

1 HIGHLIGHTED STUDENT RESEARCH

2 COMMUNITY ECOLOGY - ORIGINAL RESEARCH

3

4

5

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

7

stream organisms

8

9 Mariana P. Rocha^{1,2,*}, Luis M. Bini³, Tadeu Siqueira⁴, Jan Hjort¹, Mira

10 Grönroos⁵, Marja Lindholm¹, Satu-Maaria Karjalainen⁶ and Jani Heino^{2 1}

11

12

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

¹ **Author contribution statement**

MPR and JHeino conceived the main ideas and led the writing of the manuscript. LMB and TS contributed to the statistical analyses and writing of the manuscript. JHjort contributed to gathering the catchment variables and commented on the manuscript. MG led the field sampling and identification of insect samples. ML and SMK did the identification of diatom samples. All authors commented on and approved the final manuscript version.

We strongly believe this article would add considerably to the current discussion of the occupancy-abundance relationships, shedding more light on distribution of species in aquatic systems.

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

16 ³Department of Ecology (ICB), Universidade Federal de Goiás, Goiânia, GO, P.O. Box
17 131, 74001-970, Brazil.

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

20 ⁵University of Helsinki, Department of Environmental Sciences, Section of
21 Environmental Ecology, Niemenkatu 73, FI-15140 Lahti, Finland.

22 ⁶Finnish Environment Institute, Freshwater Center, P.O. Box 413, FI-90014 Oulu,
23 Finland.

24 *CAPES Foundation, Ministry of Education of Brazil, Brasília - DF 70040-020, Brazil

25 Correspondence to: mperezrocha@gmail.com

26

27 Number of pages: 11

28 Number of words: 7236

29

30

31

32 **Abstract**

33

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

52

53 Key words: Algae, distribution patterns, insects, occupancy-abundance relationship, OMI
54 analysis

55

56

57 **Introduction**

58

59 A positive relationship between regional occupancy (i.e. proportion of sites occupied)
60 and local abundance (i.e. mean local abundance at occupied sites) of species has been
61 widely reported for various groups of organisms (Gaston et al. 1998; Gaston and
62 Blackburn 2000; Blackburn et al. 2006). This occupancy-abundance relationship is one
63 of the strongest as well as the most extensively investigated patterns in macroecology
64 (Gaston et al. 2000; Blackburn et al. 2006). It has been observed at different scales,
65 ranging from broad biogeographic regions to small areas within a drainage basins
66 (Cowley et al. 2001; McGill and Collins 2003; Heino and Virtanen 2006; Heino and
67 Grönroos 2014).

68

69 Another commonly-observed macroecological pattern relates to the relationship
70 between body size and range size (Gaston and Blackburn 2000). For example, some
71 studies have demonstrated that algae (e.g. diatoms) are widely distributed (Soininen and
72 Heino 2005), whereas insects are typically restricted to certain regions (Hoffsten 2004).
73 Within macroorganisms, large-bodied species are, however, typically more widely
74 distributed than small-bodied species (Brown 1984; Gaston and Blackburn 2000). Thus,
75 the strength of the relationship between body size and regional occupancy may vary
76 depending on the organism group under study. Also, patterns of geographic distribution
77 may differ within smaller regions, where species occupancies may or may not be related
78 to body size (Heino and Grönroos 2014). Rather, these relationships may be complex
79 and be affected by dispersal mode (Passy 2012) or niche characteristics (Tales et al.
80 2004). For example, in passive dispersers (e.g. microalgae), large body size may limit
81 dispersal and, therefore, their regional distribution (Passy 2007, Passy 2012). On the

82 other hand, in active dispersers (e.g. aquatic insects), large body size is often associated
83 with a greater dispersal ability and colonization success (Hoffsten 2004; but see Saito et
84 al. 2015), which has also been shown generally for insects, fishes, birds and mammals
85 (Brown and Maurer 1987, 1989; Luiz et al. 2012).

86

87 Niche breadth and niche position are two other potential drivers of regional occupancy
88 and local abundance patterns. The niche breadth hypothesis predicts that an increase in
89 the variety of resources consumed and increased tolerance to environmental conditions
90 would enhance population viability and geographic range size (Brown 1984). In other
91 words, niche breadth may be positively related to a wide regional occupancy and higher
92 local abundance when species are generalists (i.e. high values of niche breadth) than
93 specialists (i.e. low values of niche breadth). However, how niche breadth relates to
94 body size in microorganisms is largely unknown. In contrast, for macroorganisms, large
95 body size is often linked to the use of a broad range of resources (Brown 1995), habitats
96 (Griffiths 2010) or tolerance to environmental conditions (Luiz et al. 2012). Besides
97 niche breadth, niche position may also account for a positive relationship between
98 regional occupancy and local abundance of species (Tales et al. 2004; Heino 2005;
99 Siqueira et al. 2009). According to this hypothesis, species having marginal niches (i.e.
100 high values of niche position or low habitat availability) are less widely distributed and
101 locally less common than species able to occur in average habitat conditions (i.e. low
102 values of niche position or high habitat availability) (Hanski et al. 1993; Venier and
103 Fahrig 1996). In other words, the niche position hypothesis predicts that species
104 utilizing common and widespread resources would be regionally widespread and locally
105 abundant (Heino and Grönroos 2014; Tonkin et al. 2016). Our simultaneous
106 understanding of the mechanisms underlying the relationships among body size, niche

107 characteristics and distribution of species is still limited (Blackburn and Gaston 2001;
108 Jenkins et al. 2007, Slatyer et al. 2013). Regional occupancy and local abundance of
109 species are, however, thought to be strongly affected by their body size, niche breadth
110 and niche position in combination (Gaston et al. 2000; Tales et al. 2004; Heino and
111 Grönroos 2014). Even amongst microorganisms (e.g. stream diatoms), some species
112 have broad niches (Cattaneo et al. 1998; Passy and Larson 2011) and others have
113 narrow niches (Pringle 1990; Passy 2007).

