Author(s): Heino, Jani; Tolonen, Kimmo

Title: Ecological niche features override biological traits and taxonomic relatedness as predictors of occupancy and abundance in lake littoral macroinvertebrates

Year: 2018

Version: Accepted version (Final draft)

Please cite the original version:

Heino, J., & Tolonen, K. (2018). Ecological niche features override biological traits and taxonomic relatedness as predictors of occupancy and abundance in lake littoral macroinvertebrates. Ecography, 41(12), 2092-2103. https://doi.org/10.1111/ecog.03968

Ecological niche features override biological traits and taxonomic relatedness as predictors of occupancy and abundance in lake littoral macroinvertebrates

Jani Heino[1] and Kimmo T. Tolonen[2]
[1]Finnish Environment Institute, Biodiversity Centre, Paavo Havaksen Tie 3, FI-90570 Oulu, Finland
[2]University of Jyväskylä, Department of Biological and Environmental Sciences, P.O. Box 35, FI-40014 Jyväskylä, Finland

Corresponding author: Jani HeinoFinnish Environment Institute, Biodiversity Centre, Paavo Havaksen Tie 3, FI-90570 Oulu, Finland. Email: jani.heino@environment.fi

Decision date: 05-Apr-2018

Abstract

The degree to which species ecological and biological traits determine their distribution and abundance has intrigued ecologists for a long time, and it has seen a revival in recent years. This topic is important because it provides information about the determinants of species rarity and their conservation implications. We examined the effects of niche breadth, niche position, biological traits and taxonomic relatedness on the interspecific occupancy-abundance relationship, as well as on occupancy and abundance, in lake littoral macroinvertebrates. We sampled 48 lakes in a boreal lake district, found altogether 155 species, and calculated regional occupancy (as the proportion of sites occupied) and local abundance (as mean abundance at occupied site) for each species. We determined niche position and niche breadth for each species using the outlying mean index analysis. Also, we calculated trait vectors and taxonomic vectors describing species trait similarity and taxonomic relatedness, respectively, using principal coordinates analysis. We found a strong positive occupancy-abundance relationship that was mostly explained by among-species variation in niche position, followed by niche breadth. Instead, trait vectors and taxonomic vectors tended to be less important in affecting occupancy and abundance than the niche features. Our results strongly suggest that niche position, a measure of habitat availability for littoral macroinvertebrates, is the chief determinant of their occupancy and abundance. This finding has important implications for ecology and conservation of species, as species with marginal niche position, a reflection of low habitat availability, are both regionally rare and locally uncommon. Such species may face double jeopardy if environmental conditions change and affect their preferred marginal habitat types.

Keywords: commonality analysis, distribution, occupancy-abundance relationship, taxonomic relatedness, traits.

Introduction

The degree to which species ecological and biological features (i.e. traits) determine their distribution and abundance has intrigued ecologists for a long time (Brown 1995), and has seen a recent revival in ecology and biogeography (McGill et al. 2006). There are basically two approaches that have been used to examine the relationships between species distribution, abundance and traits. The first is a community-level approach where sites are used as data points, and the effects of environmental variables on the community trait composition are studied (Pilière et al. 2016, Heino and Tolonen 2017a). A second approach uses species as data points, whereby one is interested in how various traits affect the across-species variation in occupancy and abundance (Quinn et al. 1997, Tales et al. 2004). The species-level approach also has a strong association with understanding the origins of the positive occupancy-abundance relationship, which is one of the most pervasive patterns in ecology (Brown 1984, Gaston 1996).

The mechanisms underlying the positive relationship between species occupancy and abundance are many (Borregaard and Rahbek 2010, Passy 2012), but can be categorised into artefactual, ecological and biological mechanisms (Gaston et al. 1997, Gaston and Blackburn 2000). These mechanisms include sampling artefact, phylogenetic non-independence, range position, aggregated spatial distributions, niche breadth, habitat availability, density-dependent habitat selection, vital rates and metapopulation dynamics (Gaston et al. 1997, Borregaard and Rahbek 2010). Here, we consider four general ‘correlates’ of occupancy and abundance that are aligned with the mechanisms listed above, namely niche breadth, habitat availability, similarity of biological traits and taxonomic relatedness that may jointly or independently affect occupancy and abundance (Fig. S1). First, the niche breadth hypothesis has traditionally received a lot of interest (e.g. Slatyer et al. 2013). This hypothesis predicts that species with large niche breadths, based on resource use or environmental tolerance, also have wide distributions and large local abundances (Brown 1984, Gregory and Gaston 2000). Second, another ecological mechanism underlying a positive occupancy-abundance relationship is related to the habitat availability hypothesis (Hanski et al. 1993, Gregory and Gaston 2000), whereby species capable of using common resources or occurring in average environmental conditions in an area are also widely distributed and locally common (Hanski et al. 1993, Venier and Fahrig 1996). This hypothesis thus entails that common species benefit from widely available habitat conditions (i.e. non-marginal niche position), whereas rare species are restricted to marginal habitats that are uncommon in a landscape (i.e. marginal niche position). Third, the degree of similarity of biological traits among species suggests that species having a similar suite of traits should also show similar levels of occupancy and abundance. This is because biological traits should be associated with use of resources and habitats, thereby affecting occupancy and abundance (e.g. Poff 1997). Fourth, as an artefactual mechanism, the degree of phylogenetic similarity among species suggests that closely related species have similar ecological and biological traits, and should thus show similar levels of occupancy and abundance (e.g. Blackburn et al. 1997). However, studies that have taken phylogenetic relatedness into account have found that it has little effect on the occupancy-abundance relationship (Quinn et al. 1997, Tales et al. 2004).

Previous studies comparing the effects of ecological and biological features of organisms on their occupancy and abundance have provided support to different hypotheses. For example, while many studies have found support for the niche breadth hypothesis (Siqueira et al. 2009, Slatyer et al. 2013), a number of recent studies have found more support for the habitat availability hypothesis (Heino and Grönroos 2014, Tonkin et al. 2016). However, relatively few studies have considered multiple ecological and biological features of organisms in affecting occupancy and abundance (Blackburn et al. 1997, Quinn et al. 1997, Cowley et al. 2001, Holt and Gaston 2003, Tales et al. 2004, Foggo et al. 2007). For example, Tales et al. (2004) found that, of the biological traits they examined, only body size explained significant variation around the relationship, being negatively correlated with local abundance and positively correlated with regional occupancy. Tales et al. (2004) also noted the prevalence of habitat availability as the most important determinant of occupancy and abundance.

