.e. niche position and niche breadth), biological trait vectors and taxonomic vectors on occupancy, abundance, and their relationship. The relationship of these two original response variables was investigated using variation in residuals of the occupancy-abundance relationship as a third response variable in our linear models. Here, the residual variation actually describes occupancy when the effect of abundance has been removed. Of the explanatory trait and taxonomic vectors, we used the first six vectors in all models. This was because the first six vectors were clearly stronger than the following ones. The directions how to interpret vectors are presented in Appendix B. We again used the R packages stats (R Core Team 2018) and ggplot2 (Wickham 2016) to perform the linear regression models and to draw associated plots.

When using linear regression models as the primary statistical method, commonality analysis (Ray-Mukherjee et al. 2014) provides additional information on the effects of the explanatory variables on the response variables. We performed commonality analyses with the R package yhat (Nimon et al. 2013). Doing so, we were able to see the individual, shared and total contributions of each explanatory variable on variation in our response variables.

3. Results

We analyzed data on 327 diatom and 117 macroinvertebrate species and found that there were large differences in their occupancies and abundances. The proportion of sites occupied varied from 2.5% to 100% for diatom species, and from 2.5% to 89% for macroinvertebrate species. The mean abundances of species at the sites they occupied varied from one to 133 for diatoms, and from one to 47 for macroinvertebrates. Appendix C presents the list of diatom and macroinvertebrate species studied in order of occupancy.

3.1. The relationship between occupancy and abundance

For both groups of organisms, we found rather strong, positive occupancy-abundance relationships. The linear regression models showed that abundance explained over 50% of variation in diatom species occupancy, and almost 40% of variation in macroinvertebrate species occupancy (Table 1). Scatter plots further showed that macroinvertebrate species indeed had more "outliers" than diatom species, when considering a pure linear relationship, which made the explanatory power of the macroinvertebrate model lower compared to the diatom model (Fig. 2).

3.2. Factors explaining variation in occupancy

The linear regression models explained approximately 60% of variation in the occupancy of diatom species and 77% of variation in the occupancy of macroinvertebrate species (Table 2). Niche position was, out of our explanatory factors, the strongest predictor of occupancy for both diatoms and macroinvertebrates. For both groups of organisms, niche position was negatively associated with occupancy. The linear regression models indicated that niche breadth was also related to the occupancy of diatom and macroinvertebrate species. For both organism groups, the relationship between niche breadth and occupancy was positive. For diatoms, trait vector 5 was also statistically significantly associated with occupancy.

The results of commonality analyses also supported the chief role of niche position on occupancy of diatom and macroinvertebrate species (unique contributions 0.531 and 0.552, respectively; Table 2). In addition, niche breadth had some unique effects on occupancy, having a larger role for the occupancy of macroinvertebrate species (0.152) than for diatom species (0.080).

3.3. Factors explaining variation in mean abundance

The linear models explained 32% of variation in mean abundance of diatom species and 34% of variation in mean abundance of macroinvertebrate species (Table 3). Both niche parameters, four trait vectors and two taxonomic vectors were statistically significantly associated with the mean abundances of diatom species. Regarding macroinvertebrates, both niche position and breadth, as well as trait vector 3 and taxonomic vector 2, were associated with mean abundances. Niche position was negatively and niche breadth positively associated with mean abundances in both groups of organisms (Table 3).

The commonality analysis indicated that niche position had a clear and comparatively strong effect on mean species abundance of diatoms (unique effect 0.146; Table 3). For diatoms, the other associated explanatory factors had smaller unique roles on mean abundance (unique effects ranging from 0.009 for taxonomic vector 2 to 0.055 for trait vector 5). The commonality analysis results further showed for macroinvertebrates that niche position and niche breadth had similar, but relatively small unique effects on mean abundance (unique effects 0.056 and 0.054, respectively). Trait vector 3 and taxonomic vector 2 also had small unique effects on mean abundance (0.053 and 0.028, respectively). Boxplots in Appendix B imply that diatom trait vector 5 may be a combination of size and colony-forming, while macroinvertebrate vector 3 is not as clearly related to a single trait but instead is probably a combination of all biological trait groups.

3.4. Factors explaining the relationship between occupancy and abundance

Modelling the residuals of the occupancy-abundance relationship provided a possibility to model their relationship, which means basically occupancy once the effect of abundance has been removed. The linear models explained 51% of variation in residuals of the occupancy-abundance relationship for diatoms, and 76% of the residual variation for macroinvertebrates (Table 4). Results of the linear regression models showed, regarding both organism groups, that niche position, niche breadth and trait vectors (trait vector 2 for diatoms and trait vector 3 for macroinvertebrates) explained the residual variation for both groups of species. For diatoms, taxonomic vectors 1 and 5 were also statistically significantly associated with the residual variation.