114

115 Strong support for the niche-based models in accounting for species regional occupancy
116 or local abundance has been found previously for aquatic organisms (Tales et al. 2004;
117 Heino 2005; Siqueira et al. 2009; Heino and Grönroos 2014). However, very few
118 studies have compared niche breadth and niche position simultaneously as explanatory
119 variables when investigating the occupancy and abundance of species (Hurlbert and
120 White 2007; Slatyer et al. 2013). Also, few studies have compared niche breadth and
121 niche position measures generated by different types of variables (i.e. local
122 environmental versus catchment variables; but see Siqueira et al. 2009). The influence
123 of catchment variables on abiotic variables at local scales (i.e. water chemistry) and
124 subsequently on aquatic organisms has not been studied thoroughly thus far (Soininen et
125 al. 2015). One could expect an influence of catchment-scale variables on local
126 environmental characteristics of streams (Hynes 1970; Corkum 1992). Thus, niche
127 measures based on catchment variables would provide similar or even better
128 explanations for the occupancy-abundance relationships than niche measures based on
129 local environmental variables. However, Siqueira et al. (2009) found that niche
130 measures based on landscape variables explained better occupancy-abundance
131 relationships in comparison to those based on local environmental variables. In addition,

132 analyzing niche variables based on different measurement scales would be an
133 interesting approach because broad-scale studies could rely on catchment-scale
134 variables, which are easier to measure and less costly to obtain compared to local-scale
135 variables (Siqueira et al. 2009, Soininen et al. 2015).

136

137 In this study, we expected that the relationships between regional occupancy, local
138 abundance and body size of species would differ between stream diatoms and stream
139 insects (Fig. 1). In general, the degree of regional occupancy should correlate negatively
140 with niche position and positively with niche breadth for both organismal groups (Fig.
141 1). Based on previous finding from streams (Heino 2005; Tales et al., 2004; Tonkin et
142 al., 2016) we expected that niche position would overcome niche breadth as the
143 predictor of regional occupancy and local abundance. However, some other studies have
144 found the opposite (Siqueira et al., 2009; Slatyer et al., 2013). On the one hand, the
145 regional occupancy and local abundance of diatoms should decrease with increasing
146 body size because larger cells are distributed passively less easily than small cells
147 (Snoeijs et al. 2002; Passy 2007). On the other hand, the regional occupancy of insects
148 should increase with increasing body size because large insects can disperse actively
149 and attain larger distribution than small species at the scale of a drainage basin
150 (Hoffsten 2004; Heino 2013). However, small species are probably more abundant
151 locally than large species in stream insect assemblages, as has been found in general in
152 the animal kingdom (Hutchinson and MacArthur 1959; Brown and Maurer 1989).
153 While we explored how environmental or catchment niches linked best with variation in
154 regional occupancy and local abundance, we did not form any specific hypotheses
155 owing to a little previous information about the topic (but see Siqueira et al. 2009).

156

157

158 Materials and Methods

159

160 Study area

161 We used diatom and insect data collected in 54 wilderness streams in the Tenojoki
162 drainage basin (centered on 70°N, 27°E; total basin area: 16386 km²) in Finland (see
163 map in Supporting information, Fig. S1). The surveys were conducted in 2012. Stream
164 waters in this drainage basin are highly oligotrophic and have pH ranging from 6.3 to
165 6.7 (Heino 2013; Schmera et al. 2013). The study area is very sparsely human-
166 populated, and human activities (e.g. forestry and agriculture) are uncommon.

167 Consequently, streams in our study area are pristine or near-pristine, providing excellent
168 possibilities to study mechanisms underlying the occupancy and abundance of species
169 in the virtual absence of human influences (Heino, Ilmonen and Paasivirta 2014). The
170 topography of the study area is dominated by rounded mountains (i.e. fells), and
171 elevation varies between 10 and 640 m above the sea level.

172

173 **Insect and diatom data**

174

175 **Insects**

176 Each of the 54 stream sites surveyed in this study was sampled using a 3-min kick-net
177 (net mesh size: 0.3 mm) sampling effort. The sample for each site consisted of six 30-s
178 subsamples divided across variations in depth, current velocity, particle size and moss
179 cover found within *ca.* 50 m². The six subsamples were pooled into a composite sample
180 representing a single sample for each stream and immediately preserved in 70 % alcohol
181 in the field. In an attempt to cover most of the different “habitat types” within a riffle

182 site (e.g. variation in depth, current velocity, particle size and moss cover), the use of
183 this kind of stratified sampling at a given site (i.e. pooling samples) is used in order to
184 cover well enough the most of the environmental variation at each site (Johnson et al.,
185 2004; Sandin and Johnson 2004). Our sampling effort has been shown to perform
186 effectively, capturing a great share of species at each surveyed stream site and only
187 missing very rare tourist species (Mykrä et al. 2006).

188

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

194

195 Based on maximum larval length, the following ranks were used to represent insect
196 body size (BS): (1) 0-0.25 cm; (2) 0.25-0.50 cm; (3) 0.5-1 cm; (4) 1-2 cm and (5) 2-4
197 cm (see Supporting Information Fig. S9 C and D for insect size class distributions). This
198 information was provided by personal communication with S. Doledéc (Université
199 Lyon, France) J. Ilmonen (Metsähallitus, Natural Heritage Services, Finland) and L.
200 Paasivirta (Salo, Finland), along with data from Tachet et al. (2010).