While freshwater species have recently been used as model organisms to examine occupancy-abundance relationships and the determinants of occupancy or abundance, most of the studies come from lotic (stream and river) environments (Tales et al. 2004, Siqueira et al. 2009, Tonkin et al. 2016, Rocha et al. 2018). The majority of studies from lotic systems have found support for the habitat availability hypothesis, with niche position emerging as the most important correlate of species occupancy and abundance (Tales et al. 2004, Heino 2005, Heino & Grönroos 2014, Tonkin et al. 2016, Rocha et al. 2018). For example, Rocha et al. recently found that niche position derived from measures of local environmental conditions overcame the effects of niche breadth and body size on stream insect and diatom distributions. In contrast, lentic (pond and lake) environments have received less interest amongst ecologists in this context, and to our knowledge only a single study examined occupancy-abundance relationships and underlying mechanisms in lentic macroinvertebrates (Verberk et al. 2010). Verberk et al. (2010) found that, in pond-dwelling macroinvertebrates, trophic position and other life-history traits explained a significant part of the variation around the occupancy-abundance relationship. They found that, among habitat specialists, many species showed synchronized life cycles, a low dispersal capacity or clustered oviposition. Habitat specialists were thus better adapted to predictable habitats, such as permanent ponds. In contrast, among habitat generalists, many species had long-lived adults, spreading reproductive effort in time and space, and were strong dispersers. Habitat generalists were thus better adapted to living in unpredictable habitats, such as temporary ponds (Verberk et al. 2010). It is also possible that there are species specialized on living in unpredictable conditions, but to our knowledge such studies on lentic macroinvertebrates are lacking in the context of the occupancy-abundance relationship. Further information about occupancy-abundance relationships is thus needed from lentic systems because they differ from lotic systems by being more discrete patches in the landscape (Heino 2013) and because they may serve as good model systems for addressing various ecological theories in the freshwater realm (Hortal et al. 2014).

Here, we build on the previous findings of the occupancy-abundance relationship through focusing on a rarely-studied organism group, namely lake littoral macroinvertebrates, in this context. We specifically aimed to determine the relative contributions of niche breadth, niche position (habitat availability), biological traits and taxonomic relatedness in accounting for interspecific variation in occupancy and abundance (Fig. S1). As previous studies in our lake study system have focused on the community-level phenomena (Heino 2013, Heino and Tolonen 2017a), we only used the species-level approach in this study. Based on previous findings, we expected that niche features, including niche breadth (Siqueira et al. 2009) and niche position (Heino 2005) would override the effects of biological traits and phylogenetic relatedness as determinants of occupancy and abundance. However, we also assumed that species traits (Verberk et al. 2010) and taxonomic relatedness (Quinn et al. 1997) would still have unique effects in accounting for variation in occupancy and abundance across species.

We asked the following specific questions: 1) Is there a positive occupancy-abundance relationship in littoral macroinvertebrates? 2) Do niche features, biological traits and taxonomic relatedness explain variation around the occupancy-abundance relationship? 3) Do niche measures, biological traits and taxonomic relatedness explain variation in occupancy? 4) Do niche measures, biological traits and taxonomic relatedness explain variation in mean local abundance? We used a novel approach in the ecological context of linear modelling, namely commonality analysis (Ray-Mukherjee et al. 2014), to account for the total, shared and unique effects of niche breadth, niche position, biological traits and taxonomic relatedness on occupancy and abundance. We also complemented the linear modelling approach using boosted regression trees (Elith et al. 2008) to disentangle the roles of the different predictor variables in affecting occupancy or abundance across species.

Material and Methods

Study area and field surveys

The study area and surveys have been detailed previously (Heino 2013, Heino and Tolonen 2017b), but we will describe the main field methods below. We focused on an area encompassing 170 km^2 that was situated in the Oulankajoki drainage basin in Finland (66^{o}22'N, 29^{o}25'E). Typical study sites varied from small calcareous lakes with high hardness and pH to bog lakes with lower hardness and pH (Heino 2013, Heino and Tolonen 2017b). Most lakes with surface area more than one ha support fish, with at least the most common species, European perch (*Perca fluviatilis*), being present in most lakes (J. Heino pers. obs.). All 48 studied lakes were close to their pristine state and were not strongly affected by human pressures. The mean surface area of the sampled lakes was 11.4 ha, and ranged from 0.8 ha to 92.5 ha.

We measured multiple environmental variables for each lake in September 2005. Water chemistry variables were pH, hardness (ºdH), conductivity (mS/m), colour (mg pt/l) and total phosphorus (µg/l). Lake morphometry variables included lake surface area (A), shoreline length (L), and shoreline development factor measured as follows: [$D_L = L/(2 \times \sqrt{(\pi \times A)}$]. Additional habitat variables were measured about one meter from the shoreline at the five sampling points and incorporated near-shore depth (cm) and cover of macrophytes. Before statistical analyses, all environmental variables, except pH, were logarithmically transformed. Furthermore, substratum particle size and organic material were also measured at the five sampling points. The benthic substratum was divided into nine classes, the percentages of which were estimated at each sampling point: wood, other coarse organic matter, fine organic matter, sand (0.25 – 2 mm), gravel (2 – 16 mm), pebble

(16 – 64 mm), cobble (64 – 256 mm), small boulder (256 – 512 mm), and large boulder (> 512 mm). Principal components analysis on these substratum variables resulted in two uncorrelated composite variables describing benthic substrates (Heino 2008): PCA axis 1 (inorganic bottom: high negative loading for fine organic matter, and high positive loadings for sand, gravel, pebble, cobble, and boulders, variance explained 45.8%) and PCA axis 2 (coarse organic matter: high positive loadings for wood and other coarse organic material, variance explained 17.4%).

Macroinvertebrate sampling was conducted in September 2005, immediately before the environmental sampling. We took five samples at the same spots as the environmental measures in each lake, starting from a random point, after which the sampling locations were evenly distributed across a 100-meter section of the littoral zone (Heino 2008). Samples were taken approximately one meter from the shoreline towards a lake's centre. A 30 cm × 100 cm sample ($0.3\ m^2$) was taken at each location. Each sample comprised a sweep of a D-shaped net (mesh size 0.5 mm) on soft organic bottoms and vegetation, whereas a similar-sized plot was kicked to disturb macroinvertebrates on stony bottoms. All five samples from each lake were pooled in the field. In the laboratory, all macroinvertebrates were separated from associated material and were identified mainly to species, but at least to genus level, including non-biting midges (Diptera: Chironomidae).

Invertebrate species trait and taxonomic relatedness data

We used information on three grouping features (Schmera et al. 2015), including body size, locomotion-substrate relation, and feeding habit (for original sources of trait information, see Tolonen et al. 2017, Heino and Tolonen, 2017a). Body size was considered a continuous variable of maximum dry weight of the aquatic stages, whereas the other two grouping features were dealt with as categorical variables (Tolonen et al. 2017, Heino and Tolonen 2017a). Body size was log-transformed before the analyses to reduce skew. The grouping feature 'locomotion-substrate relation' included five 'traits': swimmers, crawlers, burrowers, semisessiles and sessiles. The grouping feature 'feeding habit' was classified using the following species 'traits': scrapers, piercers, collector-gatherers, filterers, commensals, parasites and predators. These traits are fundamental characteristics of freshwater macroinvertebrates (Merritt and Cummins 1996, Tolonen et al. 2003) because they are related to species' life histories and vulnerability to fish predators (i.e. body size), habitat use (i.e. locomotion-substrate relation), and resource use, approximate food and feeding behavior (i.e. feeding habits). Therefore, they should be strongly correlated to occupancy and abundance variation across species. Eight out of a total of 155 detected species were removed from the present analyses owing to missing trait information. The list of taxa and their traits are shown in Heino and Tolonen (2017a).