The commonality analyses showed that niche position, with its large unique contributions, was the main predictor of residual variation for both groups of organisms (Table 4). The effects of other factors, although being statistically significantly associated, were minor, with the exception of the unique effect of niche breadth on macroinvertebrates (0.098).

4. Discussion

Our results showed relatively strong and positive occupancy-abundance relationships in our diatom and macroinvertebrate data collected from a large highly-connected lake system. These results thus corroborated both classical (e.g. Brown 1984) and more recent (e.g. Passy 2012) ideas that occupancy and abundance are strongly correlated. In addition, we showed that niche features were the chief determinants of interspecific variation in both occupancy and abundance, aligning with recent studies from various systems (Tales et al. 2004; Tonkin et al. 2016).

Occupancy was well explained by our set of explanatory variables. The linear models explained almost 60% of variation in diatom species occupancy and nearly 80% of variation in macroinvertebrate species occupancy. Thus, the occupancy of a species is closely tied to its ecological characteristics, such as habitat preferences or resource use. It was, however, evident that niche position was the main variable in predicting variation in occupancy of both diatom and macroinvertebrate species. The strong effect of niche position on occupancy of freshwater organisms has previously been noted in across-waterbodies systems (Tales et al. 2004; Heino and Soininen 2006; Rocha et al. 2018). We also found that the relationship between niche position and occupancy was always negative, basically indicating that rare species possess marginal niches. This imposes that conservation actions should be directed to regionally marginal habitats in order to secure the living conditions of regionally rare species (see also Gaston et al. 2000; Heino and Tolonen 2018; Rocha et al. 2018).

The strong impact of niche position on occupancy is, in part, surprising here, because the areal extent of our study system was comparatively small (305 km^2), the aquatic system was highly-connected (i.e. apparently free movement of organisms), and thus the ranges of environmental conditions (which were used to calculate the niche parameters in our analyses) were relatively subtle. In addition, high rates of dispersal among sites in our highly-connected study system might interfere with species sorting processes, resulting in weakly-determined species-environment relationships (Vilmi et al. 2016; Tolonen et al. 2017). The fact that niche position was also in this kind of a study setting such a strong factor in explaining interspecific variation in occupancy suggests that these patterns are perhaps universal and not strongly affected by the study systems' ecological properties (Tales et al. 2004; Tonkin et al. 2016; Heino and Tolonen 2018; Rocha et al. 2018). Thus, what we observed now from a large lake system, may be also found in similarly highly-connected aquatic systems, such as marine areas.

The linear regression models for the mean abundances resulted in lower explanatory powers than the models for occupancy. They explained over 30% of variation in mean abundances of diatom and macroinvertebrate species. This may indicate that mean abundance is a more complex variable than occupancy. The fact that abundance of macroinvertebrate species was not as well explained by the explanatory factors as their occupancy (34% vs. 77% explained) indicates that, for macroinvertebrates, the formation of variation in abundance is a more complex process than formation of variation in occupancy, and could partly be a result of other factors, such as stochasticity or random effects. The complexity was also visible in the scatter plots, where the occupancy-abundance relationship was weaker for macroinvertebrates than for diatoms. Previous research has also shown that abundance of macroinvertebrates cannot be as clearly linked to niche parameters as their occupancy (Tonkin et al. 2016).

The importance of niche position, and to a lesser extent that of niche breadth, was evident for explaining variation in our response variables. However, that was generally not the case with biological traits or taxonomic relatedness. There were small effects of the trait or taxonomic vectors, but nothing evident. This basically means that the biological traits studied here were not strongly related to occupancy and abundance. Although previous research on streams has shown that biological traits, such as body size, are connected to species regional occupancy and mean local abundance (Passy 2012; Rocha et al. 2018), we could not detect any clear patterns suggesting the importance of biological traits. Perhaps the characteristics of our study system, i.e. high connectivity, resulted in the lack of a clear relationship. Previous research, however, has shown that biological traits (i.e. organism's dispersal capacity) may affect the elevation of the abundance-occupancy relationship in highly-connected marine systems (Foggo et al. 2007). This finding was reported from a significantly larger area than what we investigated here. Consequently, perhaps we did not detect clear effects of biological traits because of the small spatial scale addressed (Brändle and Brandl 2001) or maybe we used the wrong traits (Heino and Tolonen 2018). In addition, it is possible that niche characteristics are simply just more important than biological traits for the formation of variation in occupancy and abundance of aquatic organisms (e.g. Rocha et al. 2018).