201

202 **Diatoms**

203 The same 54 stream sites surveyed for insects were also surveyed for diatoms. Five to
204 10 stones of size 10-30 cm were randomly selected at each site and algae were brushed
205 from areas of 5 cm × 5 cm on stone surfaces in order to get a brushed area of 250 cm² in
206 total for each study site. The samples from each brushing were pooled into a composite

207 sample at each site. In the laboratory, the samples were treated using acid combustion
208 (HNO_3 : H_2SO_4 , 2:1) to oxidize frustules (SFS-EN 14407, 2005). Thereafter, diatom
209 slides were prepared with Naphrax mountant. Subsamples of about 500 valves per site
210 were identified and counted using DIC microscope with $1000 \times$ magnification. Such an
211 effort is typical in studies of stream diatom communities (e.g. Kelly et al., 1998). We
212 used the following diatom cell size (BS) ranks based on Rimet and Bouchez (2012): (1)
213 $0-99 \mu\text{m}^3$; (2) $100-299 \mu\text{m}^3$, (3) $300-599 \mu\text{m}^3$, (4) $600-1499 \mu\text{m}^3$; (5) $\geq 1500 \mu\text{m}^3$ (see
214 Supporting Information Fig. S9 A and B for diatom size class distributions). Most
215 diatoms were identified to species level (c. 98%) and few to genus level (c. 2%),
216 totalizing 101 taxa that occurred at more than two sites. Again, for simplicity, we will
217 call all taxa as species.

218

219 **Local stream environmental and catchment variables**

220

221 For the local stream environmental variables, we measured current velocity (m s^{-1}) and
222 depth (cm) at 30 random spots in a riffle. Also, we measured the mean width (m) of the
223 sampling site based on five cross-channel measurements. Moss cover (%) and river
224 bottom particle size (i.e. % of particle size classes were visually estimated at ten 1 m^2
225 plots at random locations in each riffle site. For river bottom particle size, we visually
226 estimated in each plot and apply a modified Wentworth's scale (1922): sand (0.25-
227 2mm), gravel (2-19 mm), pebble (16-64 mm), cobble (64-256 mm) and boulder (256-
228 1024 mm). The mean values for each site were used in subsequent statistical analysis.
229 Shading (%) by riparian vegetation was visually estimated at each sampling site (50 m^2)
230 by a person standing in the stream. Additionally, we measured pH and conductivity
231 ($\mu\text{S/cm}^{-1}$) in the field. Also, water samples were taken, frozen and subsequently

232 analyzed for total nitrogen ($\mu\text{g/L}^{-1}$), color (Pt – Co mg/L), iron ($\mu\text{g/L}^{-1}$) and manganese
233 ($\mu\text{g/L}^{-1}$) in the laboratory, following the Finnish national standards (National Board of
234 Waters and the Environment 1981).

235

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

251

252 Taking into account the catchment features for calculating niche breadth (NB) and niche
253 position (NP) for each species, we will call them as “catchment niches” (NB-cat and
254 NP-cat), and compare them with the niche characteristics based on local environmental
255 variables, called as “local environmental niches” (NB-env and NP-env). Those
256 comparisons should provide us different information about species niches, and the

257 different niche measures (“local environmental niches” vs. “catchment niches”) may
258 thus result in different information about the occupancy-niche relationships for the
259 different organism groups.

260

261 **Data analysis: niche measurements and models**

262 In order to generate measures of NP and NB for each species, we used the sets of local
263 environmental and catchment variables separately, applying the outlying mean index
264 (OMI) analysis (Doledéc et al. 2000) to obtain our “local environmental niches” and
265 “catchment niches”. This method calculates the marginality of habitat distribution of
266 species along the measured environmental conditions (i.e. in this case environmental or
267 catchment conditions) within a region. OMI analysis measures the NP by calculating the
268 distance from the mean environmental conditions (centroid) used by a species and the
269 mean environmental conditions (niche hyperspace) of the entire surveyed area. We
270 calculated niches measures based on the full set of environmental and catchment
271 variables (see Supplementary Information Table S2 and S3 respectively). Thus, species
272 having high values of NP have marginal niches (i.e. low habitat availability), and those
273 species having low values have non-marginal niches (i.e. high habitat availability). OMI
274 also provides a measure of tolerance, or NB. Species having high values of NB occur
275 across a broad range of environmental or catchment conditions, whereas species that
276 have low values of NB occur in a limited range of conditions. NP and NB were
277 computed for each taxon, employing each set of variables (environmental or catchment)
278 separately and using the OMI analysis in the R package “*ade4*” (Chessel et al. 2012)
279 and performed in the R environment (R Development Core Team, 2013). To obtain
280 reasonable measures of NB, the OMI analysis results were derived for species occurring
281 at more than two sites in the datasets.

282

283 Our first response variable was the proportion of occupied sites. For each group of
284 organisms, we used two generalized linear models with a beta distribution (beta
285 regression) to analyse the relationship between this response variable and the predictor
286 variables. The first model, which we called “local environmental” model, included NP-
287 env, NB-env, and body size as explanatory variables. The second model, which we
288 called “catchment model”, included NP-cat, NB-cat, and body size as explanatory
289 variables (see below). We compared the performance of these models (i.e. “local
290 environmental” and “catchment” models) using the Akaike information criterion
291 (Burnham and Anderson 2002), and visual interpretation of the scatterplots as the
292 criterion to judge the importance of variables within models. We chose this modelling
293 approach (i.e. beta regression) because the proportion of occupied sites assumes values
294 in the interval [0,1]. In these cases, the use of beta regression is advisable because the
295 alternative approach, which consists of logit transforming the data and using ordinary
296 least square regression, has drawbacks. In addition to heteroscedasticity, for example,
297 proportions and rates are typically beta distributed. Thus, assuming a normal
298 distribution for estimation and hypothesis testing purposes, even after logit
299 transformation, may be inaccurate, especially when sample size is small (Ferrari and
300 Cribari-Neto, 2004).