We used taxonomic distances between species based on the path lengths in the Linnean taxonomic trees (Clarke and Warwick 1998, Winter et al. 2013) because we did not have true phylogenetic tree including all the species detected. The taxonomic tree used included seven taxonomic levels (i.e. species, genus, family, suborder, order, class and phylum), which were derived from the information given in the Fauna Europaea online database (www.faunaeuropea.org). Such information about taxonomic relatedness has been shown to be a robust proxy for true phylogeny when none is currently available, revealing main patterns in the phylogeny of freshwater macroinvertebrates (Ruhi et al. 2013, Heino and Tolonen 2017b). Unfortunately, we currently lack true phylogenetic information down to the species level, although some recent studies have analysed phylogenetic data at higher levels (Murria et al. 2018).

Data analysis

As a proxy for phylogenetic tree, we used taxonomic information (see above) and calculated taxonomic distances between species using the function 'taxa2dist' available in the *vegan* R package (version 2.0-6, Oksanen et al. 2013). This function works on a taxonomic classification matrix, with species on the leftmost column, genera on the next column, and so on. We used this function to calculate species-by-species taxonomic distance matrix. Subsequently, we used principal coordinates analysis (PCoA, Legendre and Legendre, 2012) to derive 'taxonomic vectors' describing species taxonomic relatedness. See also Diniz-Filho et al. (1998) for a similar and Diniz-Filho et al. (2015) for alternative approaches in a true phylogenetic context. Similarly, we calculated trait distances between species using Gower distance with the function 'gowdis' available in the *FD* R package (version 1.0-12, Laliberte et al. 2014). This function works with both categorical and continuous trait variables, and results in species-by-species trait distance matrix. Subsequently, we ran PCoA on the trait distances to obtain 'trait vectors' describing species trait similarity. We opted to use the PCoA approach to condense and summarise the multidimensional trait data because no single trait, but trait combinations, are likely to be related with species occupancy and abundance. Indeed, Verberk et al. (2013) aptly put it: “Selection pressures do not act independently on single traits, but rather, on species whose success in a particular environment is controlled by many interacting traits. Therefore, the adaptive value of a particular trait may differ across species, depending on the other traits possessed by the species and the constraints of its body plan”.

We used the outlying mean index (OMI) analysis to obtain measures of niche position and niche breadth for each macroinvertebrate species (Dolédec et al. 2000). This method measures the marginality of species habitat distribution (i.e., the distance between the mean environmental conditions used by a species and the average environmental conditions available in the study area). The resulting OMI index measures the “niche position” of a species, and species showing high values of OMI have marginal niches, and species showing low values have non-marginal niches. Species tolerance (Tol) is a variance term measuring the range in the distribution of a species along the sampled environmental gradients (Dolédec et al. 2000), and can be called as “niche breadth” (Tales et al. 2004). Species that show high values of tolerance occur across broad environmental ranges and have large niches, whereas species showing low values occur only across a limited range of conditions and have small niches (Tales et al. 2004, Heino 2005, Siqueira et al. 2009, Heino and Grönroos 2014). Niche position (OMI) and niche breadth (Tol) of each species were computed using the OMI analysis available in the R package *ade4* (version 1.7-10, Chessel et al. 2012). The OMI analysis was based on log-transformed species abundance data and the log-transformed and standardized environmental variables (see above). These variables thus defined the realized habitat niche position and realized habitat niche breadth for each species across the measured environmental gradients.

We used linear regression modelling (LM) to examine the relationships between proportion of sites occupied (logit-transformed) and mean local abundance at occupied sites (log-transformed), as well as between each of these two response variables and niche position, niche breadth, the six taxonomic vectors and the six trait vectors (Fig. S1). We transformed the response variables to approximate normal distributions of these variables and model residuals, and chose to use simple LMs because more complex beta regressions for proportion of sites occupied and negative binomial GLM for mean total abundance provided highly similar results to those of LMs. In addition, commonality analysis (Seibold and McPhee 1979, Ray-Mukherjee et al. 2014) could be done with LMs, and it was used to infer the unique, common and total contributions of each predictor variable to occupancy or mean local abundance. Modelling analyses were conducted and plots were produced using the R packages *stats* (version 3.4.3, R Core Team 2017), *betareg* (version 3.1-0, Zeileis et al. 2016) and *yhat* (version 2.0-0, Nimon et al. 2015).

To complement LMs, we also used Boosted Regression Tree (BRT) analysis (Elith et al. 2008) on the same response and predictor variables as in LMs. BRTs can handle various types of data, have no need for elimination of outliers, take into account non-linear relationships between response and predictor variables, and automatically consider interactions between predictor variables (Elith et al. 2008). We specifically used BRTs to obtain the contributions of each predictor variable on occupancy or mean local abundance to facilitate knowing the relative effects of the predictor variables in determining variation in the response variables. BRTs were based on the Gaussian distribution to retain the comparability of the results with those of LMs. We used the following parameters in BRTs of both occupancy and mean local abundance: tree.complexity = 5, learning.rate = 0.001, and bag.fraction = 0.5. BRTs were run using the R package *dismo* (version 1.1-4, Hijmans et al. 2016).

Results

Common macroinvertebrate groups

Our littoral samples were dominated in the numbers of species by insects (Insecta), including species of midges (Diptera), caddisflies (Trichoptera), beetles (Coleoptera) and mayflies (Ephemeroptera) (Supporting Information, Table S1), and in abundance by insects (Insecta), isopods (Isopoda) or amphipods (Amphipoda) (Heino and Tolonen 2017b). Typically, the same species were both regionally widespread and locally abundant in our study lakes, but the most abundant species varied among lakes. Generally, the most common species was the mayfly *Caenis horaria*, occurring in 46 lakes and being typically amongst the most abundant species in each lake (Heino and Tolonen 2017b). While the most common species belonged to insects, some isopods, amphipods, and gastropods also attained rather high occupancy and abundance across the lakes.

Background analyses of traits and phylogeny

Mantel correlations showed that the first six taxonomic vectors and the first six trait vectors were strongly correlated with their original species-by-species distance matrices. Thus, the Mantel correlation between original Gower trait distance and Euclidean distance on first six trait vectors was strong (r = 0.910). Also, Mantel correlation between original taxonomic distances and Euclidean distance on the first six taxonomic vectors was high (r = 0.918). These background analyses thus indicated that the taxonomic vectors or trait vectors effectively captured the information in the original taxonomic distances and traits distances, respectively. Trait distances and taxonomic distances were significantly, but not strongly correlated (Mantel r = 0.206, P < 0.001).