Similarly, taxonomic relatedness did not play a role in affecting occupancy, abundance and their relationships for freshwater diatom and invertebrate species. Previous research has neither found support for clear effects of taxonomy to variation in occupancy and abundance of freshwater organisms (Tales et al. 2004; Heino and Tolonen 2018). As biological traits are products of evolution and thus portrayed by phylogeny (Harvey 1996), it is not surprising that taxonomic relatedness neither appeared as a strong predictor of occupancy and abundance. On the other hand, some biological traits have evolved many times and can be characteristic of comparatively distant orders (Rimet and Bouchez 2012), so in that sense, biological traits may not always be as closely related to phylogeny as expected (Harvey 1996). It is also possible that the spatial scale investigated in this study was not sufficient to detect clear effects of taxonomic relatedness on occupancy and abundance. Phylogenetic signals might have been found over larger areas crossing regional species pools (Heino and Tolonen 2018).

A possible caveat of this study is that the species' niche parameters were calculated based on environmental variables collected during the same sampling as the biological samples. Because of no other suitable data existed to calculate the species-specific niche parameters, we opted to use the same dataset. There may thus be a possibility that the effects of niche parameters on occupancy and abundance may have been overestimated. Previous research however has shown that irrespective of the underlying data (i.e. same or different dataset) to calculate niche parameters, they arise as important determinants of occupancy and abundance (Heino 2005; Heino & Grönroos 2014; McCreadie and Adler 2014; Teittinen et al. 2018).

5. Conclusions

We found that niche position was the strongest predictor of occupancy and abundance of diatoms and macroinvertebrates in a freshwater system with high connectivity. This finding is consistent with previous knowledge from across-waterbody systems with presumably lower connectivity among sites (Tales et al. 2004; Heino and Soininen 2006; Tonkin et al. 2016; Heino and Tolonen 2018; Rocha et al. 2018). The fact that our highly-connected freshwater system showed similar results to comparatively weakly-connected across-waterbody systems implies that these patterns in occupancy and abundance are perhaps universal and are not strongly related to the connectivity of a study system. Furthermore, due to high connectivity, our study setting had fairly subtle environmental ranges, indicating that niche position and, to a smaller extent, niche breadth, can have strong effects on occupancy and abundance also in a situation where environmental variation is comparatively small. Importantly, these findings were evident even when controlling for biological traits and taxonomic relatedness. As species abundances and occupancies are basically the building blocks for a number of ecological indicators, the observed importance of niche parameters is alleviating regarding the use of these sorts of indicators in environmental assessment.

The effect of niche position on the response variables was negative, indicating that the more marginal the niche, the rarer the species both in terms of occupancy and abundance. In other words, rare species tended to possess marginal niches within a large lake system. This was evident for two very distinct groups of freshwater organisms, which represent different trophic positions of the food web and contribute differently to the overall functioning of the aquatic ecosystem. Thus, in order to enhance biodiversity conservation, protection of regionally marginal habitats is important for protecting regionally rare species, also in systems of high connectivity.

Acknowledgements

Annika Vilmi was supported by the Chinese Academy of Sciences President's International Fellowship Initiative (2018PS0007). We acknowledge Marja Lindholm and Mariana Perez Rocha for sharing their information on biological traits of some diatom species.

Conflict of Interest: The authors declare that they have no conflict of interest.

Supplementary material

Appendix A. A map showing the sampling sites around a large lake system of Lake Kitkajärvi.

Appendix B. Examples of how to interpret vectors.

Appendix C. The diatom and macroinvertebrate species studied, in order of occupancy.

References

Brändle M, Brandl R (2001) Distribution, abundance and niche breadth of birds: scale matters. Glob Ecol Biogeogr 10:173–177

Brown JH (1984) On the Relationship between Abundance and Distribution of Species. Am Nat 124:255–279

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

Dray S, Dufour A-B, Thioulouse J (2018) ade4: Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences. R Package version 1.7-11. URL https://CRAN.R-project.org/package=ade4

Faulks L, Svanbäck R, Ragnarsson-Stabo H, Eklöv P, Östman Ö (2015) Intraspecific Niche Variation Drives Abundance-Occupancy Relationships in Freshwater Fish Communities. Am Nat 186:272–283

Foggo A, Bilton DT, Rundle SD (2007) Do developmental mode and dispersal shape abundance–occupancy relationships in marine macroinvertebrates? Journal of Animal Ecology 76:695–702

Gaston KJ, Blackburn TM (2000) Pattern and Process in Macroecology. Blackwell, Oxford.