301

302 Our second response variable, mean local abundance at the occupied sites (after
303 rounding the mean values to integers), was also regressed against NP-env, NB-env and
304 body size (“local environmental” model) and against NP-cat, NB-cat and body size
305 (“catchment environmental”). As before, we compared the performance of these models
306 using an information-theoretic approach (Burnham and Anderson 2002). Due to data

307 overdispersion, we used negative binomial generalized linear models (Zeileis et al.
308 2008) to test the relationships between mean local abundance and the explanatory
309 variables.

310

311 For all models described above, diatom and insect body sizes (BS) were used as ranked
312 variables. Also, NP and NB were log-transformed prior to analyses. In short, the four
313 models analyzed, for diatoms and insects, were: (i) Occupancy \sim NP-env + NB-env +
314 BS; (ii) Occupancy \sim NP-cat + NB-cat + BS; (iii) Mean local abundance \sim NP-env +
315 NB-env + BS; (iv) Mean local abundance \sim NP-cat + NB-cat + BS. We ran these pairs
316 of competing models (i.e. local environmental vs. catchment) to formally avoid
317 collinearity. Also, by doing so, we avoided collinearity problems caused by the rather
318 strong relationship between NP-env and NP-cat ($r = 0.79$ and $r = 0.81$, for insect and
319 diatom datasets, respectively). For correlation coefficients showing the relationships
320 between the variables for diatoms and insects used in the models see Supporting
321 Information Table S1. Using two separate models, the variance inflation factors (VIF)
322 were always lower than 2.0 (maximum = 1.44).

323

324 To support our models and avoid biased results, we performed a re-sampling analysis
325 following a previous study (Siqueira et al. 2009). This is because OMI analysis uses
326 information on abundance and distribution, potentially leading to circularity when NB
327 and NP are used as predictors of abundance and distribution. We thus randomly divided
328 the whole dataset in two subsets; one for estimating regional occupancy, and the other
329 for niche measurements (NP and NB). Thus, with each of those acquired subsets, we ran
330 occupancy models as described above. The random split and models were repeated
331 1000 times for both insect and diatom data subsets. We reported both overall pseudo- R^2

332 values and *P*-values associated with the explanatory variables that resulted from the
333 randomizations. Thus, pseudo- R^2 and *P*-values justify, respectively, the strengthening
334 of resampling procedure in assessing which model (environmental niche or catchment
335 niche) or which predictor variable (NP-env, NB-env, NP-cat, NB-cat and body size) was
336 superior based on how often these values were significant. Analysis was performed in
337 the R environment (R Development Core Team, 2013). However, due to estimation
338 problems with the negative binomial generalized linear models when using the
339 resampling approach (the algorithm did not converge in many cases), this procedure
340 was adopted for the beta regression models only.

341

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

353

354 **Results**

355

356 **Diatoms**

357 The OMI analysis showed that the most important local environmental variables related
358 to the distribution of diatoms across the studied area were moss and shading (OMI axis
359 1), pebble, conductivity, and boulder (OMI axis 2) (Supporting information, Fig. S3A).
360 Of the catchment variables, mean slope and lake percentage (OMI axis 1), greenness
361 and mire (OMI axis 2) were the most important variables (Supporting information, Fig.
362 S3B).

363

364 Diatoms showed a strong and positive relationship between occupancy and local mean
365 abundance (pseudo $R^2 = 0.498$). The regional occupancy of diatoms was significantly
366 and negatively related with niche position and body size. We also found a positive
367 relationship between this response variable and niche breadth. These patterns of the
368 regression coefficients were found independently of the variables used to define niche
369 characteristics (i.e. local environmental or catchment variables; with pseudo $R^2 = 0.74$
370 and pseudo $R^2 = 0.57$, respectively). In general, niche position was the main predictor
371 variable (see Supporting Information Fig. S5), and the model based on local
372 environmental niches was clearly superior to the one based on catchment niches (delta
373 AIC = 67.25; Table 1).

374

375 For diatom abundance, we also found that niche position was the main explanatory
376 variable (see Supporting Information Fig. S6), and that the model based on local
377 environmental niches was superior to the one based on catchment niches (delta AIC =
378 8.98). NB-env was not a significant predictor of diatom abundance (Table 1).

379

380 Resampling analysis strengthened the role of niche position as the main predictor of
381 diatom occupancy. The role of niche breadth and body size, for both local and

382 catchment models, was less certain as we found P -values < 0.05 associated to these
383 variables in only a few cases (see Supporting Information Fig. S10 and Fig. S12 for
384 resampling procedure results).

385

386 **Insects**

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

393

394 When exploring the relationship between the response variables, we found a strong and
395 positive relationship between regional occupancy and local abundance (pseudo $R^2 =$
396 0.429). Most of the variation in regional occupancy among the 74 insect species was
397 explained by our beta regression models (local environmental niche model: pseudo $R^2 =$
398 0.73; catchment niche model: pseudo $R^2 = 0.60$). Insect occupancy was significantly and
399 positively related with niche breadth and negatively related with niche position. Niche
400 position was the main predictor of insect occupancy (see Supporting Information Fig.
401 S7), and the model based on catchment variables was clearly less important than the
402 model based on local environmental (delta AIC = 26.35; Table 1).