The OMI analysis showed that the most important environmental variables affecting the distributions of macroinvertebrates across the lakes were PC1-substrate and lake area on the first axis, and total phosphorus (tot-P) and macrophyte cover on the second axis (Supporting Information, Fig. S2). There was considerable variation in niche position (OMI) and niche breadth (Tol) values among species (Supporting Information, Table S1).

The trait PCoA vectors showed that, of functional feeding groups, gatherers had the highest scores on the first axis and piercers had the smallest scores. On the second to the sixth trait PCoA vectors, the species belonging to different functional feeding groups had the highest or the lowest scores (Fig. S3). Similarly, of locomotion-substratum associations groups, burrowers had the highest scores and crawlers the lowest scores on the first PCoA axis, whereas different groups scored the lowest or the highest on the second to sixth PCoA vectors (Fig. S4). Body size correlated most strongly with the first PCoA vector (Fig. S5).

For simplicity, we summarized patterns in the taxonomic PCoA vectors at the level of taxonomic classes, but a similar score plotting approach is valid at any taxonomic level. The taxonomic PCoA vectors showed that the species belonging to Insecta had the highest scores on the first vector and the species belonging to Gastropoda had the lowest scores (Fig. S6), agreeing with the largest phylogenetic distances between species in Insecta and those in Gastropoda in the dataset. Along the other PCoA vectors, species belonging to different taxonomic classes had the highest or the lowest scores on each vector, agreeing with smaller-scale phylogenetic differences between species (Fig. S6).

Relationships between occupancy, abundance and the predictor variables

The relationship between occupancy and mean local abundance was strong, with 66% of variation explained around this relationship (Fig. 1). LMs showed that niche position and niche breadth as well as the first trait vector explained residual variation around this positive occupancy-abundance relationship, with niche position being the most influential predictor of occupancy once the effect of mean local abundance was removed (Table 1). The unique effect of niche position was 18.6% and its total effect was 39.7%, figures which were unrivalled by the other predictor variables in accounting for variation in the residuals of the occupancy-abundance relationship.

LMs also showed that niche position was the most important variable for both occupancy and mean local abundance, whereas the taxonomic and trait vectors seemed to be less influential (Tables 2 and 3). However, taxonomic vectors pco5 and pco6 explained significant variation in occupancy, and trait vector pco1 and taxonomic vector pco5 explained significant variation in mean local abundance. Both models were rather strong, with 79% of variation explained in occupancy and 47% of variation explained in mean local abundance. In summary, niche position was by far the most important predictor, with 32% unique and 71% total contribution to occupancy, and 15% unique and 35% total contribution to mean local abundance in commonality analysis (Tables 2 and 3). The other predictor variables, except niche breadth for occupancy, had considerably lower contributions to the variation in the response variables (unique: < 2%, total: < 6%). Also, even when the unique effects of trait vectors or taxonomic vectors were combined, their contributions were much lower than that of niche position. For occupancy, the combined unique effect of the six trait vectors summed to 0.5% and that of the six taxonomic vectors to 3.1%. For mean total abundance, the combined unique effect of the six trait vectors was 4.6% and that of the six taxonomic vectors was 6.8%. To summarise, niche position clearly overcame the other predictor variable sets in the models of both occupancy and mean local abundance, and traits and taxonomy were more important in the model of mean local abundance than that of occupancy. Also, when different traits were used as categorical predictors in linear models, they did not explain much variation in occupancy and mean local abundance (models not shown), as was also suggested by plotting species of different trait groups in the occupancy-abundance plot using different symbols (Fig. 1).

Boosted regression trees (BRT) models of both occupancy and mean local abundance were strong (Table 4). BRTs generally corroborated the results of LMs in that niche position was clearly the most important variable for both occupancy (relative contribution: 67.6%) and mean local abundance (40.6%), followed by niche breadth (19.8% and 13.2%, respectively), whereas the trait (0.9-2.6 % and 3.3-6.9%, respectively) and taxonomic (0.3-1.6 % and 1.0-6.1%, respectively) eigenvectors had minor influences (Fig. 2). Note that in Fig. 2 only the eight most important predictor variables for occupancy or abundance are shown.

Discussion

We found that regional occupancy was strongly correlated with mean local abundance, and the variation around this relationship was mostly driven by niche position and niche breadth (Table 1). Regional occupancy and local abundance of species may be related to various factors that may or may not be acting in concert (Gaston et al. 1997, Borregaard and Rahbek 2010). Our findings also showed that the occupancy or abundance of lake macroinvertebrates were mostly driven by among-species differences in niche position, followed by niche breadth, whereas species trait similarity and taxonomic relatedness were clearly of lesser importance (Tables 2 and 3). Our results were robust to the statistical approach, as both the traditional linear modelling (LM) approach and a more modern machine-learning method, boosted regression trees (BRT), pointed to the prevalence of niche position as the chief determinant of occupancy and abundance.

Of the nine proposed mechanisms, sampling artefact, phylogenetic non-independence, range position and aggregated spatial distributions are related to potential artefacts in sampling design and data properties (Gaston et al. 1997, Tales et al. 2004), whereas niche breadth, habitat availability, density-dependent habitat selection, vital rates and metapopulation dynamics have underlying biological explanations (Gaston et al. 1997, Borregaard and Rahbek 2010). The artefactual mechanisms, such as sampling artefact, range position and aggregated spatial distributions may always affect, at least to some degree, occupancy-abundance relationships, but their importance may be hard to distinguish from the biological mechanisms. The fourth artefactual mechanisms, phylogenetic non-independence, has been shown to have only minor influence on occupancy and abundance patterns (Blackburn et al. 1997, Quinn et al. 1997, Tales et al. 2004), and our results corroborated this general finding. We found only minor, yet sometimes significant, effects of taxonomic relatedness on occupancy and abundance (see below). Of the biological mechanisms, density-dependent habitat selection, vital rates and metapopulation dynamics cannot be directly tested with the present data, unless they are related to the biological traits or taxonomic relatedness of species (Heino and Grönroos 2014). For example, body size might affect dispersal (Rocha et al. 2018) and thereby affect metapopulation dynamics of species (Tales et al. 2004). The two remaining biological mechanisms, niche position and niche breadth, received considerable support in our analyses, suggesting that these mechanisms are indeed driving the occupancy and abundance patterns of lentic macroinvertebrates.

Niche position has previously been shown to be an important correlate of species occupancy and abundance in aquatic organisms. Recent studies of fish (e.g. Tales et al. 2004), macroinvertebrates (e.g. Tonkin et al. 2016), and diatoms (e.g. Rocha et al. 2018) have shown that niche position overrides other ecological (e.g. niche breadth) and biological (e.g. body size) traits in affecting among-species variation in occupancy and abundance. However, other studies have also found the prevalence of niche breadth in affecting variation in occupancy and abundance (Siqueira et al. 2009, McCreadie and Adler 2014), yet the studies that have truly compared the performance of niche position and niche breadth are still few (Gregory and Gaston 2000, Slatyer et al. 2013). Therefore, it is difficult to decipher the degree to which niche position or niche breadth is a more universal determinant of occupancy and abundance, requiring further testing through a comparative analysis of different datasets (Rocha et al. 2018). Also, it is notable that almost all aquatic studies focusing on this topic are from lotic systems (Tales et al. 2004, Heino 2005, Siqueira et al. 2009, Tonkin et al. 2016), whereas lentic systems have been less examined in this context (Verberk et al. 2010). Verberk et al. (2010) also showed that the degree of specialization of species may affect their occupancy and abundance, but they did not directly test the effects of niche breadth versus niche position on occupancy and abundance in lentic macroinvertebrates. Hence, our result that niche position is an important determinant of occupancy and abundance in lentic macroinvertebrates is a novel finding. This result also concurs with the majority of studies from lotic systems (Table 5), suggesting that the lotic-lentic habitat division has little influence on the main determinants of occupancy and abundance of species.