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

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

Harvey PH (1996) Phylogenies for ecologists. J Anim Ecol 65:255–263.

Heino J (2013) The importance of metacommunity ecology for environmental assessment research in the freshwater realm. Biol Rev 88:166–178.

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, Soininen J (2006) Regional occupancy in unicellular eukaryotes: a reflection of niche breadth, habitat availability or size-related dispersal capacity? Freshw Biol 51:672–685.

Heino J, Tolonen KT (2018) Ecological niche features override biological traits and taxonomic relatedness as predictors of occupancy and abundance in lake littoral macroinvertebrates. Ecography 41, in press, doi: 10.1111/ecog.03968

Heino J, Virkkala R, Toivonen H (2009) Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. Biol Rev 84:39–54

Laliberté E, Legendre P, Shipley B (2014) FD: measuring functional diversity (FD) from multiple traits, and other tools for functional ecology. R package version 1.0-12. URL https://CRAN.R-project.org/package=FD

Lange-Bertalot H, Hofmann G, Werum M, Cantonati M (2017) Freshwater Benthic Diatoms of Central Europe: Over 800 Common Species Used in Ecological Assessment. Edited by Cantonati M, Kelly MG, Lange-Bertalot H. Koeltz Botanical Books, Germany, 942 pp.

Lecointe C, Coste M, Prygiel J (1993) Omnidia: software for taxonomy, calcula-tion of diatom indices and inventories management. Hydrobiologia 269/270:509–513

Legendre P, Legendre L (2012) Numerical Ecology, 3rd edn. Elsevier, Amsterdam

Lewis SL, Maslin MA (2015) Defining the Anthropocene. Nature 519:171–180.

Merritt RW, Cummins KW (1996) An introduction to the aquatic insects of North America, 3rd ed. Kendall/Hunt Publishing.

Nimon K, Oswald F, Roberts JK (2013) yhat: Interpreting Regression Effects. R package version 2.0-0. URL https://CRAN.R-project.org/package=yhat

Oksanen J, Blanchet FG, Friendly M, Kindt Roeland, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHM, Szoecs E, Wagner H (2018) vegan: Community Ecology Package. R package version 2.5-2. URL https://CRAN.R-project.org/package=vegan

Passy SI (2007) Diatom ecological guilds play distinct and predictable behavior along nutrient and disturbance gradients in running waters. Aquat Bot 86:171–178

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

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ray-Mukherjee J, Nimon K, Mukherjee S, Morris DW, Slotow R, Hamer M (2014) Using commonality analysis in multiple regressions: a tool to decompose regression effects in the face of multicollinearity. Methods Ecol Evol 5:320–328

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

Roberts DW (2016) labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.8-0. URL https://CRAN.R-project.org/package=labdsv

Rocha MP, Bini LM, Siqueira T, Hjort J, Grönroos M, Lindholm M, Karjalainen SM, Heino J (2018) Predicting occupancy and abundance by niche positions, niche breadth and body size in stream organisms. Oecologia 186:205–216

Rohweder J, Rogala JT, Johnson BL, Anderson D, Clark S, Chamberlin F, Runyon K (2008) Application of wind fetch and wave models for habitat rehabilitation and enhancement projects: U.S. Geological Survey, Open-File Report 2008–1200, 43 p.

Roney NE, Kuparinen A, Hutchings JA (2015) Comparative analysis of abundance–occupancy relationships for species at risk at both broad taxonomic and spatial scales. Can J Zool 93:515–519

Schmera D, Podani J, Heino J, Erös T, Poff NLR (2015) A proposed unified terminology of species traits in stream ecology. Freshwater Sci 34:823–830

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

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

Tolonen KT, Hämälainen H, Holopainen IJ, Mikkonen K, Karjalainen J (2003) Body size and substrate association of littoral insects in relation to vegetation structure. Hydrobiologia 499:179–190

Tolonen K, Vilmi A, Karjalainen SM, Hellsten S, Sutela T, Heino J (2017) Ignoring spatial effects results in inadequate models for variation in littoral macroinvertebrate diversity. Oikos 126:852–862

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 abundance-distribution relationship. OIKOS 76:564–570

Verberk WCEP, van der Velde G, Esselink H (2010) Explaining abundance–occupancy relationships in specialists and generalists: a case study on aquatic macroinvertebrates in standing waters. J Anim Ecol 79:589–601

Verberk WCEP, van Noordwijk CGE, Hildrew AG (2013) Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science. Freshwater Sci 32:531–547