403

404 The pseudo coefficients of determination of the models predicting insect local
405 abundance were substantially lower than those predicting occupancy (local
406 environmental niche model: pseudo $R^2 = 0.31$; catchment niche model: pseudo $R^2 =$

407 0.18). Considering the local environmental model, we found a negative relationship
408 between niche position and insect abundance. This response variable was also positively
409 correlated with niche breadth. However, only niche position was significantly and
410 negatively correlated with insect abundance (see Supporting Information Fig. S8) when
411 the catchment variables were used to estimate niche characteristics ($\Delta AIC = 13.17$;
412 Table 1).

413

414 Resampling analysis also strengthened the role of niche position as the main predictor
415 of insect occupancy. Again, we found P -values < 0.05 associated to niche breadth and
416 body size, for both local and catchment models, only in few cases (see Supporting
417 Information Fig. S11 and S12 for resampling procedure results).

418

419 **Discussion**

420

421 When focusing on the occupancy-abundance relationship, our findings concurred with
422 those of Gaston et al. (1997) and Gaston and Blackburn (2000), with both diatoms and
423 insects of high-latitude streams showing a positive occupancy-abundance (i.e. locally
424 abundant species tended to be regionally widespread). Also, we found that niche
425 position was always the main predictor of regional occupancy or local abundance of the
426 studied organisms. In addition, body size was a significant predictor of diatom regional
427 occupancy and local abundance. The recurrence of niche position calculated based on
428 the local environmental variables as the main predictor is an interesting result, as it
429 shows that the degree of species marginality in terms of local environmental variables is
430 important for their regional occupancy and local abundance. In the following, we will

431 consider these findings in the light of macroecological patterns (e.g. Passy 2012) and
432 the characteristics of high-latitude stream systems (e.g. Wrona et al. 2013).

433

434 We expected that body size would explain variation in the regional occupancy and local
435 abundance of stream organisms. Body size is a proxy for different attributes, including
436 resource utilization (White et al. 2007; Borregaard and Rahbek 2010; Passy 2012) and
437 dispersal ability (Hoffsten 2004; Passy 2007; Heino 2013). Our findings pointed out that
438 body size was, as predicted, a negative correlate of diatom regional occupancy and local
439 abundance, indicating what is expected for small organisms (Passy 2012). The relative
440 range of diatom body size was larger than that for insect body size and, thus, it may help
441 to explain the body size effect on diatoms only. Contrary to our results, Hoffsten (2004)
442 showed that body size (or wing size) was an important predictor of insect regional
443 occupancy. One possible explanation for this discrepancy is that niche characteristics
444 may simply be more important than body size in determining the regional occupancy
445 and local abundance of insect species in high-latitude streams. High-latitude insect
446 species also have specific adaptations to respond to seasonal and year-to-year variability
447 (e.g. Danks 2004), which may not necessarily be related to body size. For example, the
448 role of body size may be more complex in high-latitude streams (e.g. most species have
449 one generation per year, and large species even have ‘slower’ life cycles than one year),
450 although this trait is related more directly to multiple life history characteristics and
451 dispersal capacity in tropical streams (e.g. Saito et al. 2015).

452

453 Niche position and breadth have been regarded as important factors determining
454 variation in regional occupancy and local abundance among species (Tales et al. 2004;
455 Heino 2005; Siqueira et al. 2009; Heino and Grönroos 2014; Tonkin et al. 2016). Our

456 models revealed that niche position, especially when based on local environmental
457 variables, was the most important factor affecting the regional occupancy and local
458 abundance of stream organisms. Species with marginal niche positions tended to have
459 lower abundances and to be less widespread than those with non-marginal niche
460 positions (i.e. with smaller values for this variable). Previous studies have indicated that
461 the spatial configuration of the sampling sites can have a major influence on the
462 estimation of niche measures that are used to predict abundance and distribution
463 patterns (Siqueira et al. 2009). In a similar vein, Sály and Erös et al. (2016) found a
464 strong effect of sampling design on the results of variation partitioning, a method often
465 used in metacommunity ecology to infer the importance of environmental and spatial
466 variables in explaining variation in local communities. However, independently of the
467 group of organisms, our results provided consistent support for local habitat availability
468 (i.e. niche position based on local environmental variables) as a determinant of regional
469 occupancy and local abundance. These results were also robust to the spatial
470 configuration of the sampling sites for regional occupancy, as indicated by our
471 resampling analysis. However, given that the resampling procedure applied to regional
472 occupancy model generated results that agreed with those obtained with the entire
473 dataset, we are confident that local abundance models are not biased because for both
474 diatoms and insects the regional occupancy-local abundance relationship is rather
475 strong.

476

477 Previous studies in stream environments have also shown that niche breadth is a
478 significant predictor of regional occupancy and/or local abundance (Heino 2005;
479 Siqueira et al. 2009; Heino and Grönroos 2014). The difference with our result (i.e.
480 niche position more important than niche breadth) may be related to the features of

481 high-latitude streams, which are characterized by low temperatures, short growing
482 seasons and harsh winter conditions (Wrona et al., 2013; Tolonen et al. 2017). These
483 factors may dictate that species marginality (niche position) along harsh environmental
484 gradients is important when it comes to determining species regional occupancy and
485 local abundance, suggesting that environmental factors may also overwhelm dispersal
486 constraints in high-latitude streams organisms (Kärnä et al. 2015; Tolonen et al. 2017).
487 We found that occupancy-niche position and abundance-niche position relationships
488 were very strong for both diatoms and insects. Thus, diatom and insect species were
489 mostly distributed based on the availability of environmental conditions. The lower
490 level of support for the niche breadth hypothesis has also been reported in previous
491 studies (Gregory and Gaston 2000; Cowley et al. 2001; Tales et al. 2004), which might
492 in part be related to difficulties in generating adequate niche breadth measurements for
493 species (Gaston 1994). These difficulties may be related to missing, yet important,
494 variables describing niche breadth and the multidimensional nature of the niche
495 (Hutchinson 1957). However, we measured the most influential local environmental and
496 catchment variables affecting the distributions of species in northern streams (Heino and
497 Soininen 2006; Heino and de Mendoza 2016), suggesting that our niche breadth
498 measures should not be worse than the niche position measures.