Typically, biological traits should be strongly related to species occupancy and abundance, with species having similar traits occurring in similar environments (Poff 1997) and showing potentially similar levels of occupancy and abundance (Verberk et al. 2010). Despite this strong assumption that traits should mediate species distributions along environmental gradients, we failed to find strong associations of biological traits with either occupancy or abundance. This was a bit surprising because the traits we used should be strongly related with the resource and habitat use among littoral macroinvertebrates (Merritt and Cummins 1996, Tolonen et al. 2003, Statzner et al. 2008). Why traits generally failed to predict occupancy and abundance? The answer to this question may be related with spatial scale (Brändle and Brandl 2001), use of the wrong traits or inadequate trait information. First, as we focused on a relative small spatial extent (i.e. within an area of 140 km^2) and coarse local sampling (i.e. pooled samples from a 100 m stretch of the littoral zone), it is possible that all trait combinations occur in the environmentally heterogeneous littoral zones of most lakes. Hence, it is also possible that traits, at these spatial scales, do not portray the specifics of the microhabitat conditions that actually select for these traits (e.g. body size, substratum association and feeding habits). Second, it is also possible that we focused on the wrong traits affecting the distributions of littoral macroinvertebrates. This explanation is less likely, as community-level studies have shown that the traits we used characterise well the distributions of littoral macroinvertebrates along environmental gradients (Tolonen et al. 2003, Heino 2008). Third, it is also possible that the trait information we used does not portray the true responses of littoral macroinvertebrate species to environmental gradients in boreal lakes. Unfortunately, we currently lack more specific knowledge on the traits of boreal lentic macroinvertebrates, but phylogenetic information may also be useful in accounting for evolutionarily conserved traits (Harvey 1996) and their potential effects on occupancy and abundance (Quinn et al. 1997).

Phylogenetic information may be used as a proxy for missing trait information, with the assumption that biological traits are phylogenetically conserved (Harvey 1996). This may not always be the case (Gerhold et al. 2015), as some traits may show strong phylogenetic signal and others are more labile (Poff et al. 2006). We found that the effect of phylogeny was at best minor on occupancy and abundance. Why did taxonomic relatedness fail to predict occupancy and abundance? First, it is again possible that the spatial scale of the analysis prevented us from finding phylogenetic signals in occupancy or abundance. Had we sampled over very large spatial areas crossing regional species pool boundaries, we might have found clear phylogenetic signals with related species replacing each other in space and having similar determinants of occupancy and abundance. Second, it is also possible that overall biological traits of freshwater macroinvertebrates are not strongly determined by phylogenetic relations (see also Poff et al. 2006), and then there is no strong match between phylogeny and species distributions (see also Quinn et al. 1997). Third, it is possible that our proxy for phylogenetic information, taxonomic relatedness, is inadequate, and hence we could not find strong associations between phylogeny and occupancy or abundance. Even if our proxy for phylogeny may not portray the intricacies of fine-scaled phylogenetic relations (e.g. Murria et al. 2018), it does mirror major differences between species (e.g. orders, classes and families) that should be most strongly associated with trait differences among distantly-related species (see also Statzner et al. 2008). For example, freshwater snails and fingernail clams (Mollusca) are fully aquatic, whereas most aquatic insects (Insecta) have aquatic larvae and terrestrial adults. Also, the body forms of the aquatic stages of these taxa differ widely. Hence, with such profound trait differences between major taxonomic groups of littoral macroinvertebrates, we should have been able to detect taxonomy-related patterns in occupancy and abundance if they truly existed.

Despite the minor influence of biological traits and phylogenetic relatedness, some trait vectors or taxonomic vectors nevertheless explained significant variation in the response variables, particularly that in abundance. For example, trait-pco1 was a significant correlate of abundance in linear models (LMs), most likely because smaller species have larger local abundance than bigger species (Quinn et al. 1997, Tales et al. 2004). Note that body size was most strongly correlated with this trait vector in our data. Also, tax-pco5, which mainly distinguished snails (Gastropoda) and leeches (Hirudinea) was a significant predictor of abundance. This taxonomic vector may be related to some unmeasured traits differing between snails and leeches, which are also important in determining species abundance. In addition, in boosted regression trees (BRTs), tax-pco5 was the second most important variable after the niche measures in accounting for variation in abundance. However, BRTs also showed that trait-pco4 and some other trait vectors had rather important contributions (>4%) to abundance. Hence, BRTs complemented LMs in showing that variation in mean local abundance among species may also be contributed by biological traits and taxonomic relatedness, although niche position still remains as the most important correlate of abundance.

To conclude, we found a strong positive occupancy-abundance relationship in littoral macroinvertebrates that was mostly affected by niche measures, particularly niche position. The fact that niche position emerged as the most important predictor of variation in regional occupancy and local abundance suggests that species occurring in average environmental conditions do best, if this success is measured by the degree of their distributions. This finding also suggests that marginal species, i.e. those occurring in rare environmental conditions, are also likely to be rare in terms of regional occupancy and local abundance. Those species are expected to face “double jeopardy” (Lawton 1996) because they are rare on two fronts, and are thus candidates for conservation and biodiversity assessment. Our analyses well pointed out those rare species (Supporting Information, Table S1) that may be of assessment concern at least in the present study area if not beyond. Hence, predicting the occupancy and abundance of species in a comparative analysis provides important information for both ecology and conservation of species.

Declarations

*Acknowledgements – We thank all people involved with the field sampling and laboratory analysis. In particular, we would like to thank Pekka Majuri, Timo Muotka, Lauri Paasivirta and Kimmo Sivonen for help during this work. We also thank three anonymous reviewers and the associate editor, Thierry Oberdorff, for insightful comments on the previous manuscript version.

*Author contributions – JH devised the ideas, ran the statistical analyses and led the writing. KTT collected the trait information as well as contributed to the ideas and the writing.

*Conflict of interest – The authors declare no conflict of interest.

References

Blackburn, T. M et al. 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.

Blackburn, T. M. et al. 1997. Abundance-range size relationships in British birds: is unexplained variation a product of life history? – Ecography 20: 466–474.

Borregaard, M. K. and Rahbek, C. 2010. Causality of the Relationship between Geographic Distribution and Species Abundance. – Quart. Rev. Biol. 85: 3-25.