Vilmi A, Karjalainen SM, Hellsten S, Heino J (2016) Bioassessment in a metacommunity context: are diatom communities structured solely by species sorting? Ecol Indic 62:86–94

Vilmi A, Karjalainen SM, Landeiro VL, Heino J (2015) Freshwater diatoms as environmental indicators: evaluating the effects of eutrophication using species morphology and biological indices. Environ Monit Assess 187:243

Vilmi A, Alahuhta J, Hjort J, Kärnä O-M, Leinonen K, Rocha MP, Tolonen KE, Tolonen KT, Heino J (2017) Geography of global change and species richness in the North. Environ Rev 25:184–192

Vörösmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich A, Green P, Glidden S, Bunn SE, Sullivan CA, Liermann CR, Davies PM (2010) Global threats to human water security and river biodiversity. Nature 467:555–561

Waters CN, Zalasiewicz J, Summerhayes C, Barnosky AD, Poirier C, Galuszka A, Cearreta A, Edgeworth M, Ellis EC, Ellis M, Jeandel C, Leinfelder R, McNeill JR, Richter D de B, Steffen W, Syvitski J, Vidas D, Wagreich M, Williams M, Zhisheng A, Grinewald J, Odada E, Oreskes N, Wolfe AP (2016) The Anthropocene is functionally and stratigraphically distinct from the Holocene. Science 351: aad2622. doi:10.1126/science.aad2622.

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. R package version 3.0.0. Springer-Verlag New York. https://CRAN.R-project.org/package=ggplot2

Tables

Table 1. Linear regression model statistics presenting the relationship of logit-transformed occupancy and log-transformed mean abundance at occupied sites for diatom and macroinvertebrate species. Bolded p values indicate statistically significant results ($p \leq 0.05$). The explanatory power of the linear models, as indicated by R^2 values, were 0.568 for diatoms ($p < 0.001$) and 0.384 for macroinvertebrates ($p < 0.001$).

		Estimate	SE	t	p
Diatoms	(Intercept)	-4.0807	0.12281	-33.23	<0.001
	Abundance	1.80665	0.08748	20.65	**<0.001**
Invertebrates	(Intercept)	-3.7342	0.2589	-14.43	<0.001
	Abundance	1.4106	0.1667	8.46	**<0.001**

Table 2. Linear regression model statistics for occupancy models for diatoms and macroinvertebrates. Bolded p values indicate statistically significant results ($p \leq 0.05$). The explanatory powers of the linear models, as indicated by multiple R^2 values, were 0.596 ($p < 0.001$) for diatoms and 0.773 ($p < 0.001$) for macroinvertebrates. Followed by linear regression model results are the results of commonality analysis, which inform the unique, common and total contributions of each factor on occupancy of diatom and macroinvertebrate species.

		Estimate	SE	t	p	Unique	Common	Total
Diatoms	(Intercept)	-0.574	0.102	-5.614	<0.001			
	Niche position	-1.100	0.054	-20.257	**<0.001**	0.531	-0.082	0.449
	Niche breadth	0.516	0.066	7.863	**<0.001**	0.080	-0.078	0.002
	Trait vector 1	-0.304	0.201	-1.508	0.133	0.003	-0.003	0.000
	Trait vector 2	0.092	0.254	0.363	0.717	0.000	0.004	0.004
	Trait vector 3	0.231	0.342	0.674	0.501	0.001	0.009	0.009
	Trait vector 4	0.060	0.258	0.232	0.816	0.000	0.001	0.002
	Trait vector 5	1.132	0.281	4.021	**<0.001**	0.021	-0.001	0.020
	Trait vector 6	-0.587	0.336	-1.749	0.081	0.004	0.004	0.008
	Tax. vector 1	-0.002	0.003	-0.459	0.646	0.000	0.001	0.001
	Tax. vector 2	-0.002	0.003	-0.517	0.606	0.000	0.023	0.023
	Tax. vector 3	-0.001	0.004	-0.198	0.843	0.000	0.005	0.005
	Tax. vector 4	-0.005	0.005	-1.103	0.271	0.002	0.002	0.004
	Tax. vector 5	0.003	0.004	0.768	0.443	0.001	0.001	0.002
	Tax. vector 6	0.005	0.005	1.173	0.242	0.002	-0.002	0.000
Invertebrates	(Intercept)	-0.694	0.171	-4.062	<0.001			
	Niche position	-1.429	0.091	-15.737	**<0.001**	0.552	-0.037	0.515
	Niche breadth	0.890	0.108	8.249	**<0.001**	0.152	-0.111	0.040
	Trait vector 1	-0.185	0.302	-0.613	0.541	0.001	0.003	0.004