499

500 When we focused on different variables to calculate the niche measures, the “local
501 environmental niches” were more important for local abundance and regional
502 occupancy of both diatoms and insects, whereas “catchment niches” were typically less
503 important. The relationships between regional occupancy, local abundance and niche
504 characteristics thus seem to be affected by the types of variables used to generate niche
505 indices. It can be expected that local environmental niches are more important for

506 species regional occupancy and local abundance, as it is at the local scale where species
507 are ultimately selected to occur in certain sites. However, environmental variation at
508 larger scales, such as that at the catchment scale, may also affect regional occupancy
509 and local abundance (Poff 1997; Vinson and Hawkins 1998). Hence, Siqueira et al.
510 (2009), who sampled variation across catchments, also found niche indices defined by
511 landscape variables to be important for insect regional occupancy and local abundance
512 across tropical streams, whereas the distributions of high-latitude insects appear to be
513 more related to differences in local environmental factors than catchment physiographic
514 aspects.

515

516 **Conclusions**

517

518 Our results showed a strong positive occupancy-abundance relationship in both stream
519 diatoms and insects. We also showed that niche-based patterns are important predictors
520 of regional occupancy and local abundance of diatoms and insects across high-latitude
521 streams. A consistent finding was that mainly the niche position hypothesis (Hanski et
522 al. 1993) was supported in our study, whereas the niche breadth hypothesis (Brown
523 1984) was less important as the mechanism underlying variation in the regional
524 occupancy and local abundance of stream diatoms and insects. These findings did not
525 vary substantially if the species niches were characterized by local (“local
526 environmental niches”) or catchment variables (“catchment niches”). However, models
527 based on local environmental variables were superior to those based on catchment
528 variables. Body size was related to mean local abundance and regional occupancy of
529 diatoms, but it was not important for insects. Our results thus add to current
530 macroecological research by showing the strong importance of niche position for

531 occupancy and abundance in rarely-studied organisms (e.g. diatoms and insects) and
532 ecosystems (i.e. high-latitude wilderness streams).

533

534 **Acknowledgments**

535 MPR would like to thank CAPES Foundation - Brazil for her full PhD grant acquired
536 (grant number 11877-13-8) and also thank Finnish Environment Institute for receiving
537 her as a visiting researcher. TS is funded by grant #2013/50424-1, São Paulo Research
538 Foundation (FAPESP). This study was also supported by grants from the Academy of
539 Finland to JHeino and JHjort. We thank Sirkku Lehtinen, Olli-Matti Kärnä and Laura
540 Tokola for assisting with field sampling, and Sylvain Doledéc, Jari Ilmonen and Lauri
541 Paasivirta for insect body size information. Lauri Paasivirta identified the larvae of non-
542 biting midges and Jari Ilmonen those of blackflies. We thank Olli-Matti Kärnä for
543 providing the map of the study area.

544

545 **References**

546

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

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

552 Blackburn TM and Gaston KJ (2001) Linking patterns in macroecology. *J. Anim. Ecol.*
553 70: 338–352.

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

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

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

561 Brown JH (1995) *Macroecology*. University of Chicago Press.

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

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

- 567 Burnham KP and Andreson DR (2002) Model selection and multimodel inference: a
568 practical information-theoretic approach. Springer Science Press.
- 569 Cattaneo A, Asioli A, Comoli P and Manca M (1998) Organisms' response in chronically
570 polluted lake supports hypothesized link between stress and size. *Limnol.*
571 *Oceanogr.* 43: 1938–1943.
- 572 Chessel D, Dufour AB and Dray S (2012) ADE-4: analysis of ecological data. exploratory
573 and euclidean methods in environmental sciences. In: [http://CRAN.R-](http://CRAN.R-project.org/package=ADE4)
574 [project.org/package=ADE4](http://CRAN.R-project.org/package=ADE4).
- 575 Corkum LD (1992) Spatial distribution patterns of macroinvertebrates along rivers within
576 and among biomes. *Hydrobiologia* 239: 101–114.
- 577 Cowley MJR, Thomas CD, Wilson RJ, Léon-Cortés JL, Gutiérrez D and Bulman CR.
578 (2001) Density-distribution relationships in British butterflies: II. An
579 assessment of mechanisms. *J. Anim. Ecol.* 70:426–441.
- 580 Crist EP and Cicone RC (1984) A physically-based transformation of Thematic Mapper
581 data -The TM Tasseled Cap. *IEEE Trans. Geosci. Remote Sens.* 3: 256–
582 263.
- 583 Danks HV (2004) Seasonal Adaptations in Arctic Insects. *Integr. Comp. Biol.* 44: 85–94.
- 584 Doledéc S, Chessel D and Gimaret-Carpentier C (2000) Niche separation in community
585 analysis: a new method. *Ecology* 81:2914–2927.
- 586 Gaston KJ and Blackburn TM (2000) Pattern and process in macroecology. Blackwell,
587 Oxford.

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

591 Gaston KJ, Blackburn TM, Greenwood JJD, Gregory RD, Quinn RM, and Lawton JH.
592 (2000) Abundance-occupancy relationships. *J. Appl. Ecol.* 37:39–59.

593 Gaston KJ, Blackburn TM and Lawton JH (1997) Interspecific abundance-range size
594 relationships: an appraisal of mechanisms. *J. Anim. Ecol.* 66:579–601.