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

Brown, J. H. 1984. On the relationship between abundance and distribution of species. – Am. Nat. 124: 255-279.

Brown, J. H. 1995. Macroecology. – University of Chicago Press, Chicago.

Chessel, D. et al. 2012. ADE-4: Analysis of Ecological Data. Exploratory and Euclidean methods in Environmental Sciences. http://CRAN.R-project.org/package=ADE4.

Clarke, K. R and Warwick, R. M. 1998. A taxonomic distinctness index and its statistical properties. – **J. Appl. Ecol.** 35: 523–531.

Cowley, M. J. R. et al. 2001. Density-distribution relationships in British butterflies: II. An assessment of mechanisms. – **J. Anim. Ecol.** 70: 426-441.

Diniz-Filho J.A.F. et al. 1998. An eigenvector method for estimating phylogenetic inertia. – Evolution 52: 1247–1262.

Diniz-Filho, J.A.F. et al. 2015. The best of both worlds: phylogenetic eigenvector regression and mapping. – Gen. Molec. Biol. 38: 396-400.

Dolédec, S. et al. 2000. Niche separation in community analysis: a new method. – Ecology 81: 2914-2927.

Elith, J. et al. 2008. A working guide to boosted regression trees. – **J. Anim. Ecol.** 77: 802–813.

Foggo, A. et al. 2007. Do developmental mode and dispersal shape abundance-occupancy relationships in marine macroinvertebrates? – **J. Anim. Ecol.** 76: 695-702.

Gaston, K. J. 1996. The multiple forms of the interspecific abundance-occupancy relationship. – Oikos, 76, 211-220.

Gaston, K. J. and Blackburn, T. M. 2000. Pattern and Process in Macroecology. – Blackwell Science, Oxford.

Gaston, K.J. et al. 2000. Abundance-occupancy relationships. – J. Appl. Ecol. 37: 39-59.

Gaston, K. J. et al. 1997. Interspecific abundance-range size relationships: an appraisal of mechanisms. – **J. Anim. Ecol.** 66: 579-601.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. – Biometrics 27: 857-871.

Gregory, R. D. and Gaston, K. J. 2000. Explanations for commonness and rarity in British breeding birds: separating resource use and resource availability. – Oikos 88: 515-526.

Hanski, I. et al. 1993. Three explanations of the positive relationship between distribution and abundance of species. Species diversity in ecological communities: Historical and geographical perspectives (ed. by R. E. Ricklefs and D. Schluter), pp 108-116. – University of Chicago Press, Chicago.

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. 2008. Patterns of functional biodiversity and function-environment relationships in lake littoral macroinvertebrates. – **Limnol. Oceanogr.** 53: 1446-1455.

Heino, J. 2013. Does dispersal ability affect the relative importance of environmental control and spatial structuring of littoral macroinvertebrate communities? – **Oecologia** 171: 971-980.

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

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

Heino, J. and Tolonen, K.T. 2017a. Ecological drivers of multiple facets of beta diversity in a lentic macroinvertebrate metacommunity. – **Limnol. Oceanogr.** 62: 2431-2444.

Heino, J. and Tolonen, K.T. 2017b. Untangling the assembly of littoral macroinvertebrate communities through measures of functional and phylogenetic alpha diversity. – **Freshwat. Biol.** 62: 1168–1179.

Hijmans, R.J. et al. 2016. dismo: Species Distribution Modeling. R package Version 1.1-1. https://cran.r-project.org/web/packages/dismo/index.html.

Holt, A. R. and Gaston, K. J. 2003. Interspecific abundance-occupancy relationships of British mammals and birds: is it possible to explain the residual variation? – Global Ecol. Biogeogr. 12: 37-46.

Hortal, J. et al. 2014. Perspectives on the use of lakes and ponds as model systems for macroecological research. J. Limnol. 73: 46-60.

Lawton, J.H. 1996. Patterns in ecology. – Oikos 75: 145-147.

Laliberté, E. et al. 2014. FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-12. https://rdrr.io/rforge/FD/

McCreadie, J. W. and Adler, P. H. 2014. Abundance–occupancy relationships of larval black flies (Diptera: Simuliidae) in temperate Nearctic streams. – Insect Conserv Divers 7: 523–532.

Merritt, R. W. and Cummins, K. W. 1996. An Introduction to the Aquatic Insects of North America, 3rd edn. – Kendall/Hunt Publishing, Dubuque.

Murria, C. et al. (2018) Ecological constraints from incumbent clades drive trait evolution across the tree-of-life of freshwater macroinvertebrates. Ecography, in press.

Nimon, K., Oswald, F. & Roberts, J. K. (2015) yhat: Interpreting Regression Effects. R Package version 2.0-0. https://cran.r-project.org/web/packages/yhat/index.html.

Oksanen, J. et al. 2013. vegan: Community Ecology Package. R package version 2.0-6. http://CRAN.R-project.org/package=vegan.

Passy, S. I. 2012. A hierarchical theory of macroecology. – Ecol. Lett. 15: 923-934.

Pilière, A.F.H. et al. 2016. On the importance of trait interrelationships for understanding environmental responses of stream macroinvertebrates. – **Freshwat. Biol.** 61: 181-194.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. – J. N. Amer. Benthol. Soc. 16: 391-409.

Poff, N.L. et al. 2006. Functional trait niches of North American lotic insects: trait-based ecological applications in light of phylogenetic relationships. – J. N. Amer. Benthol. Soc. 25: 730-755.

Quinn, R. M. et al. 1997. Abundance-range size relationships of macrolepidoptera in Britain: the effects of taxonomy and life history variables. – Ecol. Entomol. 22: 453-461.

Ray-Mukherjee, J., Nimon, K., Mukherjee, S., Morris, D. W., Slotow, R. & Hamer, M. (2014) Using commonality analysis in multiple regressions: a tool to decompose regression effects in the face of multicollinearity. *Methods in Ecology and Evolution*, **5,** 320–328.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

Rocha, M.P. et al. 2018. Predicting occupancy and abundance by niche position, niche breadth and body size in stream organisms. – **Oecologia**, in press.

Ruhí, A. et al. 2013. Functional and phylogenetic relatedness in temporary wetland invertebrates: current macroecological patterns and implications for future climatic change scenarios. – PLoS ONE 8: e81739.

Schmera, D. et al. 2015. A proposed unified terminology of species traits in stream ecology. – Freshwat. Sci. *34: 823-830.*

Seibold, D. R. and McPhee, R. D. 1979. Commonality analysis: A method for decomposing explained variance in multiple regression analysis. – Human Comm. Res. 5: 355–365.

Siqueira, T. et al. 2009. The role of niche measures in explaining the abundance–distribution relationship in tropical lotic chironomids. – Hydrobiologia 636: 163-172.

Slatyer, R. A. et al. 2013. Niche breadth predicts geographical range size: a general ecological pattern. Ecol Lett 16: 1104–1114.

Statzner, B. et al. 2008. Predicting the abundance of European stream macroinvertebrates using biological attributes. – Oecologia 156: 65-73.