Trait vector 2	0.348	0.384	0.907	0.367	0.002	0.017	0.019
Trait vector 3	-0.450	0.371	-1.212	0.228	0.003	-0.003	0.000
Trait vector 4	0.684	0.372	1.842	0.068	0.008	-0.002	0.006
Trait vector 5	-0.070	0.393	-0.179	0.858	0.000	0.008	0.008
Trait vector 6	0.097	0.493	0.198	0.844	0.000	0.034	0.034
Tax. vector 1	0.002	0.003	0.499	0.619	0.001	0.006	0.006
Tax. vector 2	0.005	0.005	1.133	0.260	0.003	0.013	0.016
Tax. vector 3	0.013	0.007	1.763	0.081	0.007	-0.007	0.000
Tax. vector 4	0.004	0.007	0.642	0.522	0.001	0.018	0.019
Tax. vector 5	-0.011	0.009	-1.243	0.217	0.003	0.109	0.112
Tax. vector 6	0.000	0.010	-0.015	0.988	0.000	0.002	0.002

Table 3. Linear regression model statistics for mean local abundance models for diatoms and macroinvertebrates. Bolded p values indicate statistically significant results ($p \leq 0.05$). The explanatory power of the linear models, as indicated by multiple R^2 values, were 0.322 for diatoms ($p < 0.001$) and 0.339 for macroinvertebrates ($p < 0.001$). Followed by linear regression model results are the results of commonality analysis, which inform the unique, common and total contributions of each factor on abundance of diatom and macroinvertebrate species.

		Estimate	SE	t	p	Unique	Common	Total
Diatoms	(Intercept)	1.505	0.055	27.264	<0.001			
	Niche position	-0.240	0.029	-8.186	**<0.001**	0.146	-0.027	0.118
	Niche breadth	0.156	0.035	4.396	**<0.001**	0.042	-0.035	0.007
	Trait vector 1	-0.239	0.109	-2.194	**0.029**	0.011	-0.010	0.000
	Trait vector 2	0.276	0.137	2.007	**0.046**	0.009	0.013	0.022
	Trait vector 3	0.219	0.185	1.183	0.238	0.003	0.008	0.011
	Trait vector 4	0.053	0.140	0.377	0.706	0.000	0.000	0.000
	Trait vector 5	0.764	0.152	5.028	**<0.001**	0.055	0.017	0.072
	Trait vector 6	-0.491	0.181	-2.706	**0.007**	0.016	0.004	0.020
	Tax. vector 1	0.003	0.002	1.434	0.153	0.005	0.000	0.004
	Tax. vector 2	0.000	0.002	0.076	0.940	0.000	0.016	0.016
	Tax. vector 3	0.002	0.002	0.679	0.498	0.001	0.002	0.003
	Tax. vector 4	-0.005	0.003	-1.875	0.062	0.008	0.005	0.013
	Tax. vector 5	-0.005	0.002	-2.604	**0.010**	0.015	0.009	0.024
	Tax. vector 6	0.005	0.003	2.145	**0.033**	0.010	0.001	0.011
Invertebrates	(Intercept)	1.408	0.128	11.016	<0.001			
	Niche position	-0.200	0.068	-2.943	**0.004**	0.056	-0.026	0.030
	Niche breadth	0.233	0.081	2.888	**0.005**	0.054	0.034	0.088

Trait vector 1	0.056	0.226	0.248	0.804	0.000	0.043	0.044
Trait vector 2	0.276	0.288	0.960	0.339	0.006	-0.001	0.005
Trait vector 3	-0.795	0.278	-2.862	**0.005**	0.053	-0.026	0.027
Trait vector 4	0.157	0.278	0.564	0.574	0.002	0.005	0.007
Trait vector 5	-0.111	0.294	-0.377	0.707	0.001	0.017	0.018
Trait vector 6	0.615	0.369	1.667	0.099	0.018	0.030	0.048
Tax. vector 1	0.003	0.002	1.149	0.253	0.009	-0.003	0.006
Tax. vector 2	0.007	0.004	2.065	**0.041**	0.028	0.051	0.078
Tax. vector 3	0.010	0.005	1.871	0.064	0.023	-0.022	0.000
Tax. vector 4	0.000	0.005	-0.034	0.973	0.000	0.000	0.000
Tax. vector 5	-0.005	0.007	-0.734	0.465	0.004	0.064	0.067
Tax. vector 6	0.010	0.007	1.443	0.152	0.014	0.006	0.019

Table 4. Linear regression model statistics for occupancy-abundance relationship (i.e. residuals of occupancy-abundance linear model) models for diatoms and macroinvertebrates. Bolded p values indicate statistically significant results ($p \leq 0.05$). The explanatory power of the linear models, as indicated by multiple R^2 values, were 0.514 for diatoms ($p < 0.001$) and 0.757 for macroinvertebrates ($p < 0.001$). Followed by linear regression model results are the results of commonality analysis, which inform the unique, common and total contributions of each factor on the occupancy-abundance relationship of diatom and macroinvertebrate species.