595 Gregory RD and Gaston KJ (2000) Explanations for commonness and rarity in British
596 breeding birds: separating resource use and resource availability. *Oikos*
597 88:515–526.

598 Griffiths D (2010) Pattern and process in the distribution of North American freshwater
599 fish. *Biol. J. Linn. Soc.* 100: 46–61.

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

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

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

- 607 Heino J (2005) Positive relationship between regional distribution and local abundance
608 in stream insects: a consequence of niche breadth or niche position?
609 *Ecography* 28:345–354.
- 610 Heino J (2013) Environmental heterogeneity, dispersal mode and co-occurrence in stream
611 macroinvertebrates. *Ecol. Evol.* 3:344–355.
- 612 Heino J and Gönroos M (2014) Untangling the relationships among regional occupancy,
613 species traits, and niche characteristics in stream invertebrates. *Ecol. Evol.*
614 4:1931–1942.
- 615 Heino J, Ilmonen J and Paasivirta L (2014) Continuous variation of macroinvertebrate
616 communities along environmental gradients in northern streams. *Boreal*
617 *Environ. Res.* 19: 21–38.
- 618 Heino J and de Mendoza G (2016) Predictability of stream insect distributions is
619 dependent on niche position, but not on biological traits or taxonomic
620 relatedness of species. *Ecography* 39: 001–011.
- 621 Heino J and Soininen J (2006) Regional occupancy in unicellular eukaryotes: A reflection
622 of niche breadth, habitat availability or size-related dispersal capacity?
623 *Fresh. Biol.* 5: 672–685
- 624 Heino J and Virtanen R (2006) Relationships between distribution and abundance vary
625 with spatial scale and ecological group in stream bryophytes. *Freshw. Biol.*
626 51:1879–1889.

- 627 Hjort J and Luoto M (2006) Modelling patterned ground distribution in Finnish Lapland:
628 an integration of topographical, ground and remote sensing information.
629 Geogr. Ann. 88:19–29.
- 630 Hoffsten PO (2004) Site-occupancy in relation to flight morphology in caddisflies.
631 Freshw. Biol. 49: 810–817.
- 632 Holt AR and Gaston KJ (2003) Interspecific abundance-occupancy relationships of
633 British mammals and birds: is it possible to explain the residual variation?
634 Glob. Ecol. Biogeogr. 12:37–46.
- 635 Hurlbert AH and White EP (2007) Ecological correlates of geographical range occupancy
636 in North American birds. Glob. Ecol. and Biogeogr. 16: 764 –773.
- 637 Hutchinson GE (1957) Concluding remarks. Cold Spring Harbor Symposia on
638 Quantitative Biology 22: 145–159.
- 639 Hutchinson GE and MacArthur RH (1959) A theoretical ecological model of size
640 distributions among species of animals. Am. Nat. 93:117–125.
- 641 Hynes HBN (1970) The Ecology of Running Waters. University of Toronto Press,
642 Toronto.
- 643 Jenkins DG, Brescacin CR, Duxbury CV, Elliott JA, Evans JA, Grablow KR, Hillegass
644 M, Lyon BN, Metzger GA, Olandese ML, Pepe D, Silvers GA, Suresch
645 HN, Thompson TN, Trexler CM, Williams GE, Williams NC and Williams
646 SE (2007) Does size matter for dispersal distance? Glob. Ecol. Biogeogr.
647 16: 415–425.

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

651 Kelly M.G., Cazaubon A., Coring E., Dell’Uomo A., Ector L., Goldsmith B., Guasch
652 H., Hürlimann J., Jarlman A., Kawecka B., Kwandrans J., Laugaste R.,
653 Lindstrom E.A., Leitaom., Marvan P., Padisak J., Pipp E., Prygiel J., Rott E.,
654 Sabater S., Van Dam H. and Vizinet J., 1998. Recommendations for the routine
655 sampling of diatoms for water quality assessments in Europe. *J. Appl. Phycol.*,
656 10, 215–224.

657 Kärnä OM, Grönroos M, Antikainen H, Hjort J, Ilmonen J, Paasivirta L, Heino J (2015)
658 Inferring the effects of potential dispersal routes on the metacommunity
659 structure of stream insects: as the crow flies, as the fish swims or as the fox
660 runs? *J. Anim. Ecol.* 84: 1342–1353.

661 Luiz OJ, Madin JS, Robertson DR, Rocha LA, Wirtz P and Floeter SR (2012) Ecological
662 traits influencing range expansion across large oceanic dispersal barriers:
663 insights from tropical Atlantic reef fishes. *Proc. Roy. Soc. B* 279: 1033–
664 1040.

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

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

670 National Board of Waters and the Environment (1981)
671 Vesihallinnonanalyysimenetelmät. Publ. Nat. Board Wat. Finland Report
672 213:1–136.

673 National Land Survey of Finland (2010a) Basic Map 1: 20 000. PaITuli spatial data
674 service. In: <https://research.csc.fi/paituli>.

675 National Land Survey of Finland. (2010b) Topographic Database 1: 10 000. PaITuli
676 spatial data service. In: <https://research.csc.fi/paituli>.

677 National Land Survey of Finland (2000c) Digital Elevation Model. PaITuli spatial data
678 service. In: <https://research.csc.fi/paituli>.

679 Passy SI (2007) Differential cell size optimization strategies produce distinct diatom
680 richness–body size relationships in stream benthos and plankton. *J. Ecol.*
681 95:745–754.

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

683 Passy SI and Larson CA (2011) Succession in stream biofilms is an environmentally-
684 driven gradient of stress tolerance. *Microbial Ecol.* 62: 414–424.

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

687 Pringle CM (1990) Nutrient spatial heterogeneity: effects on community structure,
688 physiognomy, and diversity of stream algae. *Ecology* 71: 905–920.