Tales, E. et al. 2004. Density-range size relationship in French riverine fishes. – Oecologia 138: 360-370.

Tonkin, J. D. et al. 2016. Exploring stream communities in a tropical biodiversity hotspot: biodiversity, regional occupancy, niche characteristics and environmental correlates. – Biodivers. Conserv. 25: 975–993.

Tolonen, K.T. et al. 2003. Body size and substrate association of littoral insects in relation to vegetation structure. – Hydrobiologia 499: 179-190.

Tolonen, K.T. et al. 2017. Ignoring spatial effects results in inadequate models for variation in littoral macroinvertebrate diversity. – **Oikos** 126: 852–862.

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

Verberk, W.C.E.P. et al. 2013. Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science. – Freshwat. Sci. 32: 531-547.

Verberk, W. C. E. P. et al. 2010. Explaining abundance-occupancy relationships in specialists and generalists: a case study on aquatic macroinvertebrates in standing waters. – J. Anim. Ecol. 79: 589-601.

Zeileis, A. et al. (2016) betareg. Beta Regression. R Package version 3.1-0. https://www.rdocumentation.org/packages/betareg/versions/3.1-0

Figure Legends

Fig. 1. Relationships between occupancy and mean local abundance, with functional feeding groups (a) and locomotion and substratum associate groups (b) shown by different symbols. The relationship between occupancy and mean local abundance was very strong ($R^2 = 0.664$, $F_{1,146} = 288.7$, $P < 0.001$), but there were no clear differences among different trait groups in the degree of occupancy or mean local abundance. Species are data points in the analysis. N = 148 species.

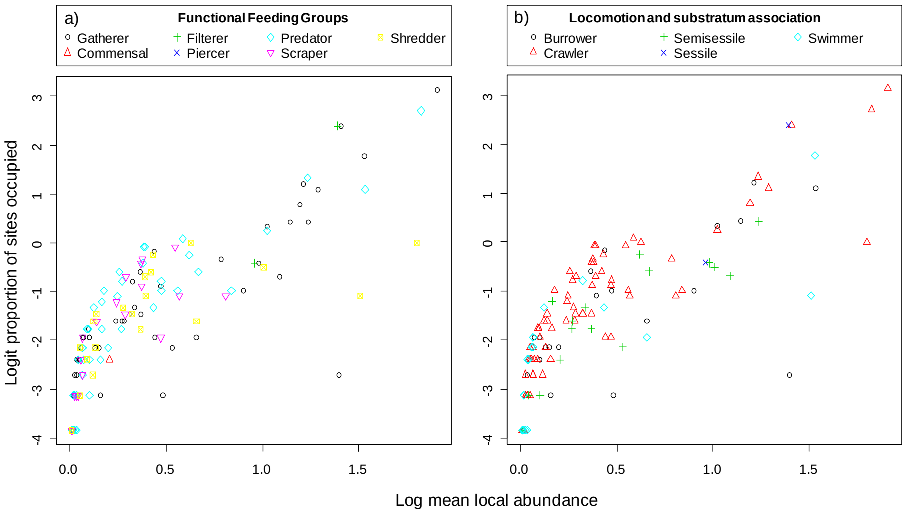

Fig. 2. Partial dependence plots of the predictor variables in boosted regression tree (BRT) analysis based on Gaussian error distribution. Shown are fitted functions and relative influences of the predictor variables (in parentheses) for logit-transformed occupancy (a) and log-transformed mean local abundance (b). For clarity, only the eight most important predictor variables (out of 14 candidate variables) are shown. Code: OMI = Niche position, Tol = Niche breadth. Trait = Trait eigenvectors from PCoA, Tax = Taxonomic eigenvectors from PCoA. N = 148 species.

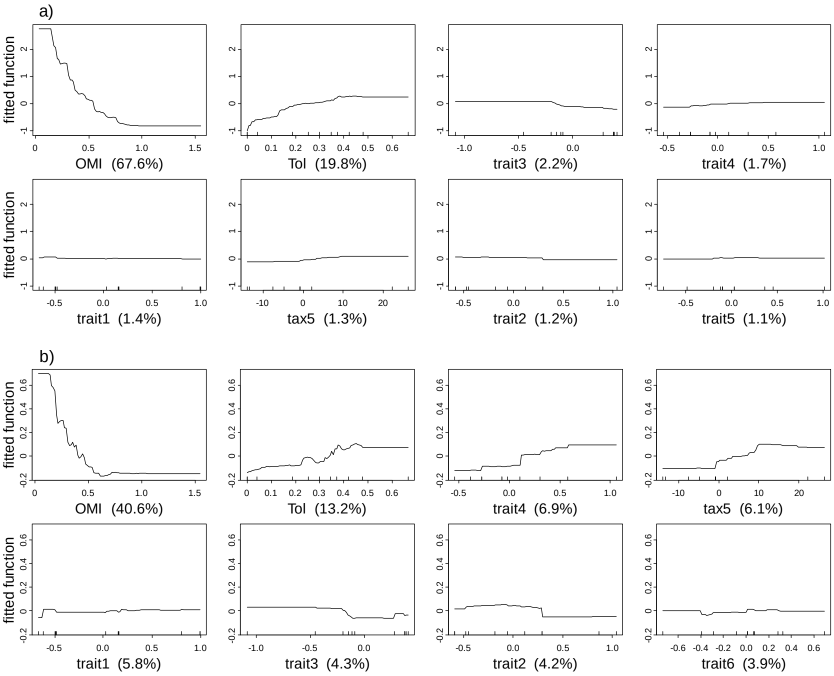

Table Legends

Table 1. Summaries of multiple regression models for explaining variation in the residuals of the occupancy-abundance relationship. Full model statistics: $R^2 = 0.520$, $F_{14,133} = 10.30$, $P < 0.001$. Also, shown are unique, common with other predictors and total effects of each predictor variable on the residuals of the occupancy-abundance relationship based on commonality analysis. N = 148 species.

	Estimate	SE	t	p	Unique	Common	Total
(Intercept)	0.612	0.204	2.997	0.003			
Niche position	-1.551	0.216	-7.179	**<0.001**	0.186	0.211	0.397
Niche breadth	0.983	0.395	2.486	**0.014**	0.022	0.209	0.231
trait-pco1	-0.319	0.119	-2.666	**0.008**	0.026	0.022	0.048
trait-pco2	-0.092	0.146	-0.634	0.527	0.001	0.006	0.007
trait-pco3	0.122	0.173	0.700	0.484	0.002	0.025	0.027
trait-pco4	-0.241	0.200	-1.202	0.231	0.005	-0.001	0.004
trait-pco5	-0.187	0.239	-0.779	0.437	0.002	0.003	0.005
trait-pco6	-0.235	0.192	-1.229	0.221	0.005	-0.005	0.000
tax-pco1	-0.004	0.003	-1.240	0.217	0.006	0.023	0.029
tax-pco2	0.007	0.004	1.798	0.074	0.012	-0.003	0.009
tax-pco3	-0.005	0.006	-0.766	0.445	0.002	0.035	0.028
tax-pco4	-0.000	0.005	-0.060	0.952	0.000	0.002	0.002
tax-pco5	-0.008	0.006	-1.401	0.164	0.007	-0.004	0.003
tax-pco6	-0.002	0.006	-0.258	0.797	0.000	0.000	0.001

Table 2. Summaries of multiple regression models for explaining variation in logit-transformed proportion of sites occupied (occupancy). Full model statistics: $R^2 = 0.794$, $F_{14,133} = 36.60$, $P < 0.001$. Also, shown are unique, common with other predictors and total effects of each predictor variable on the occupancy based on commonality analysis. N = 148 species.