		Estimate	SE	t	p	Unique	Common	Total
Diatoms	(Intercept)	0.788	0.074	10.690	<0.001			
	Niche position	-0.666	0.039	-17.009	**<0.001**	0.450	-0.059	0.391
	Niche breadth	0.235	0.047	4.954	**<0.001**	0.038	-0.037	0.001
	Trait vector 1	0.127	0.145	0.878	0.381	0.001	-0.001	0.000
	Trait vector 2	-0.406	0.183	-2.212	**0.028**	0.008	-0.002	0.006
	Trait vector 3	-0.164	0.247	-0.666	0.506	0.001	0.000	0.001
	Trait vector 4	-0.035	0.186	-0.188	0.851	0.000	0.003	0.003
	Trait vector 5	-0.249	0.203	-1.229	0.220	0.002	0.006	0.008
	Trait vector 6	0.300	0.242	1.238	0.217	0.002	-0.002	0.001
	Tax. vector 1	-0.006	0.002	-2.577	**0.010**	0.010	-0.010	0.000
	Tax. vector 2	-0.002	0.003	-0.819	0.413	0.001	0.007	0.008
	Tax. vector 3	-0.004	0.003	-1.193	0.234	0.002	0.000	0.002
	Tax. vector 4	0.003	0.003	1.008	0.314	0.002	-0.001	0.001
	Tax. vector 5	0.012	0.003	4.589	**<0.001**	0.033	-0.019	0.014
	Tax. vector 6	-0.004	0.003	-1.276	0.203	0.003	0.008	0.010
Invertebrates	(Intercept)	1.054	0.139	7.607	<0.001			
	Niche position	-1.146	0.074	-15.558	**<0.001**	0.576	0.026	0.603
	Niche breadth	0.561	0.088	6.404	**<0.001**	0.098	-0.097	0.001

Trait vector 1	-0.264	0.245	-1.078	0.283	0.003	0.005	0.007
Trait vector 2	-0.041	0.312	-0.132	0.895	0.000	0.015	0.015
Trait vector 3	0.672	0.301	2.230	**0.028**	0.012	0.007	0.019
Trait vector 4	0.463	0.302	1.534	0.128	0.006	-0.004	0.001
Trait vector 5	0.086	0.319	0.270	0.788	0.000	0.000	0.000
Trait vector 6	-0.770	0.400	-1.926	0.057	0.009	-0.005	0.004
Tax. vector 1	-0.002	0.002	-0.881	0.381	0.002	0.023	0.025
Tax. vector 2	-0.005	0.004	-1.293	0.199	0.004	-0.001	0.004
Tax. vector 3	-0.002	0.006	-0.262	0.794	0.000	0.000	0.000
Tax. vector 4	0.004	0.005	0.835	0.406	0.002	0.030	0.032
Tax. vector 5	-0.004	0.007	-0.576	0.566	0.001	0.048	0.049
Tax. vector 6	-0.015	0.008	-1.896	0.061	0.009	-0.006	0.003

Figure captions

Fig. 1. A schematic figure illustrating three species, A, B and C, and their niche positions and niche breadths across an environmental range. Species A has a marginal niche position, while species B and C have non-marginal niche positions (i.e. their niches are located close to the mean environmental conditions). Species A and C have small niche breadths, while species B has a large niche breadth (i.e. it is able to live in a broader range of environmental conditions compared to species A and C). It is noteworthy that two species can have the same niche position although the niche breadth differs (species B vs. species C). Thus, although niche position of species C is non-marginal, it is still a specialist species for those non-marginal conditions.

Fig. 2. Scatter plots describing the relationships between logit-transformed occupancy and log-transformed mean abundance at occupied site of diatoms (A) and macroinvertebrates (B).