689 R Core Team (2013) R: A Language and Environment for Statistical Computing. R
690 Foundation for Statistical Computing, Vienna. Available at: [http://www.R-](http://www.R-project.org/)
691 [project.org/](http://www.R-project.org/).

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

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

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

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

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

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

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

- 711 Snoeijls P, Busse S and Potapova M (2002) The importance of diatom size in community
712 analysis. *J. Phycol.* 38: 265–272.
- 713 Soininen J and Heino J (2005) Relationships between local population persistence, local
714 abundance and regional occupancy of species: patterns in diatoms of boreal
715 streams. *J. Biogeogr.* 32:1971–1978.
- 716 Soininen J, Bartels P, Heino J, Luoto M, Hillebrand H (2015) Toward more integrated
717 ecosystem research in aquatic and terrestrial environments. *BioScience*,
718 216: 1–9
- 719 Tachet, H., Richoux, P., Bourneau, M. and Usseglio-Polatera, P. (2010) Freshwater
720 Invertebrates, Systematics, Biology, Ecology. CNRS Editions, Paris.
- 721 Tales E, Keith P and Oberdorff T (2004) Density-range size relationship in French
722 riverine fishes. *Oecologia* 138:360–370.
- 723 Tolonen KE, Leinonen K, Marttila H, Erkinaro J and Heino J (2017) Environmental
724 predictability of taxonomic and functional community composition in high-
725 latitude streams. *Freshw. Biol.* 62: 1–16.
- 726 Tonkin JD, Arimoro FO and Haase P (2016) Exploring stream communities in a tropical
727 biodiversity hotspot: biodiversity, regional occupancy, niche characteristics
728 and environmental correlates. *Biodivers. Conserv.* 25:975–993
- 729 Venier LA and Fahrig L (1996) Habitat availability causes the species distribution-
730 abundance relationship. *Oikos* 76:564–570.

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

733 Wentworth CK (1992) A scale of grade and class terms for clastic sediments. *J. Geology*
734 30: 377–392.

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

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

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

747

748 **Figure captions**

749

750 **Fig. 1** Expected relationships between niche position (NP), niche breadth (NB) and
751 body size (BS) in predicting occupancy and abundance (A and B) of stream diatoms and
752 insects. The scenario B is predicted to be the same for both taxonomic groups. Flat line
753 = insects; dashed line = diatoms.

754

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

758

759

760

761

762 **Tables and Figures**

763 **Table 1** Regression results accounting for variation in the occupancy and abundance of
 764 diatoms and insects. Generalized liner models (assuming a beta distribution) were used
 765 to model occupancy and GLMs (assuming a negative binomial distribution) were used
 766 to model local mean abundance. Only successive pairs of AIC values are comparable.
 767 Legend: ER = expected relationship; Found = significant results found; ns = non-
 768 significant; neg = negative; pos = positive.

769

Group	Response	Model	Explanatories	Coeff.	SE	Z	P	AIC	pseudo R ²	ER	Found		
Diatoms	Occupancy	Local	(Intercept)	1.04	0.28	3.76	0.000	-185.92	0.74				
			NP-env	-2.13	0.12	-18.47	0.000	neg		yes			
			NB-env	0.58	0.20	2.87	0.004	pos		yes			
			Body size	-0.11	0.05	-2.22	0.026	neg		yes			
		Catchment	(Intercept)	0.49	0.24	2.05	0.040	-118.67	0.57				
			NP-cat	-2.24	0.18	-12.42	0.000	neg		yes			
			NB-cat	0.85	0.17	4.95	0.000	pos		yes			
			Body size	-0.22	0.07	-3.41	0.001	neg		yes			
		Diatoms	Abundance	Local	(Intercept)	3.30	0.36	9.17	0.000	597.98	0.43		
					NP-env	-0.83	0.14	-6.14	0.000	-		neg	
					NB-env	0.04	0.26	0.17	0.867	pos		ns	
					Body size	-0.19	0.07	-2.65	0.008	neg		yes	
Catchment	(Intercept)			2.83	0.27	10.68	0.000	606.96	0.37				
	NP-cat			-0.99	0.18	-5.43	0.000	-		neg			
	NB-cat			0.55	0.19	2.92	0.004	pos		yes			
	Body size			-0.26	0.07	-3.50	0.000	neg		yes			
Insects	Occupancy			Local	(Intercept)	-0.02	0.43	-0.04	0.967	-104.97	0.73		
					NP-env	-1.71	0.14	-12.65	0.000	neg		yes	
					NB-env	0.77	0.25	3.05	0.002	pos		yes	
					Body size	0.13	0.08	1.63	0.103	pos		ns	
		Catchment	(Intercept)	-0.40	0.34	-1.17	0.242	-78.62	0.60				
			NP-cat	-2.24	0.21	-10.82	0.000	neg		yes			
			NB-cat	1.24	0.25	4.90	0.000	pos		yes			
			Body size	0.08	0.09	0.92	0.360	pos		ns			
		Insects	Abundance	Local	(Intercept)	2.94	0.65	4.51	0.000	534.60	0.31		
					NP-env	-0.82	0.18	-4.59	0.000	-		neg	
					NB-env	0.92	0.37	2.53	0.011	pos		yes	

	Body size	-0.15	0.12	-1.26	0.200			neg	ns
Catchment	(Intercept)	3.36	0.48	7.06	0.000	547.77	0.18		
	NP-cat	-0.88	0.25	-3.45	0.001			-	neg
	NB-cat	0.42	0.34	1.22	0.222			pos	ns
	Body size	-0.16	0.13	-1.27	0.203			neg	ns

771

772

773

774 **Fig. 1**