	Estimate	SE	t	p	Unique	Common	Total
(Intercept)	-0.297	0.231	-1.28	0.201			
Niche position	-3.508	0.244	-14.36	**<0.001**	0.319	0.393	0.713
Niche breadth	2.035	0.447	4.55	**<0.001**	0.032	0.377	0.409
trait-pco1	0.127	0.135	0.94	0.349	0.001	-0.001	0
trait-pco2	-0.009	0.165	-0.06	0.952	0	0.018	0.018
trait-pco3	-0.213	0.196	-1.09	0.280	0.002	0.028	0.030
trait-pco4	0.134	0.227	0.59	0.556	0.001	0.004	0.005
trait-pco5	0.165	0.271	0.61	0.543	0.001	0.007	0.008
trait-pco6	-0.002	0.217	-0.01	0.991	0	0.004	0.004
tax-pco1	-0.002	0.003	-0.50	0.615	0	0	0
tax-pco2	0.002	0.004	0.56	0.574	0.001	0.029	0.030
tax-pco3	-0.008	0.007	-1.11	0.271	0.002	0.014	0.016
tax-pco4	-0.010	0.005	-1.82	0.071	0.005	-0.002	0.003
tax-pco5	0.022	0.007	3.35	**0.001**	0.017	0.036	0.053
tax-pco6	-0.014	0.007	-2.06	**0.041**	0.006	-0.002	0.004

Table 3. Summaries of multiple regression models for explaining variation in log-transformed mean local abundance at occupied sites. Full model statistics: $R^2 = 0.466$, $F_{14,133} = 8.29$, $P < 0.001$. Also, shown are unique, common with other predictors and total effects of each predictor variable on the occupancy based on commonality analysis. N = 148 species.

	Estimate	SE	t	p	Unique	Common	Total
(Intercept)	0.662	0.108	6.15	<0.001			
Niche position	-0.695	0.114	-6.11	**<0.001**	0.150	0.196	0.346
Niche breadth	0.374	0.208	1.79	0.075	0.013	0.184	0.196
trait-pco1	0.159	0.063	2.51	**0.013**	0.025	-0.001	0.024
trait-pco2	0.029	0.077	0.38	0.703	0.001	0.010	0.011
trait-pco3	-0.119	0.091	-1.30	0.196	0.007	0.003	0.001
trait-pco4	0.133	0.106	1.26	0.210	0.006	0.011	0.017
trait-pco5	0.125	0.126	0.99	0.323	0.004	0	0.004
trait-pco6	0.083	0.101	0.82	0.414	0.003	0.001	0.004
tax-pco1	0.001	0.002	0.45	0.653	0.001	0.009	0.009
tax-pco2	-0.002	0.002	-0.78	0.435	0.003	0.018	0.021
tax-pco3	-0.001	0.003	-0.33	0.744	0	0.001	0.001
tax-pco4	-0.003	0.003	-1.35	0.180	0.007	-0.005	0.002
tax-pco5	0.011	0.003	3.50	**0.001**	0.049	0.010	0.059
tax-pco6	-0.004	0.003	-1.40	0.164	0.008	-0.004	0.004

Table 4. Summary statistics of the results of the boosted regression tree (BRT) analysis for occupancy or mean local abundance. Occupancy model $D^2 = 0.923$. Abundance model $D^2 = 0.761$. N = 148 species.

Response variable	Mean total deviance	Mean residual deviance	Estimated CV deviance	SE
Occupancy	2.401	0.185	0.478	0.061
Abundance	0.201	0.048	0.104	0.017

Table 5. Comparisons of different studies on the correlates of regional occupancy and local abundance in freshwater organisms.

Study	Organism group	Habitat	Response	Main correlate	Remarks
Tales et al. (2004)	Fish	Lotic	Occupancy	Sites occupied	Body size explained some variation in abundance. Niche breadth was not important.
Tales et al. (2004)	Fish	Lotic	Abundance	Niche position	Mean local abundance and body size accounted for some variation in occupancy. Niche breadth was not important.
Heino (2005)	Macroinvertebrates	Lotic	Occupancy	Niche position	Niche breadth was also significant, but secondary to niche position.
Heino (2005)	Macroinvertebrates	Lotic	Abundance	Niche position	Niche breadth was also significant, but secondary to niche position.
Heino and Soininen (2006)	Diatoms	Lotic	Occupancy	Niche position	Niche breadth and body size were also significant.
Siqueira et al. (2009)	Macroinvertebrates	Lotic	Occupancy	Niche breadth	Two types of niche measured: local environmental niches and catchment niches.
Siqueira et al. (2009)	Macroinvertebrates	Lotic	Abundance	Niche breadth	Two types of niche measured: local environmental niches and catchment niches.
Verberk et al. (2010)	Macroinvertebrates	Lentic	Abundance	Habitat breadth	Occupancy and biological traits were also

					significant.
Heino and Grönroos (2014)	Macroinvertebrates	Lotic	Occupancy	Niche position	Niche breadth was also significant and was contingent of its measure. Species traits were not significant.
Heino and Grönroos (2014)	Macroinvertebrates	Lotic	Abundance	Niche position	Dispersal mode was also significant.
McCreadie and Adler (2014)	Blackflies	Lotic	Occupancy	Niche breadth	Niche position less important.
Tonkin et al. (2016)	Macroinvertebrates	Lotic	Occupancy	Niche position	Niche breadth also significant.
Tonkin et al. (2016)	Macroinvertebrates	Lotic	Abundance	Not significant	Niche position and niche breadth were not important.
Rocha et at. (2018)	Diatoms	Lotic	Occupancy	Niche position	Niche breadth and body size were also significant.
Rocha et al. (2018)	Diatoms	Lotic	Abundance	Niche position	Niche breadth and body size were also significant.
Rocha et al. (2018)	Insects	Lotic	Occupancy	Niche position	Niche breadth was also significant.
Rocha et al. (2018)	Insects	Lotic	Abundance	Niche position	Niche breadth was also significant.
Heino and Tolonen (2018)	Macroinvertebrates	Lentic	Occupancy	Niche position	Niche breadth was also significant, and minor contributions by taxonomic vectors.

Heino and Tolonen (2018)	Macroinvertebrates	Lentic	Abundance	Niche position	Niche breadth was also significant, and minor contributions by trait and taxonomic vectors.