Figures

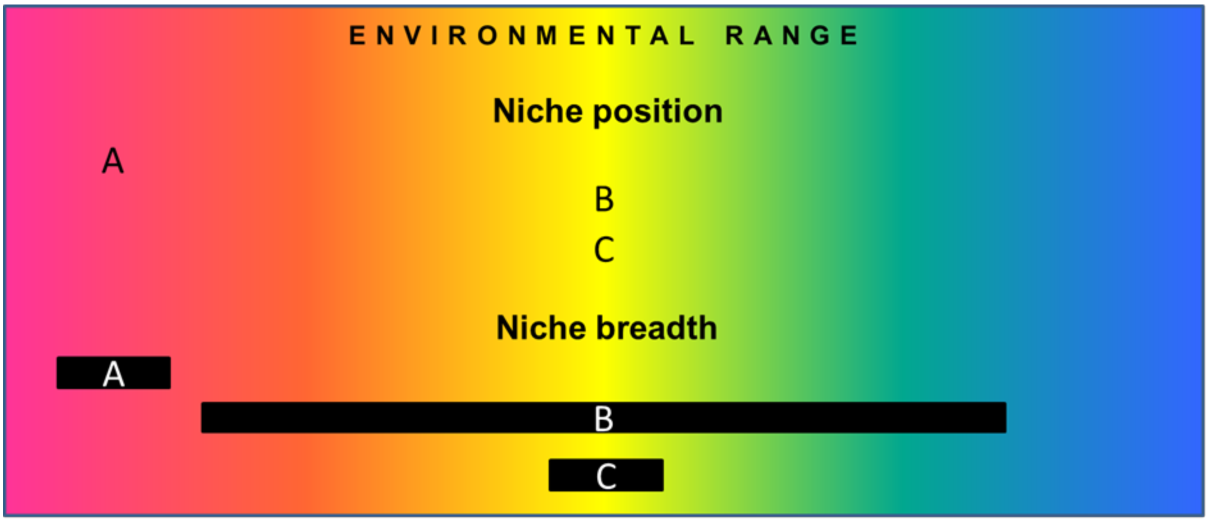
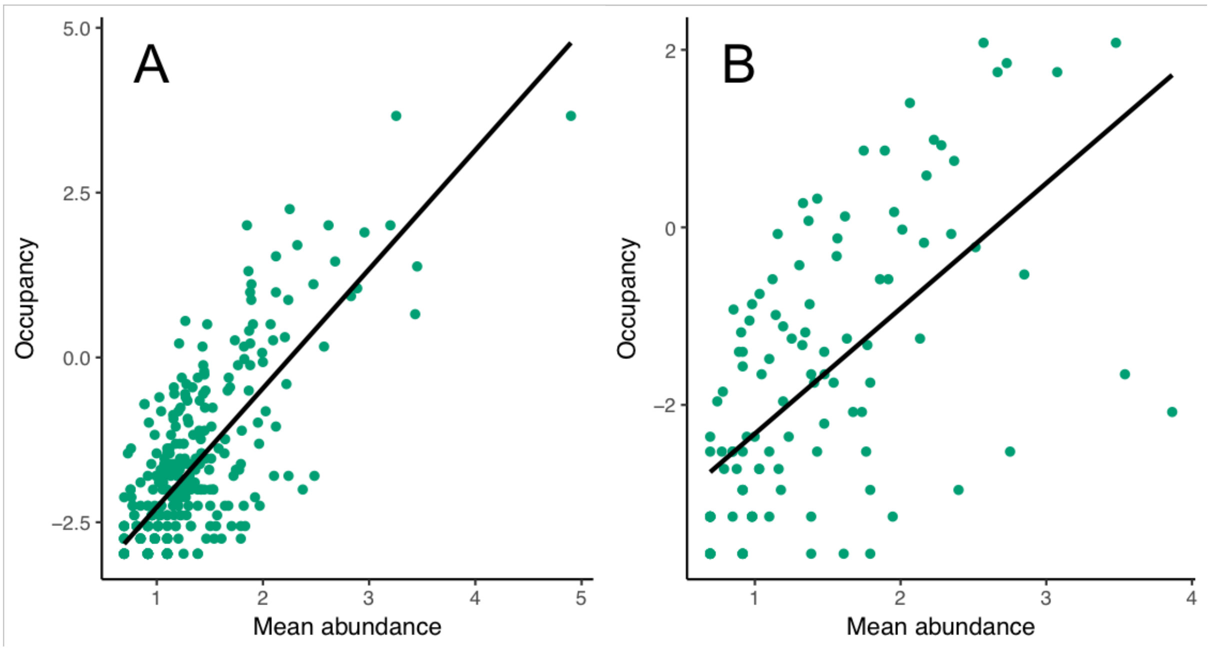


Fig. 1.

590


A
5.0
2.5
0.0
−2.5
Occupancy
1
2
3
4
5
Mean abundance
B
2
0
−2
Occupancy
1
2
3
4
Mean abundance

591 **Fig. 2.**